aruba®
NETWORKS

Implementing ArubaOS-CX Switching Volume 1

Aruba Training-Confidential

TRAINING MANUAL

EDU-ICX-ILT-v20.211

Aruba Education Services

Implementing ArubaOS-CX Switching Volume 1

April 2020

Aruba Training-Confidential

Implememting ArubaOS-CX Switching Student Guide

Copyright
© 2020 Aruba Networks, Inc. AirWave®, Aruba Networks®, Aruba Mobility Management System®, Bluescanner, For Wireless That Works®, Mobile Edge Architecture, People Move. Networks Must Follow., RFProtect, The All Wireless Workplace Is Now Open For Business, and The Mobile Edge Company® are trademarks of Aruba Networks, Inc. All rights reserved. All other trademarks are the property of their respective owners.

Open Source Code
Certain Aruba products include Open Source software code developed by third parties, including software code subject to the GNU General Public License ("GPL"), GNU Lesser General Public License ("LGPL"), or other Open Source Licenses. The Open Source code used can be found at this site:
http://www.arubanetworks.com/open_source

Legal Notice
The use of Aruba Networks, Inc. switching platforms and software, by all individuals or corporations, to terminate other vendors' VPN client devices constitutes complete acceptance of liability by that individual or corporation for this action and indemnifies, in full, Aruba Networks, Inc. from any and all legal actions that might be taken against it with respect to infringement of copyright on behalf of those vendors.

Warranty
This hardware product is protected by the standard Aruba warranty of one year parts/labor. For more information, refer to the ARUBACARE SERVICE AND SUPPORT TERMS AND CONDITIONS.
Altering this device (such as painting it) voids the warranty.

Aruba Training Confidential

SKU: EDU-ICX-ILT-
v20.21 April 2020

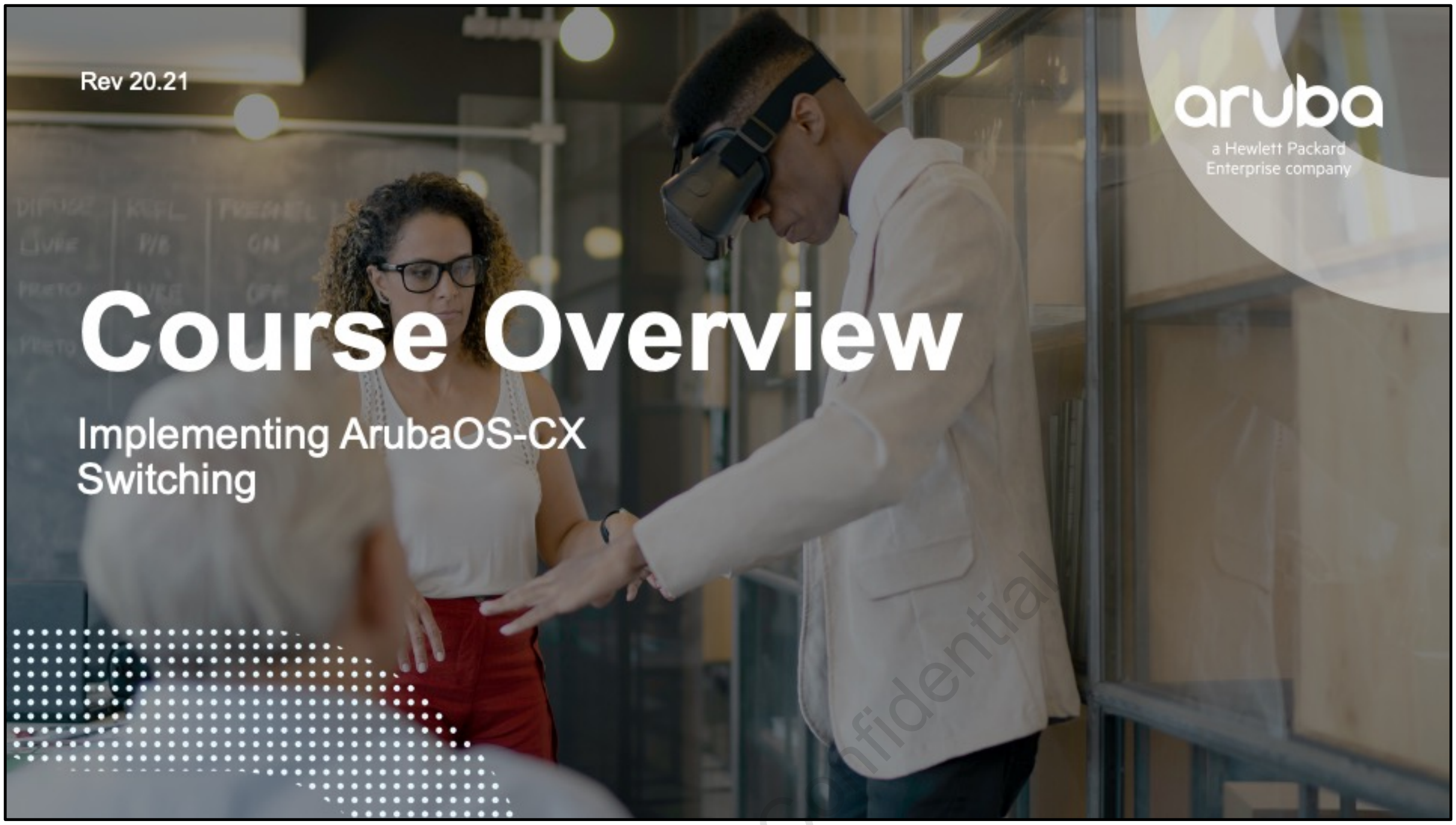
Implementing ArubaOS-CX Switching Volume 1

Table of Contents

Aruba Training-Confidential

Aruba Training-Confidential

Aruba Training-Confidential



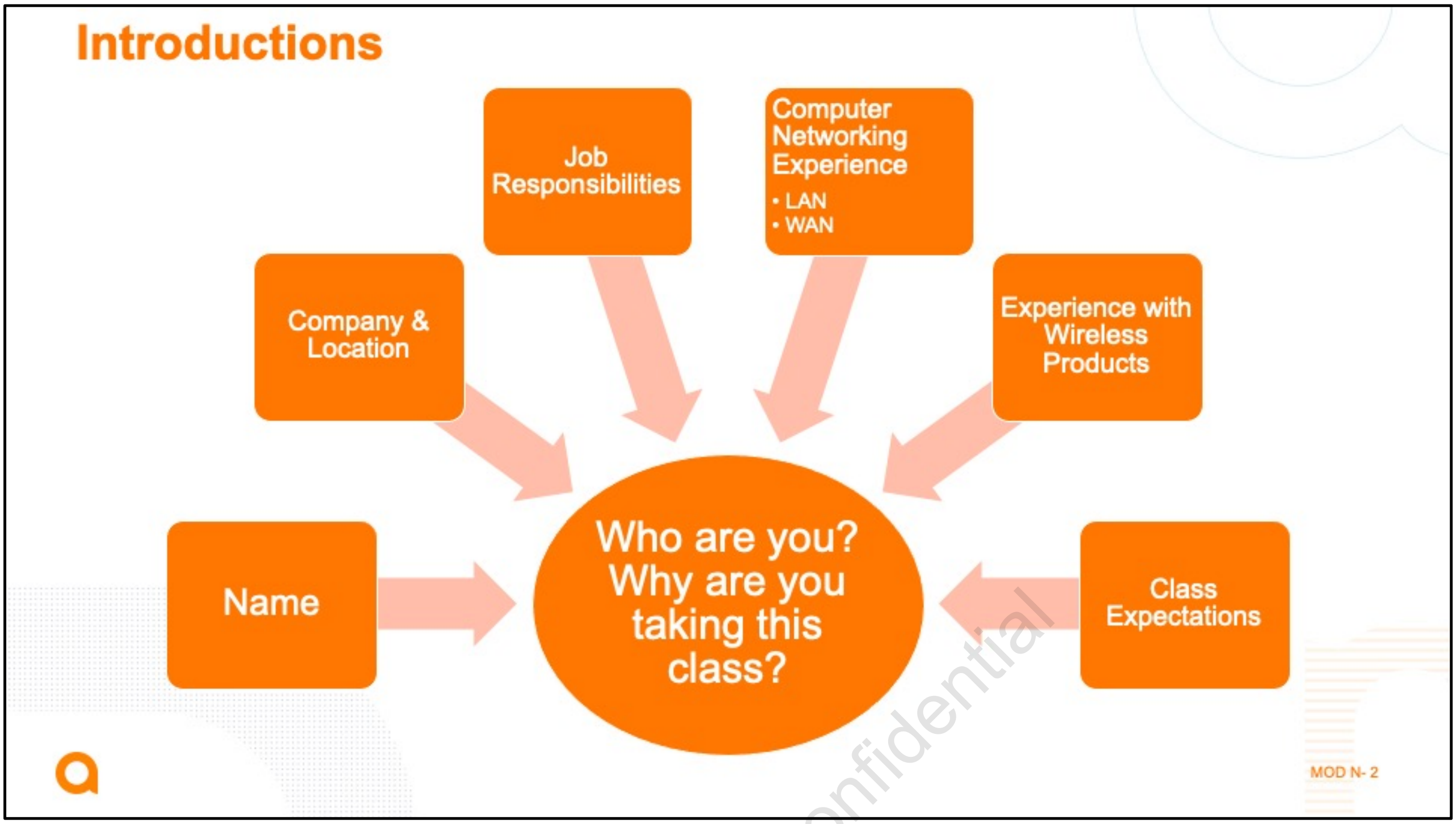
Rev 20.21

Course Overview

Implementing ArubaOS-CX Switching

Hello, and welcome to the course! Implementing ArubaOS-CX Switching.

Aruba Training-Confidential



Introductions

Job Responsibilities

Computer Networking Experience
- LAN
- WAN

Company & Location

Experience with Wireless Products

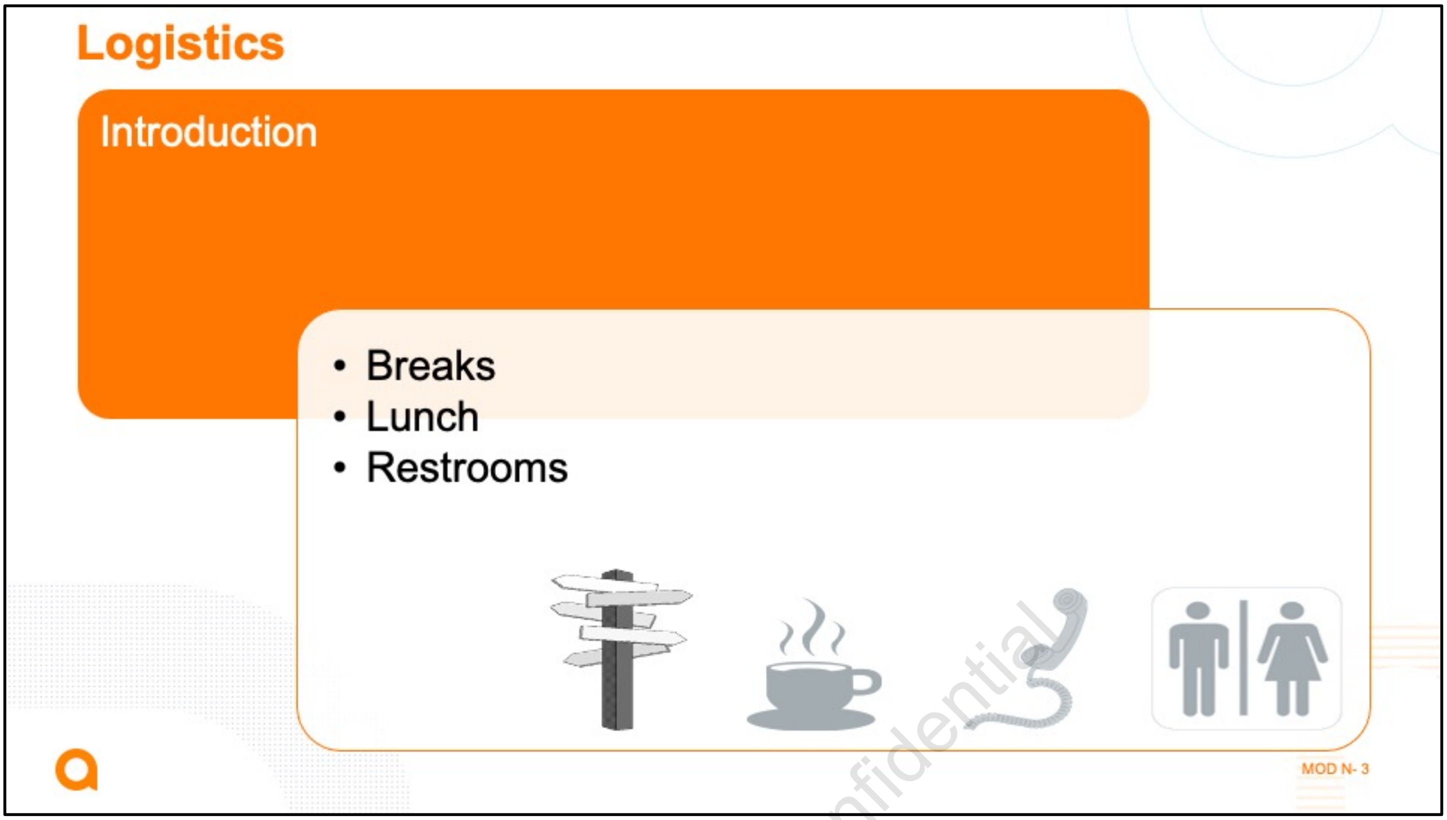
Name

Who are you? Why are you taking this class?

Class Expectations

MOD N- 2

Aruba Training-Confidential



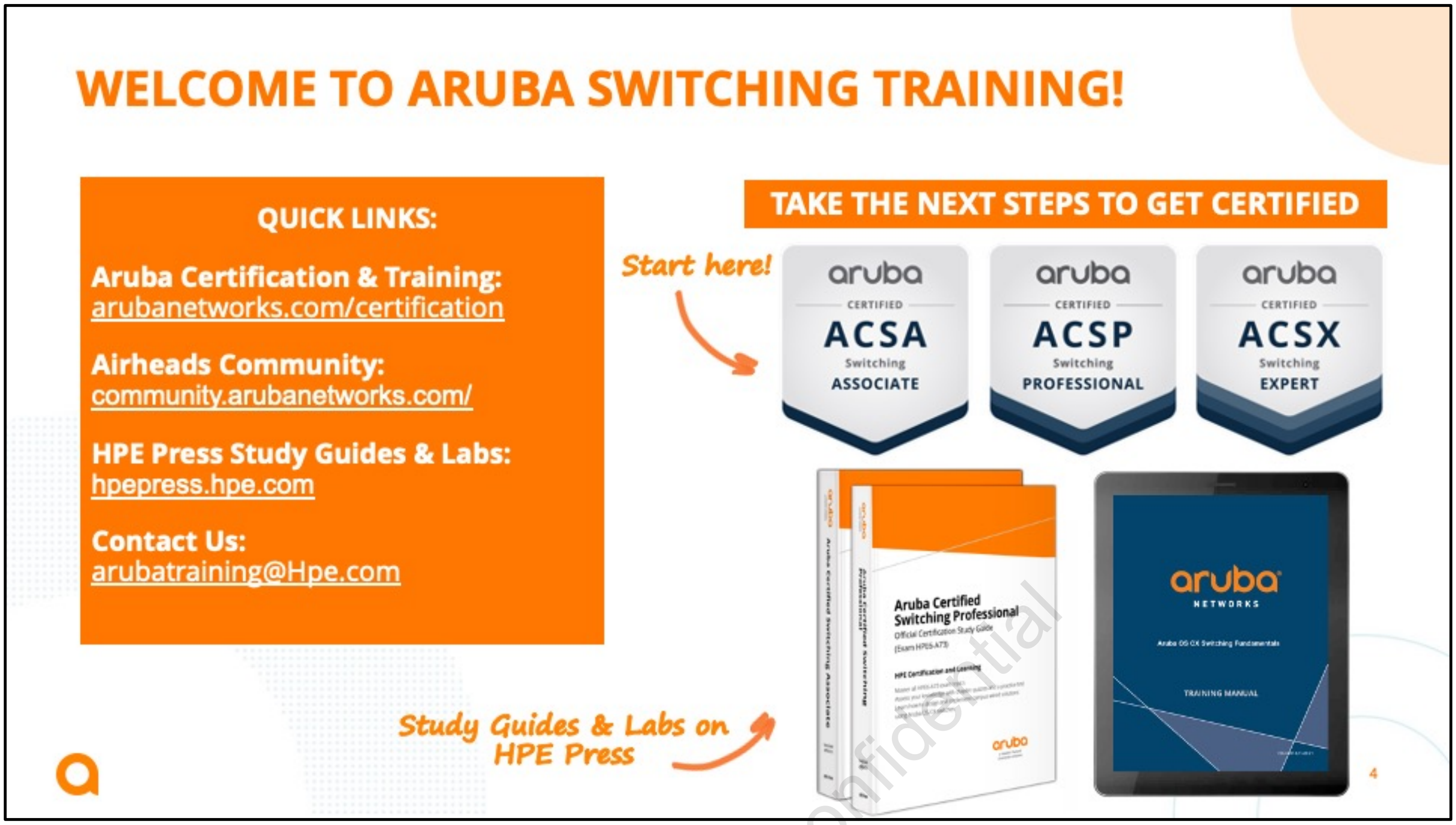
Logistics

Introduction

- Breaks
- Lunch
- Restrooms

MOD N- 3

Aruba Training-Confidential



WELCOME TO ARUBA SWITCHING TRAINING!

QUICK LINKS:

Aruba Certification & Training:
arubanetworks.com/certification

Airheads Community:
community.arubanetworks.com/

HPE Press Study Guides & Labs:
hpepress.hpe.com

Contact Us:
arubatraining@Hpe.com

TAKE THE NEXT STEPS TO GET CERTIFIED

Start here!

Study Guides & Labs on HPE Press

aruba CERTIFIED ACSA Switching ASSOCIATE
aruba CERTIFIED ACSP Switching PROFESSIONAL
aruba CERTIFIED ACSX Switching EXPERT
Aruba Certified Switching Associate
Aruba Certified Switching Professional
Aruba Certified Switching Professional Official Certification Study Guide (Exam HPE6-A73)
aruba NETWORKS
Aruba OS CX Switching Fundamentals
TRAINING MANUAL
4

Aruba Training-Confidential

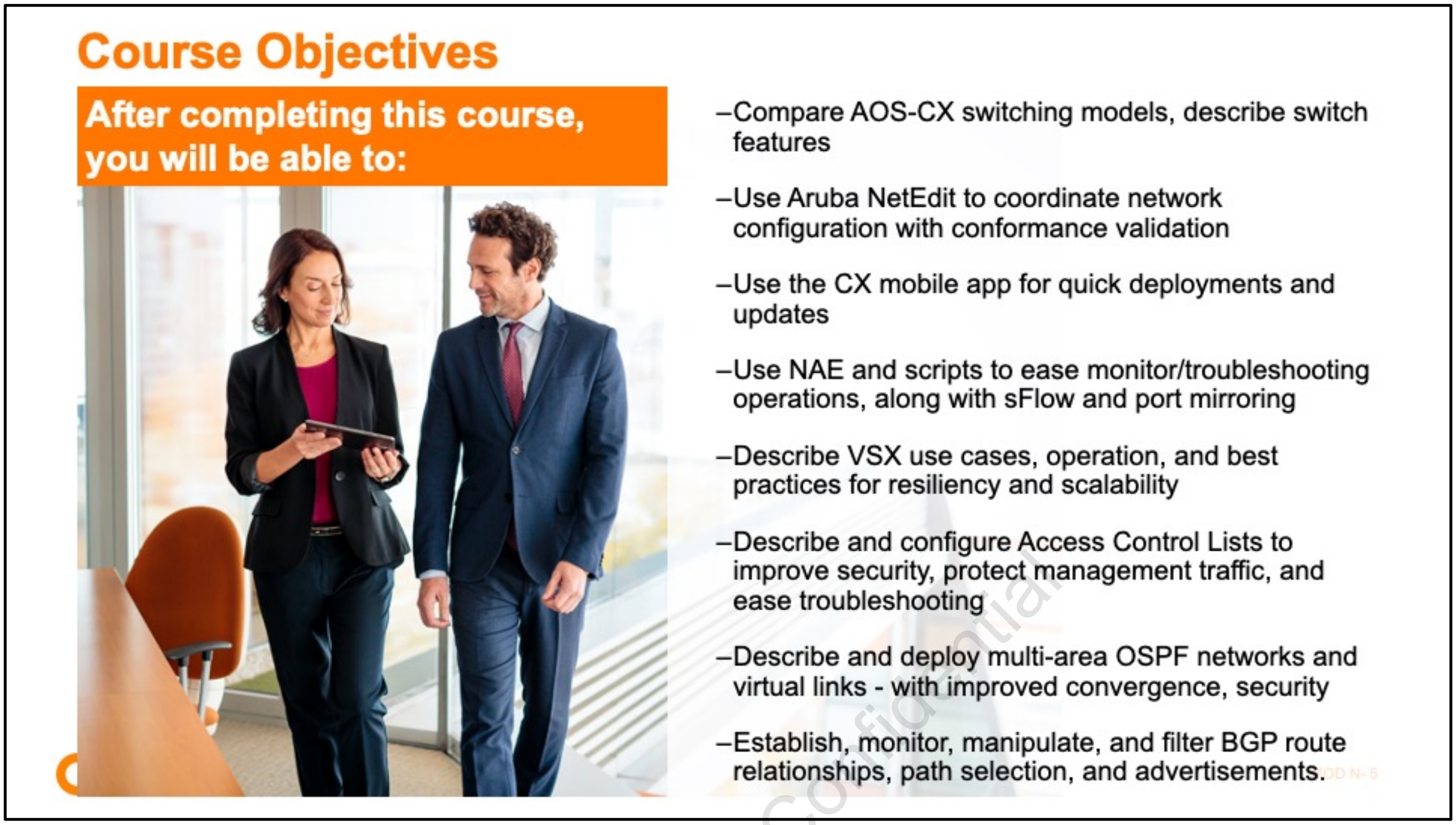
Thanks for choosing to learn about Aruba's Switching solutions! We hope you will take the next step to get certified with the appropriate Aruba Certified Switching course.

To learn more about our certification & training program visit us online at arubanetworks.com/certification

We encourage you to join the Airheads Community, a vibrant group of networking professionals where you can participate in discussions and watch free learning videos on the Airheads Learning Community.

To reinforce what you learn in this course visit HPE Press for study guides and self-directed labs.

If you have any questions or concerns we welcome you to contact our support team at arubatraining@Hpe.com



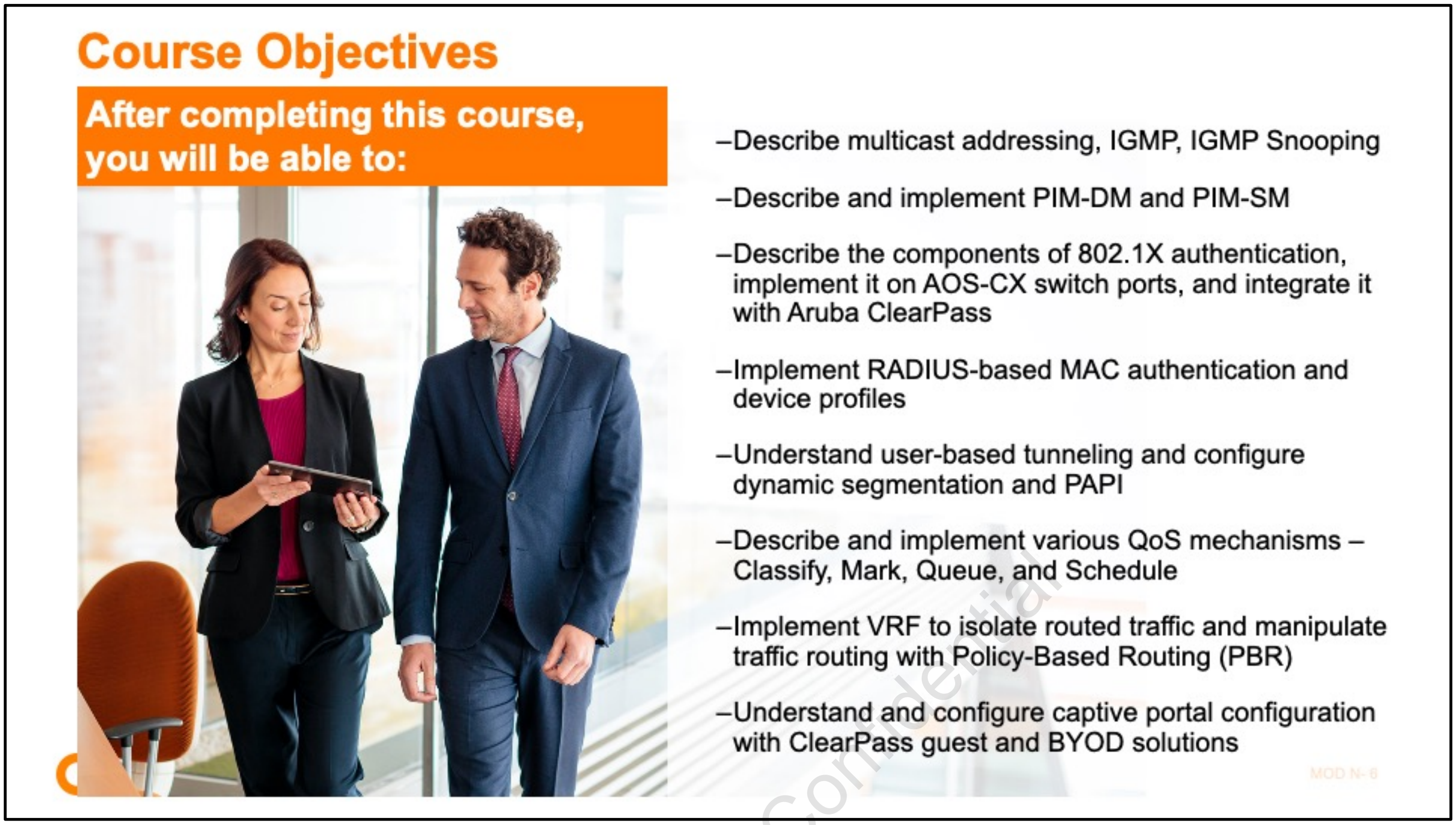
Course Objectives

After completing this course, you will be able to:

- –Compare AOS-CX switching models, describe switch features
- –Use Aruba NetEdit to coordinate network configuration with conformance validation
- –Use the CX mobile app for quick deployments and updates
- –Use NAE and scripts to ease monitor/troubleshooting operations, along with sFlow and port mirroring
- –Describe VSX use cases, operation, and best practices for resiliency and scalability
- –Describe and configure Access Control Lists to improve security, protect management traffic, and ease troubleshooting
- –Describe and deploy multi-area OSPF networks and virtual links - with improved convergence, security
- –Establish, monitor, manipulate, and filter BGP route relationships, path selection, and advertisements.

Here's the first half of the course objectives.

Aruba Training-Confidential



Course Objectives

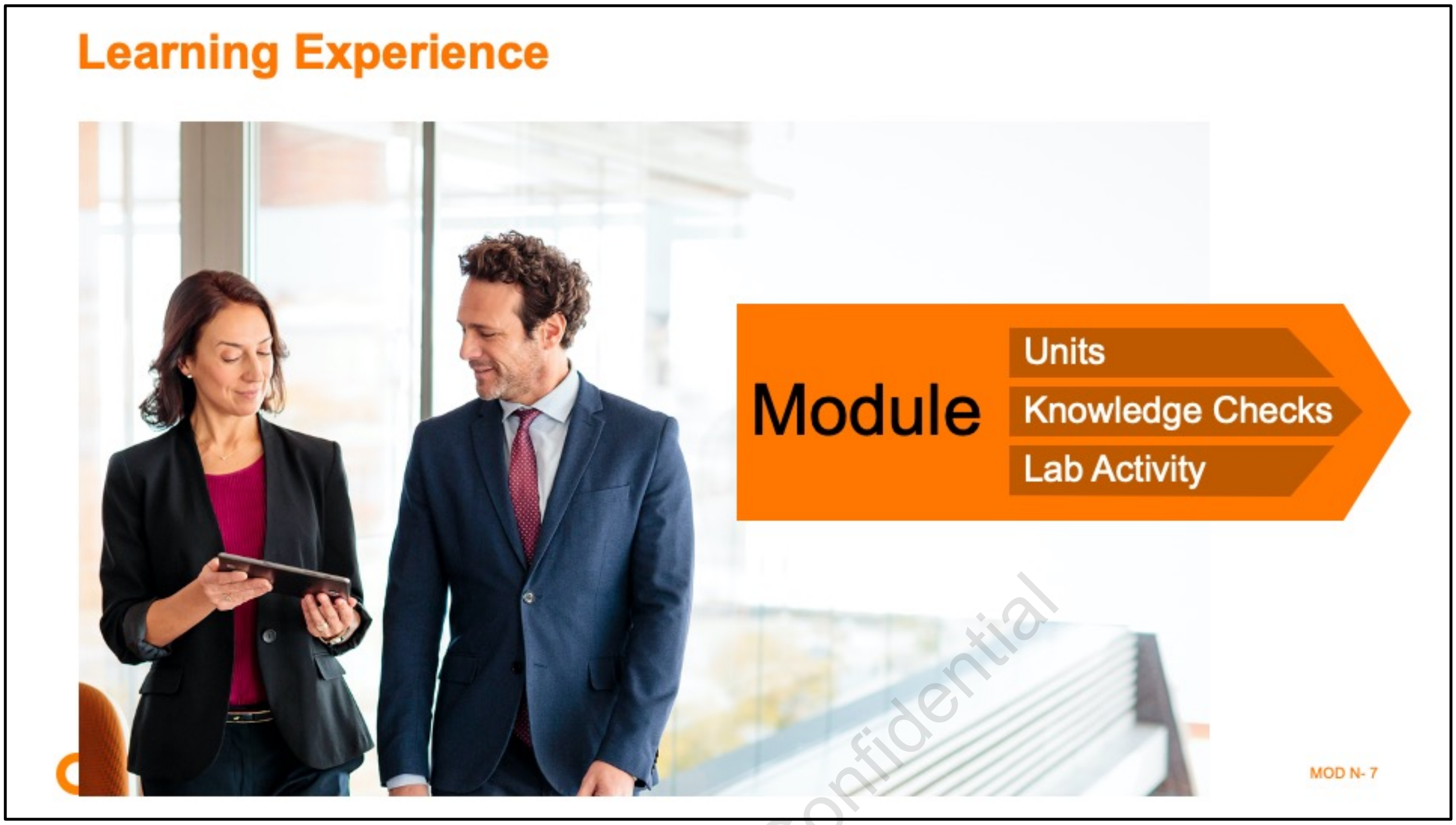
After completing this course, you will be able to:

- –Describe multicast addressing, IGMP, IGMP Snooping
- –Describe and implement PIM-DM and PIM-SM
- –Describe the components of 802.1X authentication, implement it on AOS-CX switch ports, and integrate it with Aruba ClearPass
- –Implement RADIUS-based MAC authentication and device profiles
- –Understand user-based tunneling and configure dynamic segmentation and PAPI
- –Describe and implement various QoS mechanisms – Classify, Mark, Queue, and Schedule
- –Implement VRF to isolate routed traffic and manipulate traffic routing with Policy-Based Routing (PBR)
- –Understand and configure captive portal configuration with ClearPass guest and BYOD solutions

MOD N- 6

And here’s the second half of the objectives.

Aruba Training-Confidential



Learning Experience

Module

Units

Knowledge Checks

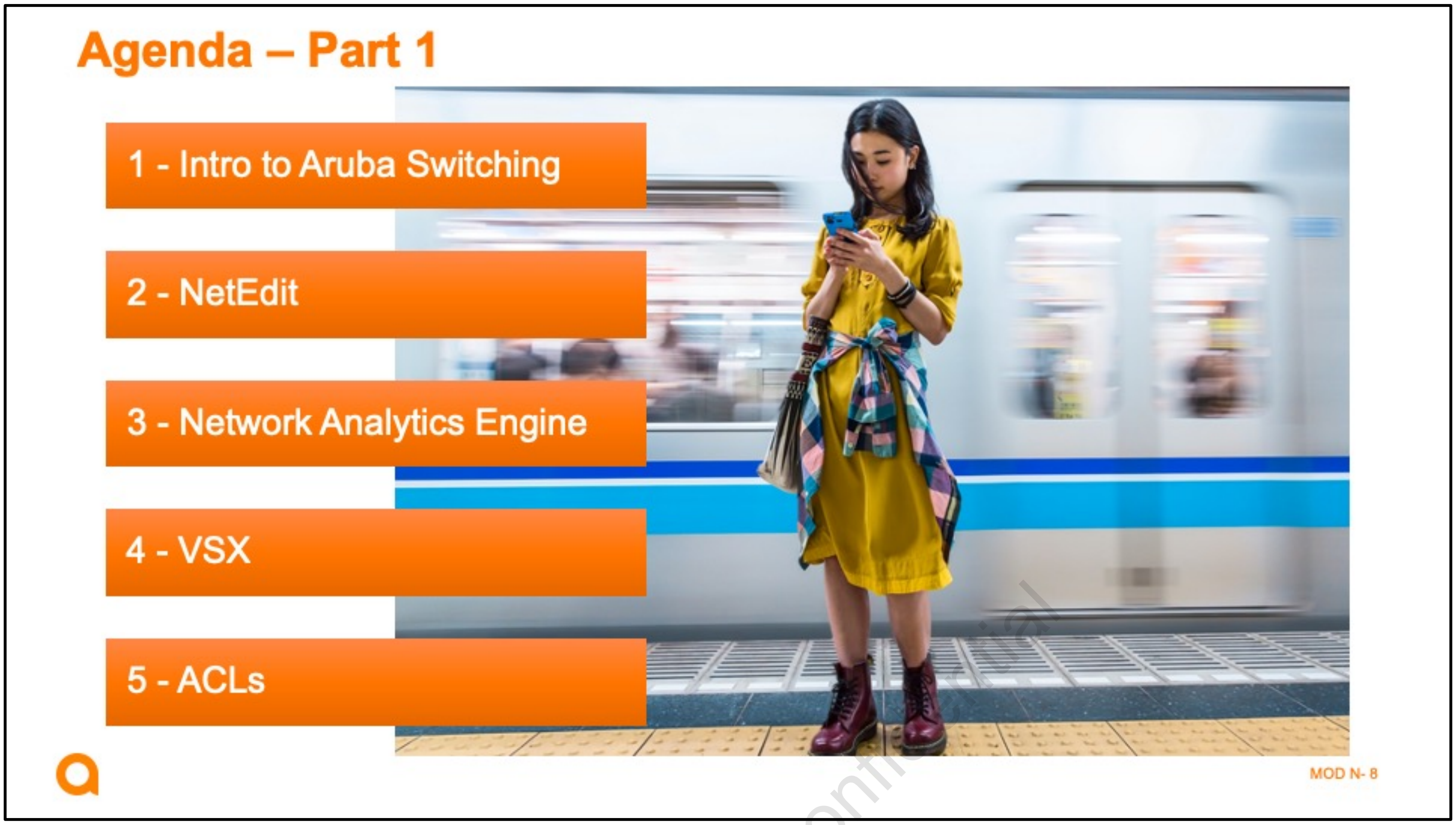
Lab Activity

MOD N- 7

The course is organized into modules, and each module is composed of presentation units and knowledge checks designed to provide essential foundational knowledge required.

Many modules include a Lab Activity - opportunity to go hands-on in a virtual application environment. What about the specific modules in this course? Let’s see what’s in store.

Aruba Training-Confidential



Agenda – Part 1

1 - Intro to Aruba Switching

2 - NetEdit

3 - Network Analytics Engine

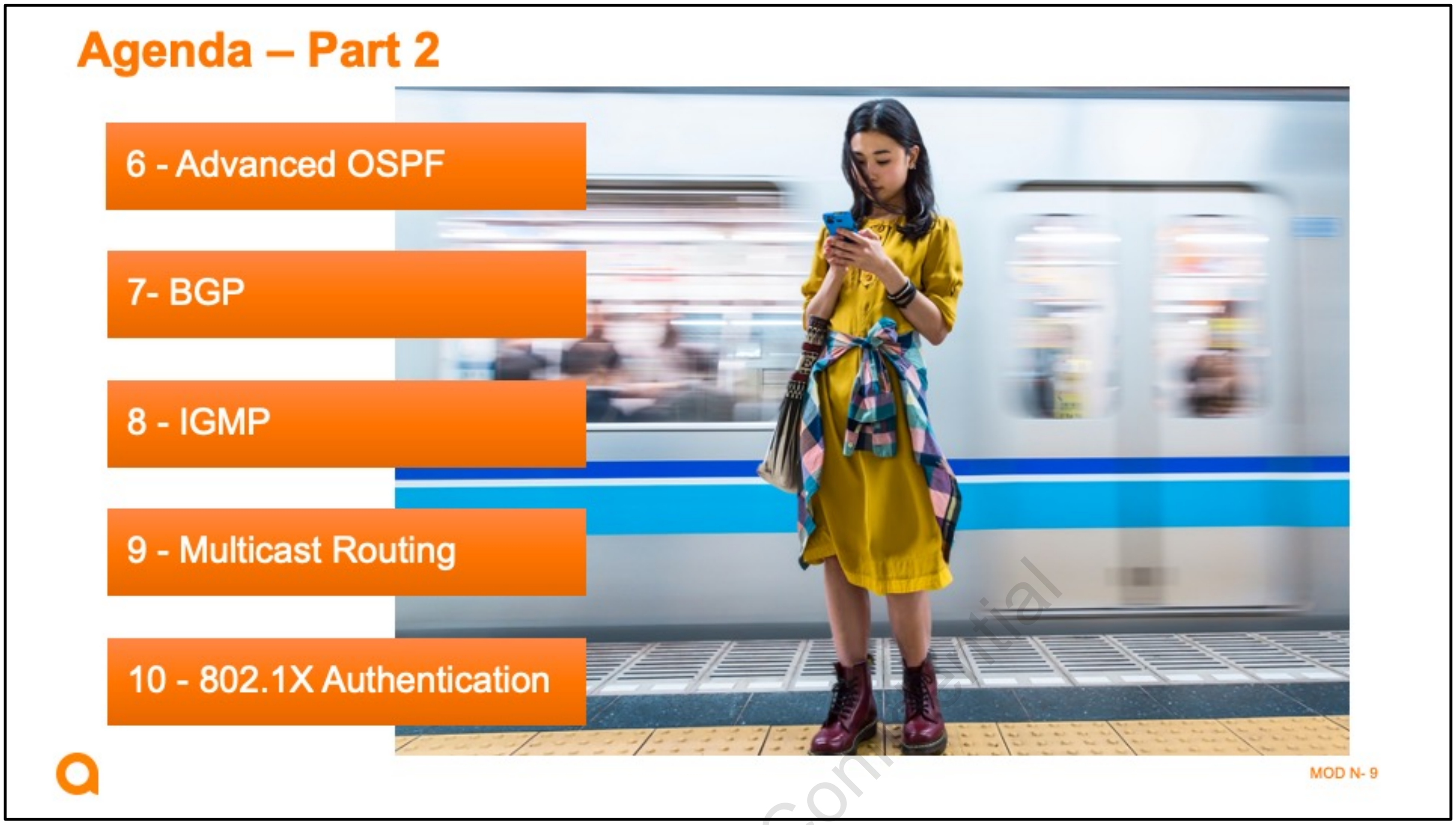
4 - VSX

5 - ACLs

MOD N- 8

Here's the first set of modules you'll explore.

Aruba Training-Confidential



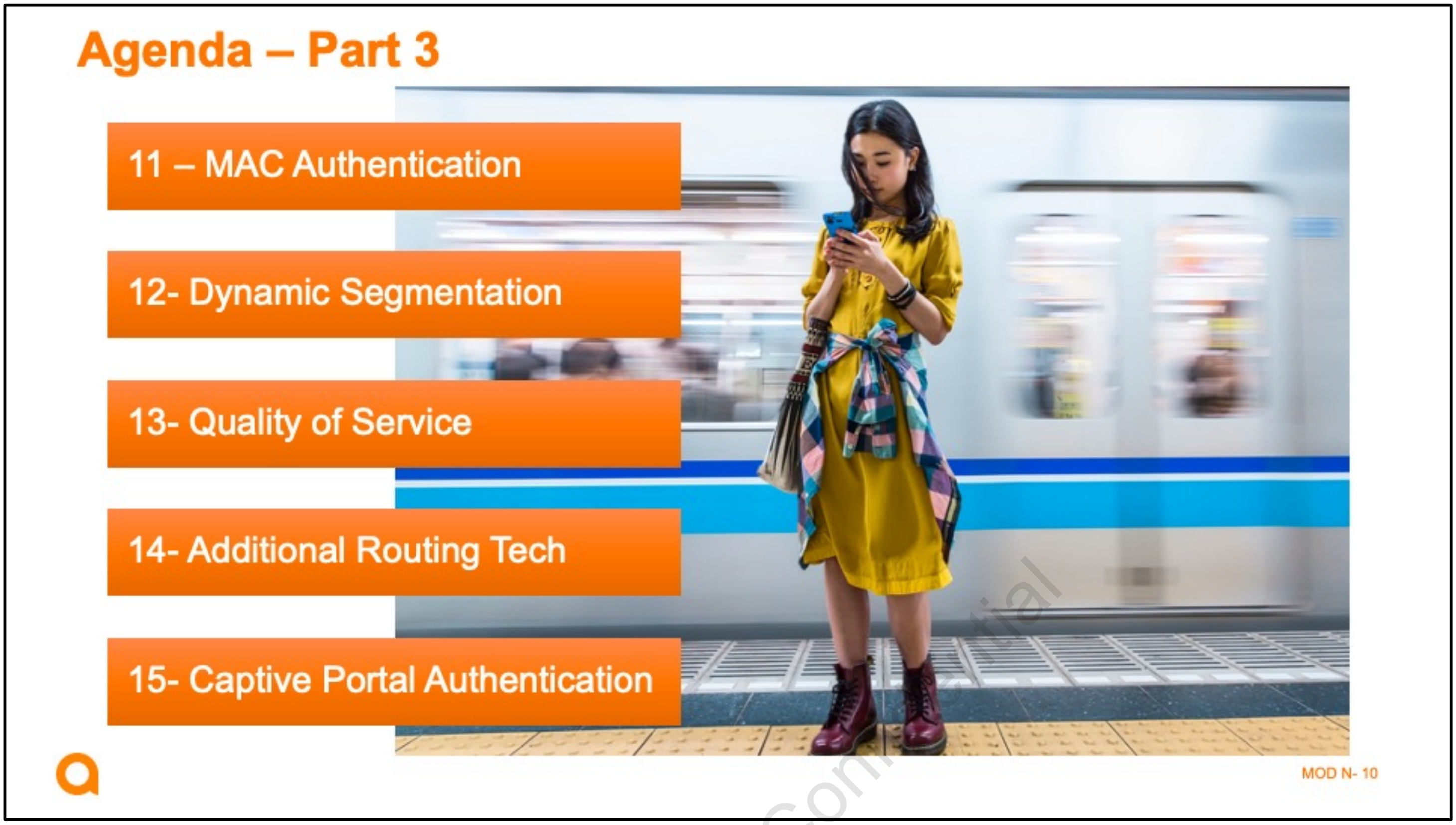
Agenda – Part 2

- 6 - Advanced OSPF
- 7- BGP
- 8 - IGMP
- 9 - Multicast Routing
- 10 - 802.1X Authentication

MOD N- 9

And here are the next set of modules.

Aruba Training-Confidential



Agenda – Part 3

- 11 – MAC Authentication
- 12- Dynamic Segmentation
- 13- Quality of Service
- 14- Additional Routing Tech
- 15- Captive Portal Authentication

MOD N- 10

And here are the final modules for this course.

Aruba Training-Confidential

Resources

- Validated Reference Design (VRD) Guides
  - https://community.arubanetworks.com/t5/Validated-Reference-Design/tkb-p/Aruba-VRDs
- Aruba Support
  - http://support.arubanetworks.com/
- Airheads
  - http://community.arubanetworks.com/
- Aruba Solution Exchange
  - http://ase.arubanetworks.com

MOD N- 11

You may find these resources quite valuable – especially if you intend to expand your knowledge and participate in the active Aruba community. Click on the Resources tab at the top of the course here at any time, for links to these and other resources.

Aruba Training-Confidential

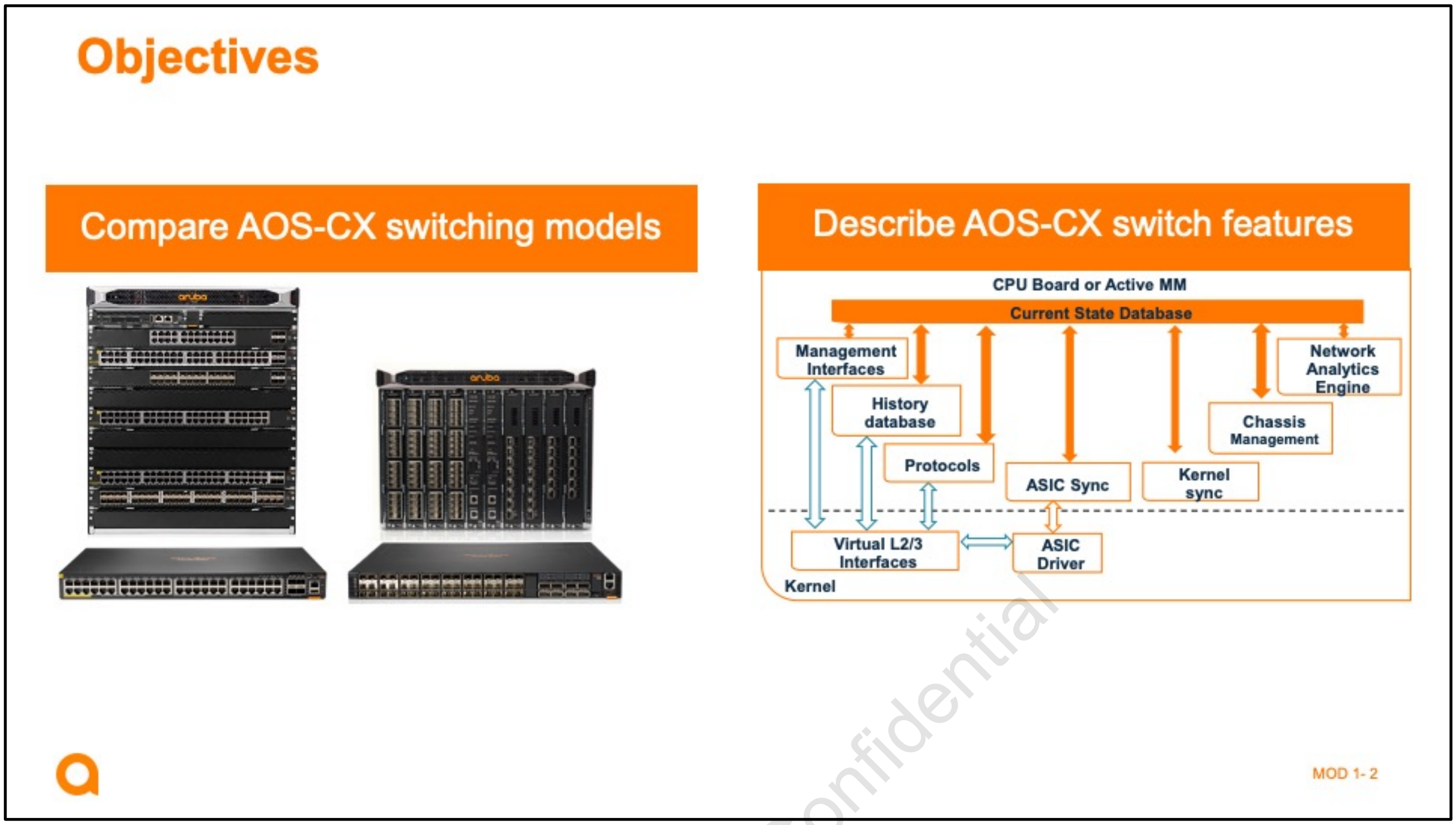
Rev # 20.21

Introduction to Aruba Switching

Implementing ArubaOS-CX Switching

Hello everyone, and welcome to the course! Implementing ArubaOS-CX Switching. This is Module 1.

Aruba Training-Co



Objectives

Compare AOS-CX switching models

Describe AOS-CX switch features

CPU Board or Active MM

Current State Database

Management Interfaces

History database

Protocols

ASIC Sync

Kernel sync

Chassis Management

Network Analytics Engine

Virtual L2/3 Interfaces

ASIC Driver

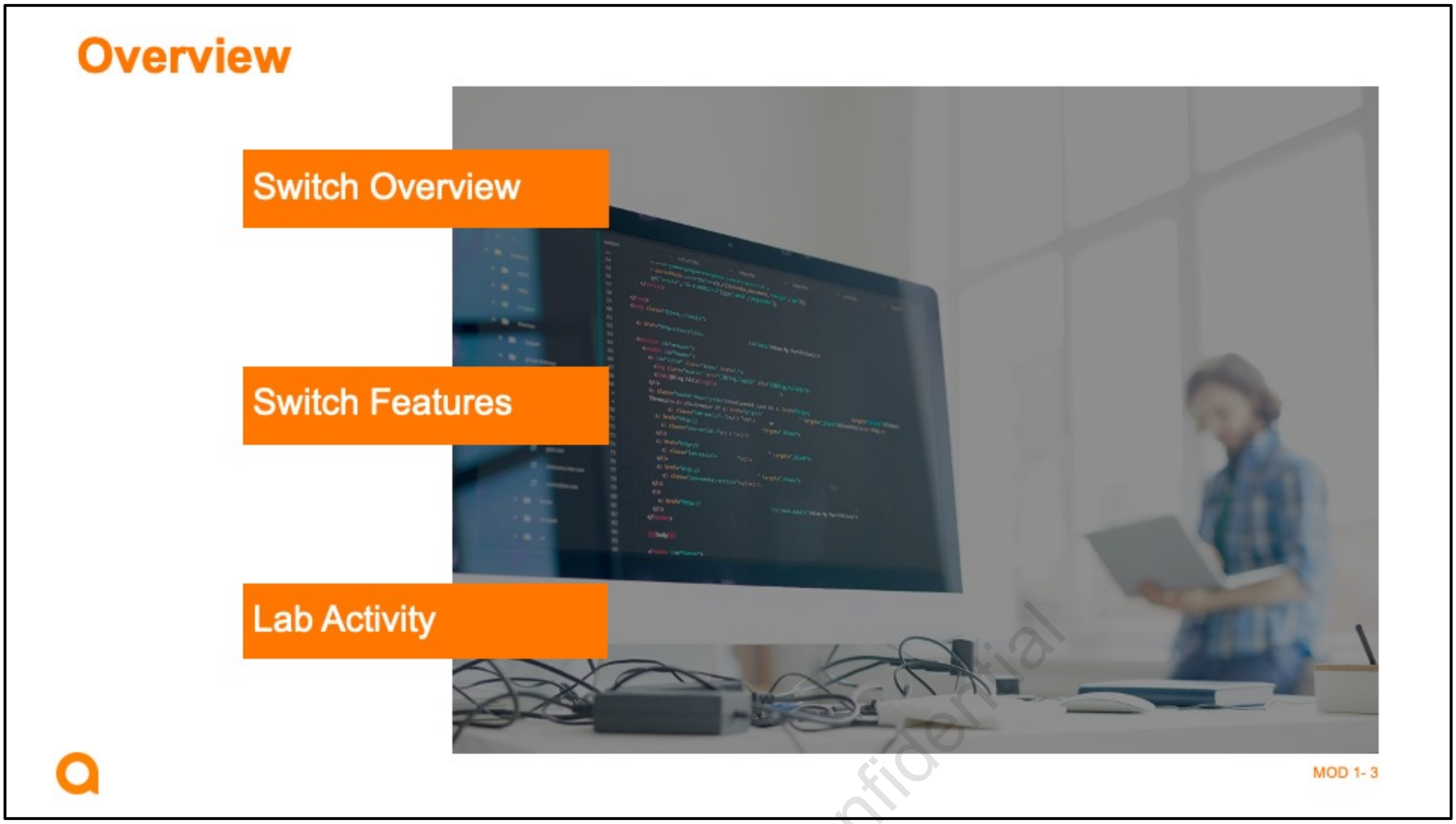
Kernel

MOD 1- 2

When you complete this module, you should be able to:

- Compare and contrast the different AOS-CX switching models
- Describe the important features of the AOS-CX switching product line

Aruba Training-Confidential



Overview

Switch Overview

Switch Features

Lab Activity

MOD 1- 3

Aruba Training-Confidential

This module introduces key AOS-CX products and features, and a lab activity

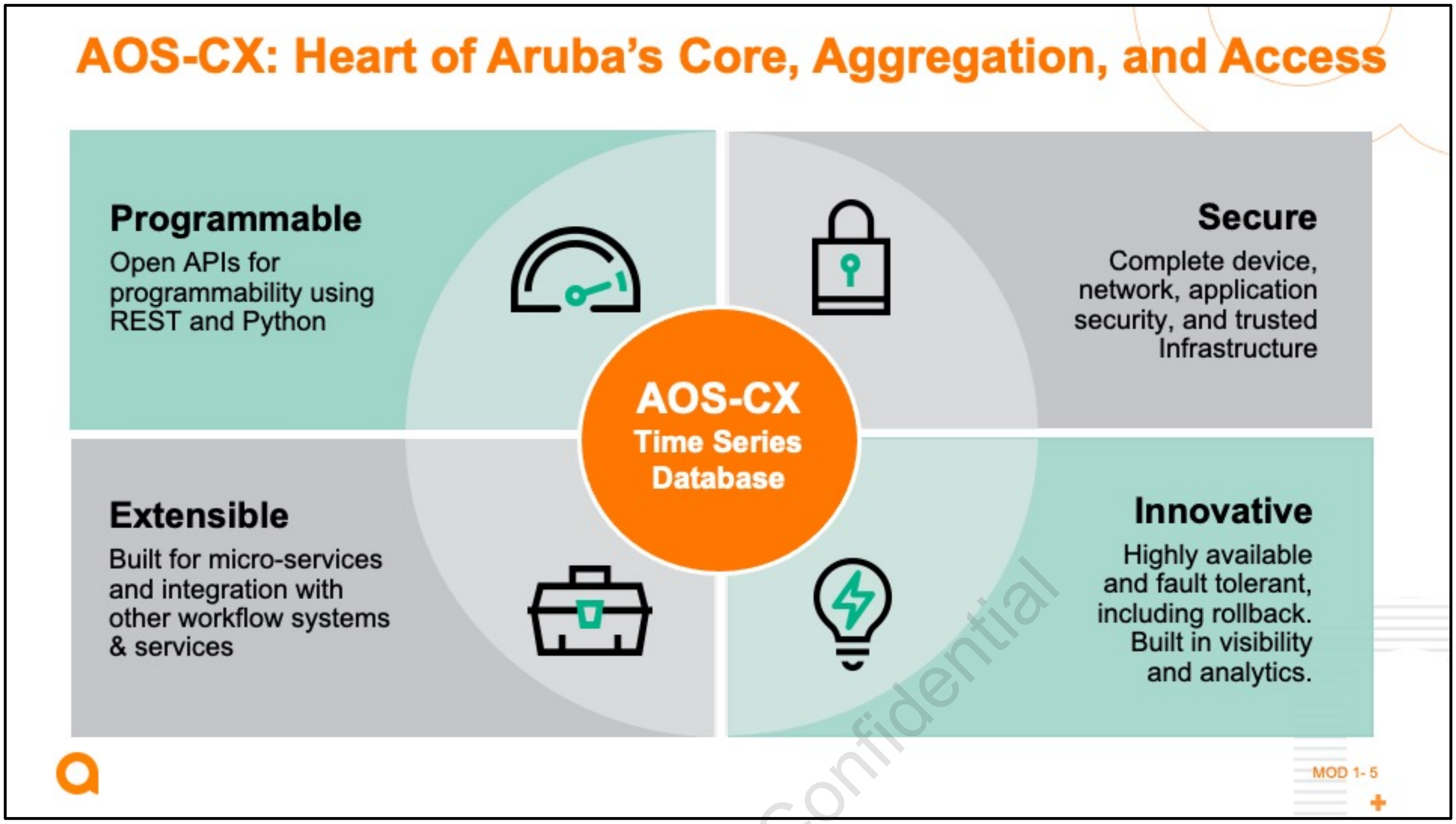
Note: Data center designs and features are covered in other Aruba-related courses. This course focuses on campus solutions.

AOS-CX Switch Overview

aruba
a Hewlett Packard Enterprise company

Let’s explore the AOS-CX switches.

Aruba Training-Confidential



AOS-CX: Heart of Aruba's Core, Aggregation, and Access

Programmable
Open APIs for programmability using REST and Python

Secure
Complete device, network, application security, and trusted Infrastructure

AOS-CX
Time Series Database

Extensible
Built for micro-services and integration with other workflow systems & services

Innovative
Highly available and fault tolerant, including rollback. Built in visibility and analytics.

MOD 1- 5

Aruba Training-Confidential

The AOS-CX software forms the heart of Aruba's strategy for campus core and aggregation switching.

| Programmable

AOS-CX is based on open, REST-API features that enable external applications to securely view and configure all system components. Developers can use the familiar Python language to script REST API interaction, and to create NAE agents for system monitoring and troubleshooting.

| Extensible

AOS-CX is extensible - built to support micro-services and integrate with other workflow systems and services.

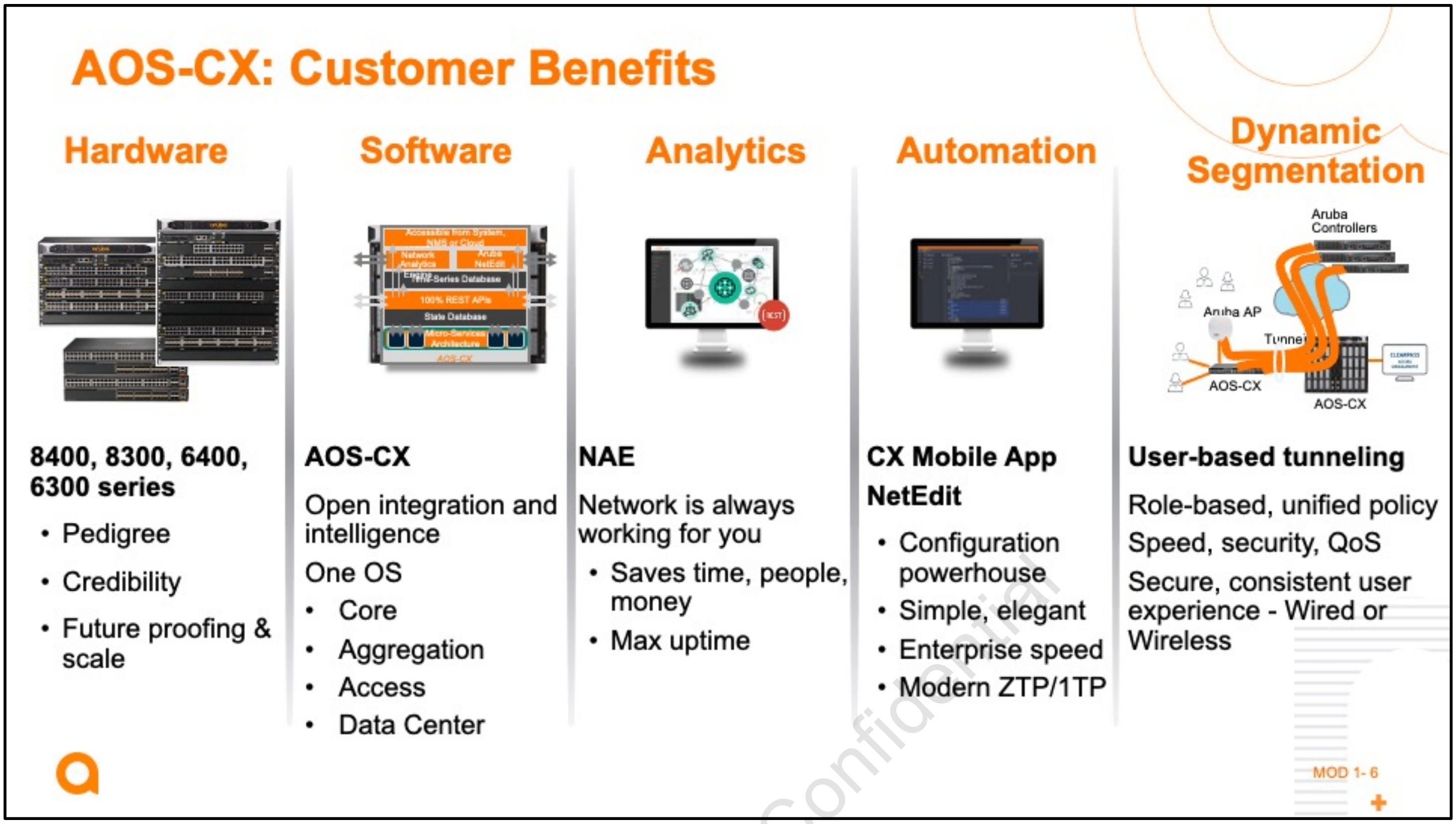
| Innovative

AOS-CX takes an innovative approach to high availability and fault-tolerance. The software architecture supports resiliency; for example, a process that fails can roll back to a good state using the stateful database. AOS-CX has even been honored by the industry for its innovation. CRN awarded the 8400 switch Product of the Year 2017 in large part due to NAE. NAE takes the guesswork out of troubleshooting and optimization and will be a game-changer for many customers.

| Secure

Finally, the AOS-CX is secure and forms the bedrock for a trusted infrastructure. In addition to standard security features such as access control lists (ACLs), the AOS-CX supports a trusted boot process. The software will not boot unless every hardware component checks out as authentic and free from malware and rootkits.

Note: Configuration version roll back may cause issues when going between major AOS-CX software versions.


Accessible from System, NMS or Cloud
Network Analytics Engine
Aruba NetEdit
Time-Series Database
100% REST APIs
State Database
Micro-Services Architecture
AOS-CX
REST
Aruba Controllers
Aruba AP
Tunnel
AOS-CX
AOS-CX

AOS-CX: Customer Benefits

Hardware

8400, 8300, 6400, 6300 series

- Pedigree
- Credibility
- Future proofing & scale

Software

AOS-CX

Open integration and intelligence

One OS

- Core
- Aggregation
- Access
- Data Center

Analytics

NAE

Network is always working for you

- Saves time, people, money
- Max uptime

Automation

CX Mobile App

NetEdit

- Configuration powerhouse
- Simple, elegant
- Enterprise speed
- Modern ZTP/1TP

Dynamic Segmentation

User-based tunneling

Role-based, unified policy

Speed, security, QoS

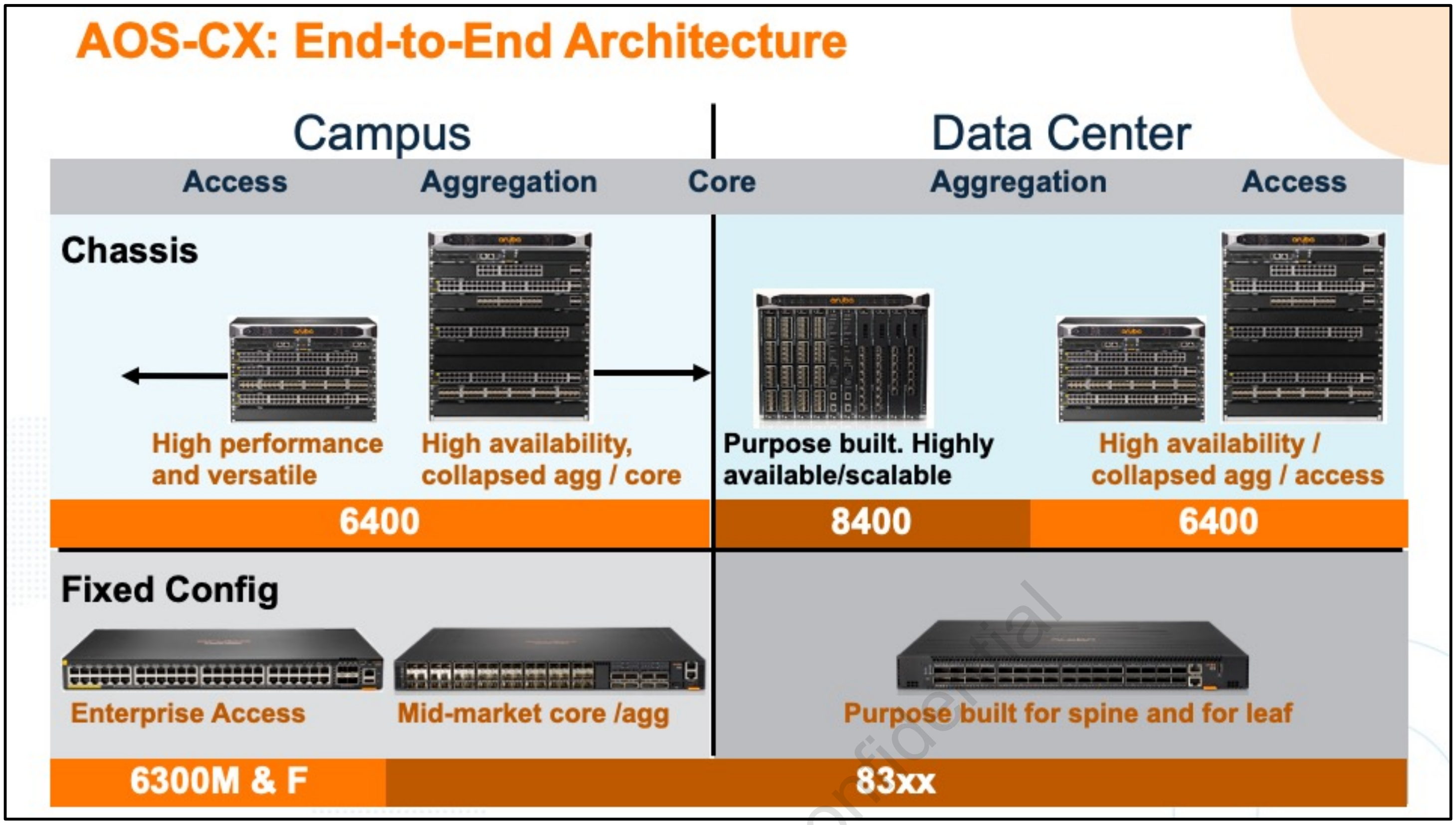
Secure, consistent user experience - Wired or Wireless

MOD 1- 6

Aruba Training-Confidential

By deploying the AOS-CX switches in a campus or data center network, you get all the benefits of an advanced operating system. This includes:

- | New hardware that brings together Gen 7 HPE ASICs with Virtual Output Queuing, Smart Rate ports, and up to Class 6 PoE
- | The AOS-CX modern architecture enables a seamless REST interface, with support for the exclusive Network Analytics Engine
- | A set of advanced tools that cover zero-touch or one-touch deployment via the CX mobile app, and advanced configuration management with NetEdit
- | Proven technologies like dynamic segmentation that unifies wired and wireless access, enhances and simplifies policy enforcement, and protects IOT and other critical devices



AOS-CX: End-to-End Architecture

Campus | Data Center

Access | Aggregation | Core | Aggregation | Access

Chassis

High performance and versatile

High availability, collapsed agg / core

Purpose built. Highly available/scalable

High availability / collapsed agg / access

6400 | 8400 | 6400

Fixed Config

Enterprise Access

Mid-market core /agg

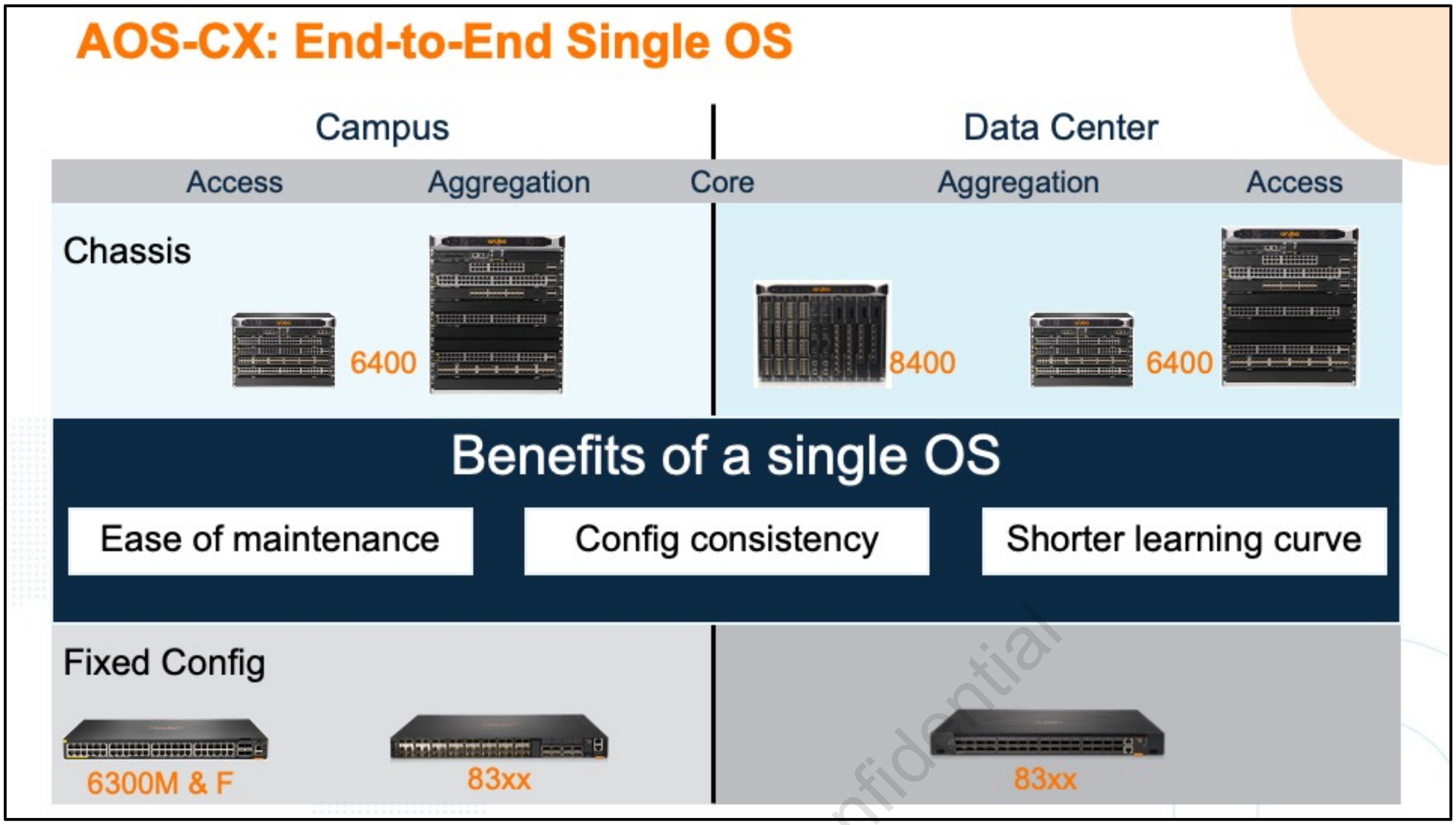
Purpose built for spine and for leaf

6300M & F | 83xx

Aruba Training-Confidential

Aruba offers various switches to support all common use cases – Chassis-based 6400-series Campus core, aggregation, and access switches on the left, and Data Center switches on the right. At the bottom you see the fixed configuration switches – 6300M and F and 83xx.

You can integrate the 6300 and 6400 switch series with existing AOS-CX portfolio, for an end-to-end solution with specialized hardware in each role and for each size and feature requirement.



AOS-CX: End-to-End Single OS

Campus | Data Center

Access | Aggregation | Core | Aggregation | Access

Chassis

6400 | 8400 | 6400

Benefits of a single OS

Ease of maintenance | Config consistency | Shorter learning curve

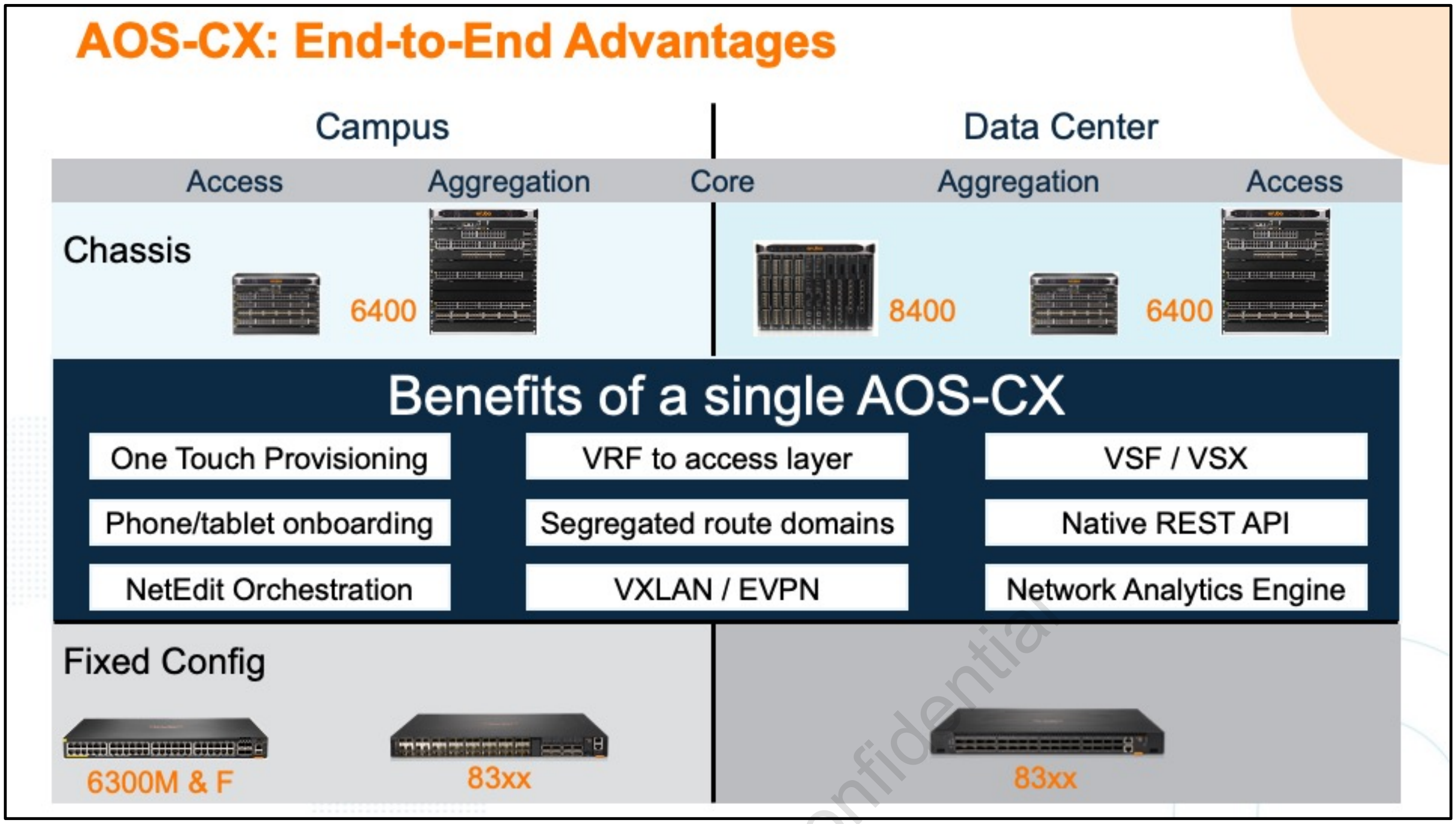
Fixed Config

6300M & F | 83xx | 83xx

Aruba Training-Confidential

The benefits of having a single operating system across the entire network are many:

- Ease of maintenance
- Configuration consistency: same CLI and WebUI
- Shorter/flatter learning curve



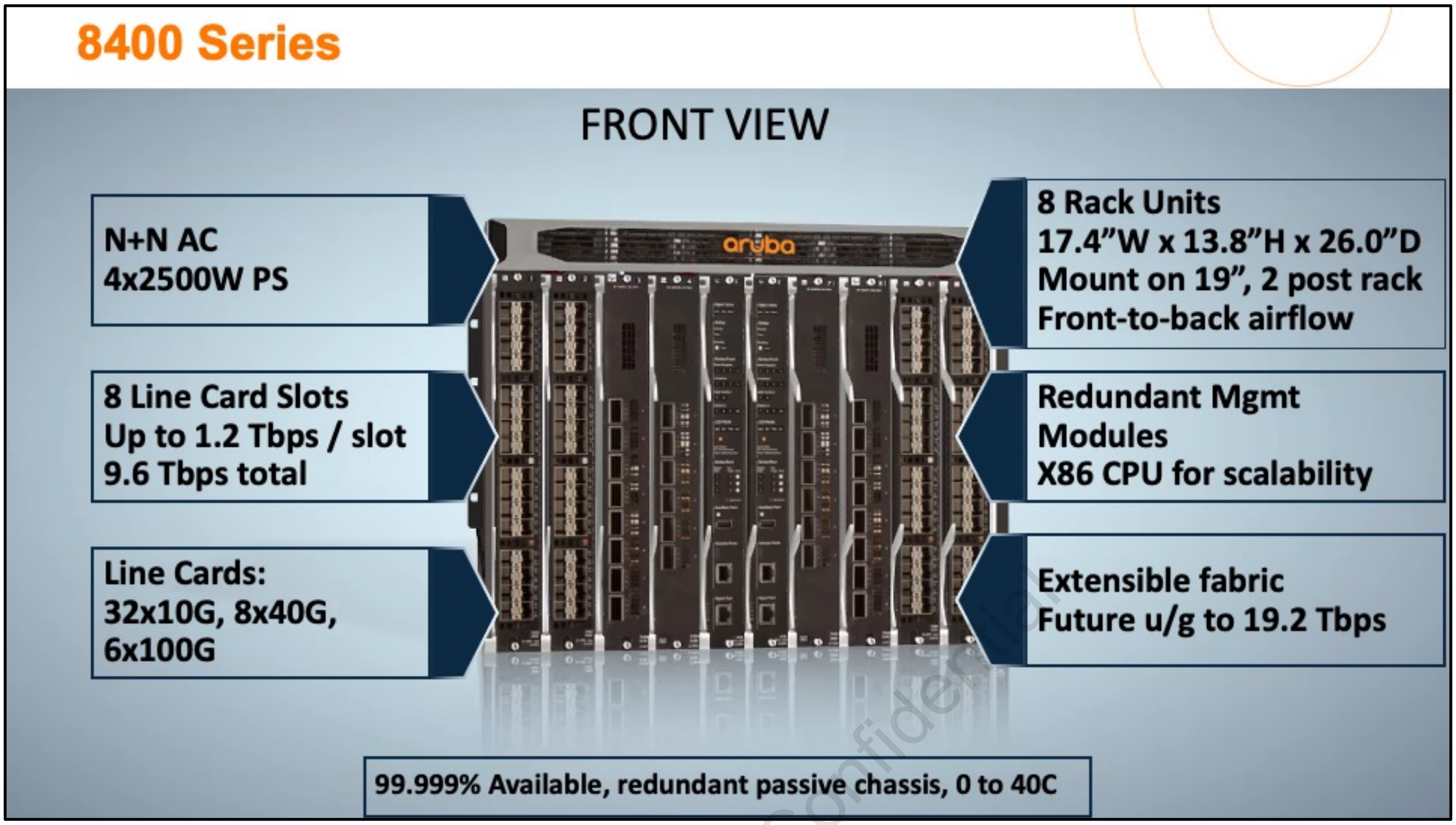
Aruba Training-Confidential

But even more important, because the end-to-end solution is AOS-CX, the benefits are even greater:

- CX Mobile App for One Touch Provisioning
- Easy Phone or tablet-based switch onboarding into NetEdit for configuration orchestration and auditing
- VRF all the way to the access layer
- Simple IP segmentation, to provide segregated routing domains for different types of traffic including IOT, VOIP and specific applications
- VXLAN/EVPN (except for the 8320 all platforms will provide full support in the near future)
- VSF and VSX in the right parts of the network
- Full/native REST API
- Network Analytics Engine
- Network administrator can decide “which agents are needed where”

And finally, all these benefits are available for enterprise networks of all types and sizes.

Let’s examine this switch product line in a bit more detail



8400 Series

FRONT VIEW

N+N AC
4x2500W PS

8 Line Card Slots
Up to 1.2 Tbps / slot
9.6 Tbps total

Line Cards:
32x10G, 8x40G,
6x100G

8 Rack Units
17.4"W x 13.8"H x 26.0"D
Mount on 19", 2 post rack
Front-to-back airflow

Redundant Mgmt
Modules
X86 CPU for scalability

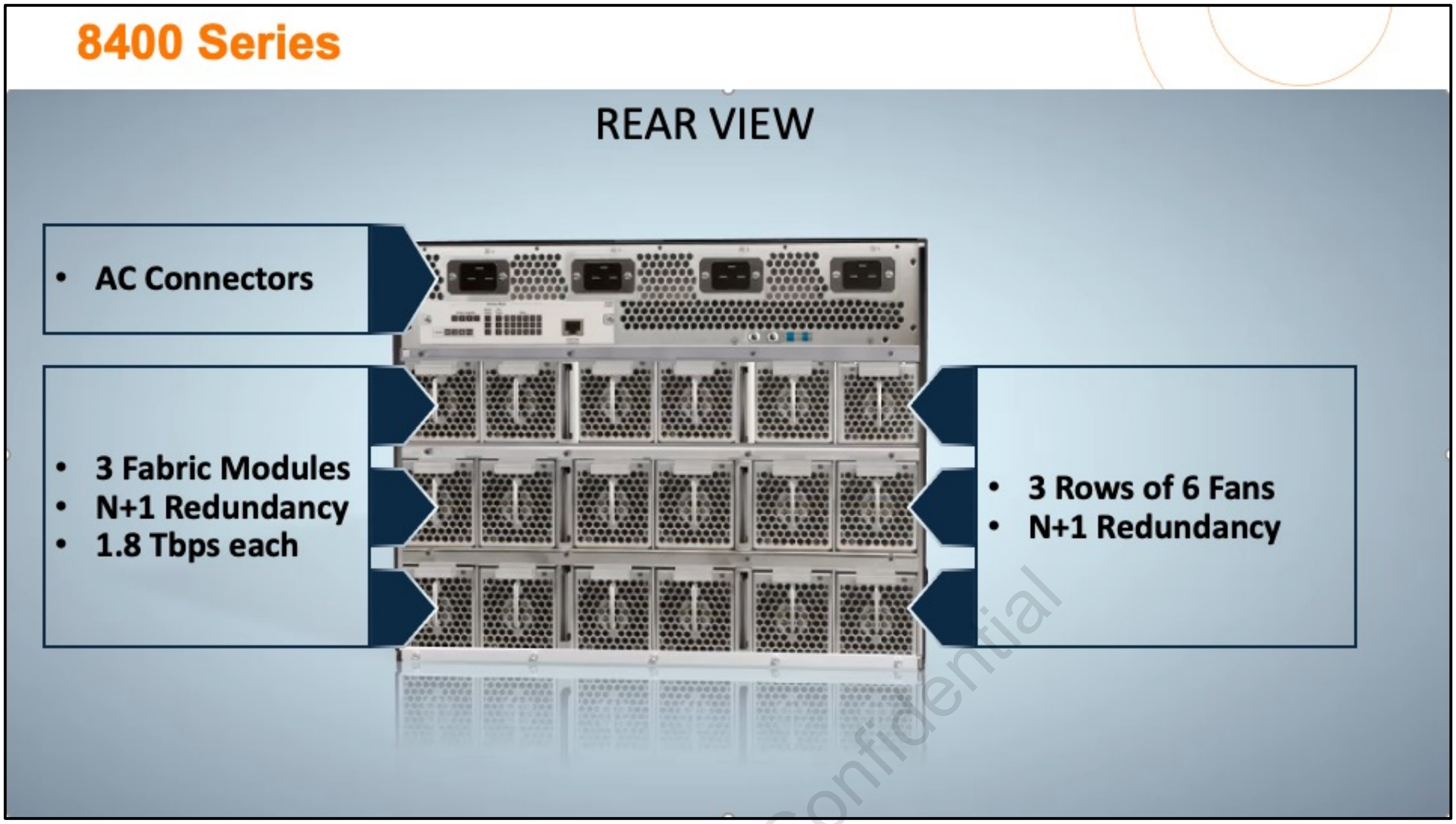
Extensible fabric
Future u/g to 19.2 Tbps

99.999% Available, redundant passive chassis, 0 to 40C

The Aruba 8400 is targeted for large campus networks. It is built to provide high availability and resiliency in every part of the hardware, supporting 99.999% availability. | The front of this modular switch houses eight slots for interface modules or line cards (LC). Options for LCs include 32-port 10GbE modules, 8-port 40GbE modules, and 6-port 100GbE modules—all amply supported by the 1.2Tbps capabilities of the LC slot. Refer to the most recent datasheet for the latest list of supported modules and precise specifications.

The 8400 also supports redundant management modules (MMs), built on an X86 CPU for scalability. Each management module provides a micro-USB slot and an RJ45 console port, as well as an Ethernet port for out-of-band management (OOBM).

Aruba Training-Confidential



8400 Series

REAR VIEW

- AC Connectors

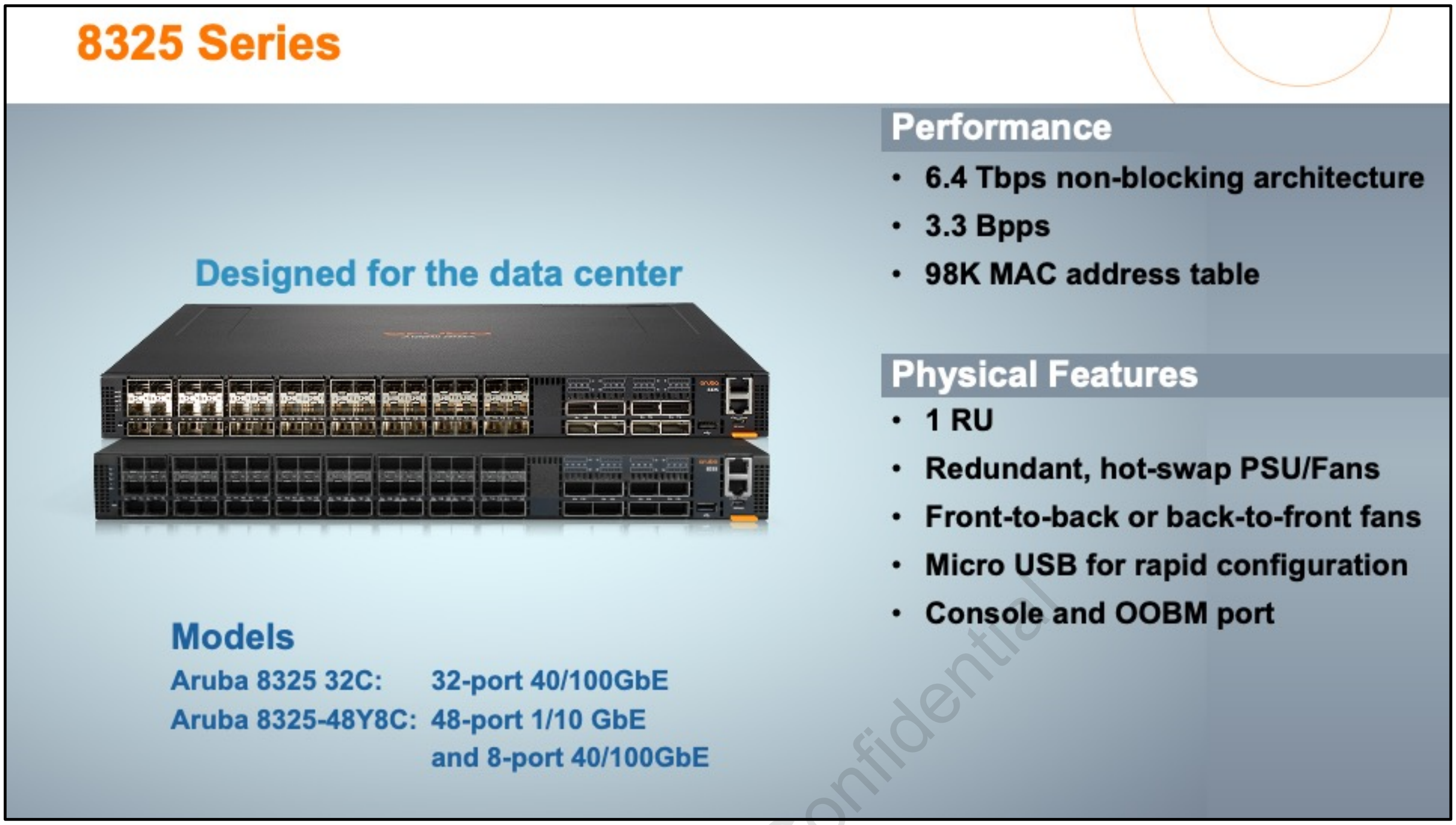
- 3 Fabric Modules
- N+1 Redundancy
- 1.8 Tbps each

- 3 Rows of 6 Fans
- N+1 Redundancy

Aruba Training-Confidential

The rear of the 8400 has slots for three fabric modules, supporting 1.8Tbps each. The fabric modules deliver high throughput between line cards and N+1 redundancy. | On the rear of the chassis are three rows of six fans, again providing N+1 redundancy, and a front to back airflow.

To further bolster availability, the 8400 supports up to four power supplies. If customers select high line voltage supplies (200V to 240V), each supply delivers 2500W. Two power supplies can power a fully loaded chassis, three supplies provide N+1 redundancy, and four supplies provide N+N redundancy. Low-line voltage (110V-120V) supplies are also available and can be appropriate for customers with less full chassis or no redundancy requirements.



8325 Series

Designed for the data center

Performance

- 6.4 Tbps non-blocking architecture
- 3.3 Bpps
- 98K MAC address table

Physical Features

- 1 RU
- Redundant, hot-swap PSU/Fans
- Front-to-back or back-to-front fans
- Micro USB for rapid configuration
- Console and OOBM port

Models

Aruba 8325 32C:	32-port 40/100GbE
Aruba 8325-48Y8C:	48-port 1/10 GbE and 8-port 40/100GbE

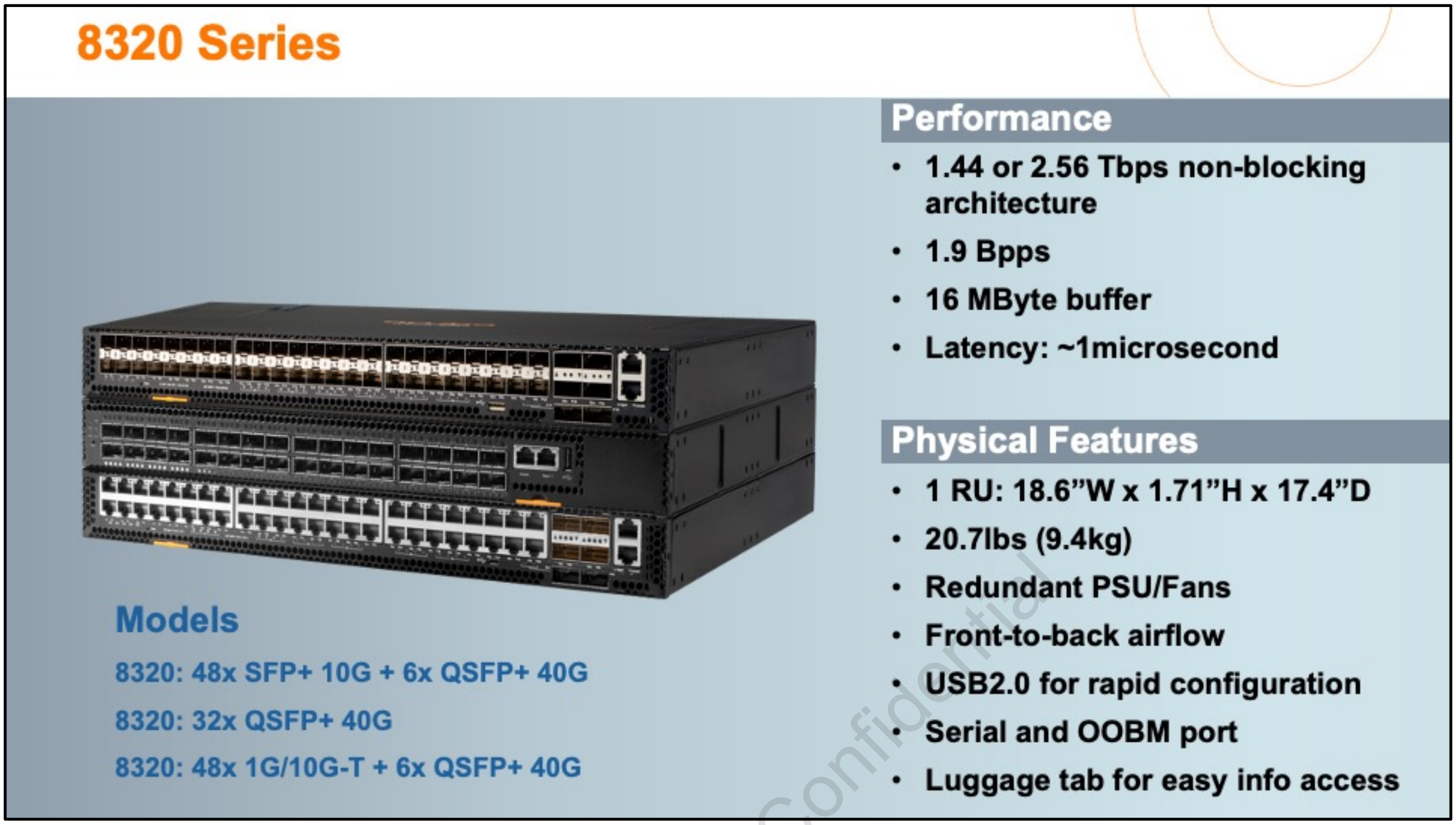
Aruba Training-Confidential

The Aruba 8325 Series is designed to bring the benefits of AOS-CX software to the data center. As of the publication of this course, the series consists of two models. The 8325 32C switch provides 32 40/100 GbE ports. The 8325-48Y8C offers 48 flexible 1/10/25 GbE ports and eight 40/100 GbE ports. The switches support full line rate for all of these high-speed ports with a maximum 6.4 Tbps non-blocking architecture and 3.3 Bpps forwarding. (Of course, the capacity is also limited by the bandwidth on the ports as well.) The 8325 series provides a very large MAC address table of 98,304.

The switches deliver hardware redundancy with redundant, hot-swappable PSUs and fans.

To work within the data center hot-aisle/cold-aisle design, the switches support deployment with either front-to-back or back-to-front airflow. Note the airflow option is two different models of the switch, beware of what you order.

They also offer similar management.



8320 Series

Performance

- 1.44 or 2.56 Tbps non-blocking architecture
- 1.9 Bpps
- 16 MByte buffer
- Latency: ~1microsecond

Physical Features

- 1 RU: 18.6”W x 1.71”H x 17.4”D
- 20.7lbs (9.4kg)
- Redundant PSU/Fans
- Front-to-back airflow
- USB2.0 for rapid configuration
- Serial and OOBM port
- Luggage tab for easy info access

Models

8320: 48x SFP+ 10G + 6x QSFP+ 40G

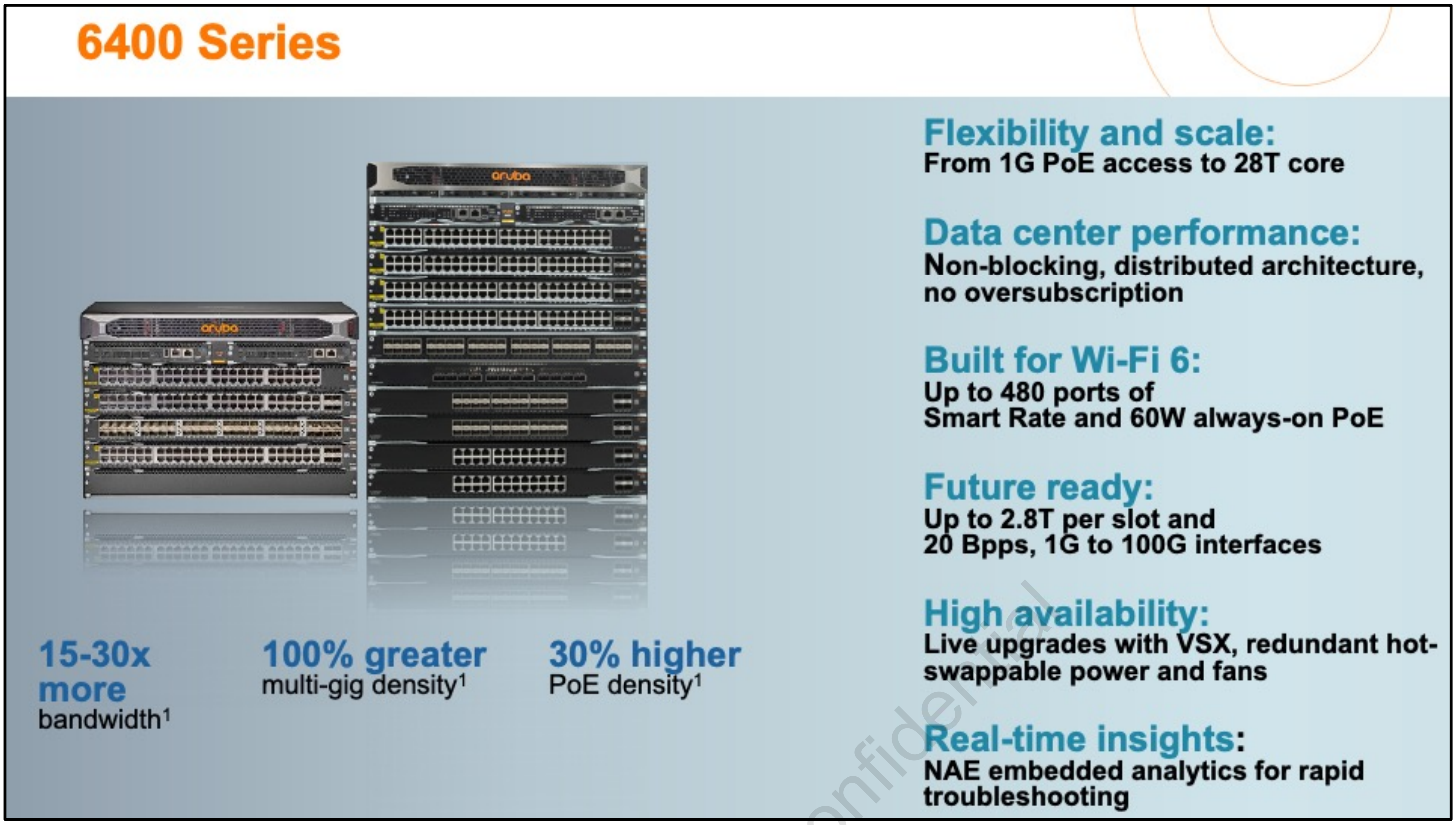
8320: 32x QSFP+ 40G

8320: 48x 1G/10G-T + 6x QSFP+ 40G

Aruba Training-Confidential

The Aruba 8320 series of fixed-port switches is intended for the core aggregation layer or mid-sized campus cores. The series includes several different modules with different combinations of 10GbE and 40GbE ports. For the 10GbE ports, one model supports fiber ports and another model copper ports. The switches feature a non-blocking architecture that delivers the full line rate for every port. When both inbound and outbound traffic is considered, that translates to 1.44 Tbps capacity on the models with a mix of 10GbE and 40GbE ports and 2.56 Tbps on the model with 32 40GbE ports. The switches support up to 1905 million packets per second (Mpps). They have a large 16MB buffer and deliver low latency under one microsecond.

The 8320 switches are also built for resiliency. They offer redundant power supplies and fans. The front of the switches provides several management features, including a USB2.0 slot, a serial port, and an Ethernet out-of-band management port. A luggage tab can be pulled out and provides information such as the switch serial number.



6400 Series

Flexibility and scale:
From 1G PoE access to 28T core

Data center performance:
Non-blocking, distributed architecture, no oversubscription

Built for Wi-Fi 6:
Up to 480 ports of Smart Rate and 60W always-on PoE

Future ready:
Up to 2.8T per slot and 20 Bpps, 1G to 100G interfaces

High availability:
Live upgrades with VSX, redundant hot-swappable power and fans

Real-time insights:
NAE embedded analytics for rapid troubleshooting

15-30x more bandwidth[1]

100% greater multi-gig density[1]

30% higher PoE density[1]

Aruba Training-Confidential

The first is the AOS-CX 6400 series. This family consists of two chassis. There's a five slot and a ten slot and they can support very high-density port counts, with flexibility and scale, giving Aruba a significant competitive advantage.

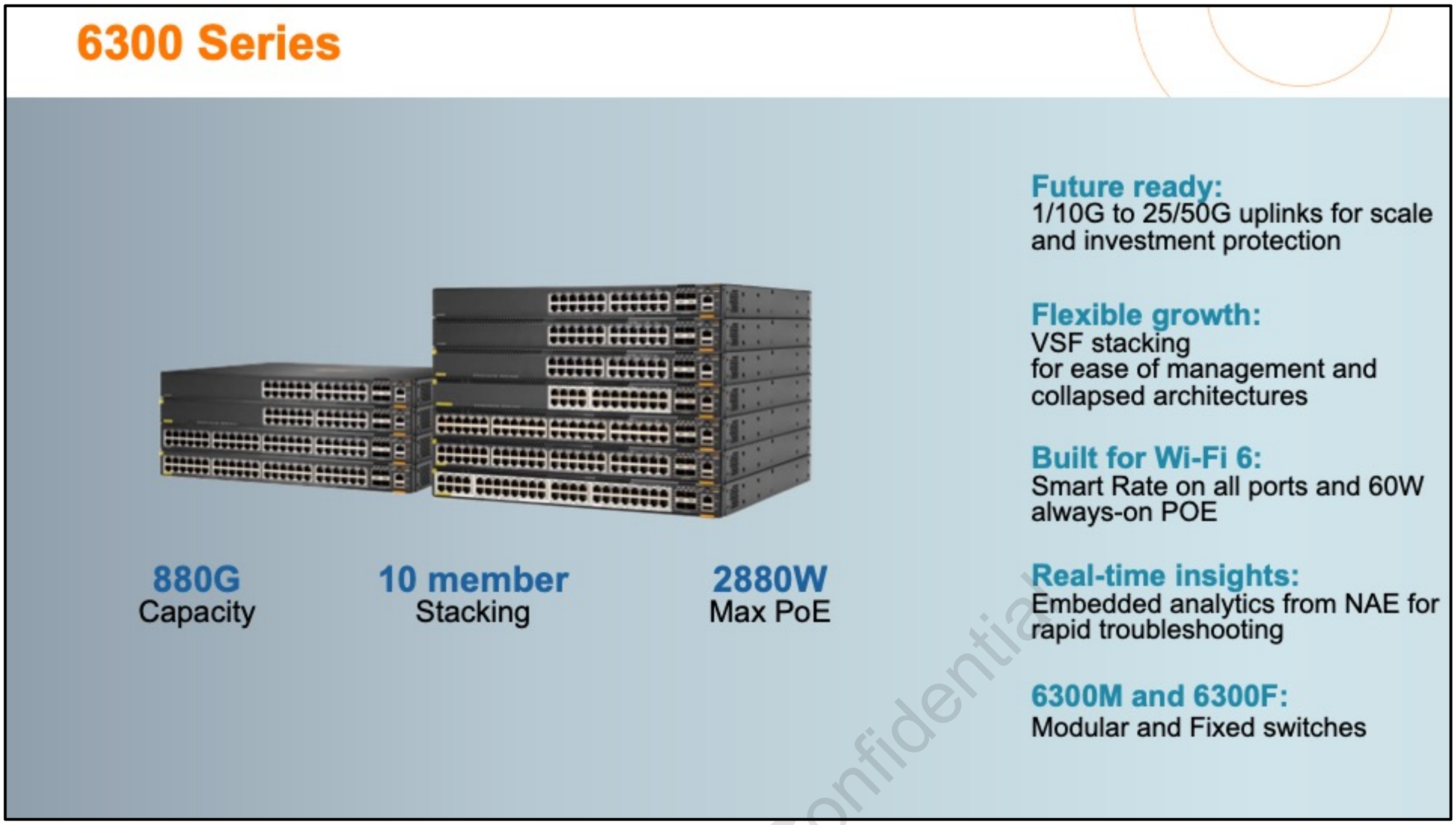
Aruba built the AOS-CX 6400 products using a lot of internal intellectual property with custom-built ASICs and the AOS-CX operating system. With a non-blocking distributed architecture, with no oversubscription, this greatly simplifies networking. This means network admins don't have to plan around what they're plugging into which port on what slot because the switches support a complete non-blocking fabric.

The port counts vary in terms of the types of connectivity offered with up to 60 Watts always-on PoE. You can support all the Wi-Fi 6 and IoT devices that are coming out. The 6400 also supports high density Smart Rate with 1/2.5/5 gig ports.

It also provides line cards that have 24 ports of 10GbE, both in Gigabit Ethernet, as well as SFP+. There's also a 48-port 25 gig card as well as a 12-port 40/100 gig card. This is a big breadth of different connectivity options, allowing this platform to be used all the way from the access in the campus to the core and also in the data center in some cases. That's a great value proposition for some customers who can simplify their design and standardize on a single platform across their entire network.

For high availability (HA), the AOS-CX 6400 supports VSX Live Upgrades and also has redundant management cards, fans, power supplies, etc.

Aruba is also extending the Aruba Network Analytics Engine (NAE) onto these platforms. If you are already familiar with 8400 and 8325, that is a key differentiator in terms of providing visibility and telemetry to help with troubleshooting.



6300 Series

880G
Capacity

10 member
Stacking

2880W
Max PoE

Future ready:
1/10G to 25/50G uplinks for scale and investment protection

Flexible growth:
VSF stacking
for ease of management and collapsed architectures

Built for Wi-Fi 6:
Smart Rate on all ports and 60W always-on POE

Real-time insights:
Embedded analytics from NAE for rapid troubleshooting

6300M and 6300F:
Modular and Fixed switches

Aruba Training-Confidential

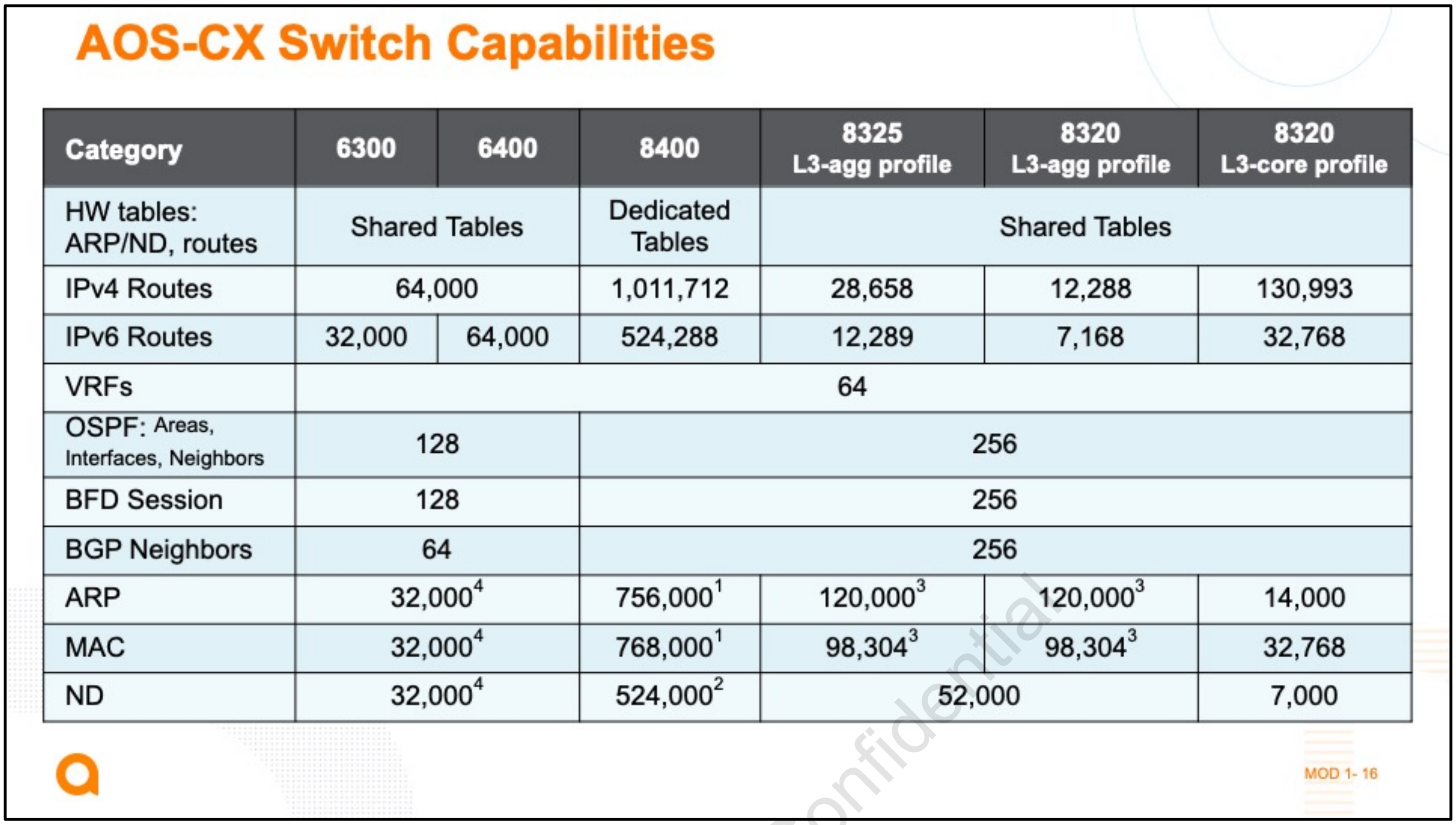
The second newest AOS-CX hardware platform is the 6300 series, which is essentially built on the same architecture as the 6400. It's a stackable form factor – you can stack up to 10 switches in a single stack which is perfect for different sized wiring closets. The AOS-CX 6300 also supports of variety of different port types, such as PoE (up to 60 watts), Smart Rate - there's actually 11 different models in this family. Several are fixed and several are modular - the modular ones support modular power supplies and modular fans. These are great for people looking for additional power redundancy. The fixed models are ideal for smaller spaces or areas where that redundancy isn't as big of an issue.

It also supports ten-unit VSF stacking. There's a 24-port 10GbE SFP+ model as well which you can use in the aggregation layer. This means you can do a collapsed access aggregation with just this family of products where you're using stacking to simplify management and deployment.

The AOS-CX 6300 switch series will also support real time analytics, NAE, as mentioned for the 6400.

There are 11 different SKUs in the Aruba 6300 switch series family, seven modular (6300M) and four fixed (6300F) with two air-flow of front to back and left to back respectively. Each switch provides a different number/type of interfaces/ports but they all share the same operating system AOS-CX.

With fixed SKUs there are no hot-swappable power supplies and fans, everything is fixed. Modular SKUs have a hot-swappable power supplies and fans. Aruba 6300M SKUs support fan direction front to back (same as 2930M) and two PSU slots (PSU sharable with 3810/2930M).



AOS-CX Switch Capabilities

Category	6300	6400	8400	8325 L3-agg profile	8320 L3-agg profile	8320 L3-core profile
HW tables: ARP/ND, routes	Shared Tables		Dedicated Tables	Shared Tables		
IPv4 Routes	64,000		1,011,712	28,658	12,288	130,993
IPv6 Routes	32,000	64,000	524,288	12,289	7,168	32,768
VRFs	64					
OSPF: Areas, Interfaces, Neighbors	128		256			
BFD Session	128		256			
BGP Neighbors	64		256			
ARP	32,000[4]		756,000[1]	120,000[3]	120,000[3]	14,000
MAC	32,000[4]		768,000[1]	98,304[3]	98,304[3]	32,768
ND	32,000[4]		524,000[2]	52,000		7,000

MOD 1- 16

Aruba Training-Confidential

This table compares AOS-CX platform scalability as of the 10.4 release.

As a large campus core switch, the Aruba 8400 can scale to support a great number of MAC addresses, routes, ARP entries, and other such entries. The scalability depends on an intersection of the hardware and software capabilities. It is important to notice that 6300 and 6400 provides exactly the same scale capability as expected as this is the same ASIC used on 6300 than on 6400. Notice as well that scale for IPv6 is the same than for IPv4 on those new platforms. 8400 remains the only platform with very large-scale numbers. For 8320, 8325 and 8400 the scale numbers are the same in 10.4 than in 10.3.

You can configure an Aruba 8300 switch to operate in one of two modes, depending on how you intend to deploy the switch. The mode affects the limits for various types of entries, allocating memory differently based on the where the memory is needed. Mobile First mode is designed for when the switch connects to Mobility Controllers (MCs) that support a great many wireless devices. This mode allocates the greater part of the memory to MAC addresses and ARP addresses, enabling the switch to act as default gateway for large VLANs with many devices. Route mode allocates less memory to MAC and ARP addresses and more to IPv4 and IPv6 routes. Use this mode when the switch needs to learn more routes, but is default gateway for smaller subnets.

Number of IPv4 Devices/Clients limited to 81,000 for ARP Resolution – 8400 Egress MAC Table Limit

Number of IPv6 Devices/Clients limited to 55,000 for ND Resolution – 8400 FEC Table Limit

Number of Devices/Clients limited to 43,000 (8320) / 47,000(8325) for ARP/ND Resolution – Egress MAC Table Limit

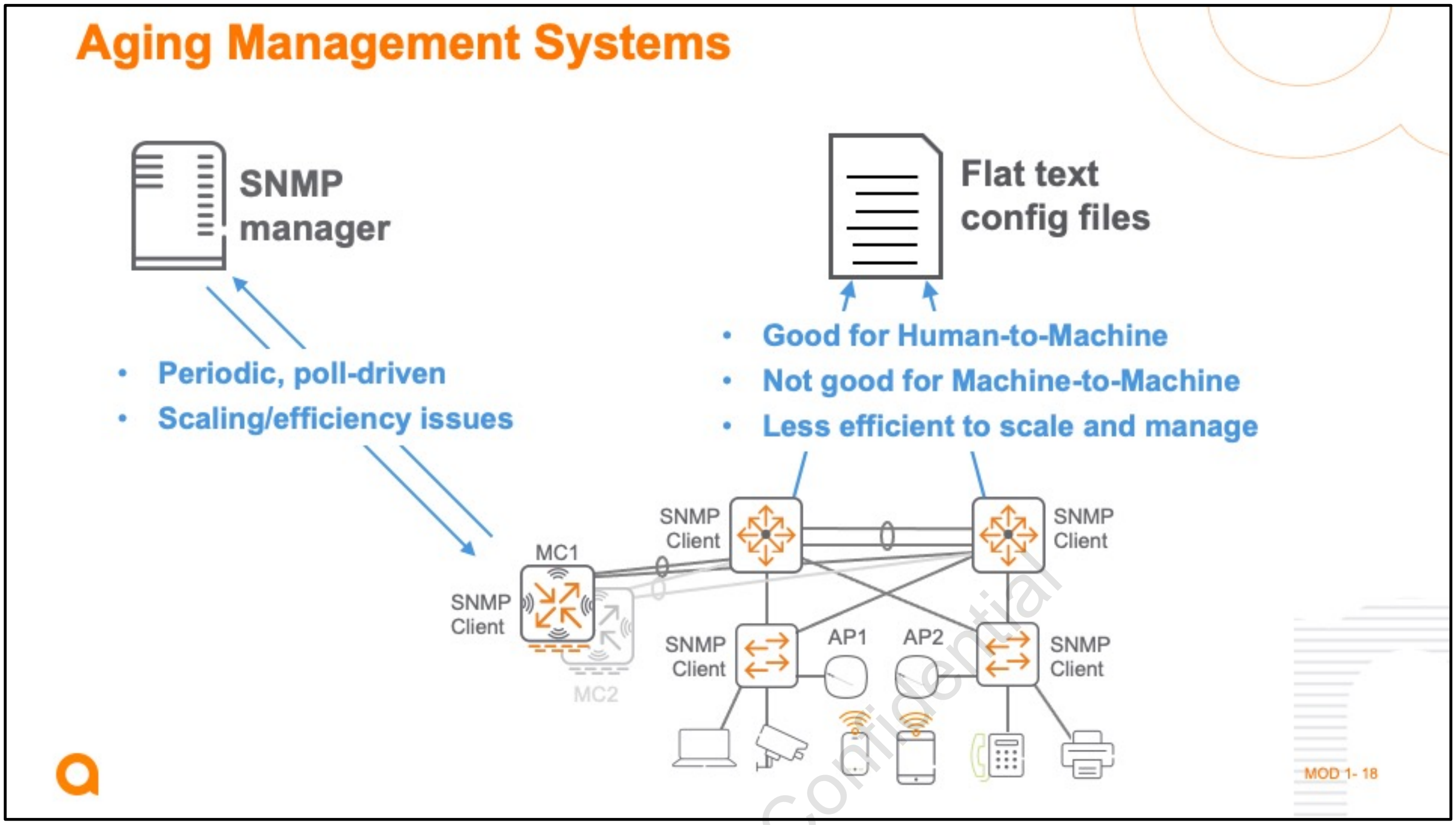
Restrictions are currently the same for the AOS-CX 6300 and 6400 switches

Features

aruba
a Hewlett Packard Enterprise company

This section introduces some of the important features of the AOS-CX operating system and switches.

Aruba Training-Confidential



Aging Management Systems

SNMP manager

- Periodic, poll-driven
- Scaling/efficiency issues

Flat text config files

- Good for Human-to-Machine
- Not good for Machine-to-Machine
- Less efficient to scale and manage

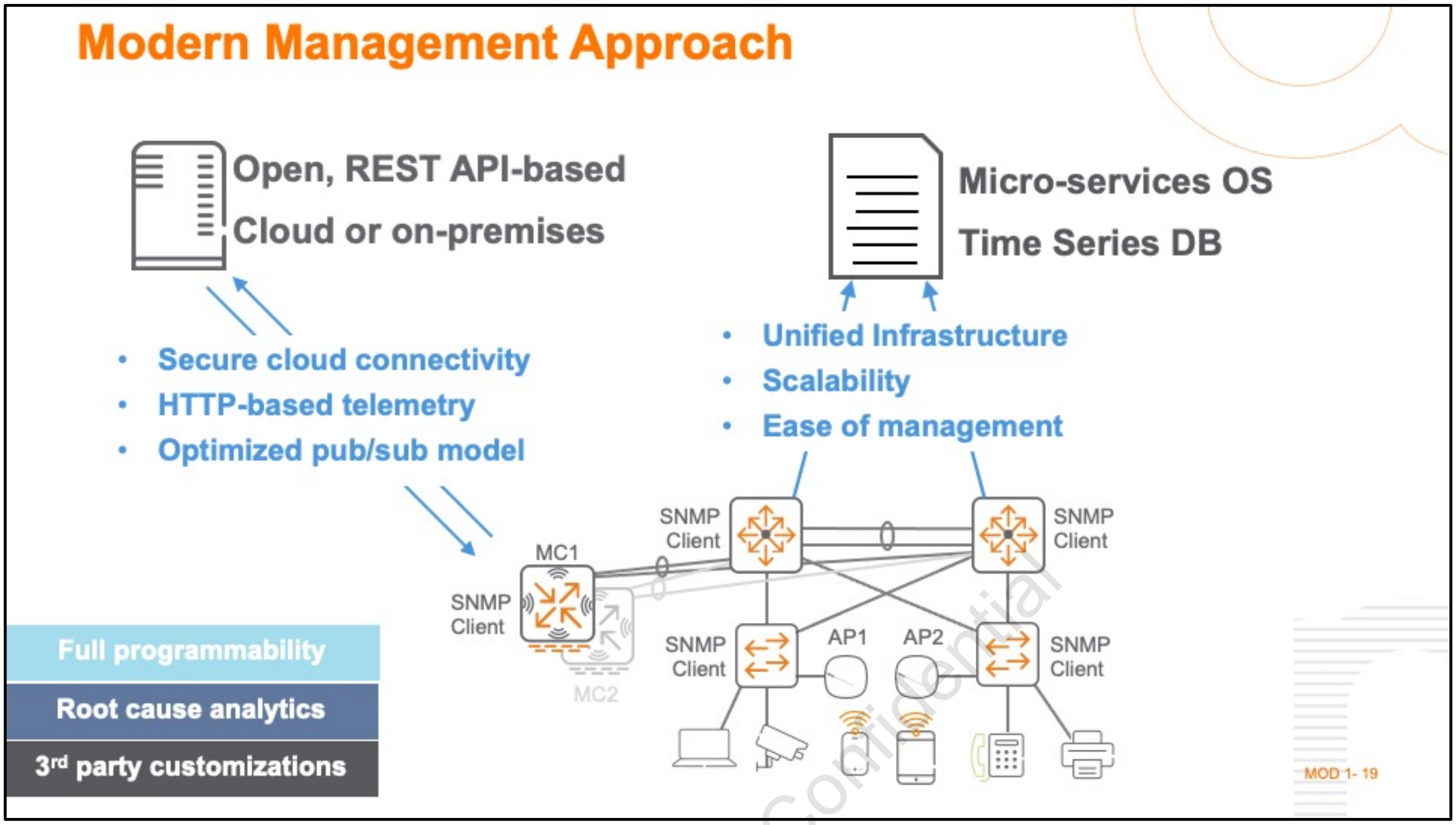
MOD 1- 18

The figure shows a typical 2-tier network architecture, often duplicated several times in a large campus deployment – dozens or hundreds of switches, multi-layer switches, and controllers. For many years, Network Admins and have relied on the old Simple Network Management Protocol (SNMP). It still works OK, but its “poll driven approach to telemetry is a bit dated. It remains heavily used, but there are opportunities to find a more streamlined, universal approach to configuration, logging, and diagnostics.

And the old Command Line Interface (CLI) with flat-text configuration files is getting a bit old-fashioned as well. The old CLI is great for Human to Machine interaction, but not so great for Machine-to-Machine interaction. It just doesn’t lend itself to the scalability and ease of management requirements of modern campus designs.

Wouldn’t it be nice to have a unified infrastructure, with a single operational model across the infrastructure.

Aruba Training-Confidential



Modern Management Approach

Open, REST API-based
Cloud or on-premises

- Secure cloud connectivity
- HTTP-based telemetry
- Optimized pub/sub model

Micro-services OS
Time Series DB

- Unified Infrastructure
- Scalability
- Ease of management

Full programmability
Root cause analytics
3rd party customizations

MOD 1- 19

Aruba Training-Confidential

Aruba CX switches are built on a microservices-based operating system with an open REST API available for configuration and management functions.

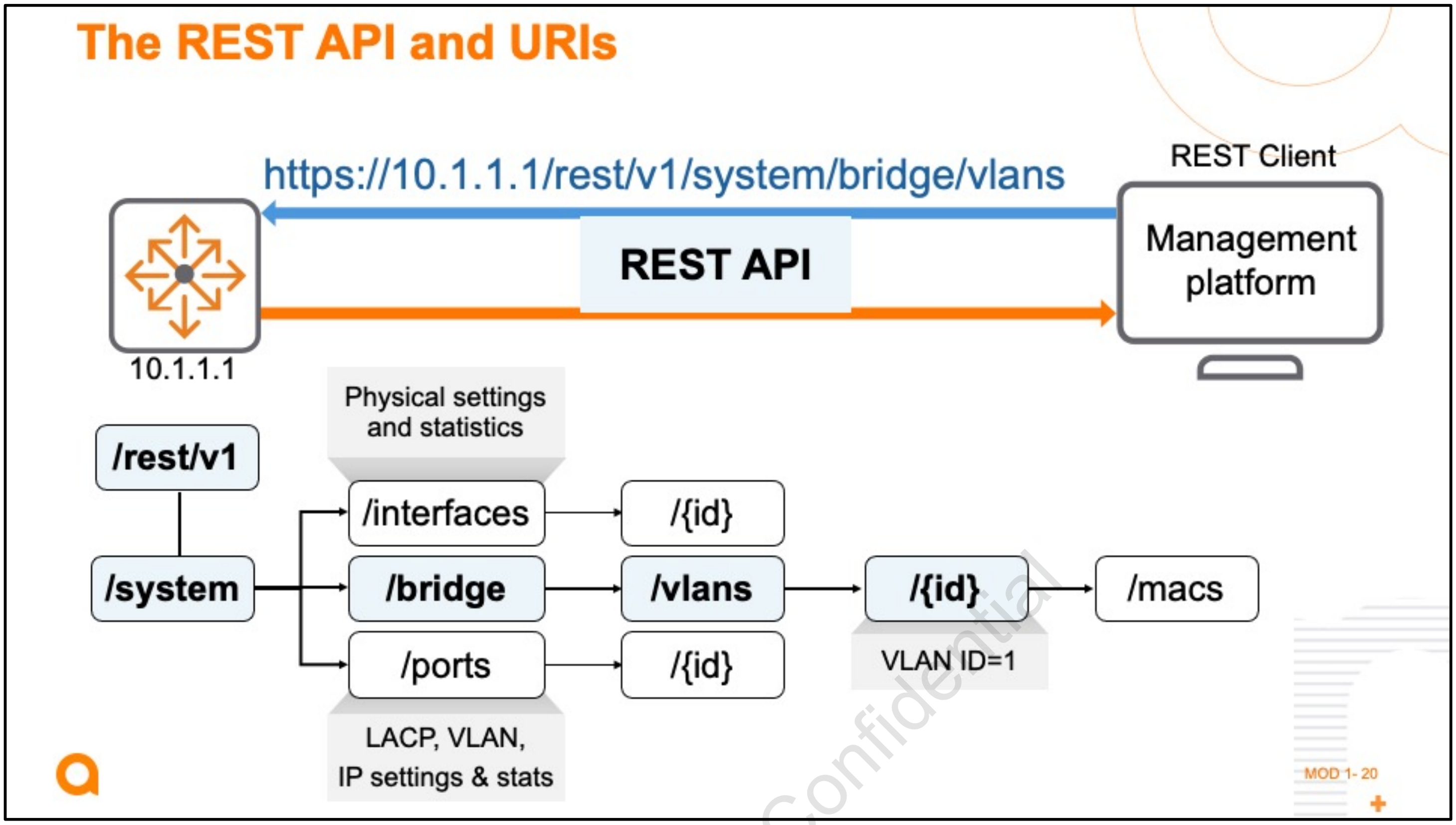
You get easy, secure cloud connectivity, with modern HTTP-based telemetry that uses a more optimized publisher/subscriber model, as opposed to the cumbersome polling method of SNMP.

This facilitates zero-downtime upgrades and moving to a more centralized control and management plane – in-cloud or on-premises. It also facilitates more automation related to typical network administration tasks.

Everything is fully programmable, with intelligent root case analytics and third-party customizations.

And scalability? A single customer can manage up to 100,000 network devices and 1 million client devices.

Several of these advantages are due to the efficient, hierarchic nature of the REST API.



The REST API and URIs

https://10.1.1.1/rest/v1/system/bridge/vlans

REST Client

Management platform

REST API

10.1.1.1

Physical settings and statistics

/rest/v1

/system

/interfaces

/bridge

/ports

/{id}

/vlans

/{id}

/{id}

/macs

VLAN ID=1

LACP, VLAN, IP settings & stats

MOD 1- 20

Aruba Training-Confidential

The REST API defines resources, each of which is a Uniform Resource Identifier (URI) that a client can access. For example, a switch that has the IP address 10.1.1.1, has the URL 10.1.1.1/rest/v1/system/bridge/vlans which lists VLANs on the switch. A REST client can contact this URL to obtain information about this list or add to it.

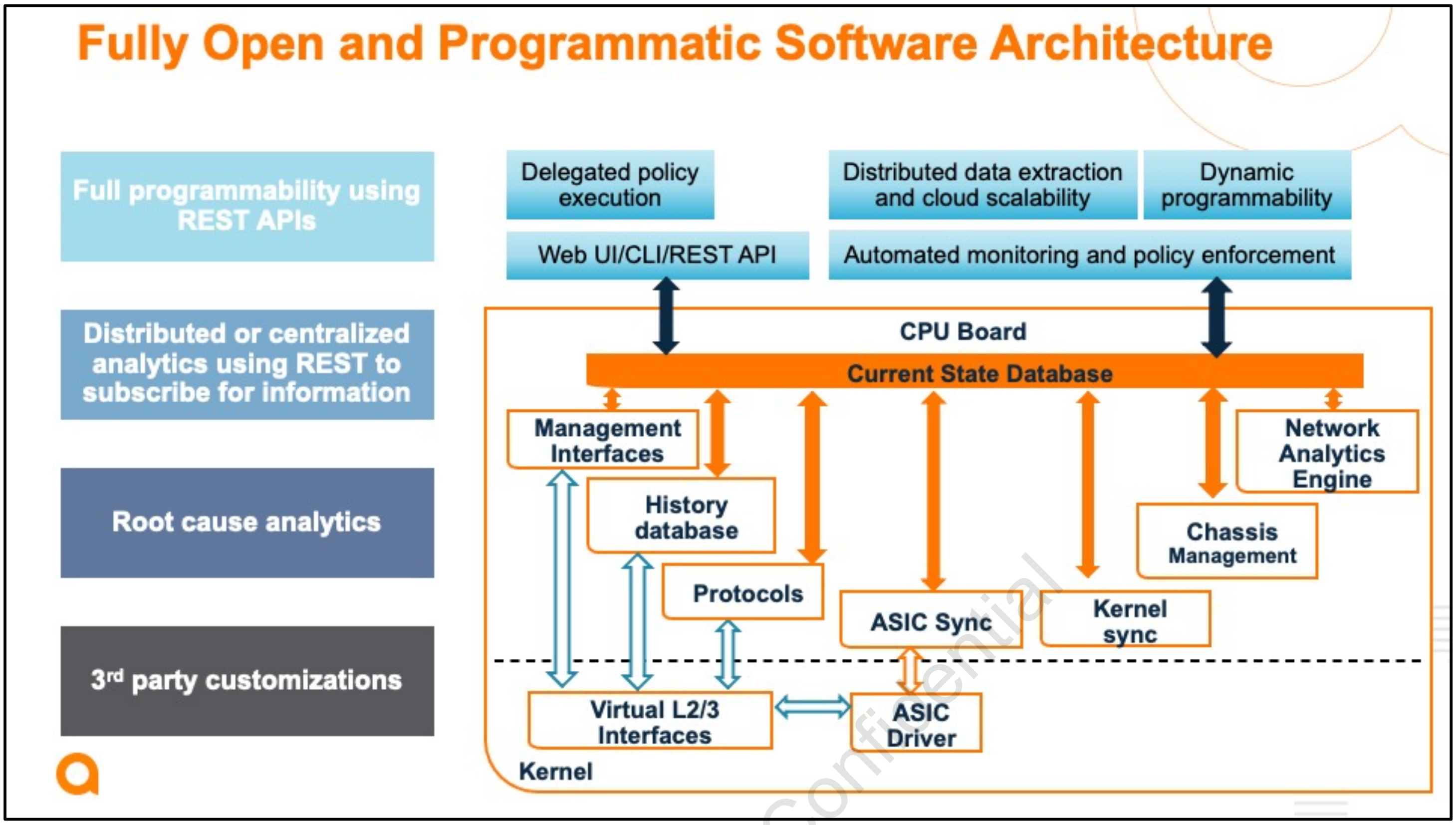
| As you see here, the URIs for resources follow a hierarchical structure. This means that a resource can also be a container for other resources. For example, the system/vlans resource contains multiple VLAN resources, each identified with an ID. The same concept holds true for the system/interfaces resource, which is a container for multiple interface IDs, and for several other resources on the switch.

The "interfaces" resource stores physical and lower-level information - admin status, UDLD, and more.

The ports resource stores VLAN and IP/IPv6 protocol information. This is where you configure things like LACP, IP, OSPF, PIM, IGMP, and VLAN settings.

ArubaOS-CX is architected to optimize these programmatic functions, enabling the scripting, automation and orchestration features we expect from modern systems.

Let’s zoom in on that internal switch architecture.



Fully Open and Programmatic Software Architecture

Full programmability using REST APIs

Distributed or centralized analytics using REST to subscribe for information

Root cause analytics

3rd party customizations

Delegated policy execution

Distributed data extraction and cloud scalability

Dynamic programmability

Web UI/CLI/REST API

Automated monitoring and policy enforcement

CPU Board

Current State Database

Management Interfaces

History database

Protocols

ASIC Sync

Kernel sync

Chassis Management

Network Analytics Engine

Virtual L2/3 Interfaces

ASIC Driver

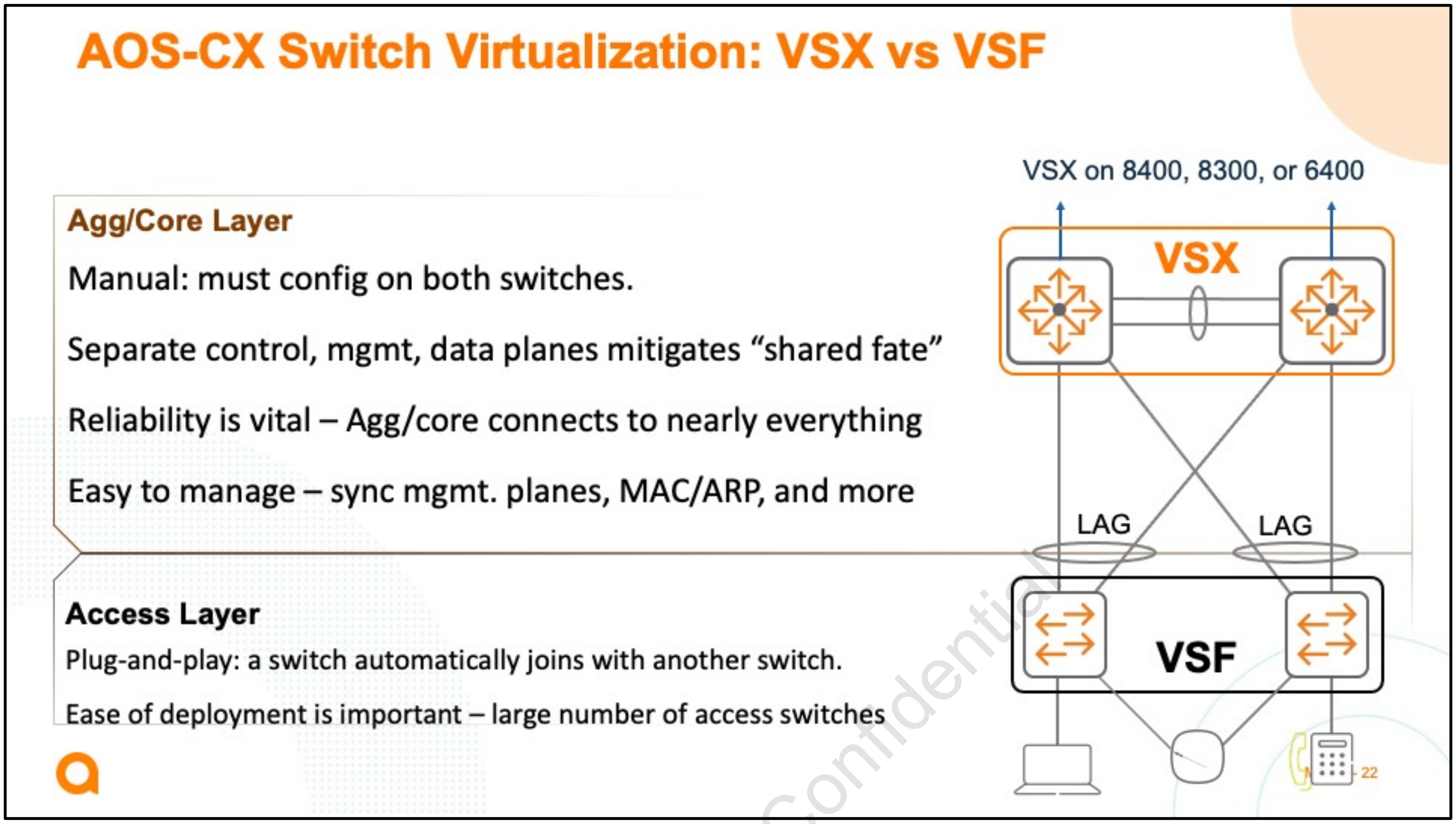
Kernel

Aruba Training-Confidential

Inside you have a CPU board and Kernel, with all management, history, protocol functions, and network analytics based around a stateful database. Of course, you have high-speed communications down to the Application Specific Integrated Circuits (ASIC) hardware for network connectivity.

Plus, you can access all this modern internal architecture using the same Command Line and Graphical interfaces you know and love, but are also ready to take advantage of that REST API for cloud-like services, automation, and more.

You also get modern switch virtualization technologies like VSX and VSF.



AOS-CX Switch Virtualization: VSX vs VSF

Agg/Core Layer

Manual: must config on both switches.

Separate control, mgmt, data planes mitigates "shared fate"

Reliability is vital – Agg/core connects to nearly everything

Easy to manage – sync mgmt. planes, MAC/ARP, and more

Access Layer

Plug-and-play: a switch automatically joins with another switch.

Ease of deployment is important – large number of access switches

Aruba Training-Confidential

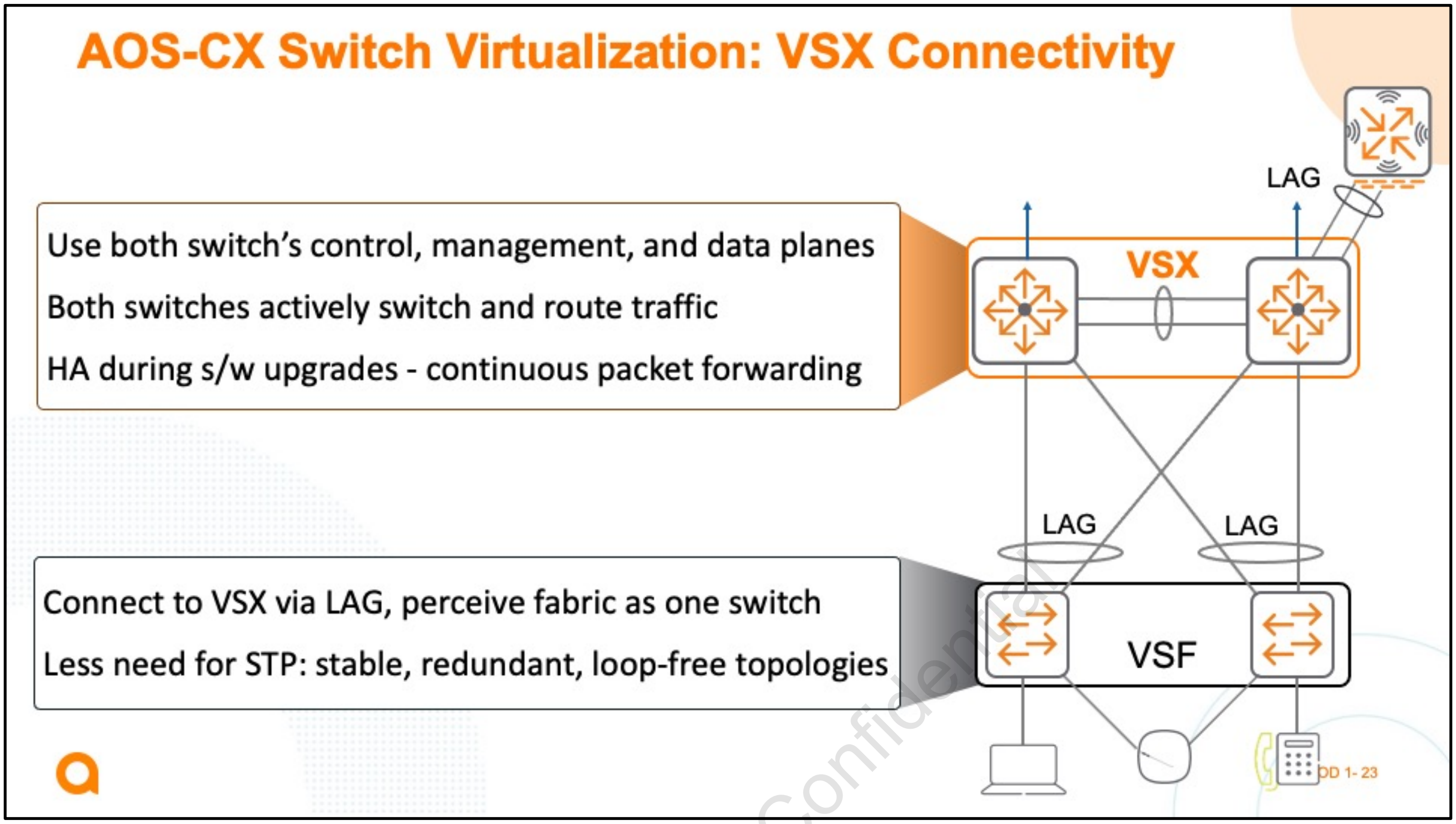
VSX and VSF are similar to each other, and even to earlier virtualization techniques like Virtual Router Redundancy Protocol (VRRP) – multiple devices are perceived as a single device. This improves resiliency while reduces configuration and management overhead. But there are distinct differences between VSX and VSF, as summarized in the figure.

You use **VSX** at the campus network aggregation and core layers. You use **VSF** primarily at the access layer. The differences in features and function are intentional – because priorities differ between access and core.

Configuring VSF and backplane stacking is a plug-and-play affair - a switch, at default settings, automatically unites with another switch. This ease of deployment is especially important at the access layer – you may have dozens, or hundreds of access switches.

You must manually configure VSX on both switches. VSX maintains separate control plane and management planes, and much of the data plane. This mitigates the "shared fate" nature of common planes, for improved uptimes – you get high availability and redundancy. This is especially important when aggregating all those access layer switches. And it remains easy to manage, because you can synchronize the management planes, and optionally, other features - MAC and ARP addresses, and more.

Let's zoom in on VSX.



AOS-CX Switch Virtualization: VSX Connectivity

Use both switch’s control, management, and data planes

Both switches actively switch and route traffic

HA during s/w upgrades - continuous packet forwarding

Connect to VSX via LAG, perceive fabric as one switch

Less need for STP: stable, redundant, loop-free topologies

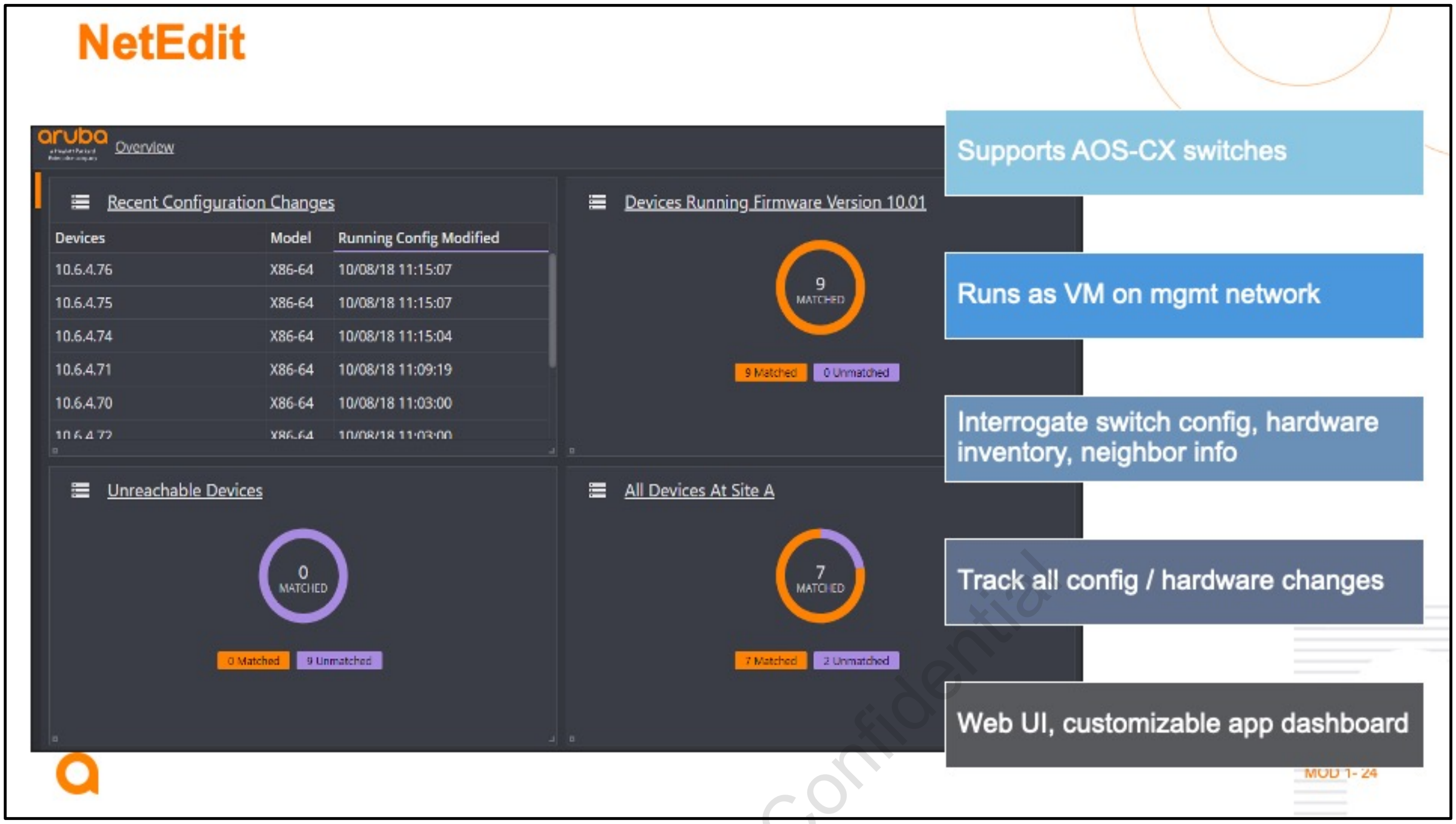
Aruba Training-Confidential

The VSX fabric leverages both switch’s control, management, and data planes, while appearing to other devices as a single switch. They support an active-active path for Layer 2 traffic, and for Layer 3 unicast and multicast traffic. In other words, both switches actively switch and route traffic.

VSX maintains the AOS-CX default behavior - ports are disabled and operate at Layer 3. Finally, VSX delivers high availability during software upgrades, with near zero downtime and continuous packet forwarding.

Access switches and mobility controllers connect to core switches via LAG connections. If configured for LACP, they perceive the VSX fabric as a single switch. This mitigates reliance on STP to ensure stable, redundant, loop-free topologies.

You’ll explore these technologies more deeply in a later module.



NetEdit

Recent Configuration Changes

Devices	Model	Running Config Modified
10.6.4.76	X86-64	10/08/18 11:15:07
10.6.4.75	X86-64	10/08/18 11:15:07
10.6.4.74	X86-64	10/08/18 11:15:04
10.6.4.71	X86-64	10/08/18 11:09:19
10.6.4.70	X86-64	10/08/18 11:03:00
10.6.4.72	X86-64	10/08/18 11:03:00

Devices Running Firmware Version 10.01

Unreachable Devices

All Devices At Site A

- Supports AOS-CX switches
- Runs as VM on mgmt network
- Interrogate switch config, hardware inventory, neighbor info
- Track all config / hardware changes
- Web UI, customizable app dashboard

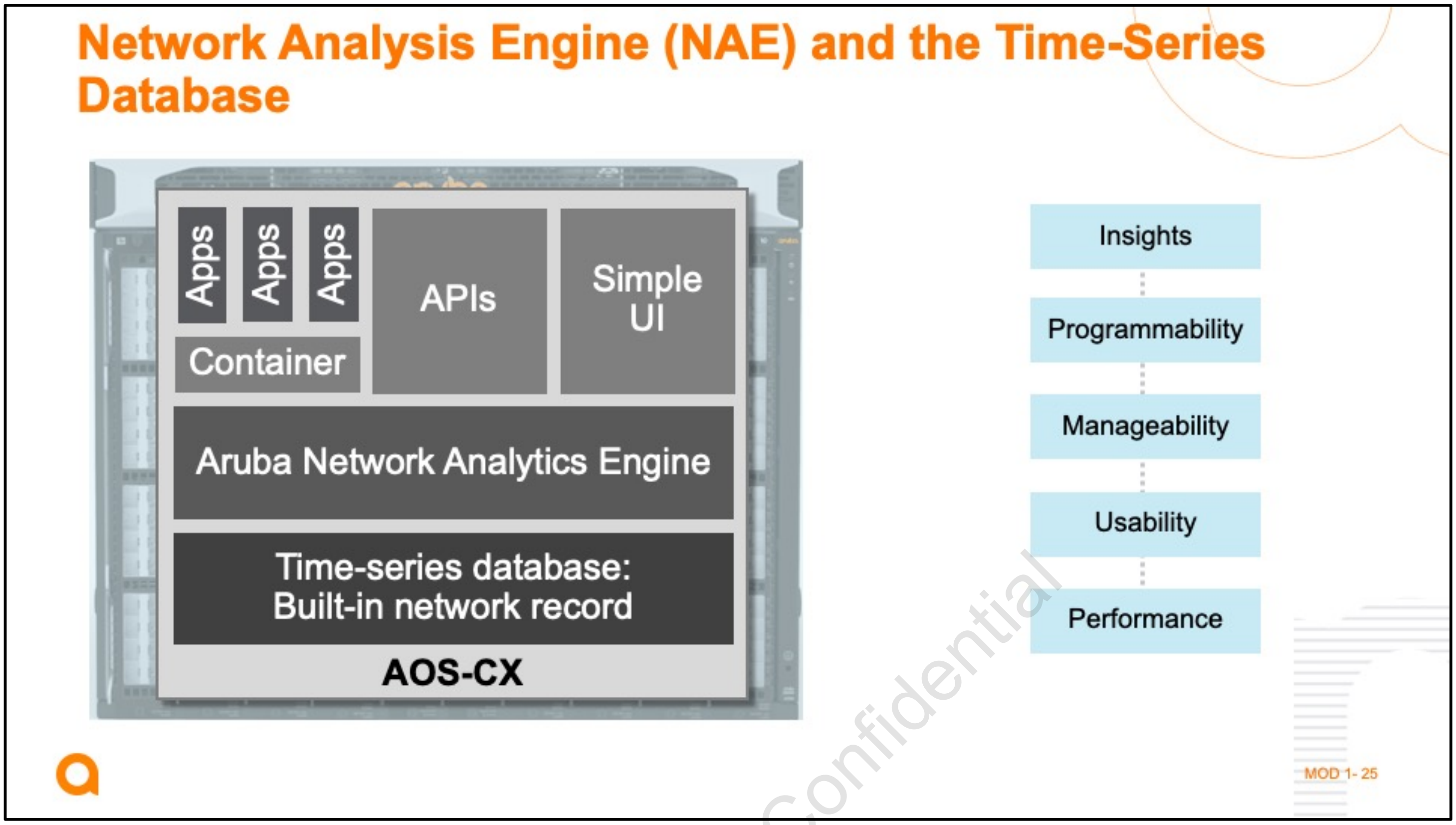
NetEdit supports switches running AOS-CX. NetEdit runs as an Open Virtualization Application (OVA) virtual machine (for example, VMware’s ESXi, KVM, Hyper-V, etc.) on a server. When NetEdit learns of a new switch, it interrogates its configuration, hardware inventory, and neighbor information. NetEdit ensures a complete history of all devices, tracking all configuration and hardware changes.

A user-friendly web-based UI allows network admins to interpret the status of network devices easily. The customizable application dashboard provides quick visibility into information and metrics that matter most to you.

NetEdit provides automation for network configuration such as search, plan, edit, validation, deployment, and audit. It provides intelligent assistance and continuous validation to help ensure that device configurations are consistent, compliant, and error free.

You can use NetEdit without retraining by leveraging your existing knowledge and experience with switch configurations. This enables you to automate switch configuration change workflows without having to have programming knowledge.

Aruba Training-Confidential



Network Analysis Engine (NAE) and the Time-Series Database

Apps

Apps

Apps

Container

APIs

Simple UI

Aruba Network Analytics Engine

Time-series database:
Built-in network record

AOS-CX

Insights

Programmability

Manageability

Usability

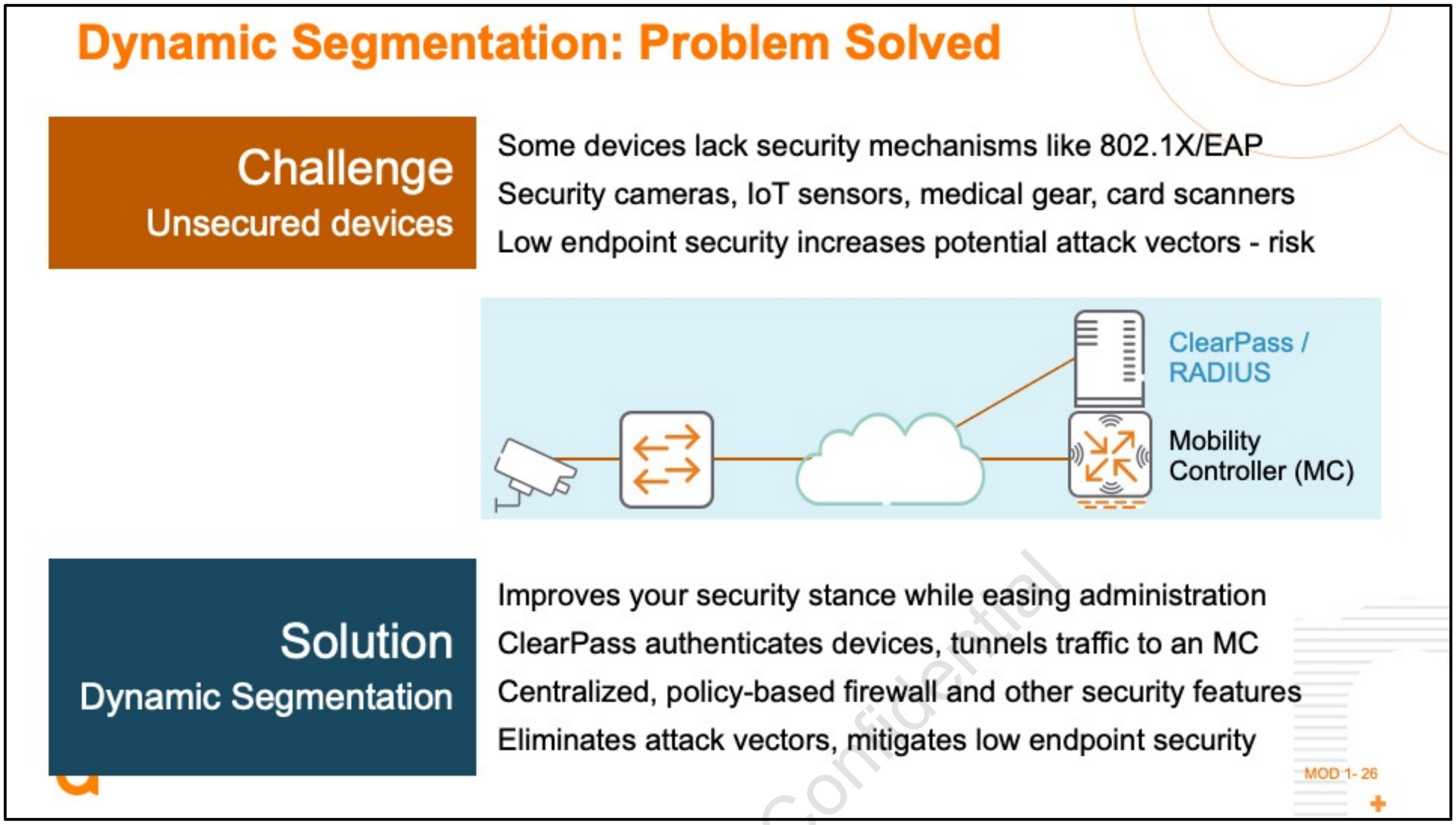
Performance

MOD 1- 25

Aruba Training-Confidential

The AOS-CX includes the time-series database, which establishes a built-in network record for the state of various components on the switch. NAE agent monitors the current state of one or more attributes in the Configuration and State Database and then writes the state to a Time-Series Database for an ongoing record They can establish baselines of normal states and detect significant deviations from those. When they detect issues, they can trigger alerts and actions and collect further data. Customers can script, and they can create scripts to obtain information about NAE agents from the REST API.

Together these components create a switch that delivers critical insights that can make the company more secure and competitive. Also, you can program the switch to act on those insights, permitting customers to transform their operations with automated workflows and business application integration.



Dynamic Segmentation: Problem Solved

Challenge
Unsecured devices

Some devices lack security mechanisms like 802.1X/EAP
Security cameras, IoT sensors, medical gear, card scanners
Low endpoint security increases potential attack vectors - risk

Solution
Dynamic Segmentation

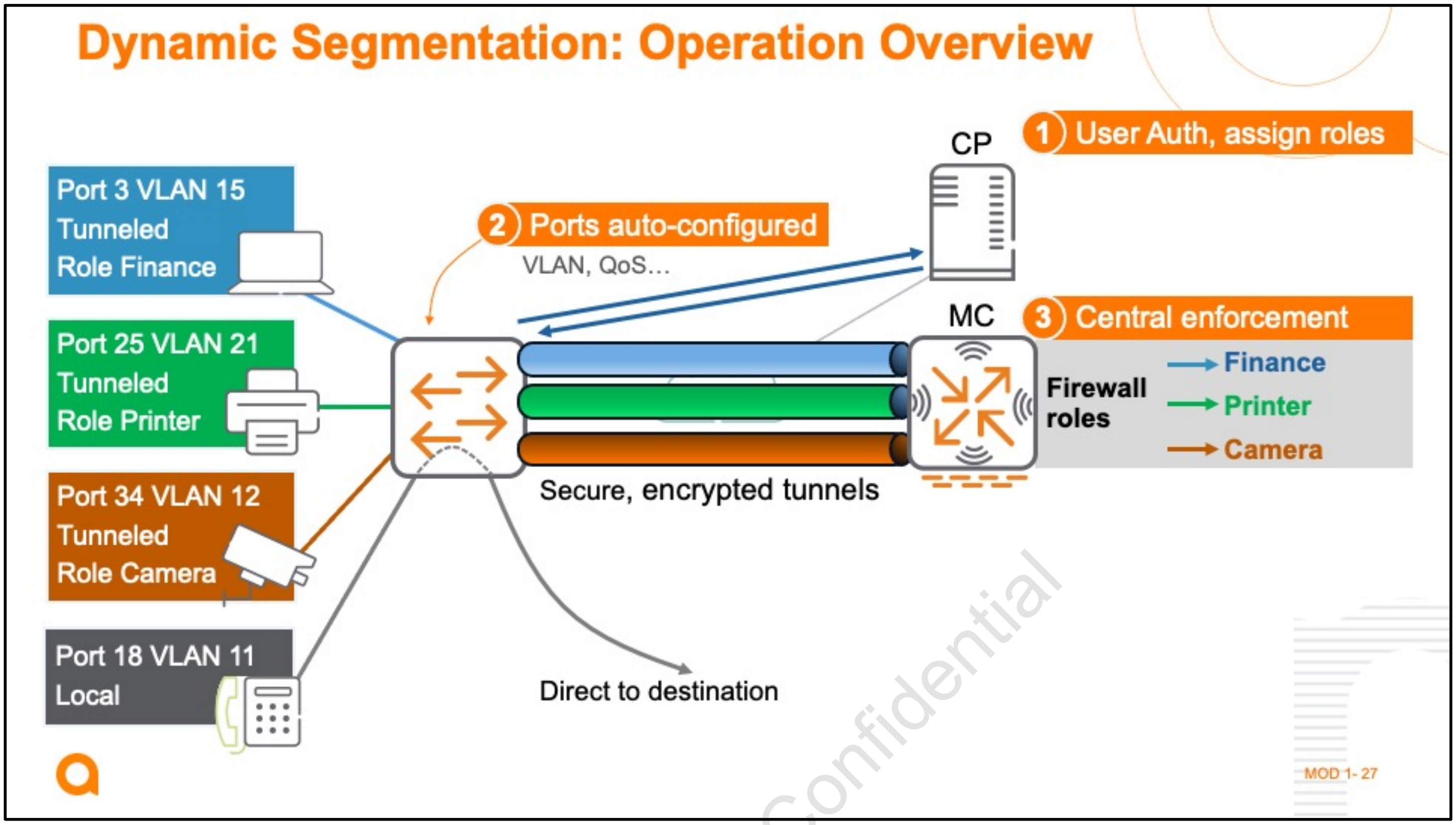
Improves your security stance while easing administration
ClearPass authenticates devices, tunnels traffic to an MC
Centralized, policy-based firewall and other security features
Eliminates attack vectors, mitigates low endpoint security

MOD 1- 26

Aruba Training-Confidential

All modern desktops, laptops, tablets, and smart phones support robust security mechanisms like 802.1X/EAP. But many devices lack these security mechanisms. This includes specialty devices, and some devices that rely on Power over Ethernet (PoE), such as security cameras, payment card readers, Internet of Things (IoT), and medical devices.This lack of endpoint security can pose a serious risk to your infrastructure.

| A switch with dynamic segmentation improves your security stance while easing administration. It authenticates these devices using ClearPass, and tunnels the traffic to a mobility controller, which centralizes formerly disparate Access Lists (ACLs) and firewall rules into single, intuitive policy system. You can continue to use these devices without compromising network security.



Dynamic Segmentation: Operation Overview

Port 3 VLAN 15
Tunneled
Role Finance

Port 25 VLAN 21
Tunneled
Role Printer

Port 34 VLAN 12
Tunneled
Role Camera

Port 18 VLAN 11
Local

CP

1 User Auth, assign roles

2 Ports auto-configured
VLAN, QoS…

MC

3 Central enforcement

Firewall roles

→ Finance
→ Printer
→ Camera

Secure, encrypted tunnels

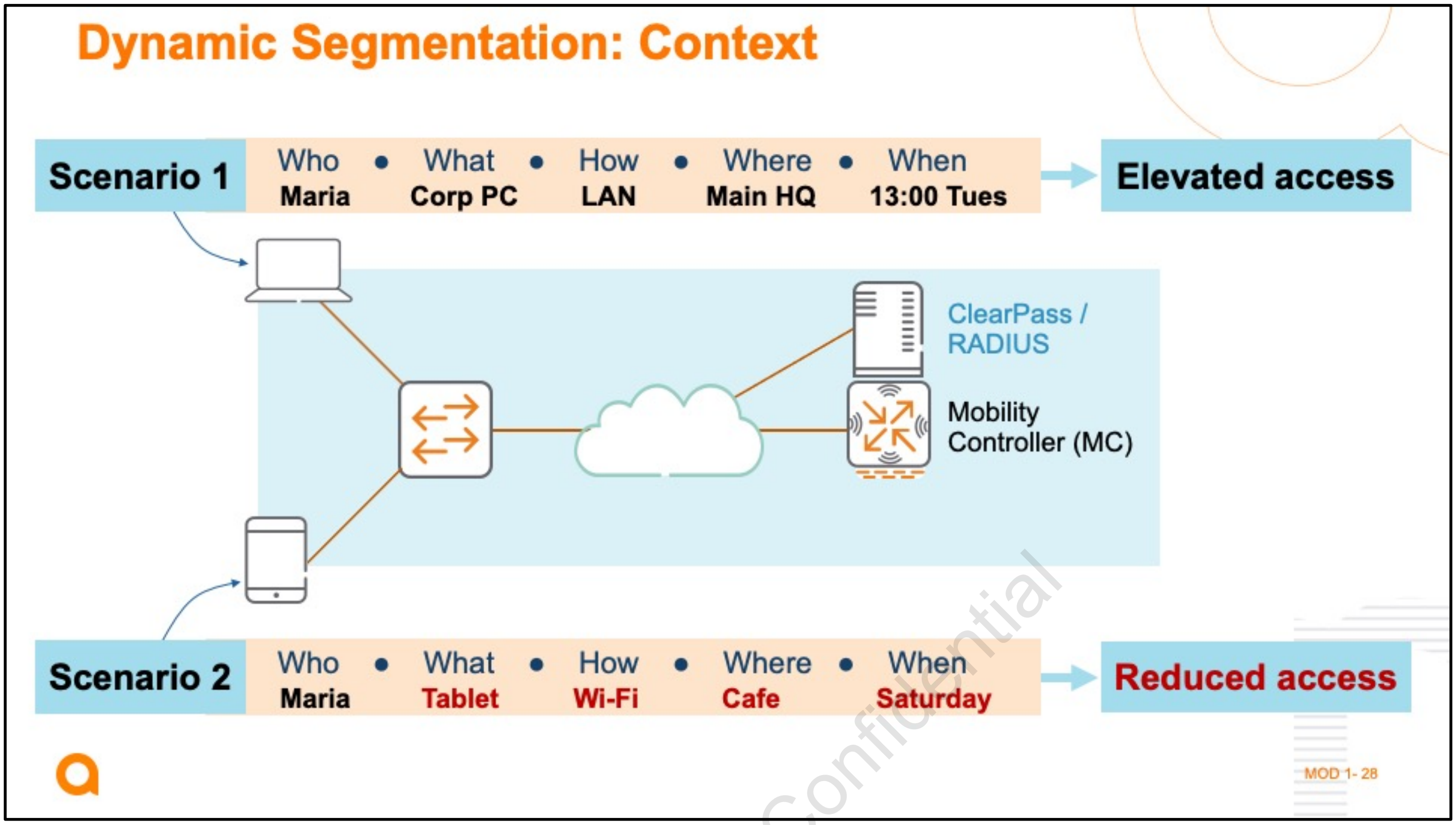
Direct to destination

MOD 1- 27

ClearPass profiles each device, deems appropriate endpoints as “acceptable”, and sends accept messages to the controller. The messages include tunnel information, assigned VLAN, and secondary role. In the figure, a user was profiled into the Finance role, and so their port is configured for VLAN 15, and their traffic is securely tunneled to the controller. The controller applies centralized policy enforcement to limit finance traffic as appropriate. Printers and cameras are similarly treated, as shown in the figure.

ClearPass can also indicate if device traffic should be locally forwarded by the switch, directly toward its ultimate destination, without tunneling to the controller. You sacrifice some of that centralized PEF enforcement, but you accommodate devices that are uber-sensitive to delay and jitter.

Aruba Training-Confidential



Dynamic Segmentation: Context

Scenario 1 Who • What • How • Where • When
Maria Corp PC LAN Main HQ 13:00 Tues → **Elevated access**

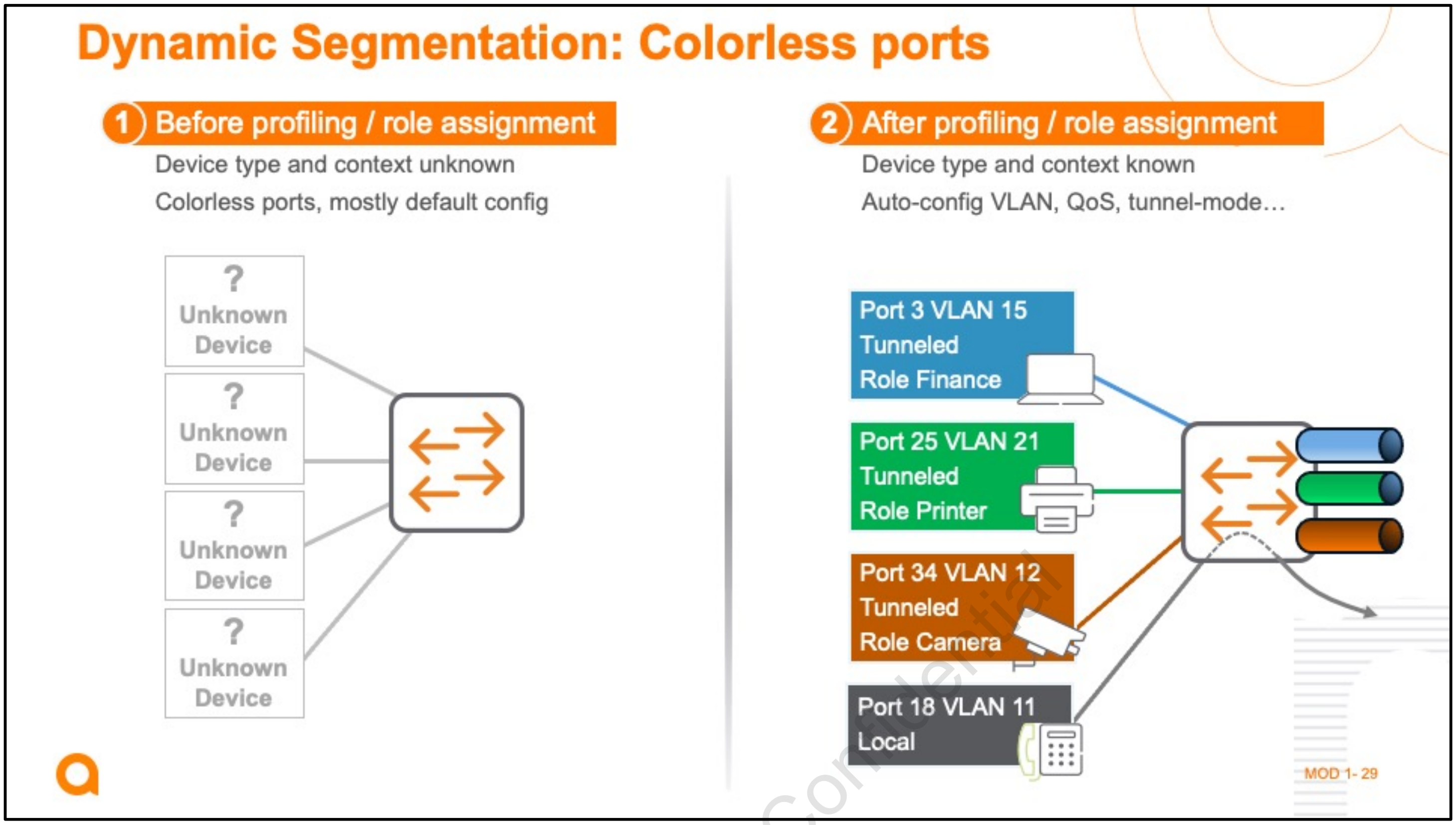
Scenario 2 Who • What • How • Where • When
Maria Tablet Wi-Fi Cafe Saturday → **Reduced access**

MOD 1- 28

ClearPass can leverage device profiling to auto-determine device types, and a rich set of contextual “who-what-how-where-when” information to assign policies. Maria’s login credentials might give her elevated access when using her corporate PC connected via wireless at the main building, during business hours. But when she uses her personal tablet connected from a local café on Saturday, she may have reduced access.

Among other things, this helps you to implement colorless ports on your switches.

Aruba Training-Confidential



Dynamic Segmentation: Colorless ports

1) Before profiling / role assignment

Device type and context unknown

Colorless ports, mostly default config

2) After profiling / role assignment

Device type and context known

Auto-config VLAN, QoS, tunnel-mode…

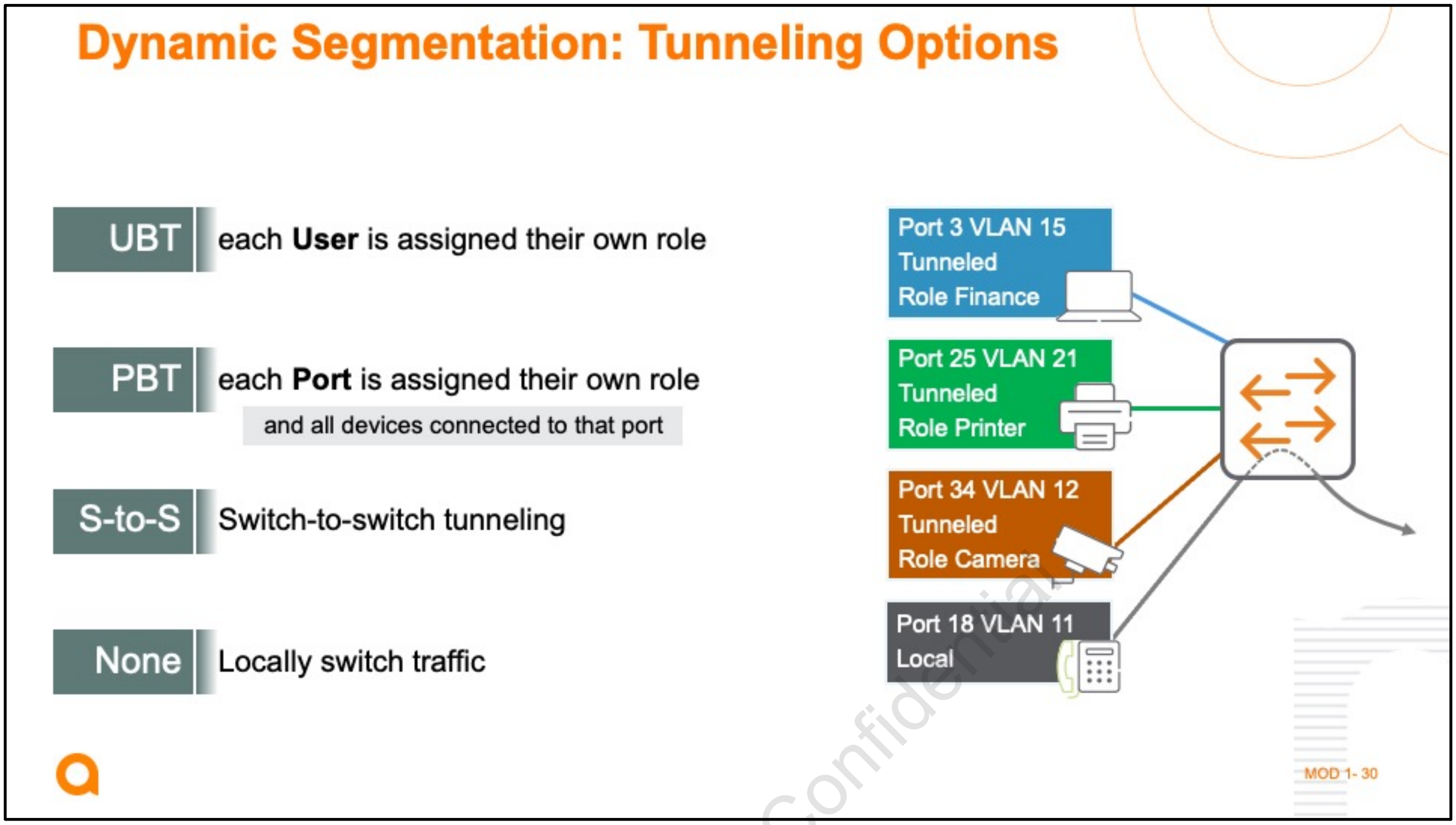
? Unknown Device
? Unknown Device
? Unknown Device
? Unknown Device
Port 3 VLAN 15 Tunneled Role Finance
Port 25 VLAN 21 Tunneled Role Printer
Port 34 VLAN 12 Tunneled Role Camera
Port 18 VLAN 11 Local
MOD 1- 29

Aruba Training-Confidential

A colorless port is basically a port with default characteristics that are then dynamically changed based on device characteristics and authentication details. Before profiling, context assessment, and role assignment the device type and context is largely unknown.

Manually configuring VLANs, QoS, ACLs, and more is tedious, time-consuming, and error prone – and you are now free from much of this!

Leave the ports at default settings. Connected ^devices are automatically profiled context is assessed, and roles are assigned. Today the finance person connected to port 3 on the switch. Tomorrow they may connect to port 12. It doesn't matter – dynamic segmentation "colors" the port automatically.



Dynamic Segmentation: Tunneling Options

UBT each **User** is assigned their own role

PBT each **Port** is assigned their own role
and all devices connected to that port

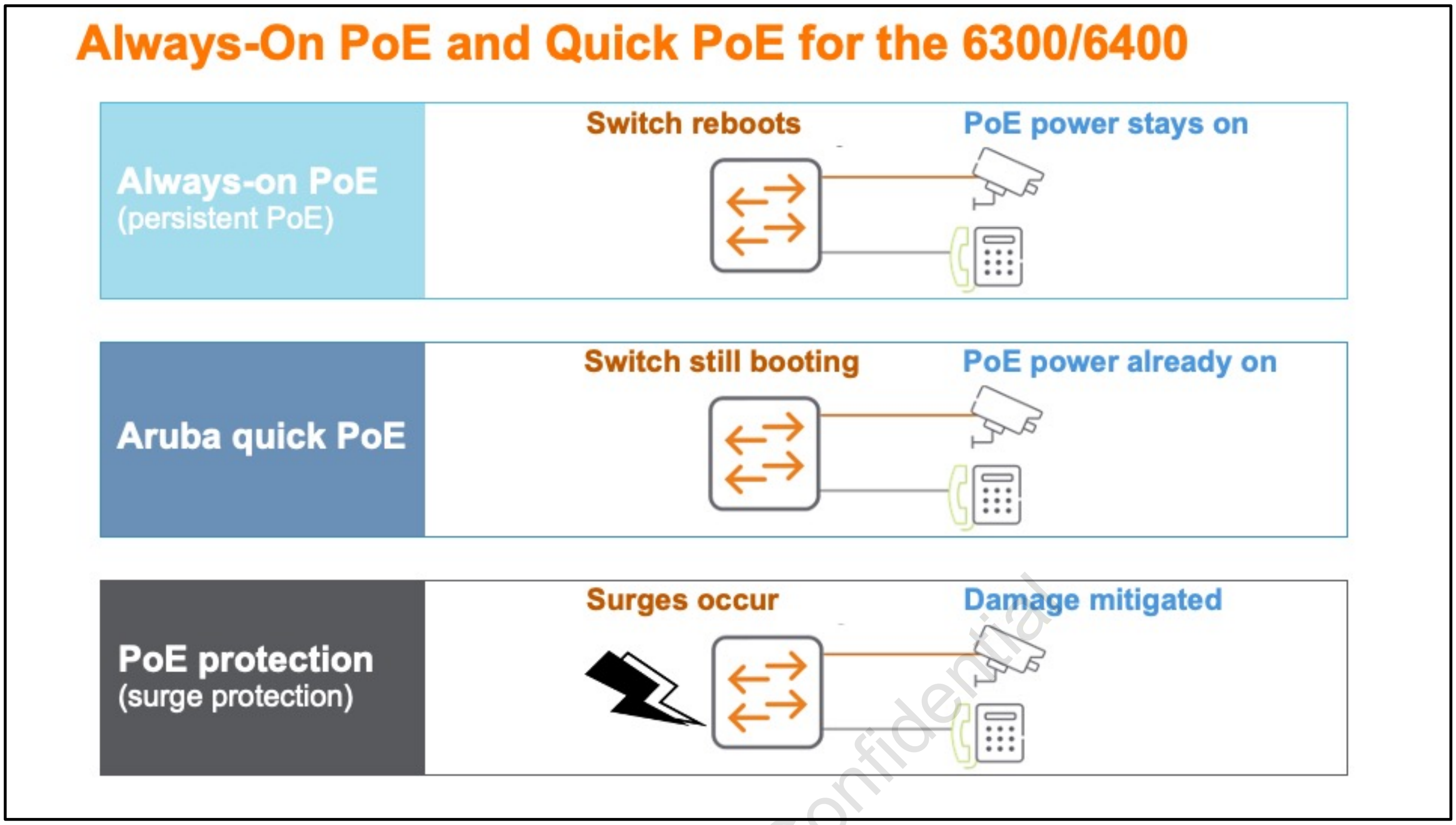
S-to-S Switch-to-switch tunneling

None Locally switch traffic

The figure introduces you to tunneling options related to dynamic segmentation

- **User-Based Tunnel (UBT)**: each user is assigned their own role
- **Port-Based Tunnel(PBT)**: each port (and all the devices connected to the same port) are assigned the same role (PBT is not currently supported in AOS-CX 10.4 but there are plans to add it in a future release)
- **Switch-to-switch tunneling:** planned release in AOS-CX 10.5
- **None**: Exempt certain traffic from tunneling by performing local switching/forwarding (like voice, for example)

Aruba Training-Confidential



Always-On PoE and Quick PoE for the 6300/6400

Aruba Training-Confidential

Another valuable feature of the ArubaOS-CX 6300/6400 switches relates to Power over Ethernet (PoE).

Always-on PoE (persistent PoE)

Switches have supported PoE standards (IEEE 802.3af, 802.3at) for years, but now we want to maintain PoE power across planned and unplanned software upgrades – especially for planned upgrades. Aruba calls this feature Always-On PoE, which some may refer to as Persistent PoE. This feature ensures that end devices do not power cycle on switch reboot. This is especially important for some IoT and Zigbee-based devices, which may behave oddly or take a long time to reboot after a switch upgrade.

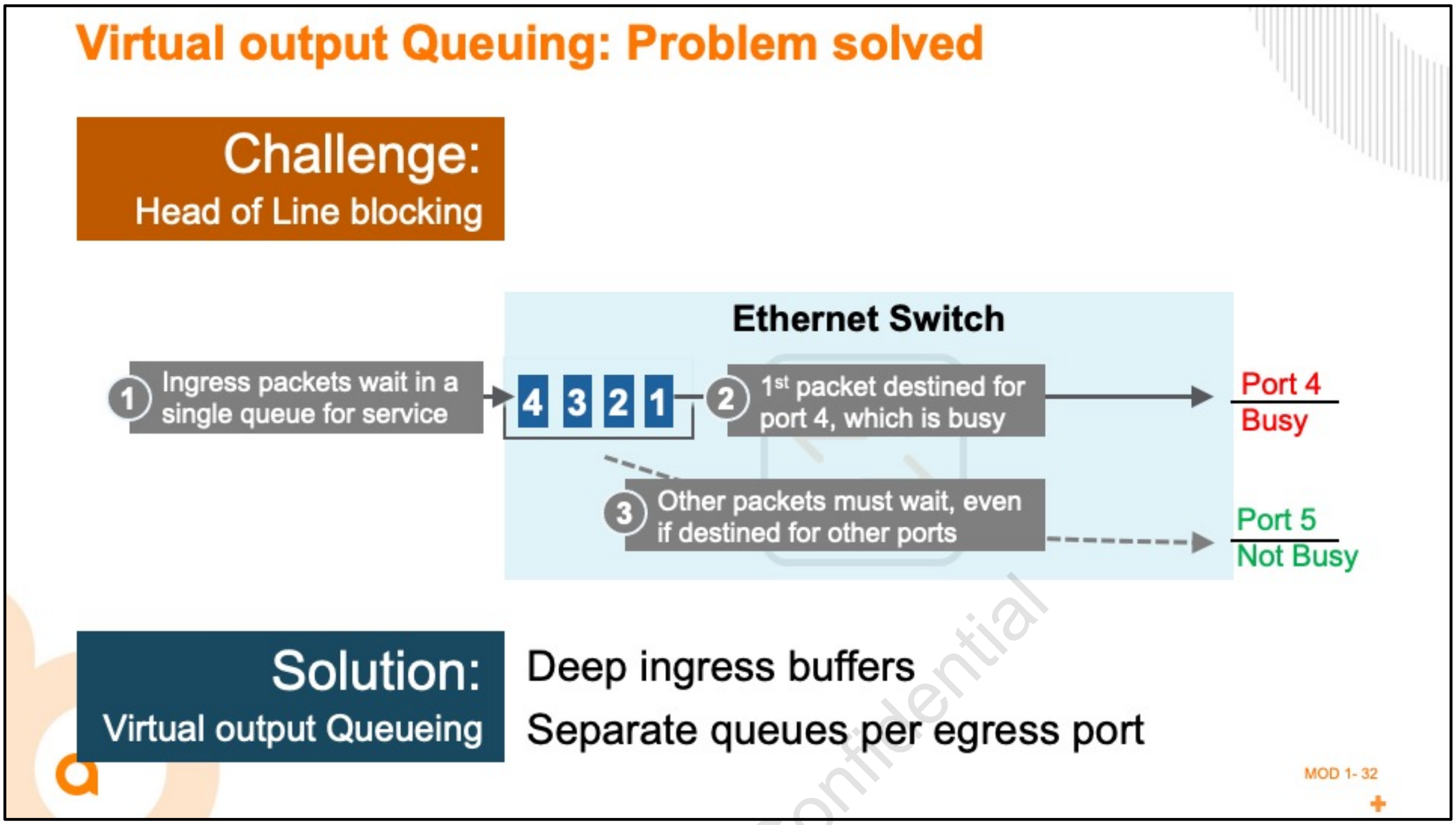
Aruba quick PoE

Quick PoE (also known as Fast PoE) can power devices the moment you connect switch power, without having to wait for the AOS-CX to completely boot. This is especially important for PoE-based lighting and other IoT devices.

The typical use case for this would be a device power outage. Aruba 6300 Module SKUs are to be release with IEEE802.3bt or 4-pair PoE enabled. The Fixed SKUs are 2-Pair PoE, but the mainboard is hardware-designed to support a 4-pair PoE controller and is 4-pair (60W per port) ready.

PoE protection (surge protection)

One method of surge protection relies on a very quick reacting electronic component called a Shockley Diode. When there is a negative surge, current flows from The switch Power Supply Unit (PSU) to RJ Lineside, thus protecting the PSE chip and downstream devices.



Virtual output Queuing: Problem solved

Challenge: Head of Line blocking

Ethernet Switch
1 Ingress packets wait in a single queue for service
4 3 2 1
2 1st packet destined for port 4, which is busy
3 Other packets must wait, even if destined for other ports
Port 4
Busy
Port 5
Not Busy

Solution: Virtual output Queueing

Deep ingress buffers

Separate queues per egress port

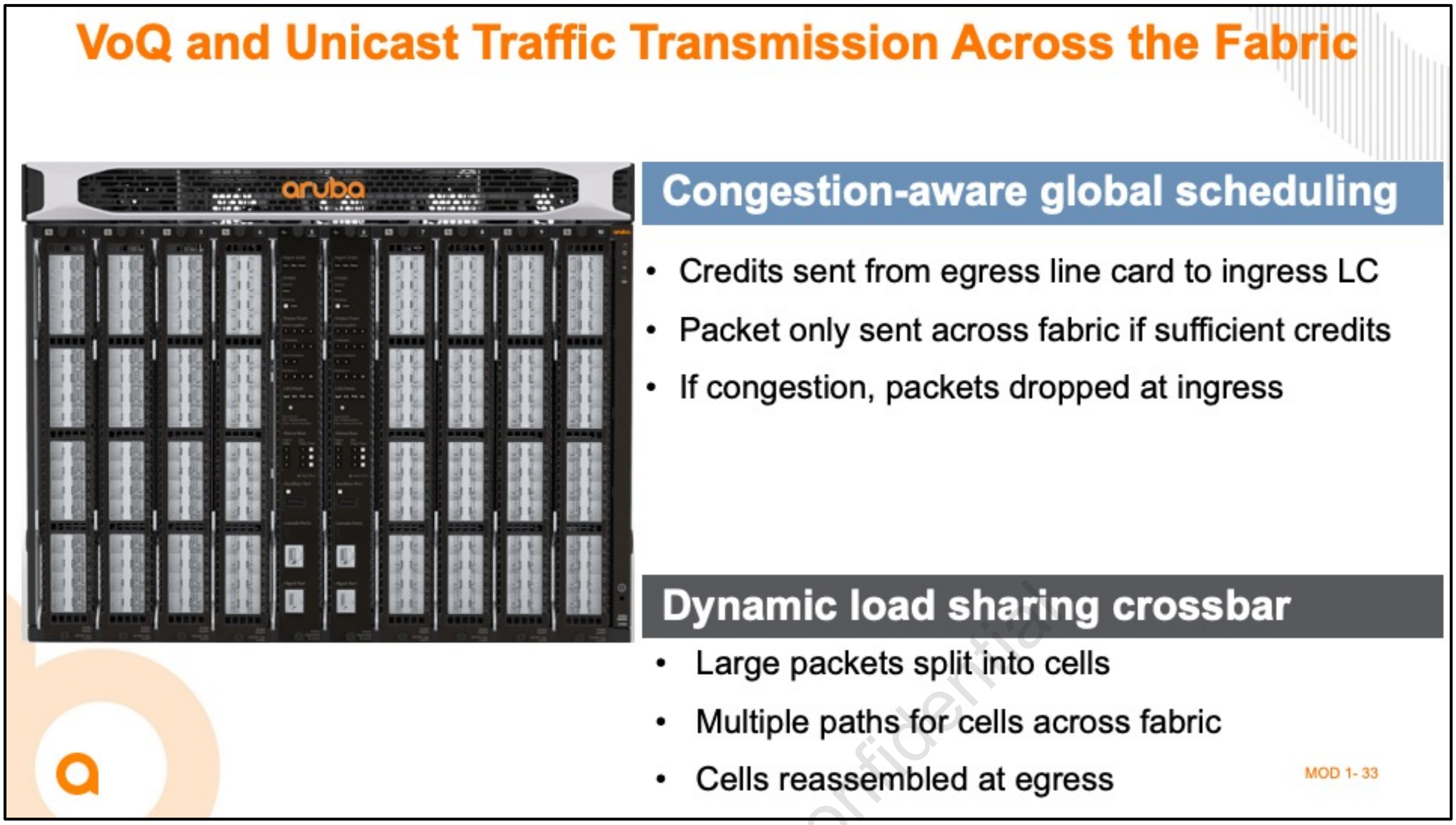
MOD 1- 32

Let’s take a closer look at how traffic crosses the switch’s backplane from one line card through the fabric to another line card(LC).

The figure shows that 4 packets have arrived at some interface, sitting in a queue, waiting for service. If the ingress buffer used a single queue, Head Of Line (HOL) blocking could delay traffic. This occurs when the first packet in the queue (at the “head of the line”) is destined out a congested port, it delays all packets behind it, even though those that are destined to non-congested ports.

| AOS-CX switches use an intra-switch queuing method called Virtual output Queuing (VoQ). VOQ prevents this problem by providing deep ingress buffers with separate queues for each egress port. Physically the ingress buffer consists of internal and external DRAM banks with most traffic typically being buffered in the internal DRAM.

Aruba Training-Confidential



VoQ and Unicast Traffic Transmission Across the Fabric

Congestion-aware global scheduling

- Credits sent from egress line card to ingress LC
- Packet only sent across fabric if sufficient credits
- If congestion, packets dropped at ingress

Dynamic load sharing crossbar

- Large packets split into cells
- Multiple paths for cells across fabric
- Cells reassembled at egress

MOD 1- 33

Aruba Training-Confidential

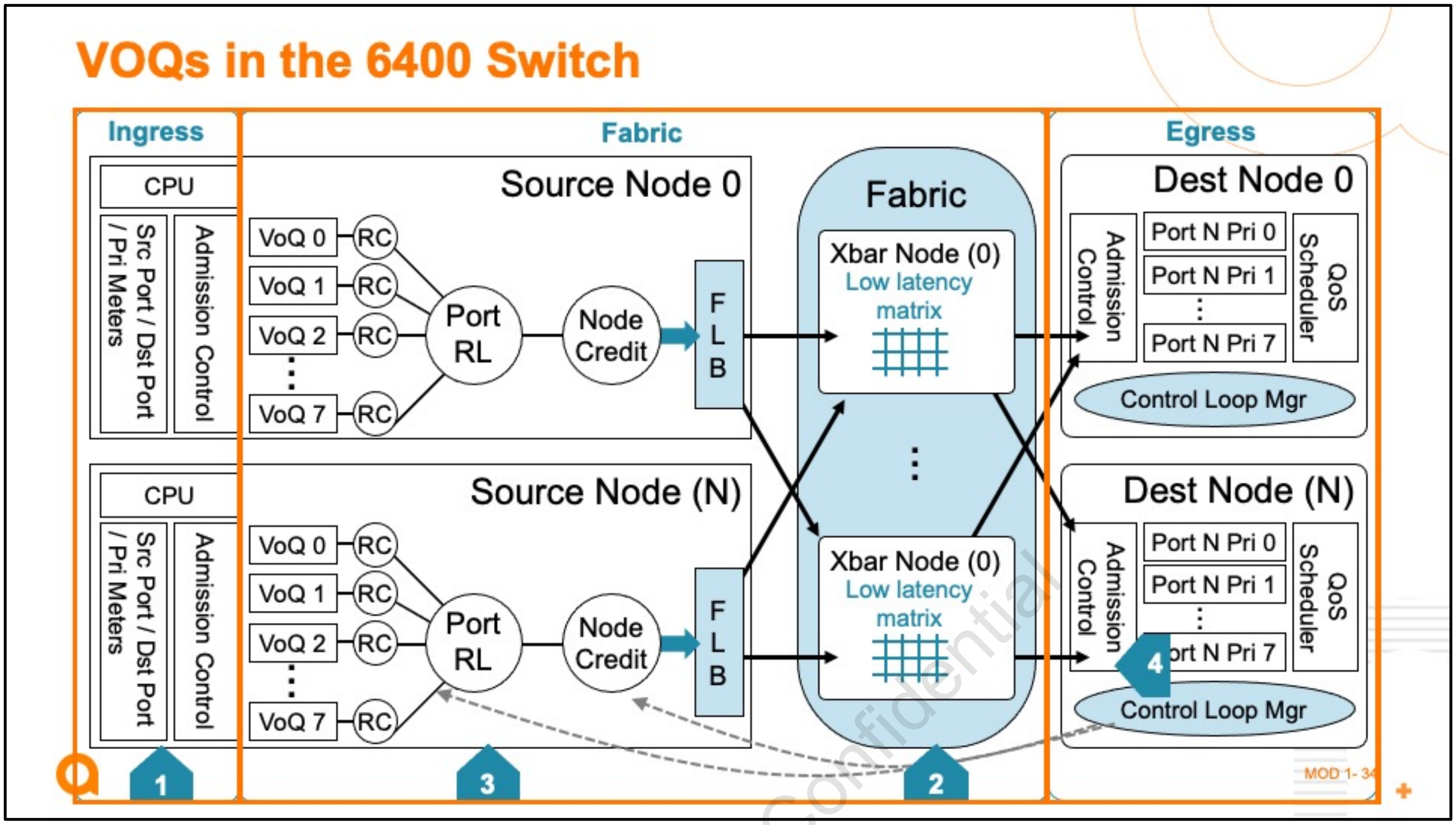
Congestion-aware global scheduling

On AOS-CX switches, traffic is scheduled for transmission across the fabric based on messages between the ingress and egress line cards. The egress line card sends credits to the ingress line card, indicating how many bits the egress buffer can receive based on current queue fullness. In this way, the scheduling is aware of global congestion.

If the ingress line card receives enough credits, it forwards the packet through the fabric to the egress card. If packets cannot be forwarded due to lack of room in the egress buffer, they are dropped at ingress. The ports have shallower egress queues because by the time traffic reaches that point, it should be rapidly forwarded. Packets are always dropped at the ingress buffer, never at the egress.

Dynamic load sharing crossbar fabric

The fabric uses a dynamic load share crossbar architecture to connect line cards together. Larger packets are split into cells, which are dispatched across different paths based on both random assignment and a protocol that is aware of path congestion. Remember that each line card has 12 connections to each fabric card, so several paths are available. The egress card reassembles the cells into a packet.



VOQs in the 6400 Switch

This figure shows VOQ for the 6400 switches. Note that VOQ is supported on **ALL** AOS-CX switches.

Aruba Training-Confidential

Overview

In a VOQ architecture, the egress side selects which traffic will cross the fabric. In the 6400 Switch Series, this block is called the Traffic Regulator, which sends small messages across the fabric to tell ingress VOQs how much they can send into the fabric

The objective is to keep just enough packets in egress queues to keep the port fully busy given the port speed. Regarding how VOQs are selected, the port's schedule-profile is also a factor. For example, with Deficit Weighted Round Robin (DWRR), heavier weighted queues get more selections.

| Ingress

So, the ingress pipeline processes the frame with information like egress port(s) and queue, drop eligibility (color). Then the Ingress Admission Control puts the frame's descriptor record in the respective VOQ – the frame itself is now in the packet buffer

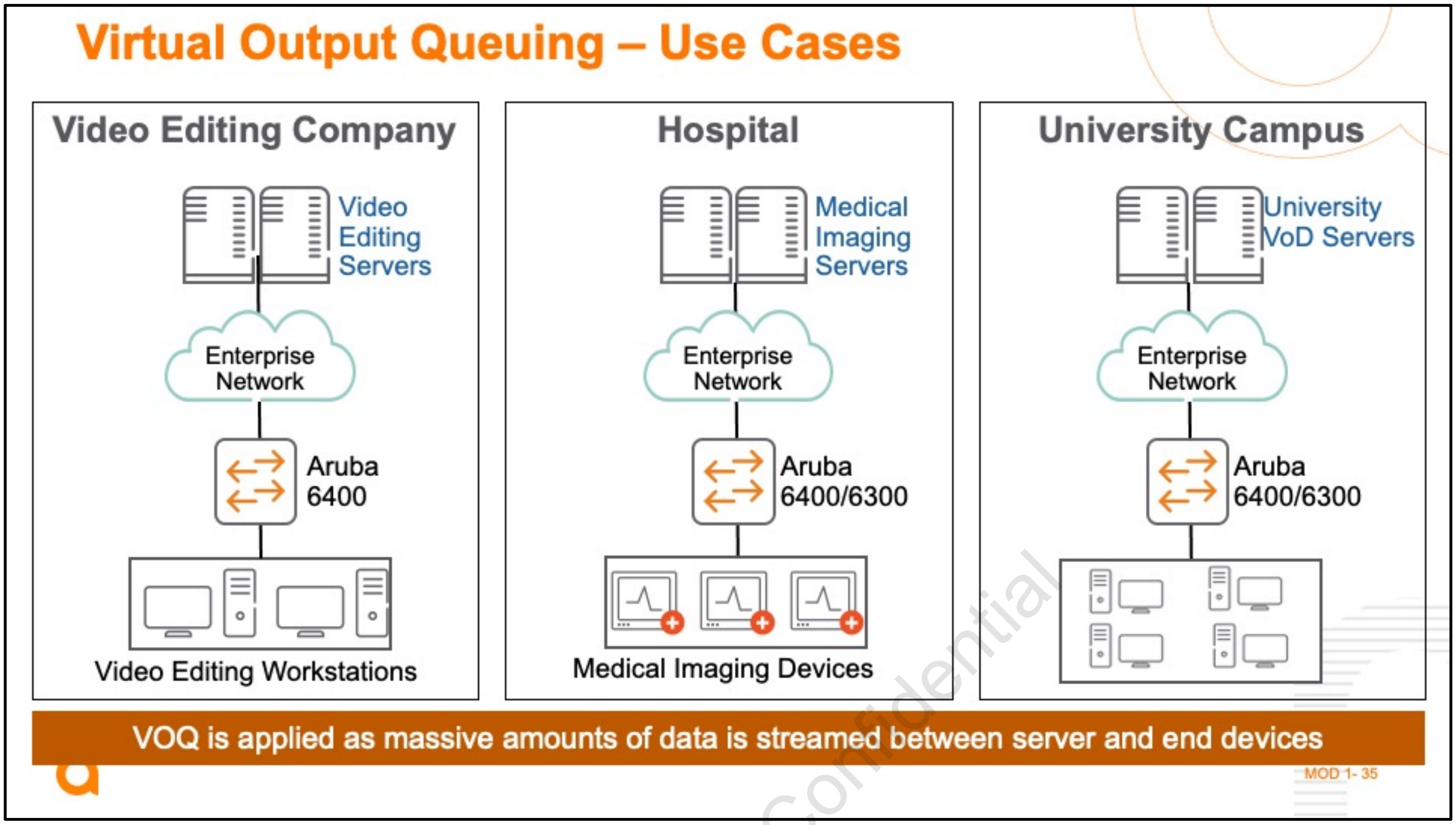
| Fabric

The process starts at the Control Loop Manager in the egress line card (destination node). Along with the Admission Control Process, it monitors all egress queues in the module, including congestion. Depending on the queue's available space, it sends "credit" messages to the ingress line cards (source node)

The Fabric Load Balancer(FLB) in the source node transmits the frames with most credits across the fabrics. Packets with the same credit level are selected using a source/destination node hashing algorithm . In multi-ASIC fabrics packets do not have a pre-assigned dedicated path

| Egress

The Egress Admission Control places frames in their corresponding queue for transmission



Virtual Output Queuing – Use Cases

Video Editing Company

Video Editing Servers

Enterprise Network

Aruba 6400

Video Editing Workstations

Hospital

Medical Imaging Servers

Enterprise Network

Aruba 6400/6300

Medical Imaging Devices

University Campus

University VoD Servers

Enterprise Network

Aruba 6400/6300

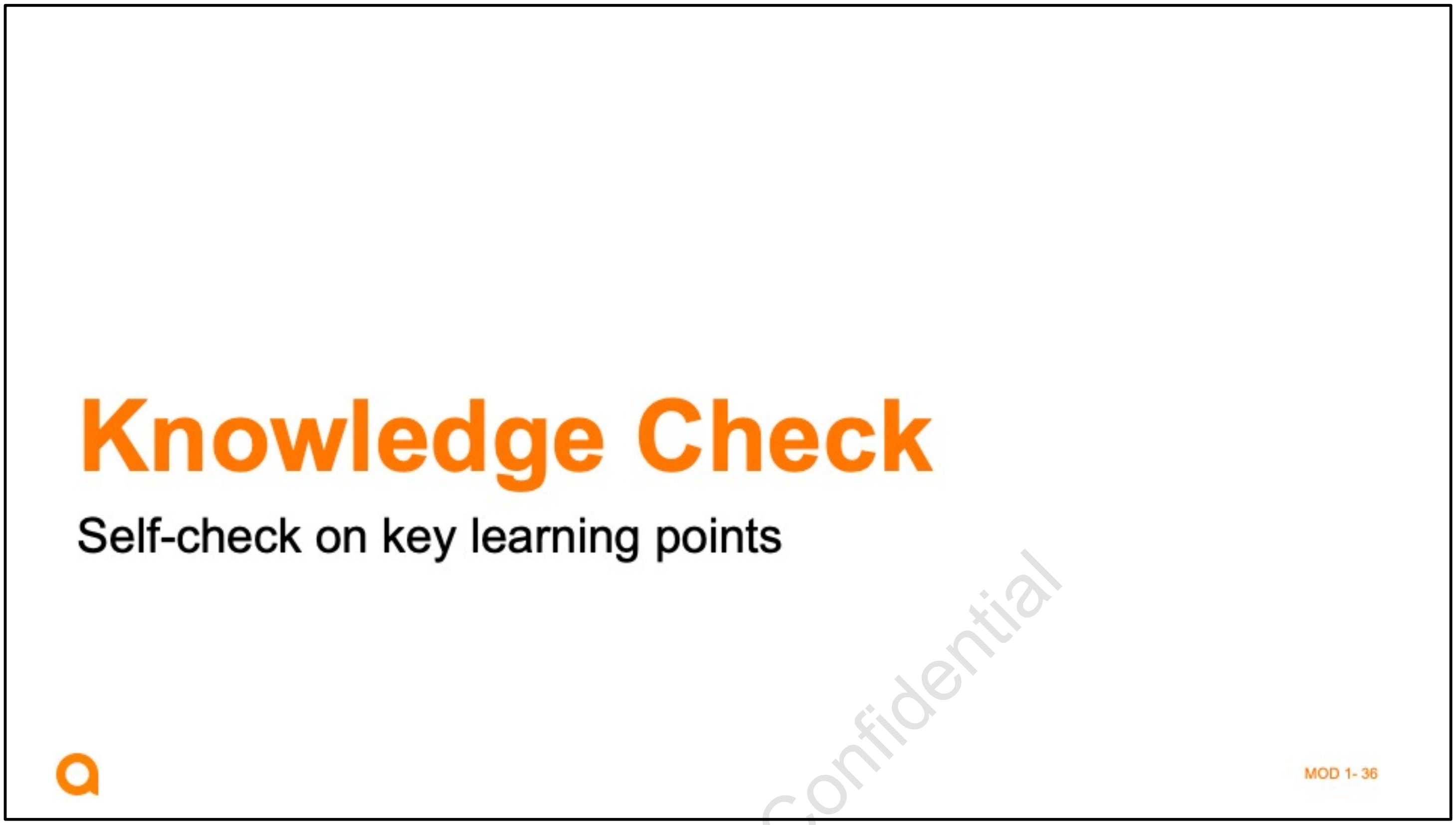
VOQ is applied as massive amounts of data is streamed between server and end devices

MOD 1- 35

Aruba Training-Confidential

The figure shows three use cases for Virtual Output Queuing:

- **Video editing company:** Many employees use video editing workstations to send and receive massive amounts of video traffic. Too many simultaneous, high-bandwidth users could congest the network. During times of congestion, the 6400 switch uses VOQ to help ensure that video traffic remains unaffected.
- **Hospital:** Sophisticated medical imaging equipment send massive amounts of imaging data between the imaging department and data-handling servers. With multiple machines in use during a busy night in the emergency room, congestion is likely. VOQ ensures that patient images are not dropped due to a busy network.
- **University campus:** Suppose that too many students are live streaming video from the university’s on-demand servers. This could cause congestion at the switch fabrics if other line cards are also streaming traffic or sending large amounts of file transfers.



Knowledge Check

Self-check on key learning points

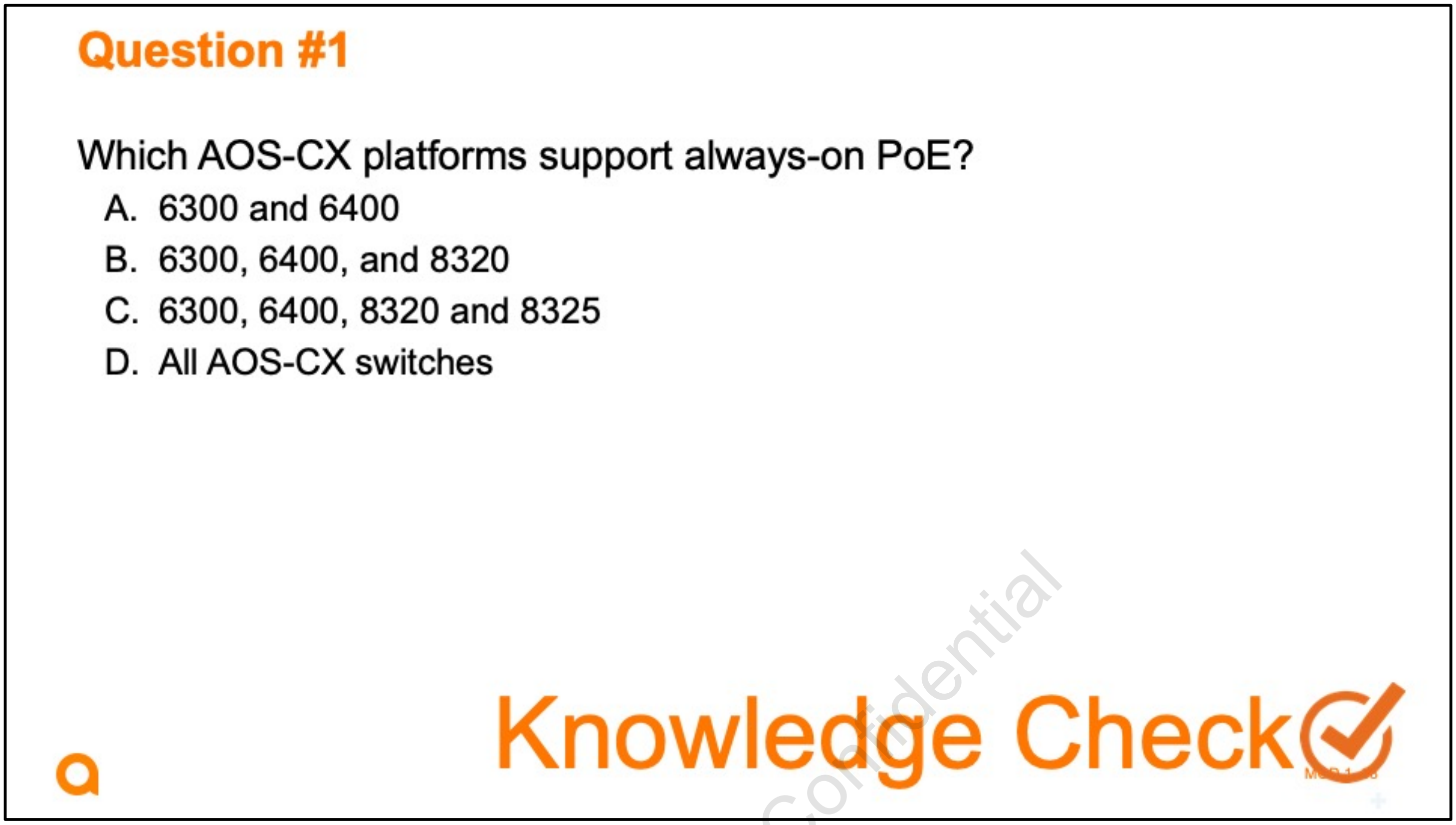
MOD 1- 36

Aruba Training-Confidential

Lab Activity

Its time for a lab activity

Aruba Training-O



Question #1

Which AOS-CX platforms support always-on PoE?

A. 6300 and 6400

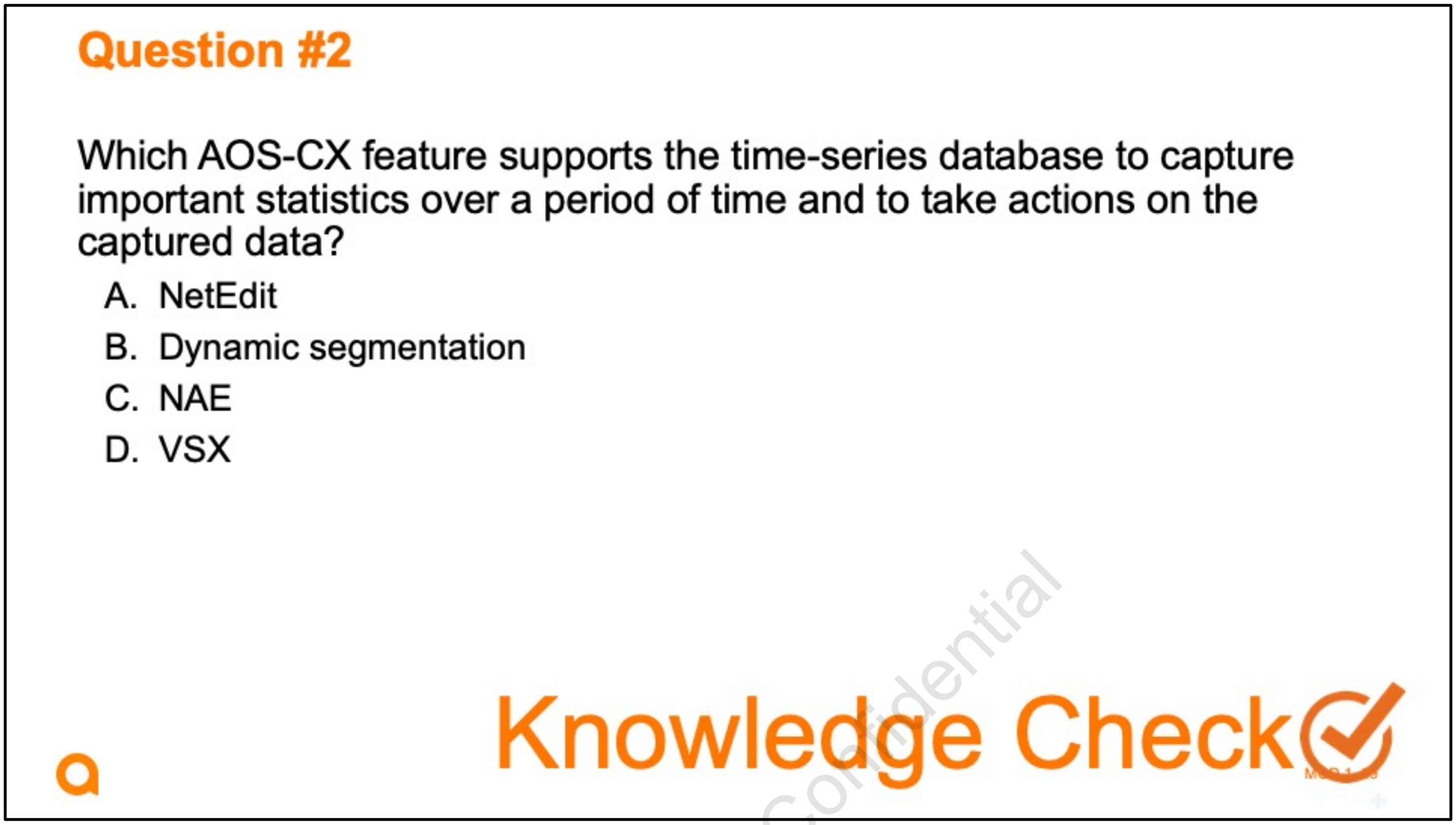
B. 6300, 6400, and 8320

C. 6300, 6400, 8320 and 8325

D. All AOS-CX switches

Knowledge Check

Aruba Training-Confidential



Question #2

Which AOS-CX feature supports the time-series database to capture important statistics over a period of time and to take actions on the captured data?

A. NetEdit

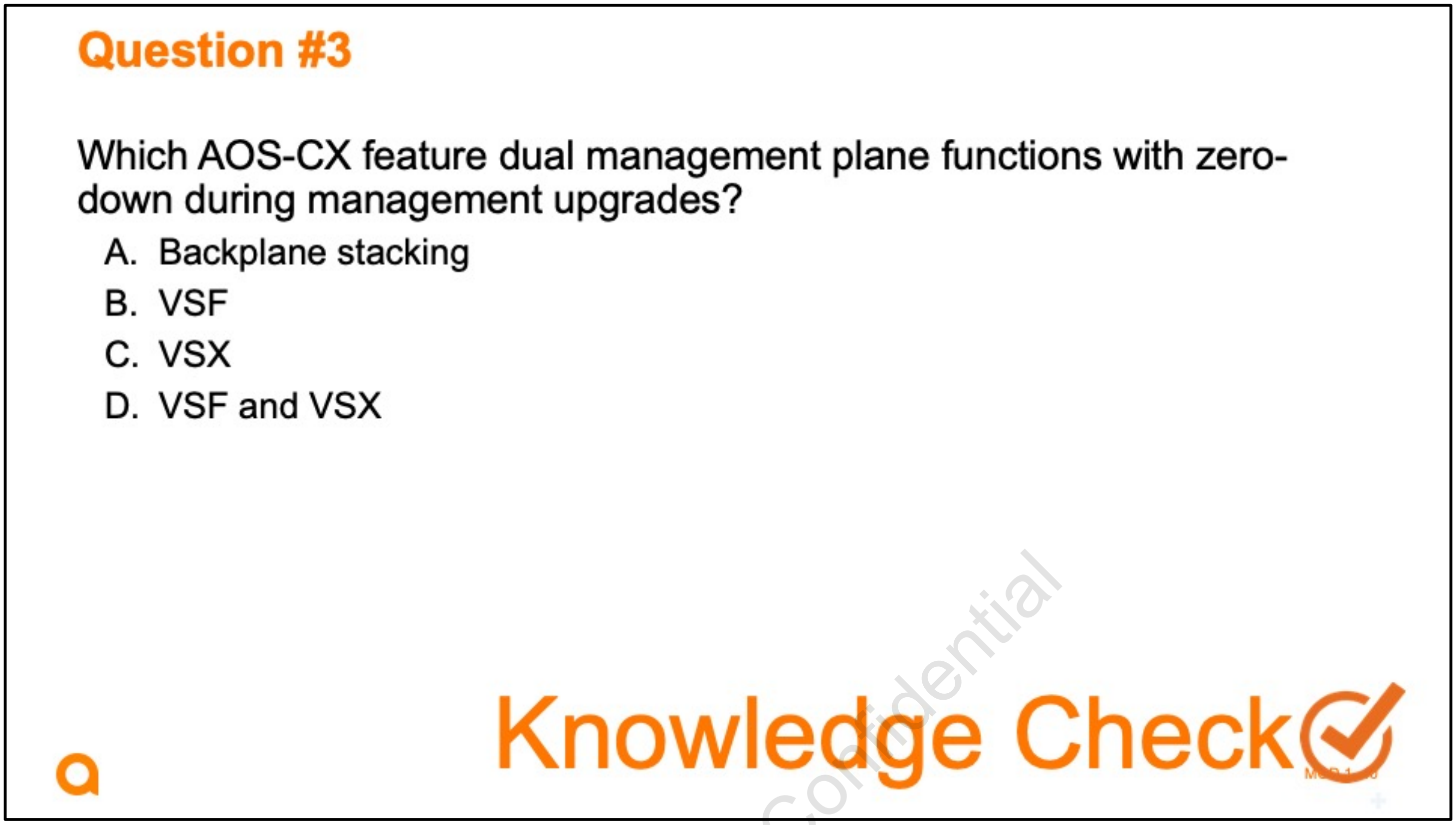
B. Dynamic segmentation

C. NAE

D. VSX

Knowledge Check

Aruba Training-Confidential



Question #3

Which AOS-CX feature dual management plane functions with zero-down during management upgrades?

A. Backplane stacking

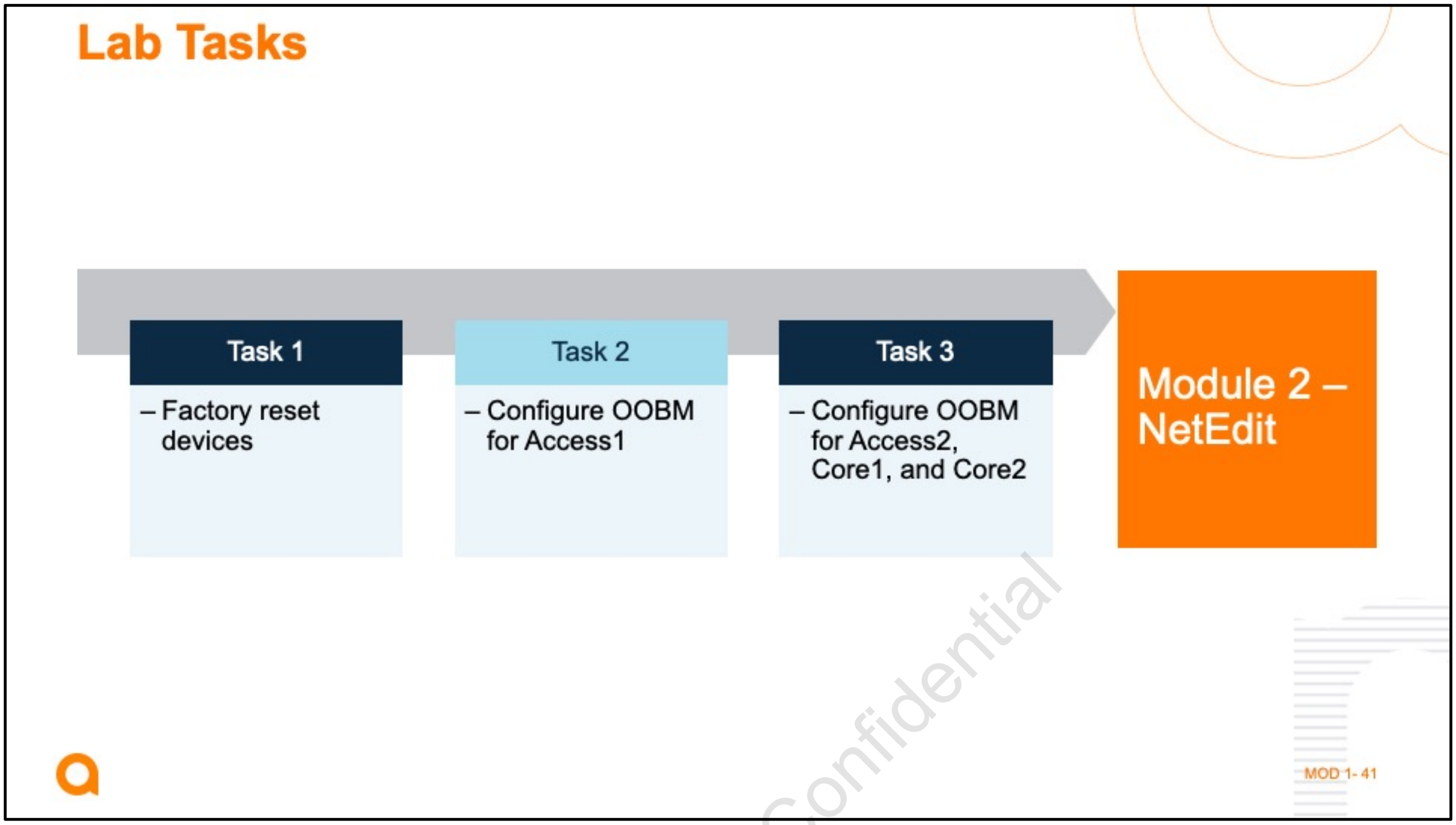
B. VSF

C. VSX

D. VSF and VSX

Knowledge Check

Aruba Training-Confidential



Lab Tasks

Task 1

– Factory reset devices

Task 2

– Configure OOBM for Access1

Task 3

– Configure OOBM for Access2, Core1, and Core2

Module 2 – NetEdit

Aruba Training-Confidential

The figure provides a brief review of lab tasks. Please see your lab guide for details.

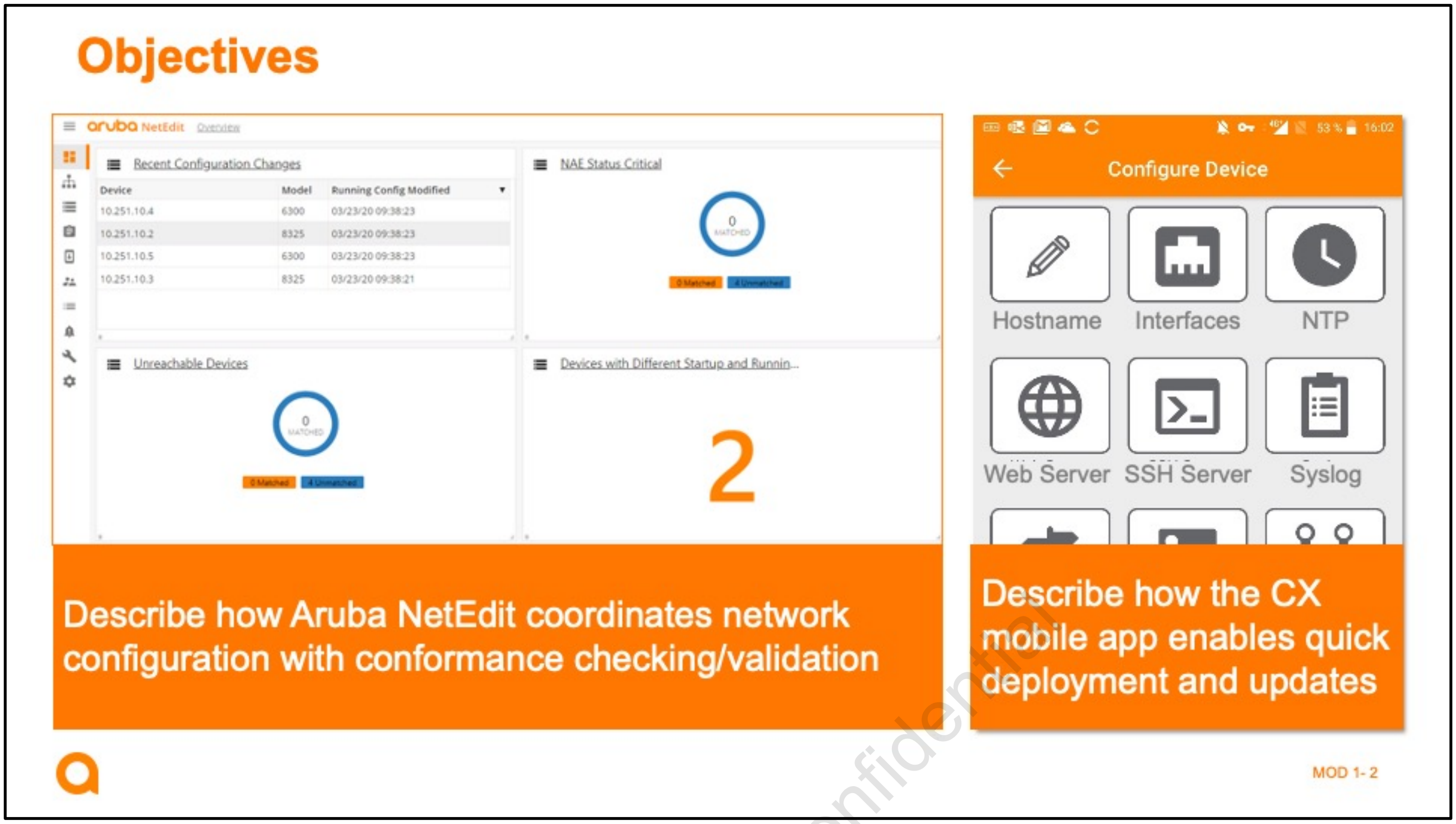
Rev # 20.21

NetEdit

Implementing ArubaOS-CX Switching

Welcome back wise learners. This is Module 2 – NetEdit.

Aruba Training-Confidential



Objectives

Device	Model	Running Config Modified
10.251.10.4	6300	03/23/20 09:38:23
10.251.10.2	8325	03/23/20 09:38:23
10.251.10.5	6300	03/23/20 09:38:23
10.251.10.3	8325	03/23/20 09:38:21

Describe how Aruba NetEdit coordinates network configuration with conformance checking/validation

Describe how the CX mobile app enables quick deployment and updates

MOD 1- 2

Aruba Training-Confidential

After completing this module, you should be able to:

Describe how the Aruba NetEdit tool coordinates network configuration with conformance checking and validation

Describe how the CX mobile App provides for quick switch deployments and updates

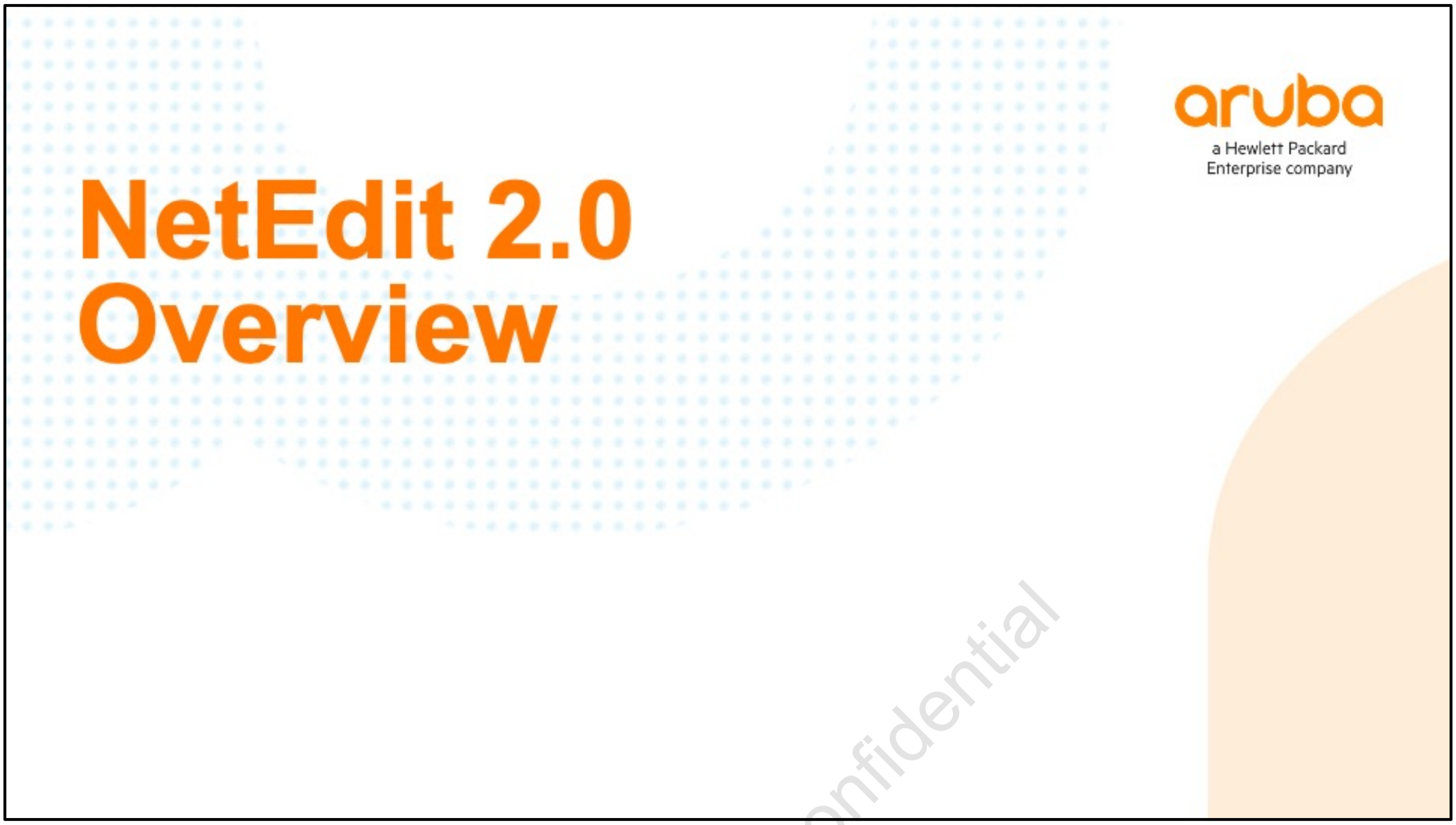
Overview

- NetEdit Overview
- Initial Setup
- Key Cababilities
- AOS-CX Mobile App
- Lab Activity

MOD 1- 3

Aruba Training-Confidential

Aruba NetEdit tool coordinates network configuration with conformance checking and validation. The CX mobile App version 10.4 updates for quick deployment and update switches. You'll go through the sections shown here in this module to become quite knowledgeable about netedit.

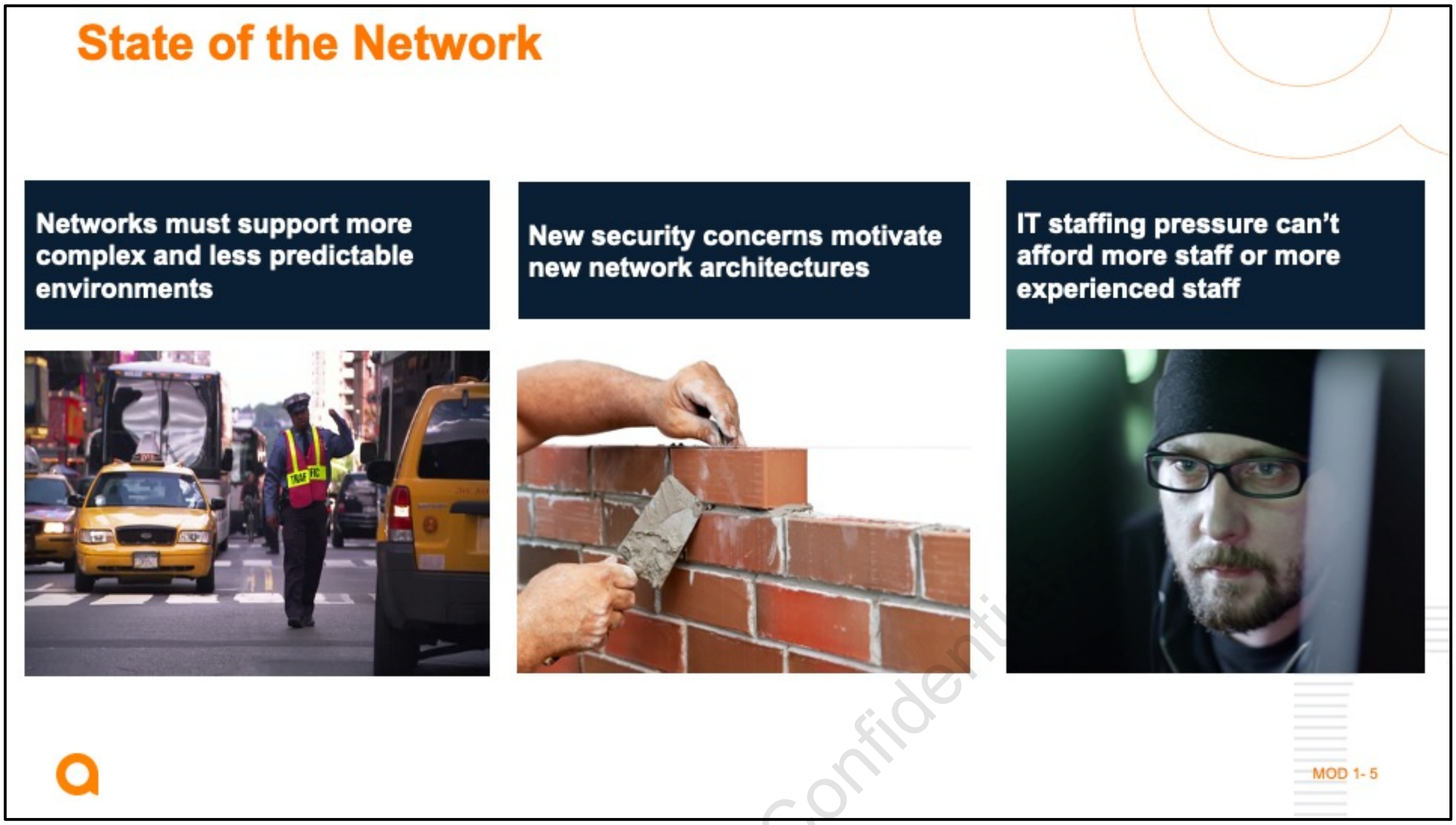


NetEdit 2.0 Overview

aruba
a Hewlett Packard Enterprise company

This section provides an overview of common management problems and how to use NetEdit to deal with those issues.

Aruba Training-Confidential



State of the Network

Networks must support more complex and less predictable environments

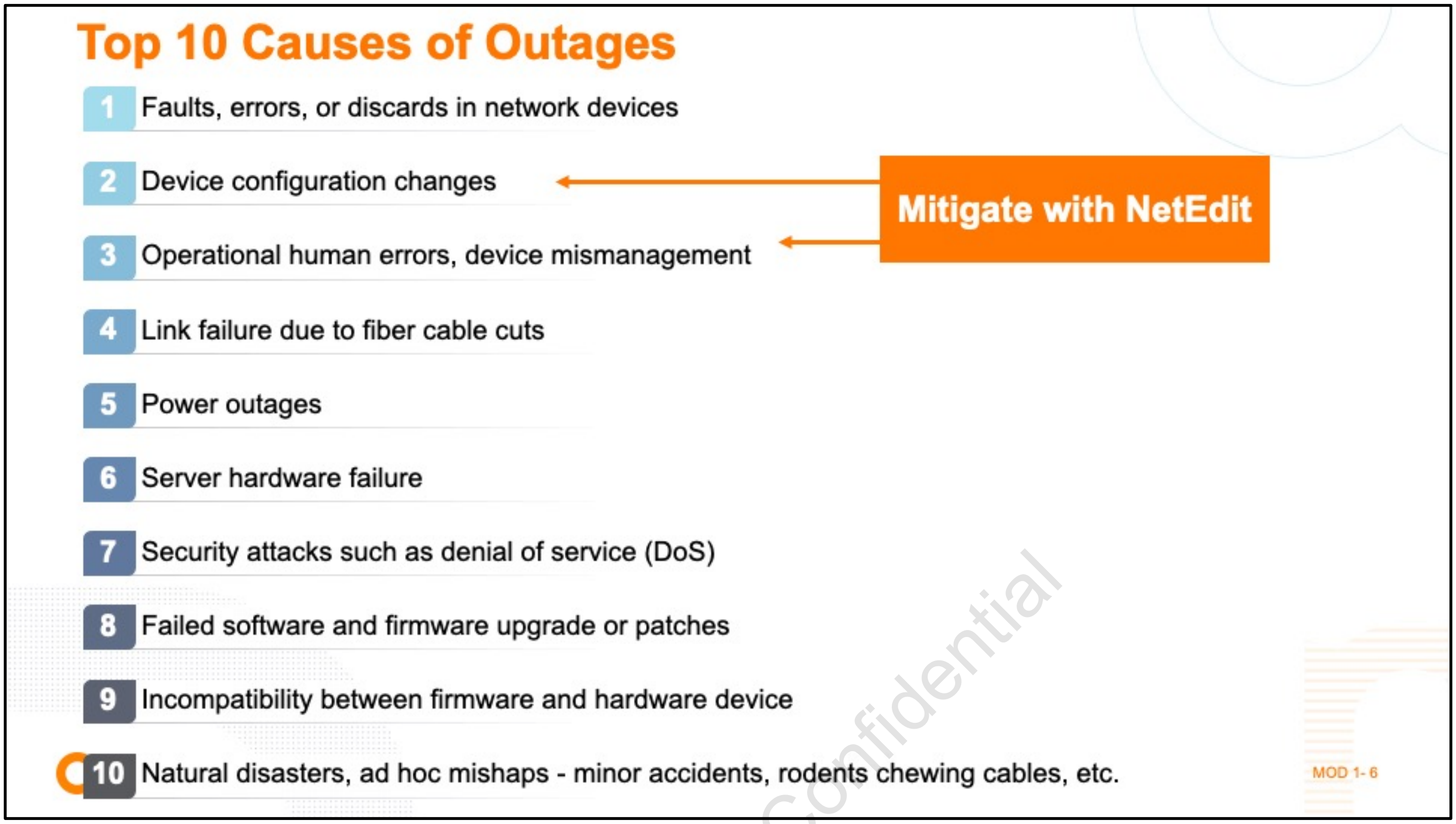
New security concerns motivate new network architectures

IT staffing pressure can't afford more staff or more experienced staff

MOD 1- 5

NetEdit arms IT teams with the power to smoothly coordinate end-to-end service rollouts, automate rapid network-wide changes, and ensure policy conformance after network updates. This intelligent assistance and continuous validation help assure that network-wide configurations are consistent and compliant.

Aruba Training-Confidential



Top 10 Causes of Outages

1 Faults, errors, or discards in network devices

2 Device configuration changes

3 Operational human errors, device mismanagement

4 Link failure due to fiber cable cuts

5 Power outages

6 Server hardware failure

7 Security attacks such as denial of service (DoS)

8 Failed software and firmware upgrade or patches

9 Incompatibility between firmware and hardware device

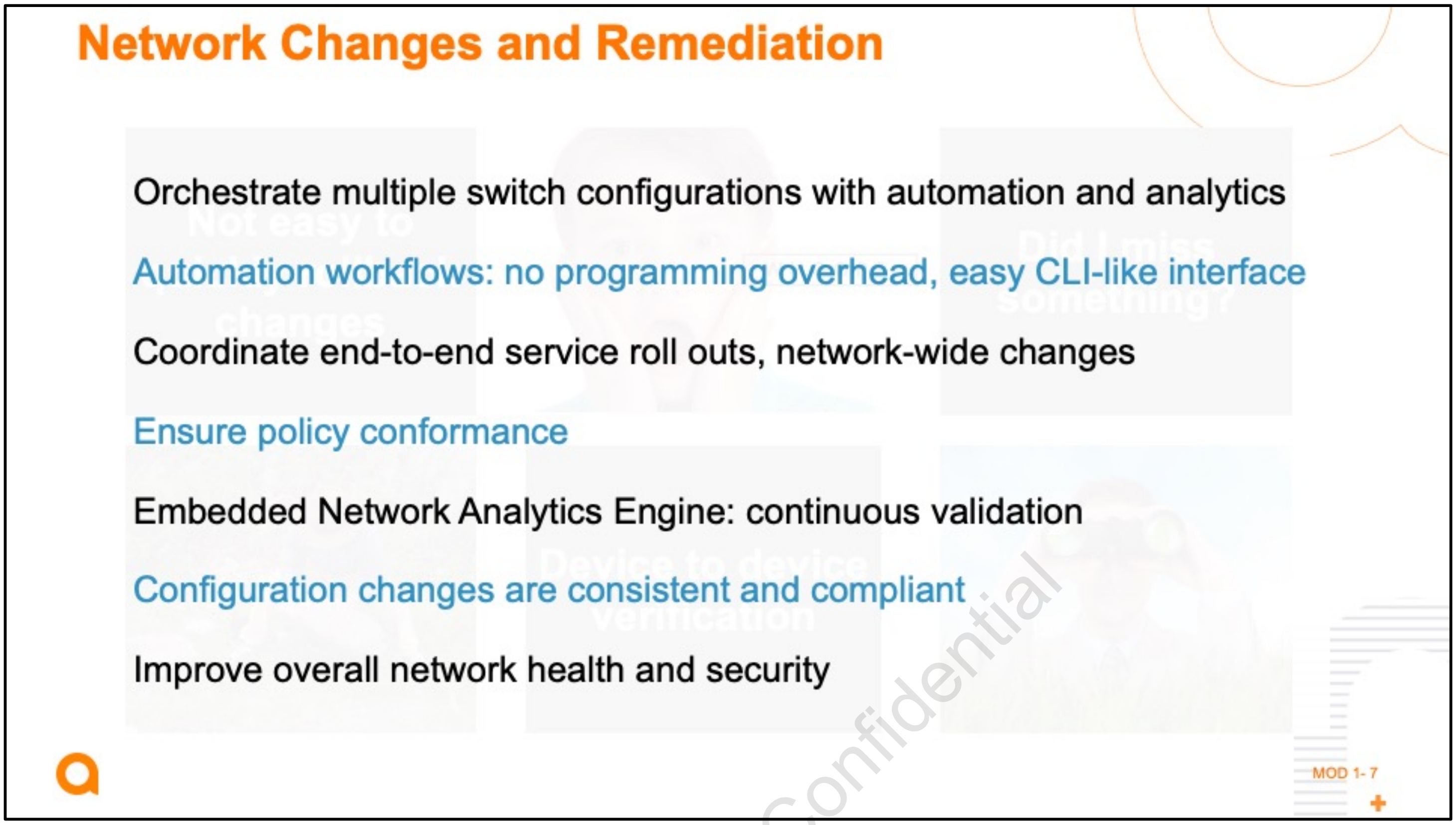
10 Natural disasters, ad hoc mishaps - minor accidents, rodents chewing cables, etc.

Mitigate with NetEdit

MOD 1- 6

Based on a Network World survey, most causes of network outages or performance issues are caused by the frequency of network changes. From the list of top 10 causes of outages shown here, Aruba NetEdit remediates device configuration changes (#2) and operational human errors and mismanagement of devices (#3) from the list.

Aruba Training-Confidential



Network Changes and Remediation

Orchestrate multiple switch configurations with automation and analytics

Automation workflows: no programming overhead, easy CLI-like interface

Coordinate end-to-end service roll outs, network-wide changes

Ensure policy conformance

Embedded Network Analytics Engine: continuous validation

Configuration changes are consistent and compliant

Improve overall network health and security

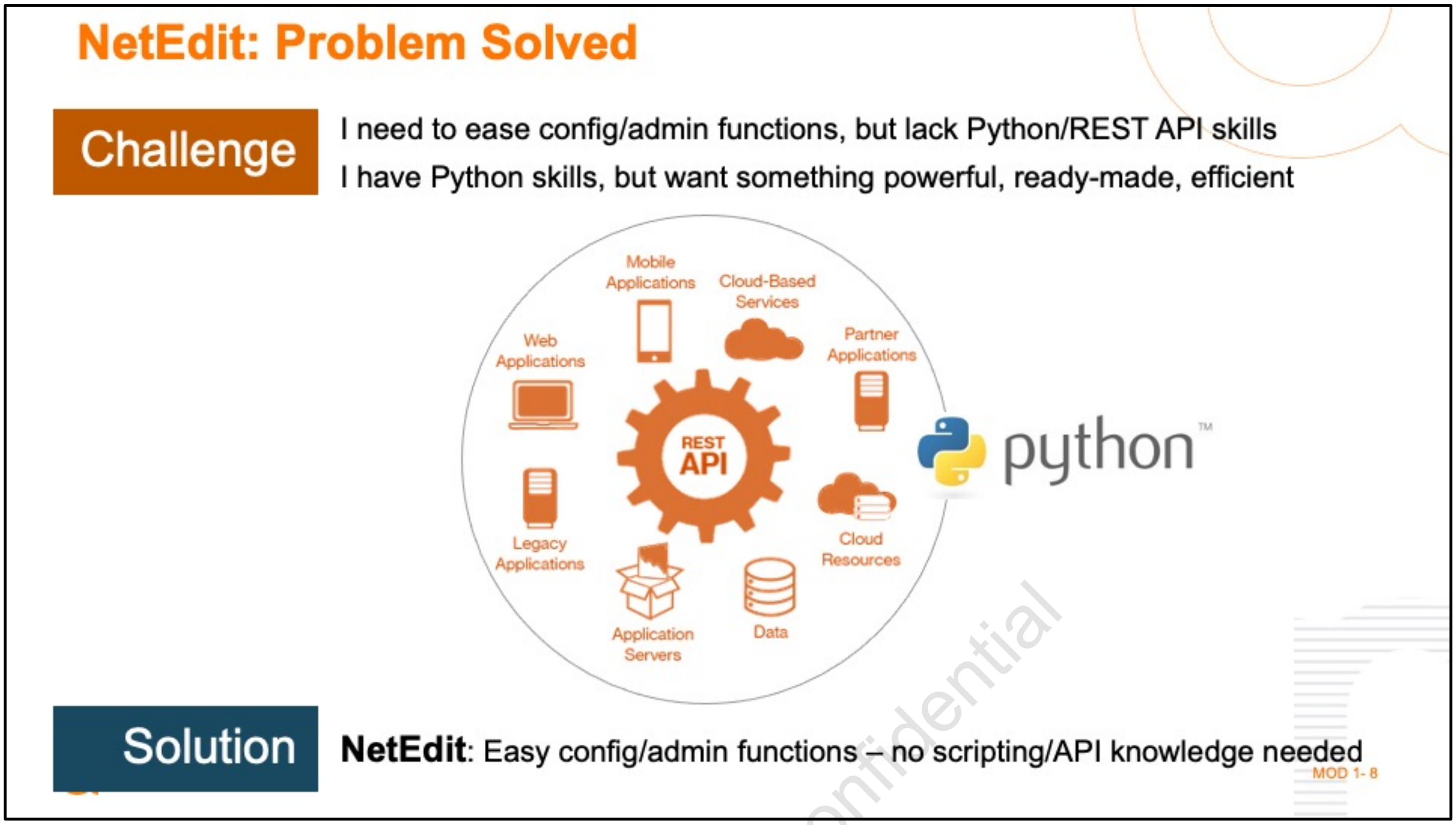
MOD 1- 7

Aruba Training-Confidential

According to Network World, 69% of admins still use the CLI to verify and roll back changes. This requires a constant device to device verification. Plus, it is easy to forget to run a command, and it's not easy to roll back once a change has been pushed.

Additional challenges include device proliferation brought on by IoT and BYOD, and the need for reliable, secure mobile access to any user, anywhere. The high rate of change to support evolving business requirements is another issue. Non-stop adds, moves and changes can overwhelm what are often short-staffed IT teams, which increases the risk of operator error and, subsequently, performance issues or downtime.

In response to these challenges, Aruba NetEdit empowers IT teams to orchestrate multiple switch configurations with automation and analytics - deployments are consistent, conformant, and free of errors. Automation workflows allow for changes without the overhead of programming, with a user-friendly, CLI-like interface. You can smoothly coordinate end-to-end service roll outs, rapidly initiate network-wide changes, and ensure policy conformance after network updates. With embedded analytics delivered from the Aruba Network Analytics Engine (NAE), this intelligent assistance and continuous validation assure that network-wide configuration changes are consistent and compliant, ultimately improving the overall health and security of the network.



NetEdit: Problem Solved

Challenge

I need to ease config/admin functions, but lack Python/REST API skills

I have Python skills, but want something powerful, ready-made, efficient

Solution

NetEdit: Easy config/admin functions – no scripting/API knowledge needed

MOD 1- 8

Aruba Training-Confidential

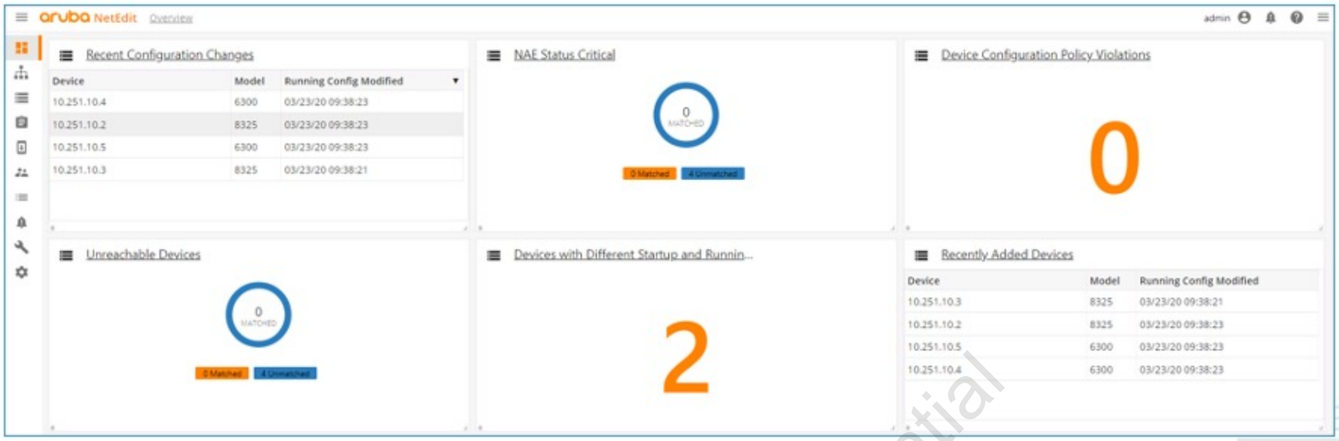
By all means, if you have the skill set to use APIs then use them. However, there are network admins that do not have these skill sets but would like to get some advantages of scripting/API without the overhead of programming, and that's where NetEdit can assist. Even for those that do have the skills to go it alone, many find that NetEdit is powerful, ready-made, and efficient.

NetEdit Functionality

Supports switches running AOS-CX

Runs as a VM

Track all config / hardware changes



Recent Configuration Changes

Device	Model	Running Config Modified
10.251.10.4	6300	03/23/20 09:38:23
10.251.10.2	8325	03/23/20 09:38:23
10.251.10.5	6300	03/23/20 09:38:23
10.251.10.3	8325	03/23/20 09:38:21

Recently Added Devices

Device	Model	Running Config Modified
10.251.10.3	8325	03/23/20 09:38:21
10.251.10.2	8325	03/23/20 09:38:23
10.251.10.5	6300	03/23/20 09:38:23
10.251.10.4	6300	03/23/20 09:38:23

Interrogate switch config, hardware inventory, and neighbor information

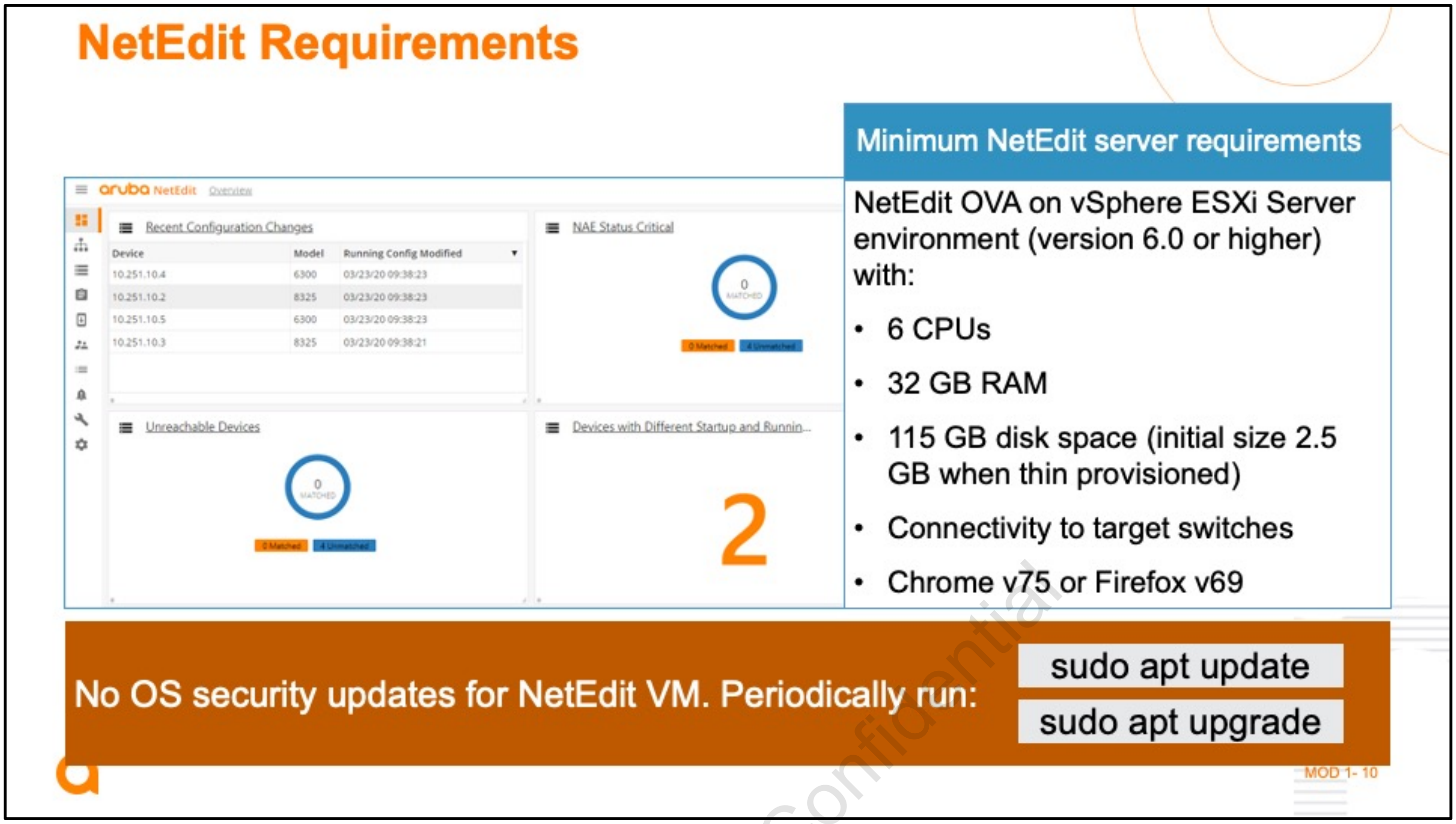
Configuration consistency

Easy web UI, customizable dashboard

Aruba Training-Confidential

NetEdit supports all switches running AOS-CX. NetEdit runs as an Open Virtualization Application (OVA) virtual machine (for example, VMware's ESXi, KVM, Hyper-V, etc.) on a server.

To start, simply enter appropriate subnets within NetEdit's web-based user interface. NetEdit then automatically discovers and imports configuration files for each Aruba CX switch. As NetEdit learns of new switches, it interrogates their configuration, hardware inventory, and neighbor information – you can easily interpret the status of each device. A customizable application dashboard provides quick visibility into metrics of highest interest. The Network tab provides a deeper view into network health, along with red/yellow/green statuses for every deployed switch.



NetEdit Requirements

Device	Model	Running Config Modified
10.251.10.4	6300	03/23/20 09:38:23
10.251.10.2	8325	03/23/20 09:38:23
10.251.10.5	6300	03/23/20 09:38:23
10.251.10.3	8325	03/23/20 09:38:21

Minimum NetEdit server requirements

NetEdit OVA on vSphere ESXi Server environment (version 6.0 or higher) with:

- 6 CPUs
- 32 GB RAM
- 115 GB disk space (initial size 2.5 GB when thin provisioned)
- Connectivity to target switches
- Chrome v75 or Firefox v69

No OS security updates for NetEdit VM. Periodically run:

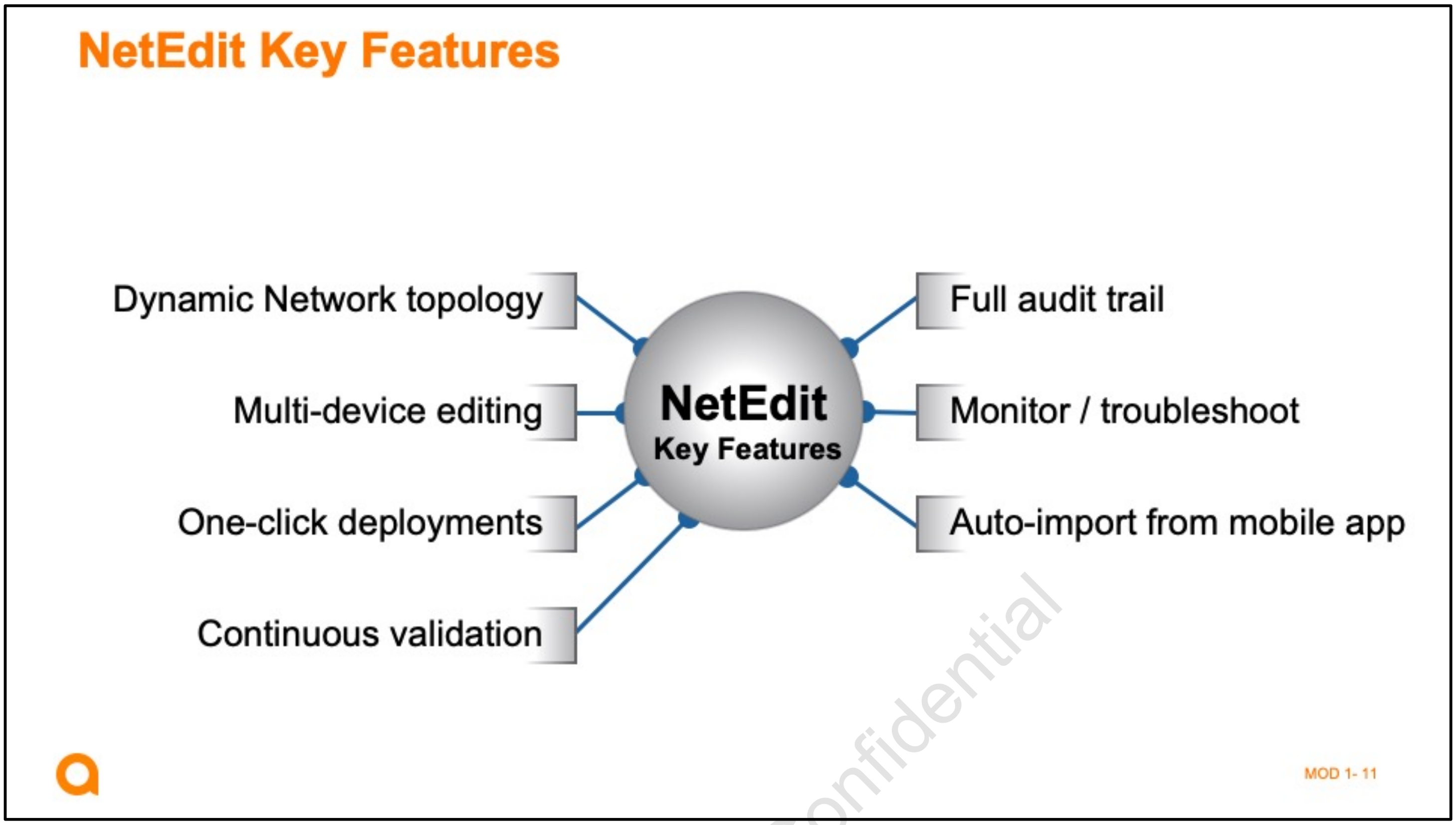
```
sudo apt update
sudo apt upgrade
```

Aruba Training-Confidential

The figure shows **minimum** NetEdit server requirements. For more detail about OVA installation, please refer to the NetEdit Installation Guide.

Important: Aruba does not provide OS security updates to the NetEdit VM. To continue to benefit from important security updates offered by Debian, please run the following items periodically:

- sudo apt update
- sudo apt upgrade



NetEdit Key Features

NetEdit
Key Features

Dynamic Network topology

Multi-device editing

One-click deployments

Continuous validation

Full audit trail

Monitor / troubleshoot

Auto-import from mobile app

MOD 1- 11

Aruba Training-Confidential

The figure shows key NetEdit capabilities. Please click on each feature above for more details.

Dynamic Network Topology

The NetEdit Network tab provides a holistic view of your network topology, including Aruba CX switches, Aruba access points, and third-party switches.

The Health Summary panel reveals a real-time snapshot of Aruba switch health. You see devices with inconsistencies, errors, security policy violations, and more.

Dynamic, tailored network views are triggered based on which layers you select. These layers offer more visibility into Aruba CX device status and configuration, including what is contributing to a performance or compliance issue. Supported layers include application, client service, device, routing, bridging, segmentation, and “other” (which includes all NAE agents not designated to a specific layer).

For further simplicity, NetEdit automatically discovers new network infrastructure devices using the Link Layer Discovery Protocol (LLDP), using REST APIs for Aruba CX switches and SNMP for Aruba wireless and third-party devices. Newly connected switches appear automatically in the Network tab, so you can automate switch configuration change workflows without programming.

Multi-device Editing

NetEdit enables you to easily change configurations on multiple devices at once using complete details of each configuration file. For example, you can centrally define the NTP or the RADIUS server address for all relevant switches, or set an ACL entry on only the access switches.

NetEdit also taps into your existing CLI knowledge - no retraining or new skills are required. Predictive assistance dramatically reduces the time it takes to enter commands. You get command completion, syntax highlighting, and validation.

Plus you can implement common configurations, such as switch-to-switch virtual extensible LANs (VXLANs), by using only a few prompt-driven commands.

Aruba Training-Confidential

One-Click Deployments with Auto-Verification

Use NetEdit to stage a deployment (or rollback) on multiple switches, and validate correct switch operation afterward. For instance, you can determine whether a change is working properly before deploying it more widely, or you can quickly back out changes if there is a problem.

Upon deployment, NetEdit automatically collects network and services state information before and after a change, and intelligently displays the difference. This lets you decide whether to keep or roll back the change within your allocated window.

Continuous Validation

Continuously monitor and ensure conformance for both corporate and regulatory policies. Perform validation checks on all configuration changes, including those made outside of NetEdit, such as using the switch's CLI or through Ansible.

For example, it's simple to verify that all management IP addresses are on the management subnetwork, or that all routers running OSPF are logging adjacency changes. Validation tests can easily be customized and extended based on your existing knowledge of configuration commands.

Full Audit Trail

NetEdit records all hardware and software versions, and other configuration changes. Search through and view all changes, or groups of changes.

This allows you to track changes to the hardware, software and configurations using an automated versioning feature, regardless of whether changes are made in NetEdit or through other means. You can also perform these rollbacks selectively, based upon factors such as the location of the switches or the dates of the changes.

Monitoring and Troubleshooting

Integration with the Aruba Network Analytics Engine (NAE) provides real-time access to advanced network analytics and built-in automation for troubleshooting – quickly detect and resolve network issues.

NAE natively gathers switch telemetry via agents (based on Python scripts). Proactively set rules to monitor and collect specific traffic or events of interest - monitor CPU utilization above a desired threshold, a configuration mismatch, or an OSPF reachability problem.

NetEdit subscribes to the status of the NAE agent, collecting data when an issue of interest occurs. Integrations with tools like ServiceNow, TOPdesk and Slack provide for fast notifications, prompting operators to take corrective action.

Upon clicking into NetEdit, the Network tab highlights the devices that match the notification criteria. Then you quickly drill into the impacted device or service and begin troubleshooting, with full, time-stamped diagnostic details.

Thus, NetEdit and NAE significantly reduce the number of manual troubleshooting tasks. It also produces less network overhead load, so performance is not impacted while collecting network-wide telemetry.

Note: Configuration version roll back may cause issues when going between major AOS-CX software versions.

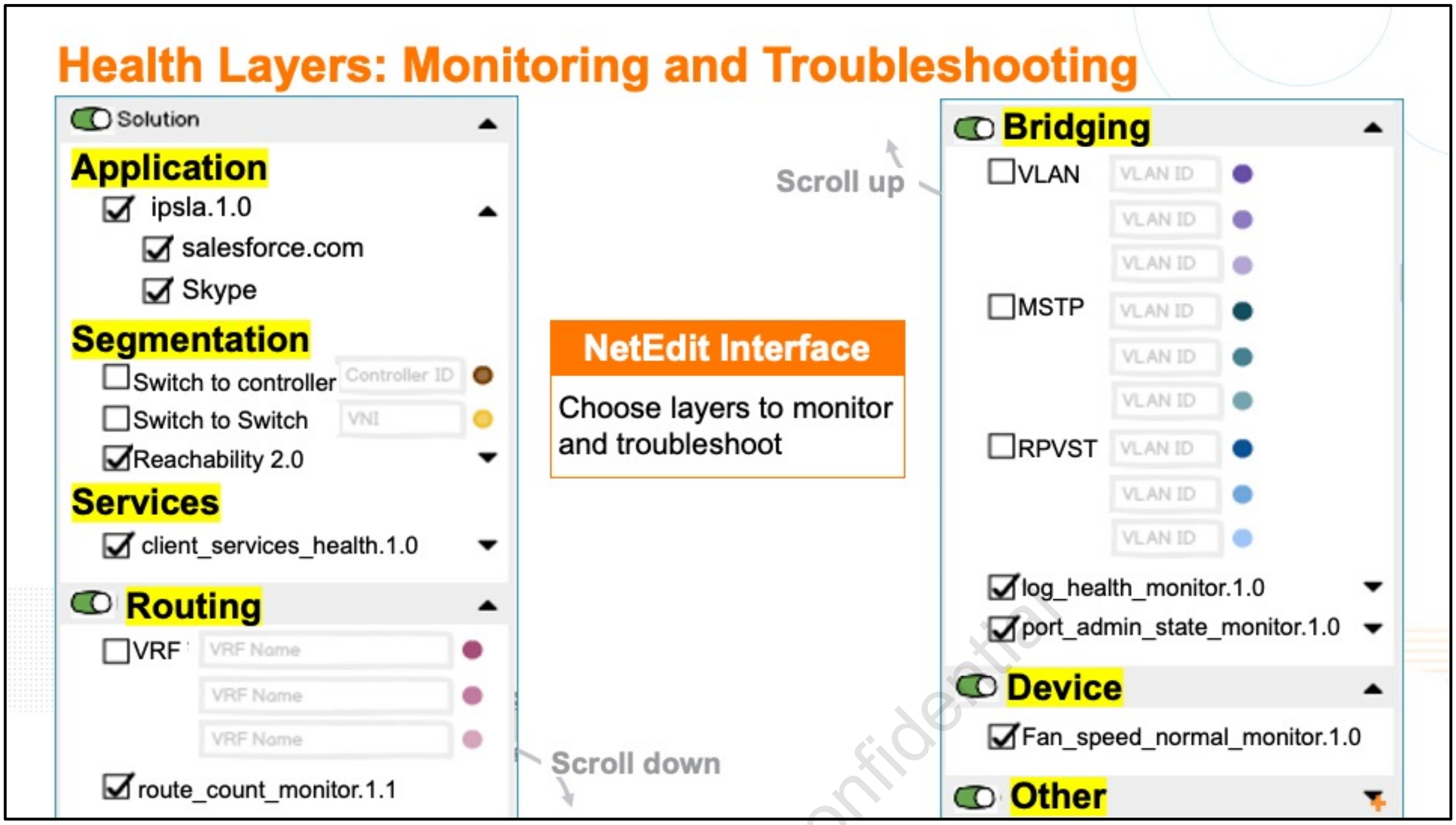
Automatic Imports from the Aruba CX Mobile App

For further simplicity, Aruba offers a mobile app that lets you install Aruba CX

Aruba Training-Confidential

switches from your iOS or Android device. Switches that are connected to the network via the Aruba CX Mobile App are automatically imported into NetEdit for policy conformance verification.

Aruba Training-Confidential



Health Layers: Monitoring and Troubleshooting

Solution

Application

- ☑ ipsla.1.0
  - ☑ salesforce.com
  - ☑ Skype

Segmentation

- □ Switch to controller — Controller ID
- □ Switch to Switch — VNI
- ☑ Reachability 2.0

Services

- ☑ client_services_health.1.0

Routing

- □ VRF — VRF Name, VRF Name, VRF Name
- ☑ route_count_monitor.1.1

NetEdit Interface

Choose layers to monitor and troubleshoot

Scroll up

Scroll down

Bridging

- □ VLAN — VLAN ID, VLAN ID, VLAN ID
- □ MSTP — VLAN ID, VLAN ID, VLAN ID
- □ RPVST — VLAN ID, VLAN ID, VLAN ID
- ☑ log_health_monitor.1.0
- ☑ port_admin_state_monitor.1.0

Device

- ☑ Fan_speed_normal_monitor.1.0

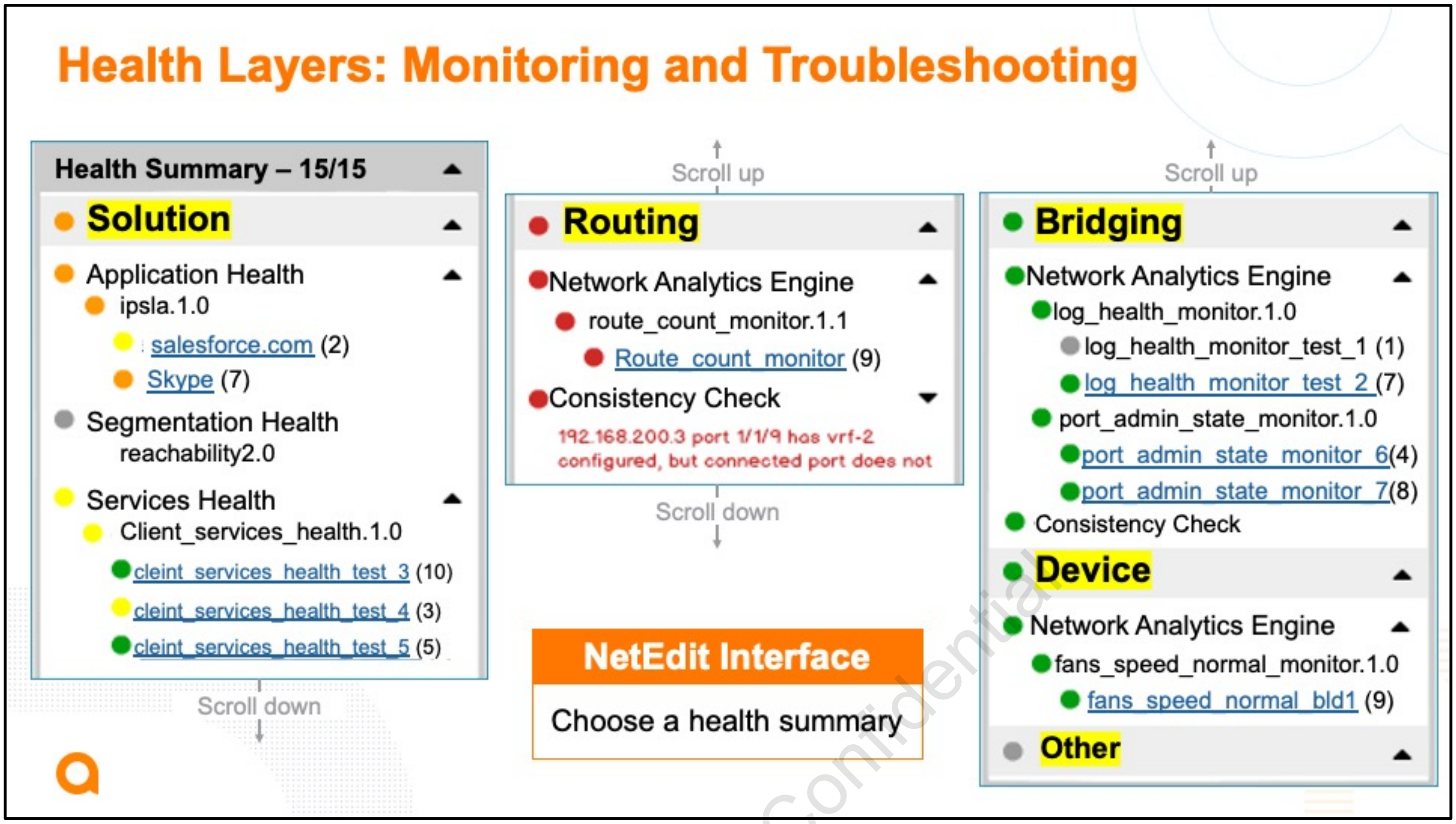
Other

You control the layered view of additional network information for monitoring and troubleshooting. The figure shows the NetEdit GUI. The top portion is shown on the left. You would scroll down to see the rest, which is shown on the right. This is where you choose the appropriate layers to work with:

- Application
- Segmentation
- Client Service
- Routing
- Bridging
- Device
- Other

NetEdit also supports Aruba Network Analytics Engine (NAE) status aggregation of data.

Aruba Training-Confidential



Health Layers: Monitoring and Troubleshooting

Health Summary – 15/15

- ● **Solution**
  - ● Application Health
    - ● ipsla.1.0
      - ● salesforce.com (2)
      - ● Skype (7)
  - ● Segmentation Health
    - reachability2.0
  - ● Services Health
    - ● Client_services_health.1.0
      - ● cleint_services_health_test_3 (10)
      - ● cleint_services_health_test_4 (3)
      - ● cleint_services_health_test_5 (5)

Scroll down

Scroll up

- ● **Routing**
  - ● Network Analytics Engine
    - ● route_count_monitor.1.1
      - ● Route_count_monitor (9)
  - ● Consistency Check
    - 192.168.200.3 port 1/1/9 has vrf-2 configured, but connected port does not

Scroll down

NetEdit Interface

Choose a health summary

Scroll up

- ● **Bridging**
  - ● Network Analytics Engine
    - ● log_health_monitor.1.0
      - ● log_health_monitor_test_1 (1)
      - ● log_health_monitor_test_2 (7)
    - ● port_admin_state_monitor.1.0
      - ● port_admin_state_monitor_6(4)
      - ● port_admin_state_monitor_7(8)
  - ● Consistency Check
- ● **Device**
  - ● Network Analytics Engine
    - ● fans_speed_normal_monitor.1.0
      - ● fans_speed_normal_bld1 (9)
- ● **Other**

Aruba Training-Confidential

You can see health summaries at the network, multiple devices, and single devices. You see feature status, NAE agent state, and consistency validation for things like VLANs or maybe a VRF mismatch on a link.

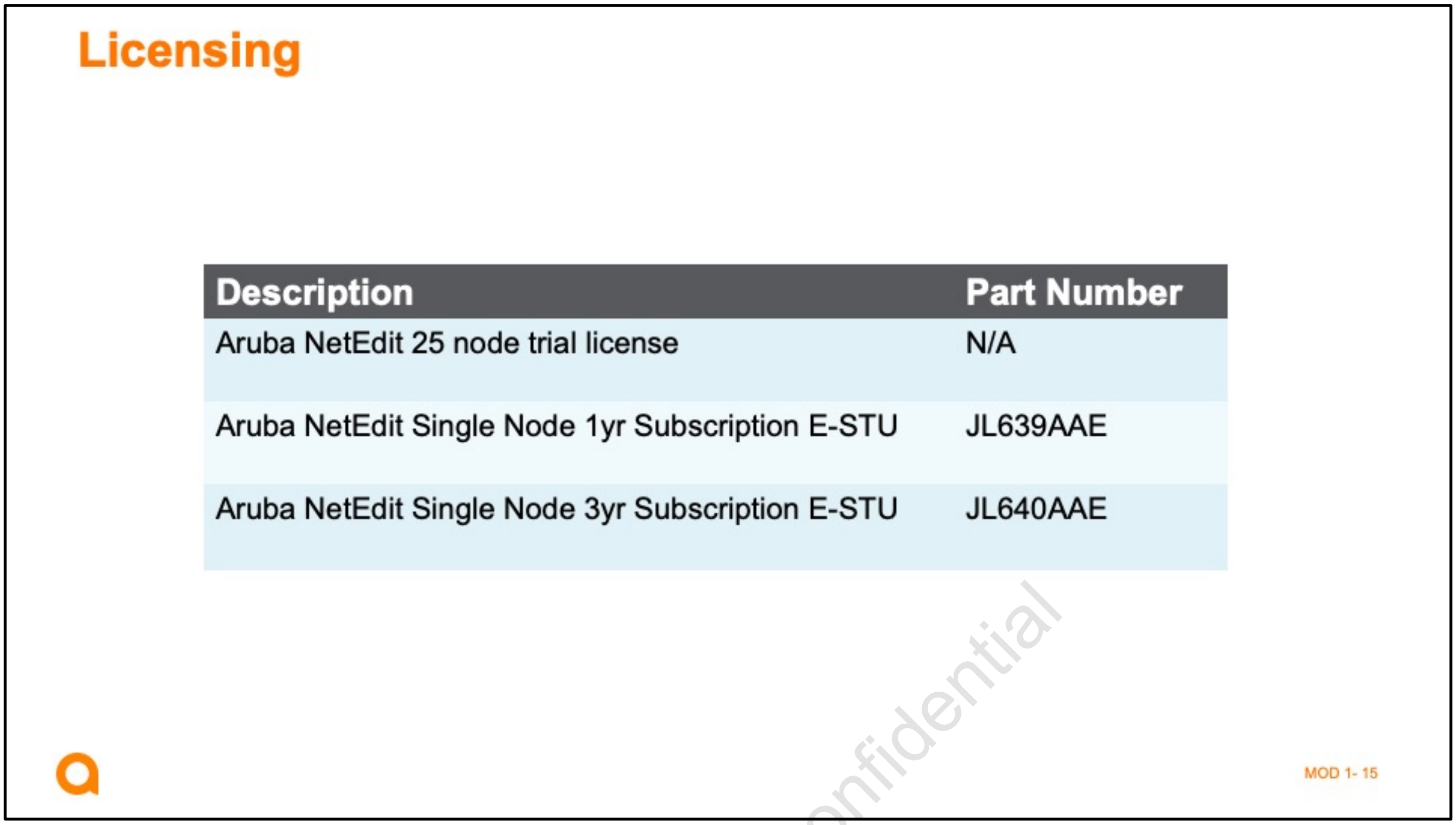
Plus you get subscription-based updates

aruba
a Hewlett Packard
Enterprise company

Initial Setup

This section introduces you to what you need prepare on your AOS-CX switches in order for NetEdit to successfully integrate with them as well as how NetEdit can discover devices.

Aruba Training-Confidential



Licensing

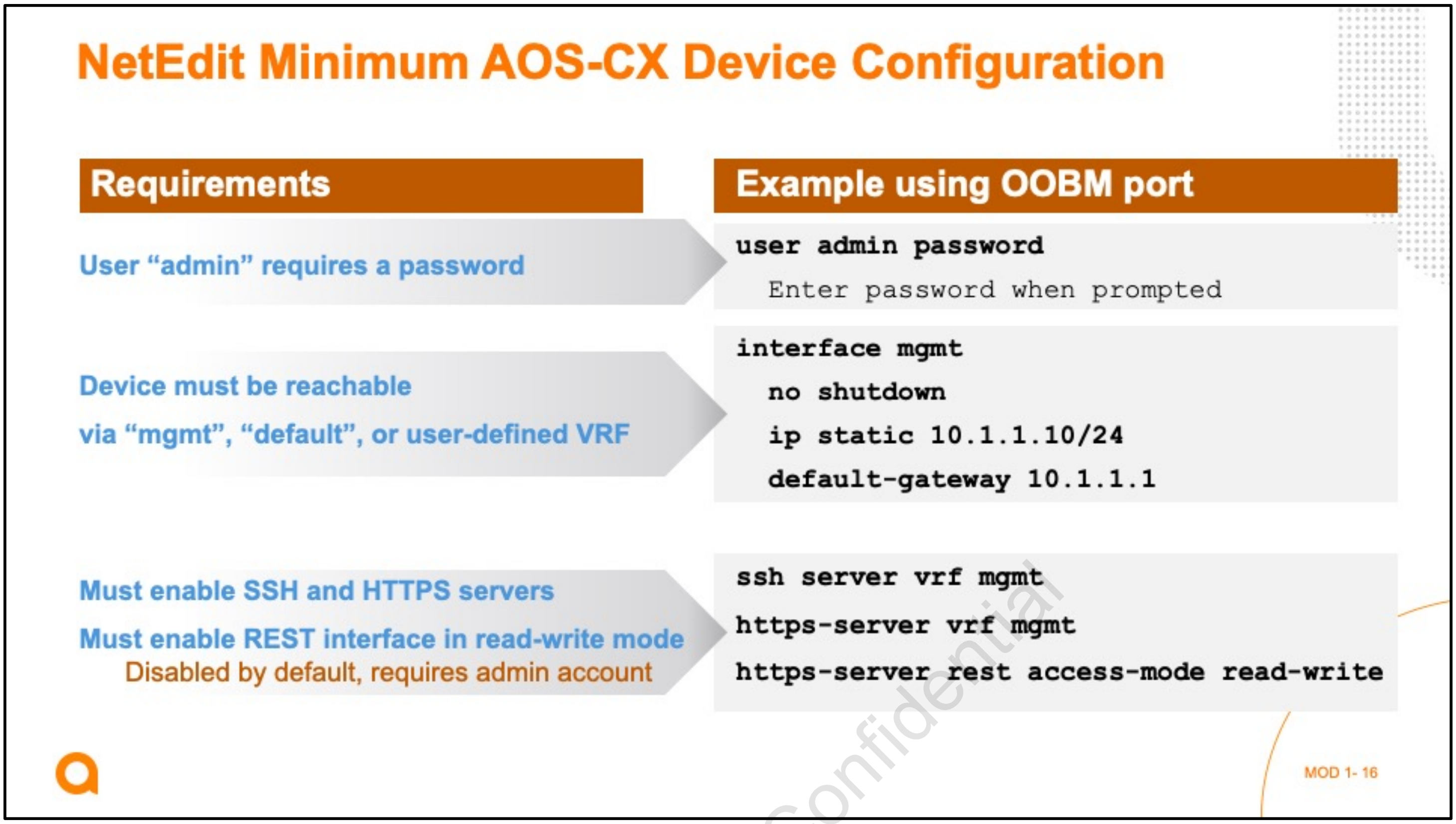
Description	Part Number
Aruba NetEdit 25 node trial license	N/A
Aruba NetEdit Single Node 1yr Subscription E-STU	JL639AAE
Aruba NetEdit Single Node 3yr Subscription E-STU	JL640AAE

MOD 1- 15

Aruba Training-Confidential

NetEdit is currently available on a trial basis for up to 25 nodes. There are also licensing options for one-year and three-year subscriptions for nodes 26 and upwards.

Note: Aruba does not officially offer support for the trial version of NetEdit. You must purchase licenses to get NetEdit support.



NetEdit Minimum AOS-CX Device Configuration

Requirements	Example using OOBM port

User "admin" requires a password	`user admin password`
	`Enter password when prompted`

Device must be reachable
via "mgmt", "default", or user-defined VRF
	`interface mgmt`
	`no shutdown`
	`ip static 10.1.1.10/24`
	`default-gateway 10.1.1.1`

Must enable SSH and HTTPS servers
Must enable REST interface in read-write mode
Disabled by default, requires admin account
	`ssh server vrf mgmt`
	`https-server vrf mgmt`
	`https-server rest access-mode read-write`

MOD 1- 16

To communicate with Aruba NetEdit, the AOS-CX switch requires some minimum configuration.

The REST interface is disabled by default, and like HTTPS, must be enabled. Likewise, access requires a switch account with administrative access.

Note: These commands are covered in the Aruba associate-level switching course.

Aruba Training-Confidential

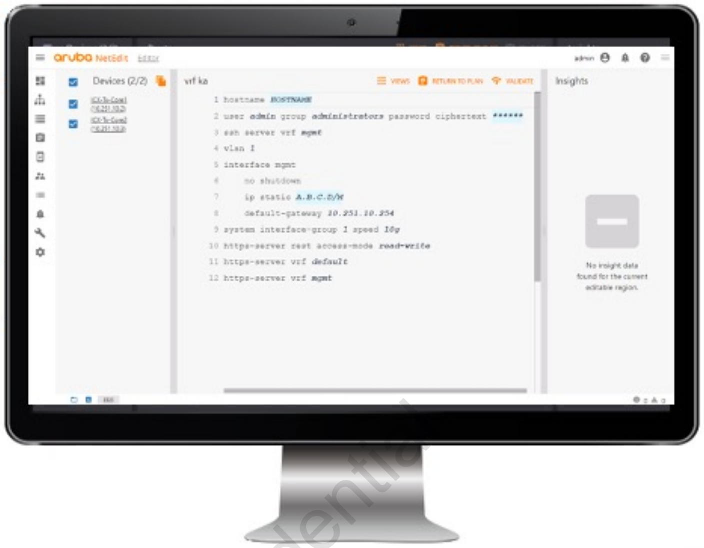
NetEdit Access

Two username/passwords

- One for Linux access
- One for NetEdit access

First-time login

- Username admin
- Blank password
- Change password, min. 8 characters



```
hostname HOSTNAME
user admin group administrators password ciphertext ******
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static A.B.C.D/M
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

No insight data found for the current editable region.

MOD 1- 17

NetEdit has two sets of usernames and passwords: one for Linux access and one for NetEdit access. When you log into NetEdit for the first time, you enter a username of admin and a blank password. You are then prompted to change the password, which must minimally be 8 characters in length.

Aruba Training-Confidential

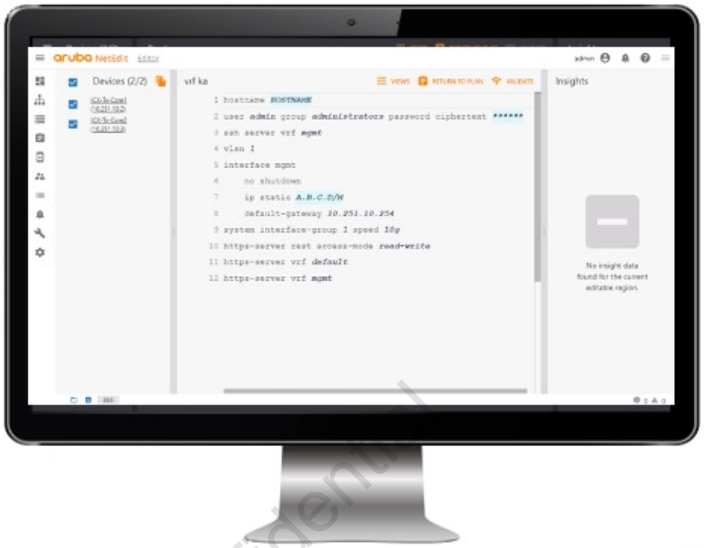
NetEdit Use

Save device credentials

- SNMP v2/3
- SSH
- Rest API (AOS-CX Devices only)

Discover managed subnets

- Subnets managed by NetEdit
- Excluded addresses
- Applies the credentials to a device



```
 1 hostname HOSTNAME
 2 user admin group administrators password ciphertext ******
 3 ssh server vrf mgmt
 4 vlan 1
 5 interface mgmt
 6     no shutdown
 7     ip static A.B.C.D/M
 8     default-gateway 10.251.10.254
 9 system interface-group 1 speed 10g
10 https-server rest access-mode read-write
11 https-server vrf default
12 https-server vrf mgmt
```

MOD 1- 18

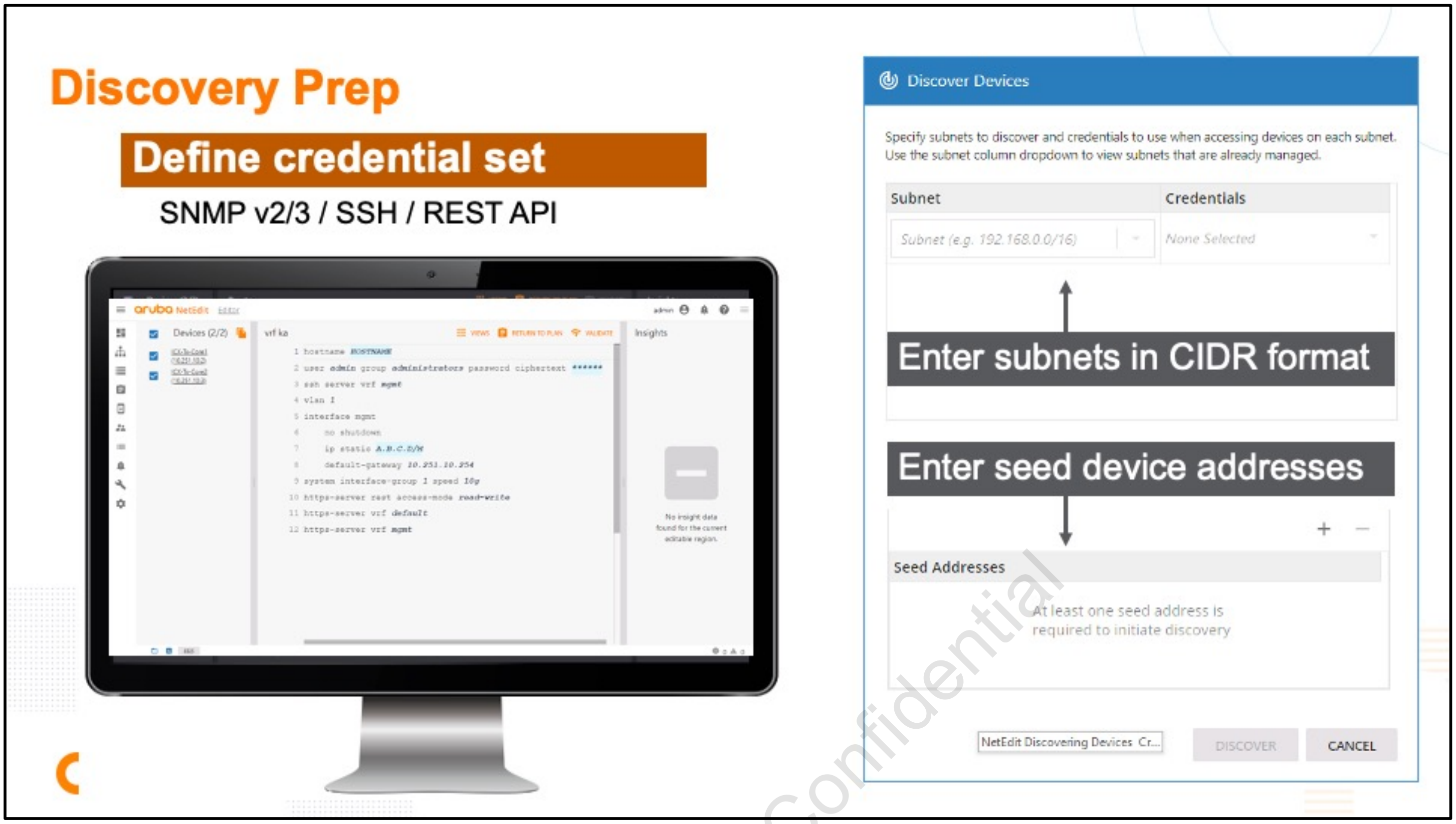
Aruba Training-Confidential

NetEdit has two sets of usernames and passwords: one for Linux access and one for NetEdit access. When you log into NetEdit for the first time, you enter a username of admin and a blank password. You are then prompted to change the password, which must minimally be 8 characters in length.

In NetEdit 2.0 there is no longer the batch adding and single devices adding. NetEdit now allows for users to save credentials within NetEdit and discover devices within a subnet.

NetEdit will now also discover and display 3rd party devices that are using standard SNMP MIB's, and you can enter SSH credentials for 3rd party devices.

The managed subnets determine what devices fall under the NetEdit management umbrella. You can exclude specific addresses to make this umbrella as large or small as you like.



Discovery Prep

Define credential set

SNMP v2/3 / SSH / REST API

```
hostname HOSTNAME
user admin group administrators password ciphertext ******
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static A.B.C.D/M
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

Specify subnets to discover and credentials to use when accessing devices on each subnet. Use the subnet column dropdown to view subnets that are already managed.

Subnet	Credentials
Subnet (e.g. 192.168.0.0/16)	None Selected

Enter subnets in CIDR format

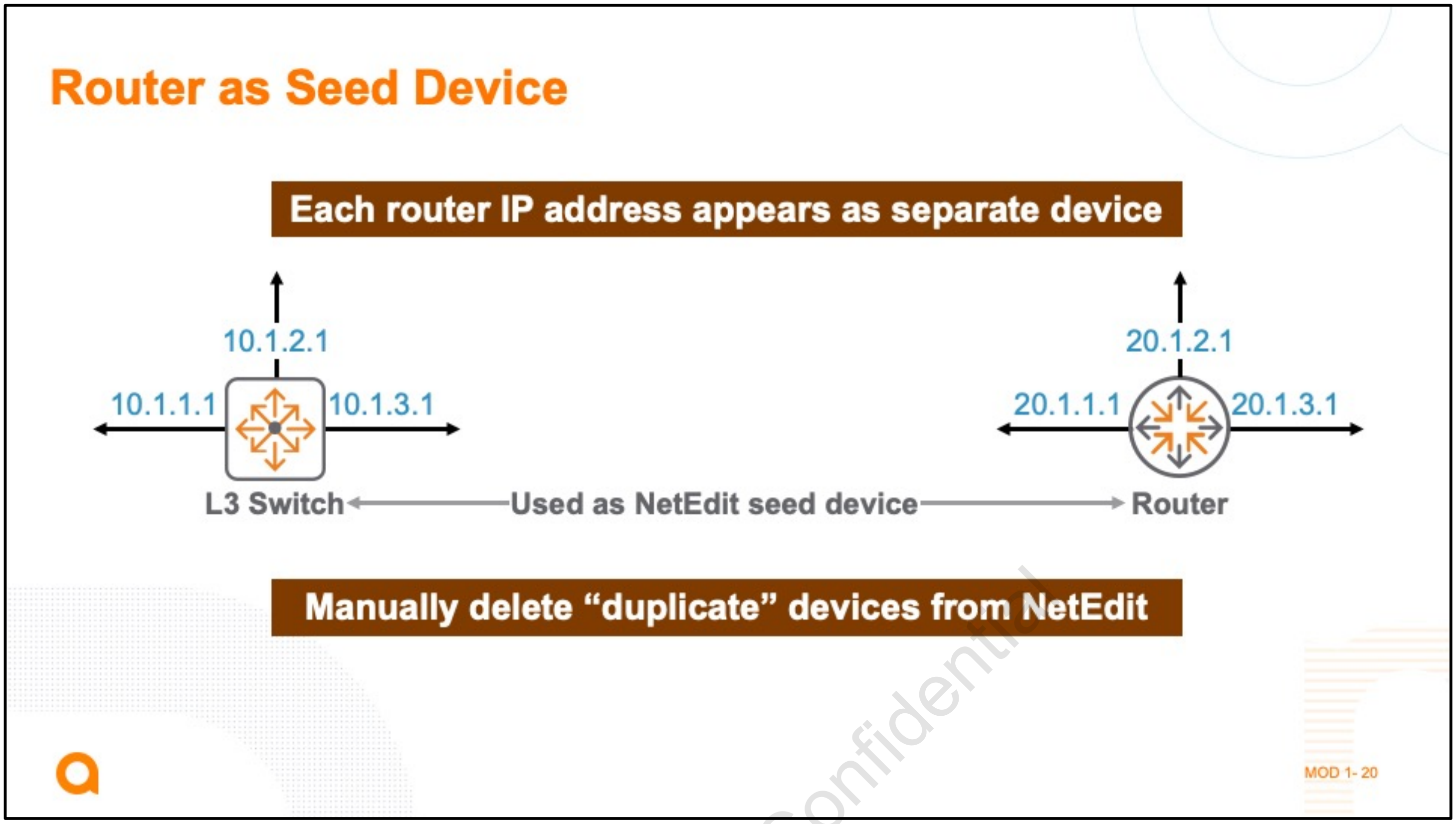
Enter seed device addresses

At least one seed address is required to initiate discovery

So you first specify credential sets, then define a subnet for discovery and the credential set to use. Each subnet must be associated with credentials that specify the REST and/or SNMP credentials that NetEdit should use to read data from the devices in that subnet.

Then define one seed device in the subnet that will then discover other devices based on LLDP. Assuming good connectivity, NetEdit finds all connected subnets that seed device. For subnets to discover that are not connected, provide a seed device in each separate group of connected subnets.

Aruba Training-Confidential



Router as Seed Device

Each router IP address appears as separate device

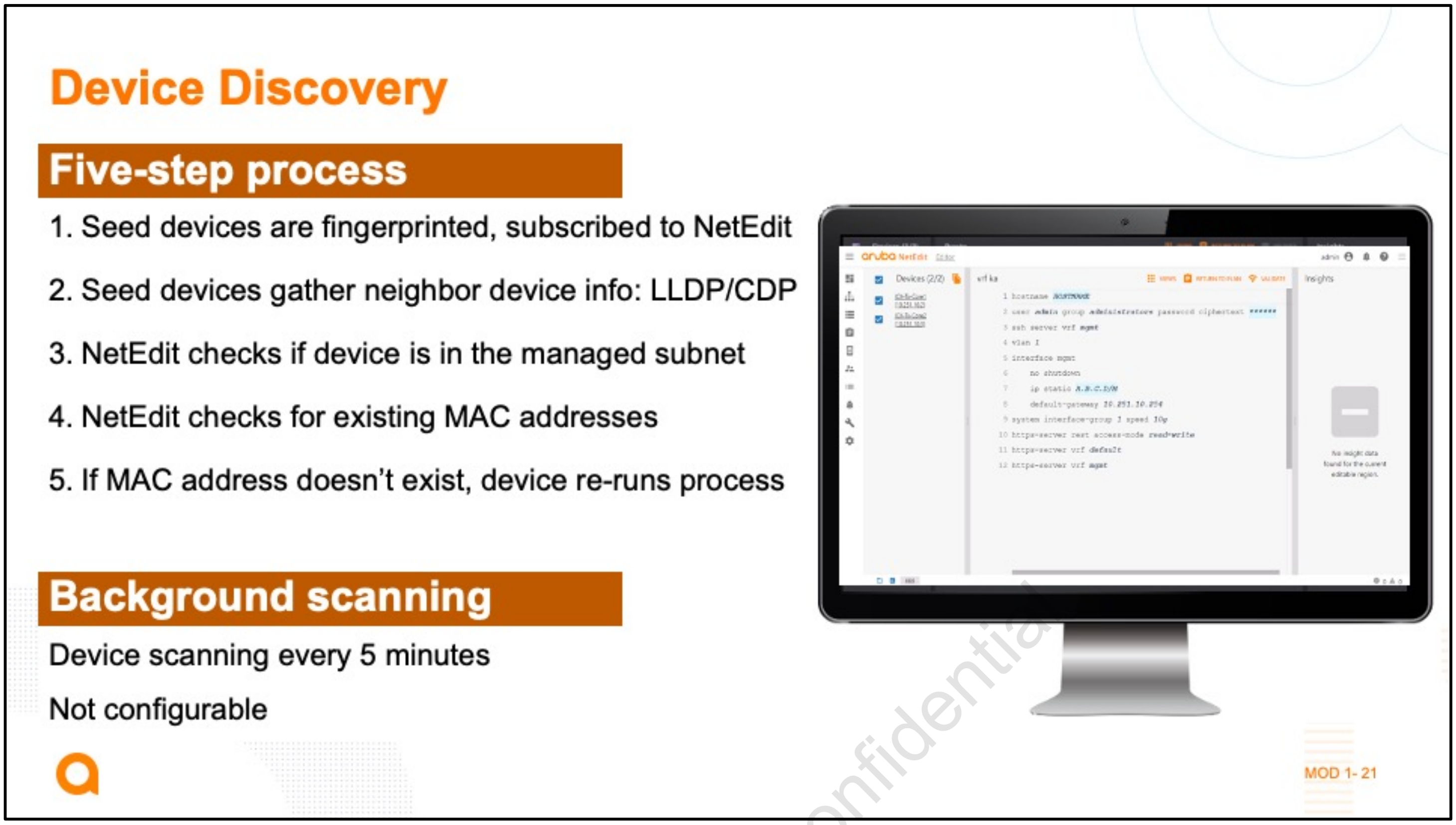
Used as NetEdit seed device

Manually delete "duplicate" devices from NetEdit

MOD 1- 20

Aruba Training-Confidential

When discovery is initiated using a router or Layer 3 switch attached to multiple subnets as a seed device, each router IP address on a target subnet for NetEdit may appear as a separate device in NetEdit. Specifying multiple IP addresses from such a router as seeds will definitely cause such "duplicate" devices to appear in NetEdit. If this happens, manually identify and delete the extraneous "duplicate" devices from NetEdit using the Devices page.



Device Discovery

Five-step process

1. Seed devices are fingerprinted, subscribed to NetEdit
2. Seed devices gather neighbor device info: LLDP/CDP
3. NetEdit checks if device is in the managed subnet
4. NetEdit checks for existing MAC addresses
5. If MAC address doesn't exist, device re-runs process

Background scanning

Device scanning every 5 minutes

Not configurable

```
hostname HOSTNAME
user admin group administrators password ciphertext ******
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static A.B.C.D/M
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

Discovery is a five-step process. Seed devices are fingerprinted, subscribed to NetEdit. These seed devices gather neighbor device info via LLDP and/or CDP. NetEdit checks if the device is in a managed subnet, and checks for existing MAC addresses. If the MAC address doesn't already exist, the process re-runs.

And NetEdit does background ^scanning every five (5) minutes - not configurable.

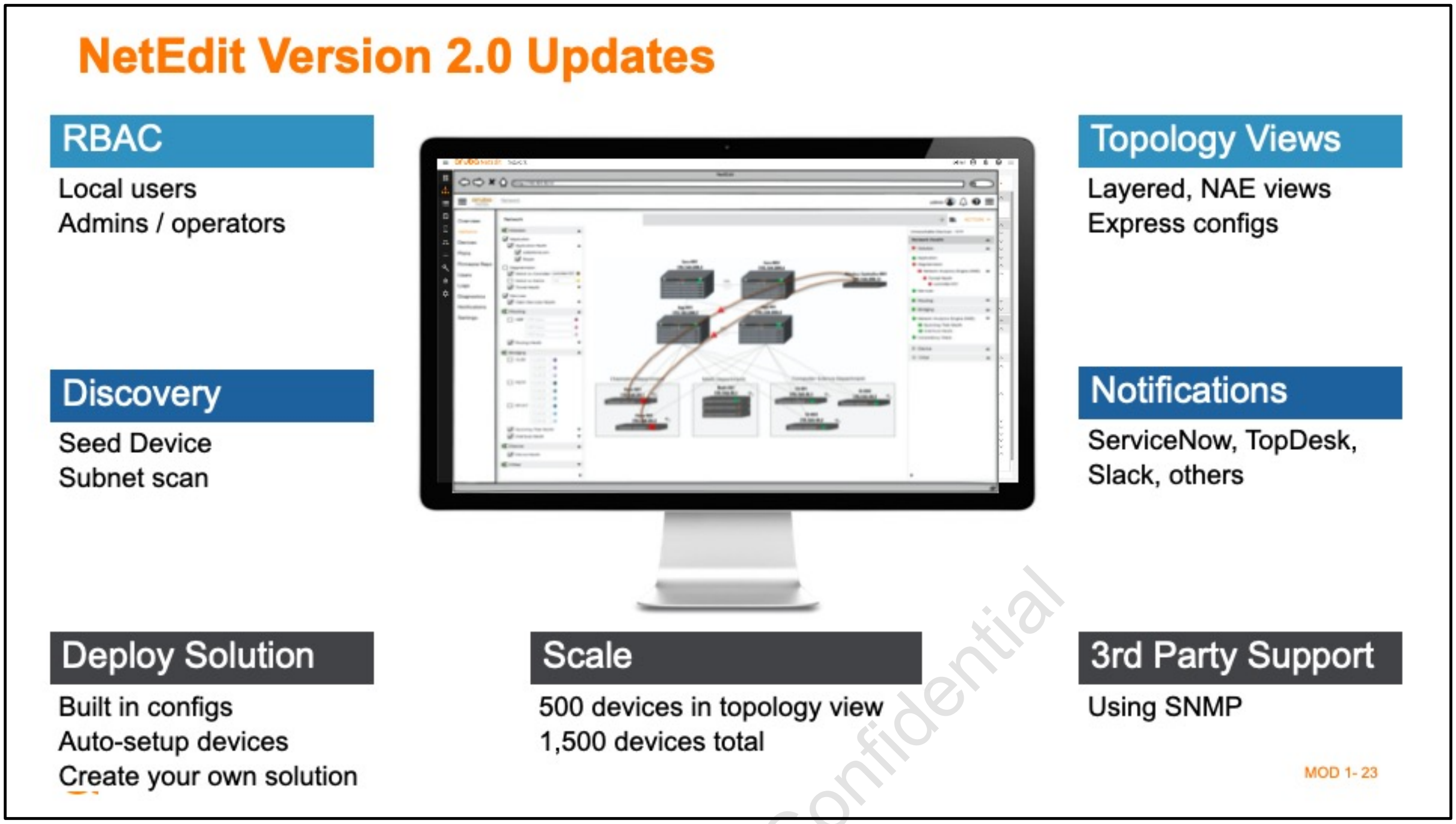
You should now be able to see the devices on your network and configure them using NetEdit.

Aruba Training-Confidential

Key Capabilities

This section introduces some of the key capabilities of NetEdit as well as introducing the implementation of device configuration plans.

Aruba Training-Confidential



NetEdit Version 2.0 Updates

RBAC
Local users
Admins / operators

Discovery
Seed Device
Subnet scan

Deploy Solution
Built in configs
Auto-setup devices
Create your own solution

Scale
500 devices in topology view
1,500 devices total

Topology Views
Layered, NAE views
Express configs

Notifications
ServiceNow, TopDesk,
Slack, others

3rd Party Support
Using SNMP

MOD 1- 23

Aruba Training-Confidential

The figure summarizes key features introduced in NetEdit 2.0. Click on each feature to explore.

Role based Access Control (RBAC)

Two levels of users now—admins and operators

Topology Views

The layered view shows different network layers - all devices with VLAN 100, all devices running OSPF, and etcetera.

The NAE View shows an aggregation of NAE scripts. If scripts are tagged properly, NetEdit will show which device are running the scripts and which are not. You can also see NAE statistics.

Discovery

Discovery works based on a seed device within a given range of devices. It leverages LLDP information to discover new devices – just provide proper login credentials. NetEdit will continue to discover devices based on LLDP information

Notifications

Get notified based on network changes. ServiceNow, TopDesk, and Slack will all send a URI with a link to the error in NetEdit – just click and get right to the information you need.

Deploy Solution

You get built-in “Express configurations”, to automatically configure devices with just a few parameters. You can also create your own.

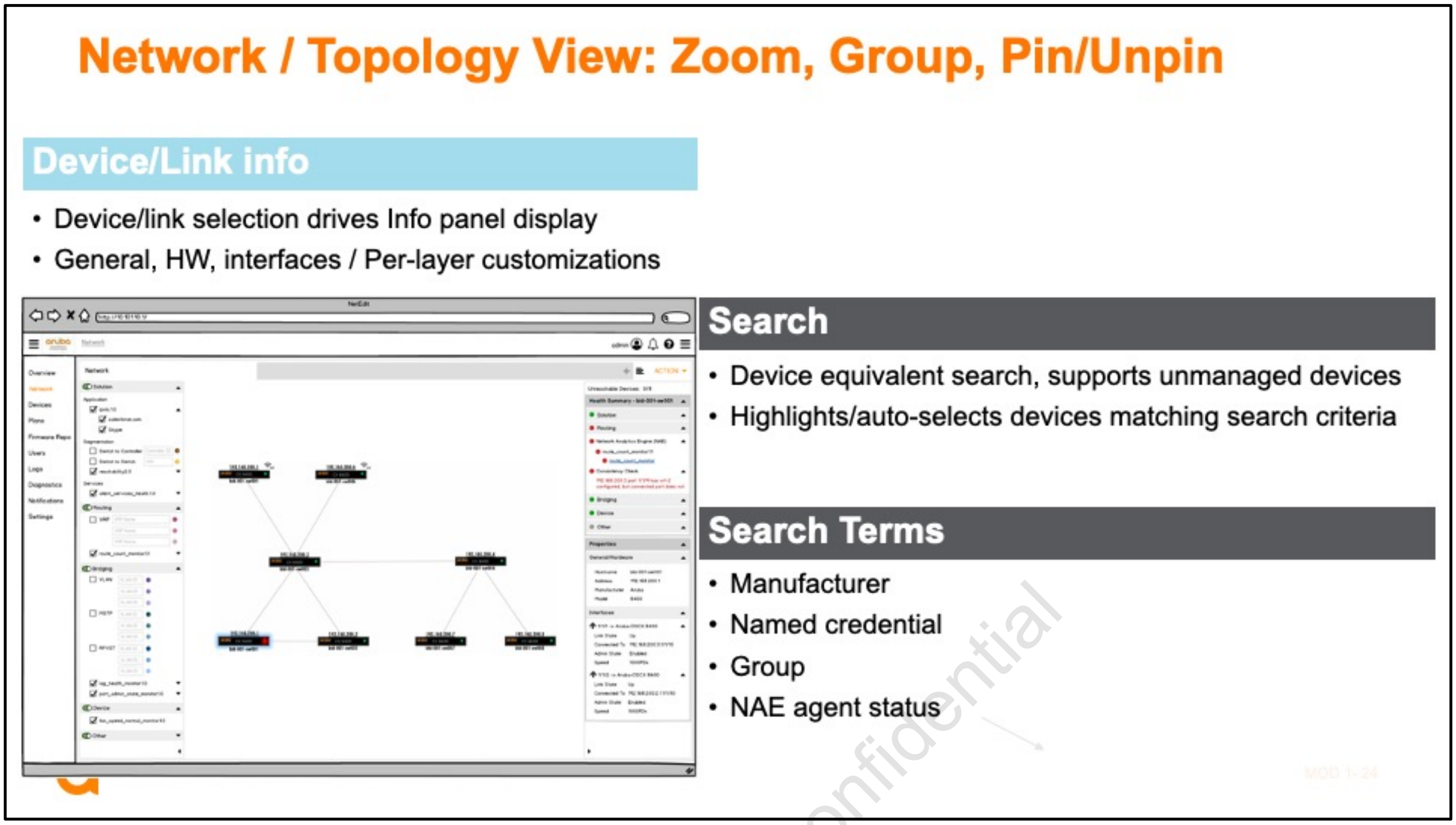
Scalability

Scalability has increased. Up to 500 AOS-CX switches are supported at one time in topology view and 1,500 devices total.

Third-party support

NetEdit supports any third-party devices that use SNMP.

Aruba Training-Confidential



Network / Topology View: Zoom, Group, Pin/Unpin

Device/Link info

- Device/link selection drives Info panel display
- General, HW, interfaces / Per-layer customizations

Search

- Device equivalent search, supports unmanaged devices
- Highlights/auto-selects devices matching search criteria

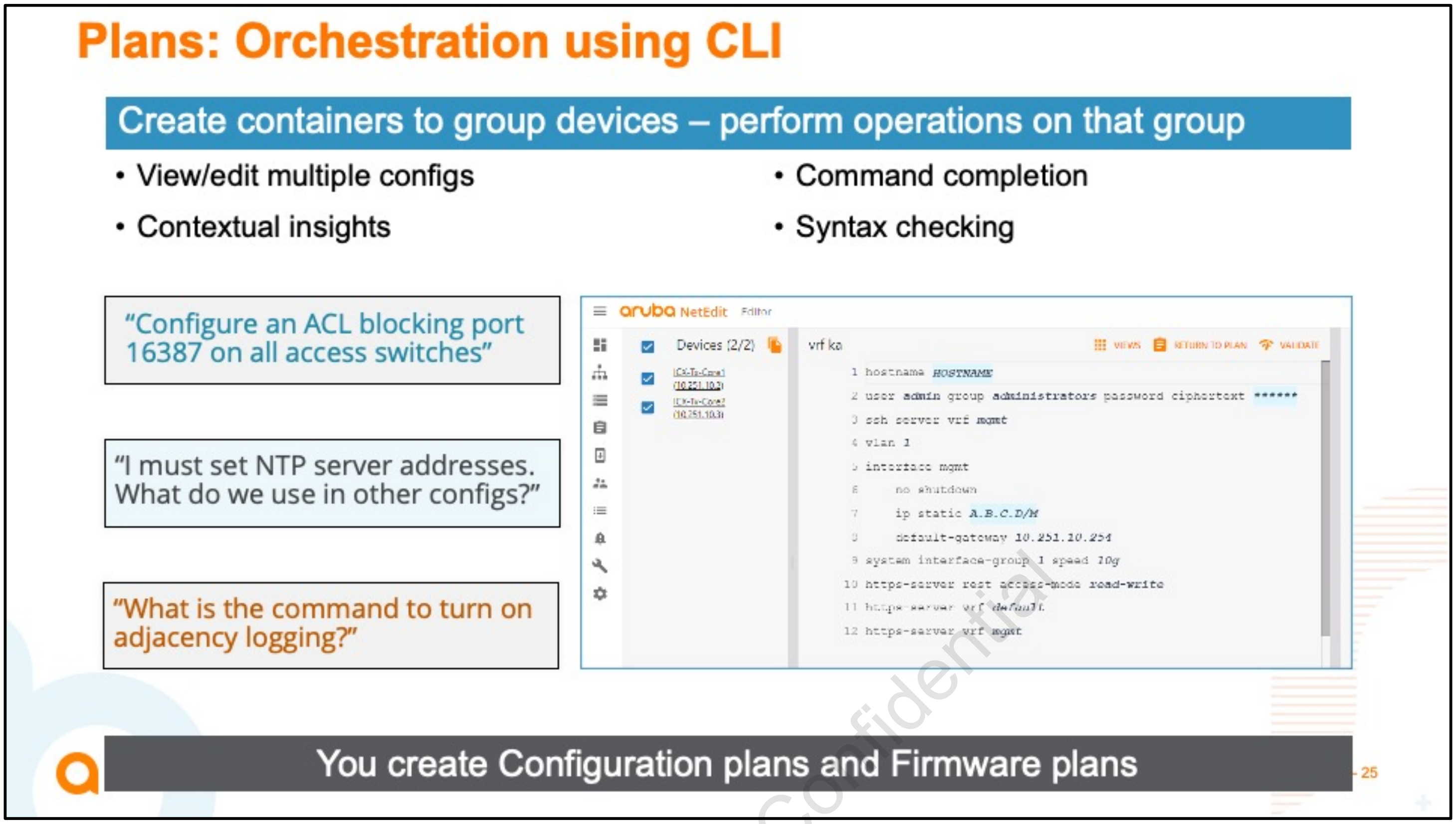
Search Terms

- Manufacturer
- Named credential
- Group
- NAE agent status

MOD 1-24

New in NetEdit 2.0 is a network view (topology page) where you can see the interconnections of both AOS-CX and non-AOS-CX devices that NetEdit learned. Please note that the information on the screen is searchable.

Aruba Training-Confidential



Plans: Orchestration using CLI

Create containers to group devices – perform operations on that group

- View/edit multiple configs
- Contextual insights
- Command completion
- Syntax checking

"Configure an ACL blocking port 16387 on all access switches"

"I must set NTP server addresses. What do we use in other configs?"

"What is the command to turn on adjacency logging?"

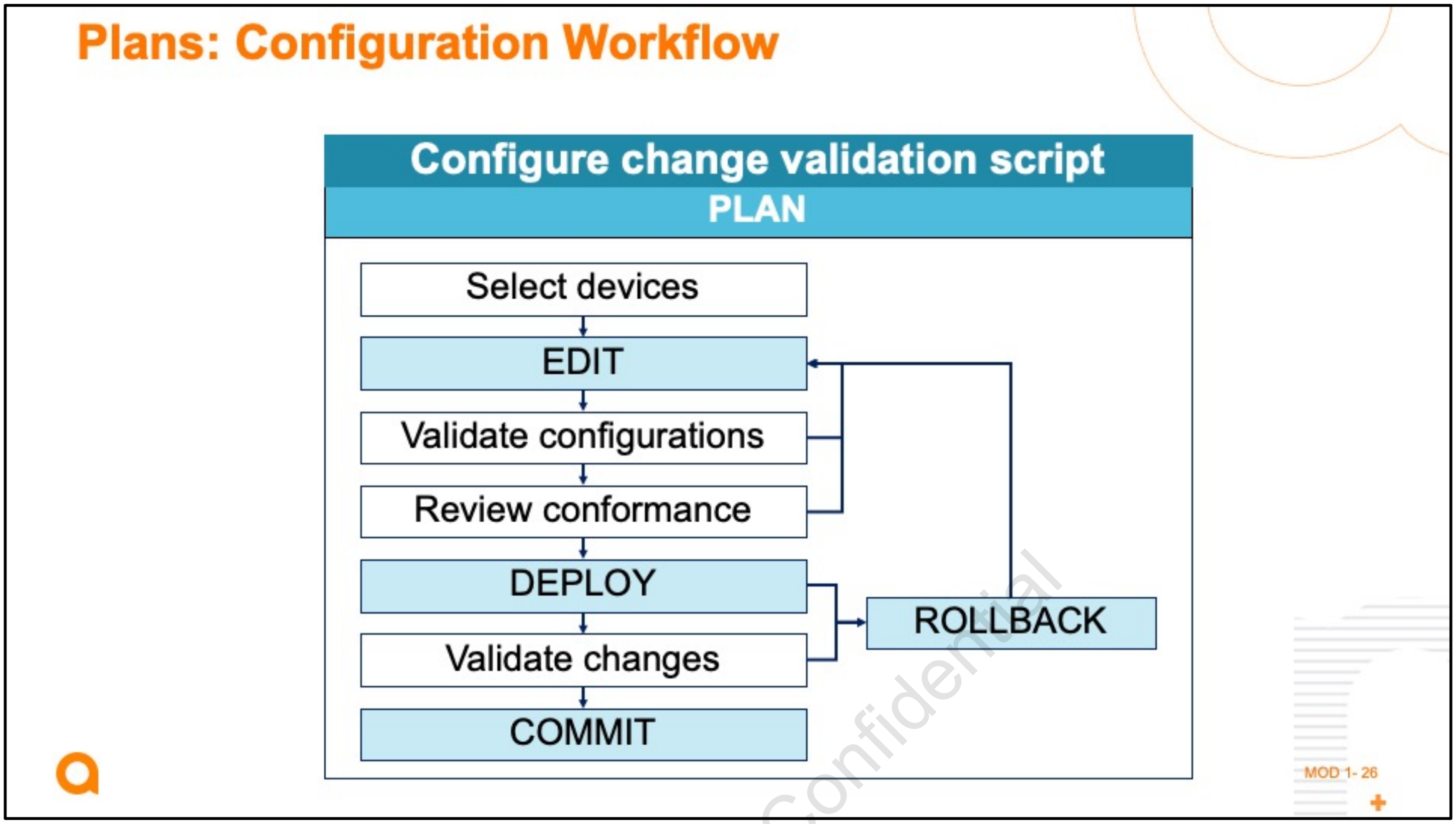
```
hostname HOSTNAME
user admin group administrators password ciphertext ******
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static A.B.C.D/M
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

You create Configuration plans and Firmware plans

25

A plan enables you and the NetEdit application to create a container to group devices, then perform operations on that group. These include edit, deploy, commit or rollback device configuration or deploy firmware versions. There are seven different plan types. You can define the configuration plan and firmware plan, while the NetEdit application defines the remaining plan types.

Aruba Training-Confidential



Plans: Configuration Workflow

Configure change validation script

PLAN

Select devices

EDIT

Validate configurations

Review conformance

DEPLOY

Validate changes

COMMIT

ROLLBACK

MOD 1- 26

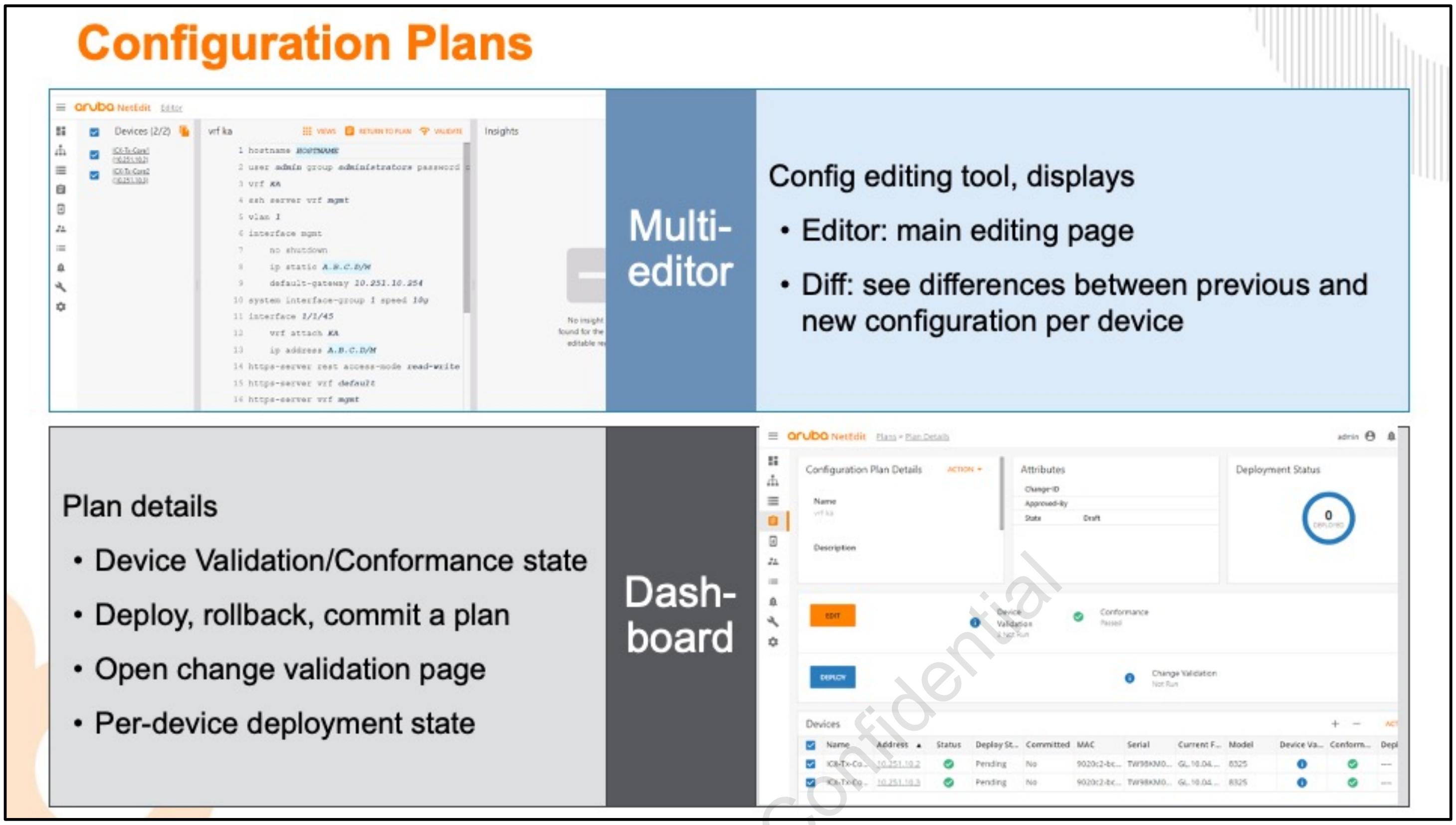
Aruba Training-Confidential

The configuration plan is a core feature of NetEdit. It allows you to create configuration plans and edit either running or start-up configurations for a set of devices.

The figure shows that you first select devices, then use the Edit feature to change multiple configurations at once, using complete knowledge of each configuration file. Take advantage of contextual insights that are automatically displayed. For example, you can use Edit to set the NTP or the RADIUS server address for all relevant switches or to set an ACL entry on all access switches. The Edit feature also includes command completion, syntax highlighting and validation.

You validate your configurations and review them for conformance to corporate policy.

Then use the Deploy feature to stage a deployment (or a rollback) on multiple switches and validate the correct operation of switches after changes. For instance, you can determine whether a change is working properly before deploying it more widely, or you can quickly remove changes if there is a problem. Deploy automatically collects network and services state information before and after a change, and intelligently displays the difference. This lets you decide whether to keep or rollback the change within your allocated change window.



Configuration Plans

```
hostname HOSTNAME
user admin group administrators password c
vrf KA
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static A.B.C.D/M
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
interface 1/1/45
    vrf attach KA
    ip address A.B.C.D/M
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

Multi-editor

Config editing tool, displays

- Editor: main editing page
- Diff: see differences between previous and new configuration per device

Plan details

- Device Validation/Conformance state
- Deploy, rollback, commit a plan
- Open change validation page
- Per-device deployment state

Dash-board

You have learned that a configuration plan helps you to edit the running or startup-configuration for a set of devices. You can also view the validation status for all device configurations in the plan. This plan will deploy candidate configurations to one or more devices running the configuration.

You can view the conformance status for all device candidate configurations in a configuration plan. You can also view a change validation report (see the Change Validation section for more information). Additional features include writing the deployed running configuration, deployment rollback and delete plans from the Action menu on the plan's page.

Aruba Training-Confidential

The configuration plan capabilities include the following:

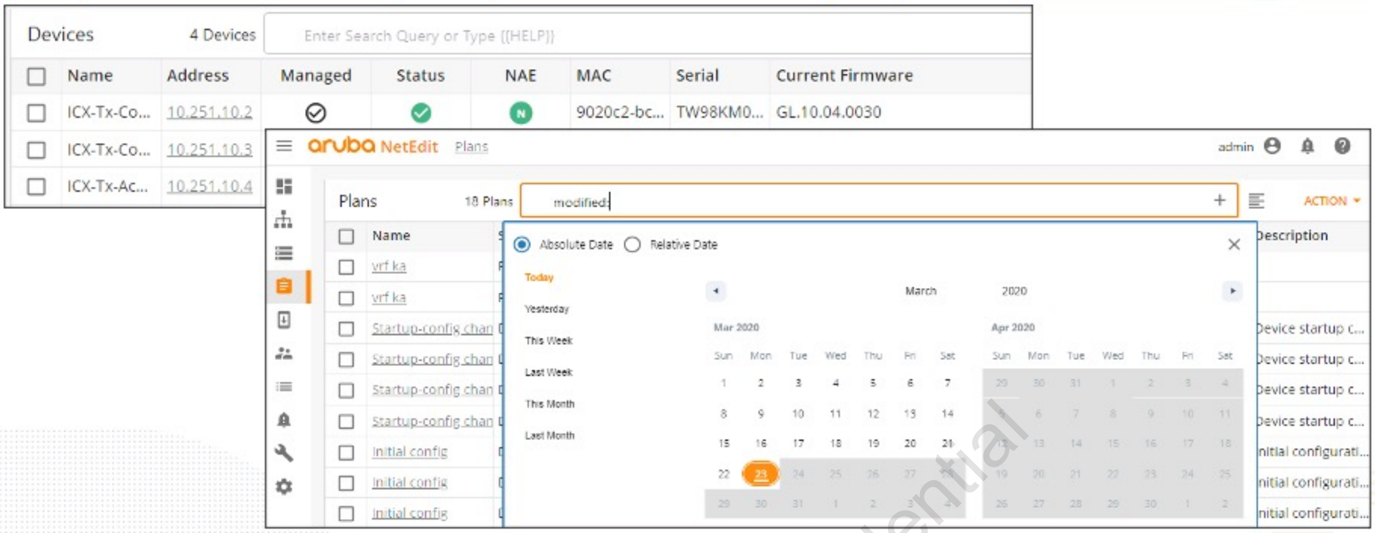
- Create a plan with the intent of editing running or startup configurations for a set of devices.
- Edit configurations for one or more devices in the plan (see Multi-Edit section for information).
- View validation status for all device configurations in the plan.
- View conformance status for all device candidate configurations in the plan.
- Deploy candidate configurations to one or more device(s) running configuration.
- View a change validation report (see Change Validation section for more information) for the plan.
- Write (Commit) the deployed running configuration to startup.
- Rollback (i.e. Undo) the deployment as long as the device(s) running configuration is what was deployed from the plan and as long as the commit operation (write to startup) has not been executed from within the plan.
- Plans can be deleted from the Action menu in the plans page.

Analyze Configurations and Plans

Custom tags for devices and plans

Correlate software version, hardware, feature usage

Find inconsistencies and errors



Devices 4 Devices Enter Search Query or Type {{HELP}}

□	Name	Address	Managed	Status	NAE	MAC	Serial	Current Firmware
□	ICX-Tx-Co...	10.251.10.2	✓	✓	N	9020c2-bc...	TW98KM0...	GL.10.04.0030
□	ICX-Tx-Co...	10.251.10.3						
□	ICX-Tx-Ac...	10.251.10.4						

aruba NetEdit Plans admin

Plans 18 Plans modified:

□	Name	Description
□	vrf ka	
□	vrf ka	
□	Startup-config chan	Device startup c...
□	Startup-config chan	Device startup c...
□	Startup-config chan	Device startup c...
□	Startup-config chan	Device startup c...
□	Initial config	Initial configurati...
□	Initial config	Initial configurati...
□	Initial config	Initial configurati...

● Absolute Date ○ Relative Date

"Show configs of any access switch in building 5 that changed within the last week."

"Are there any configurations that violate our security hardening policy."

Use the search feature to quickly find the group of switches that require configuration changes or investigation. Identify inconsistencies and errors, such as security policy violations. For example, you can identify configurations of all switches of a specific type, in a particular location, where the configuration has changed within a selected time period.

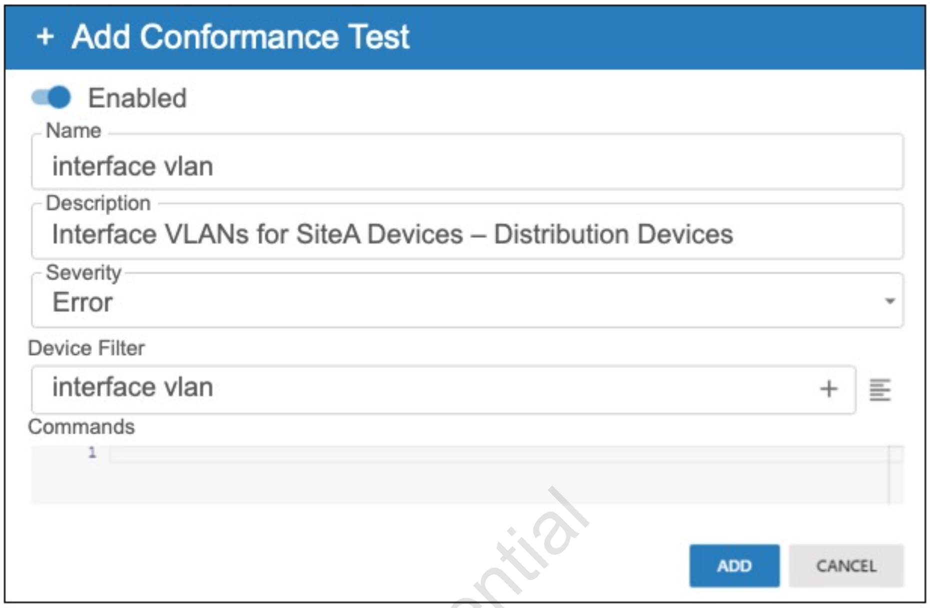
Aruba Training-Confidential

Conformance: Consistent Configurations

Customize rules for each network

Ensure devices support company policy

Audit changes fit the policy



+ Add Conformance Test

Enabled

Name: interface vlan

Description: Interface VLANs for SiteA Devices – Distribution Devices

Severity: Error

Device Filter: interface vlan

Commands

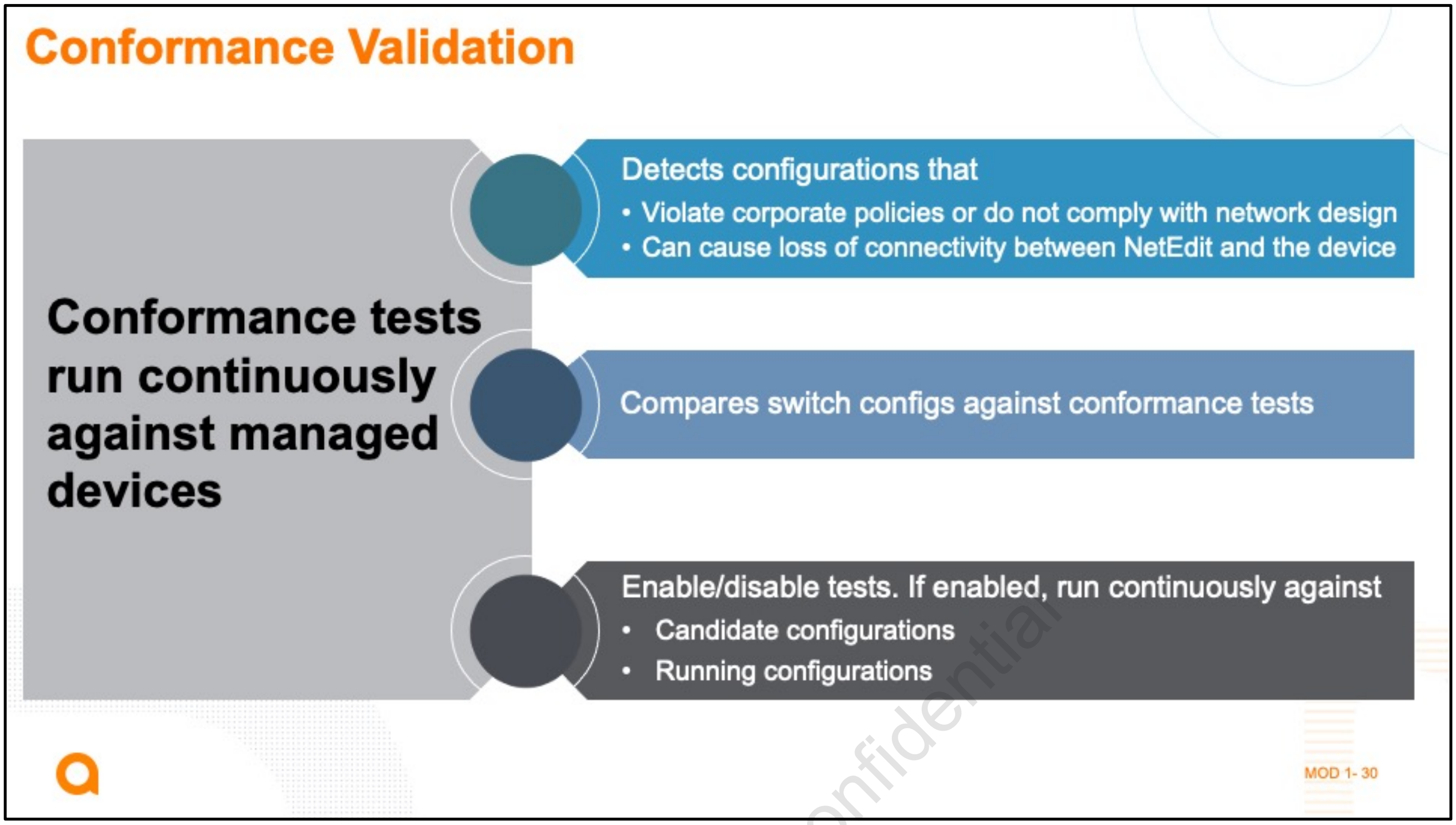
"Legal requires specific wording for login banners. Ensure all switches have that banner"

"Regulatory compliance: All switch configs must support FIPS-approved mode of operation"

Aruba Training-Confidential

Conformance validation detects configurations that violate your corporate policies or network design by comparing switch configurations against conformance tests. Define policies by creating and enabling conformance tests in **Settings > Conformance**. The tests run continuously against both candidate and running configurations. Conformance test results for running configurations are searchable on the device search page and displayed on both the dashboard and plan details page. Conformance test results for candidate configurations are displayed in the editor and on the plan details page.

The Audit feature records all hardware and software versions, as well as other configuration changes. You can then search and view all changes, or groups of changes. This allows you to track all changes to hardware, software, and configurations with automated versioning - whether made through NetEdit or directly on the switch. You can immediately rollback to any previous configuration. You can perform these rollbacks selectively, based upon factors such as the location of the switches or the date of the changes.



Conformance Validation

Conformance tests run continuously against managed devices

Detects configurations that

- Violate corporate policies or do not comply with network design
- Can cause loss of connectivity between NetEdit and the device

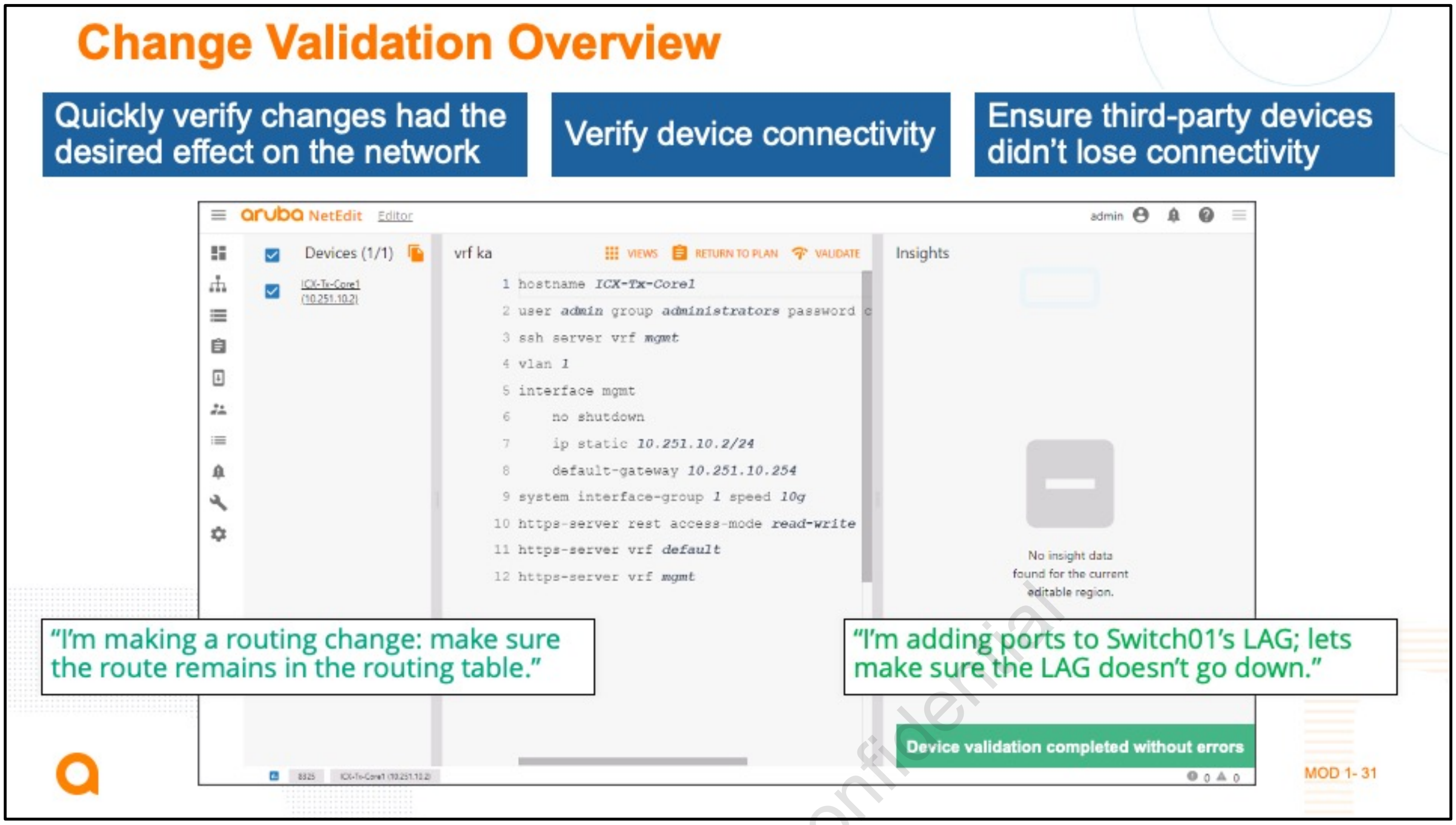
Compares switch configs against conformance tests

Enable/disable tests. If enabled, run continuously against

- Candidate configurations
- Running configurations

Aruba Training-Confidential

Conformance tests run continuously against managed devices. When conformance validation detects a violation in a device's current running and startup configurations it flags the device. You can identify it through a dashboard tile that displays the conformance violating devices.



Change Validation Overview

Quickly verify changes had the desired effect on the network

Verify device connectivity

Ensure third-party devices didn't lose connectivity

```
hostname ICX-Tx-Core1
user admin group administrators password c
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static 10.251.10.2/24
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

No insight data found for the current editable region.

"I'm making a routing change: make sure the route remains in the routing table."

"I'm adding ports to Switch01's LAG; lets make sure the LAG doesn't go down."

Device validation completed without errors

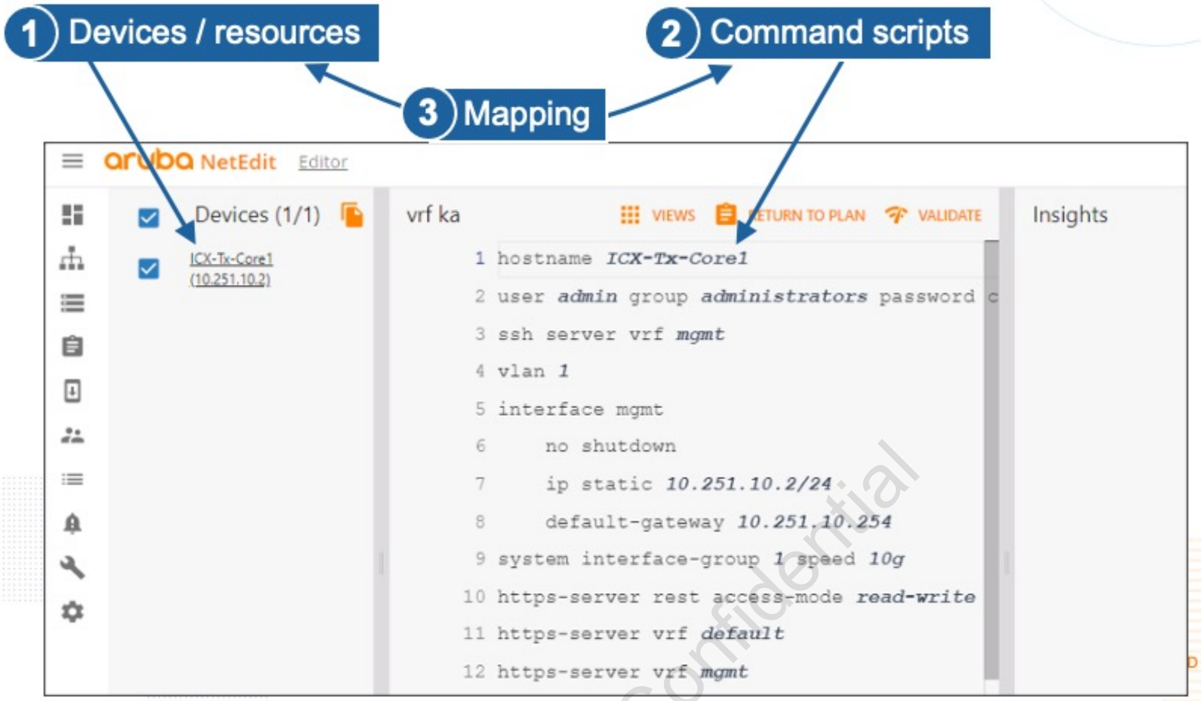
MOD 1- 31

Aruba Training-Confidential

Use Change Validation to determine whether a configuration change gave the desired result. The tool automatically collects network state information, before and after deployment, and presents it in a side-by-side difference report. Network state is gathered using the SSH protocol.

To view the change validation settings or state for a plan, select Change Validation in the plan details page. The parameters are shown before the state is gathered. You can refresh the state information after deployment by selecting the REFRESH button. This is necessary when the network state stabilizes over time, and the difference report will be updated. You can verify expected changes in network state, or detect unexpected by reviewing this report.

Configure Change Validation: Three Parts



```
hostname ICX-Tx-Core1
user admin group administrators password c
ssh server vrf mgmt
vlan 1
interface mgmt
    no shutdown
    ip static 10.251.10.2/24
    default-gateway 10.251.10.254
system interface-group 1 speed 10g
https-server rest access-mode read-write
https-server vrf default
https-server vrf mgmt
```

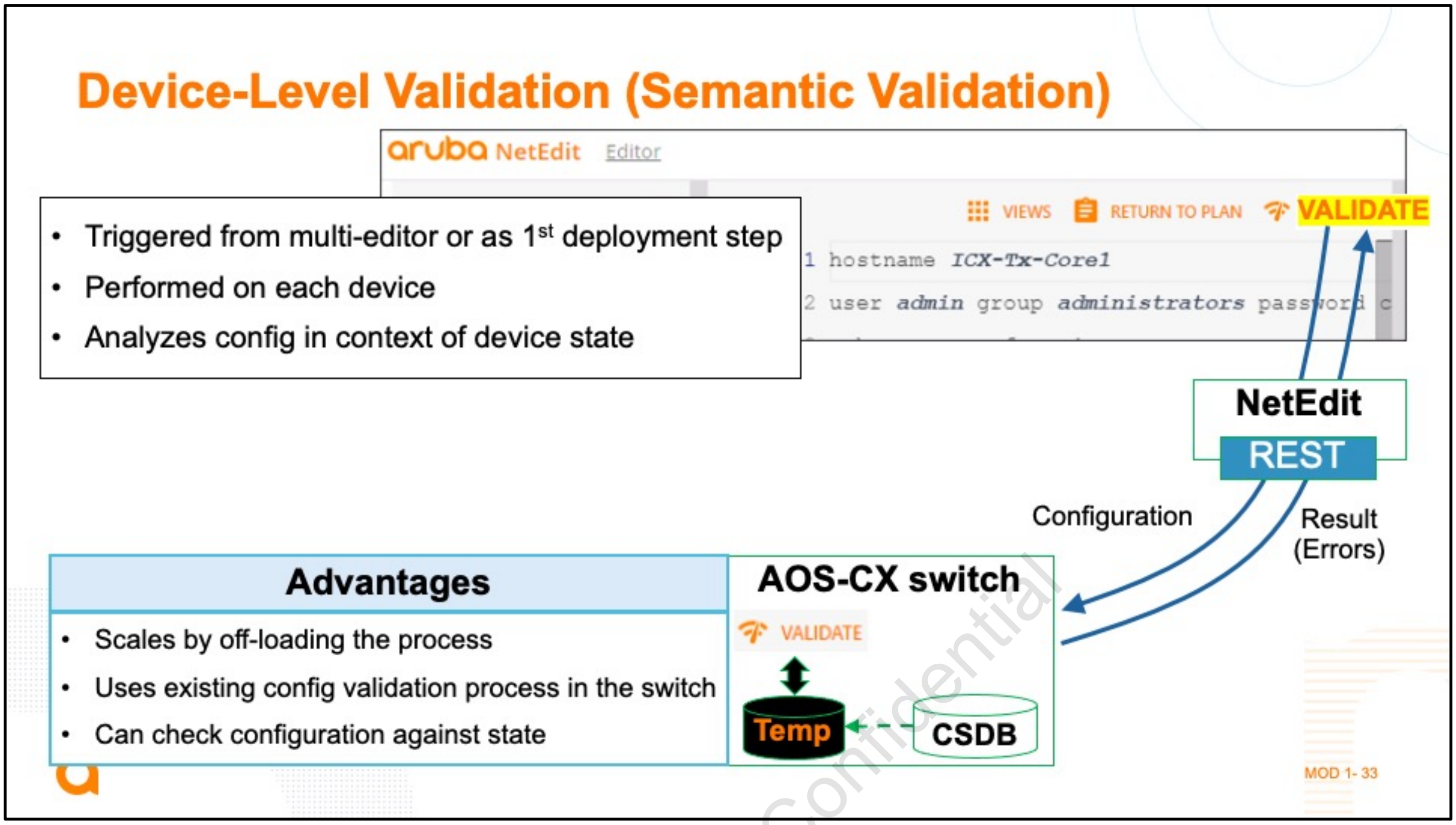
Aruba Training-Confidential

The configuration consists of three parts, the **devices/resources** from which to collect state information; the **command scripts** that display the state information; and a **mapping** of the command scripts to the set of devices/resources where those commands will be executed.

Command Scripts are a series of show and configuration commands that you can create and update

Command Mappings are sets of Command Scripts that run during Plan Deploy. Remember, you see the before and after comparison after applying changes. The “After” deploy can be refreshed.

Notice the options in orange at the top of the Command Scripts window – VIEWS, RETURN TO PLAN, and VALIDATE. Let’s look at that validate option next.



Device-Level Validation (Semantic Validation)

- Triggered from multi-editor or as 1st deployment step
- Performed on each device
- Analyzes config in context of device state

```
hostname ICX-Tx-Core1
user admin group administrators password c
```

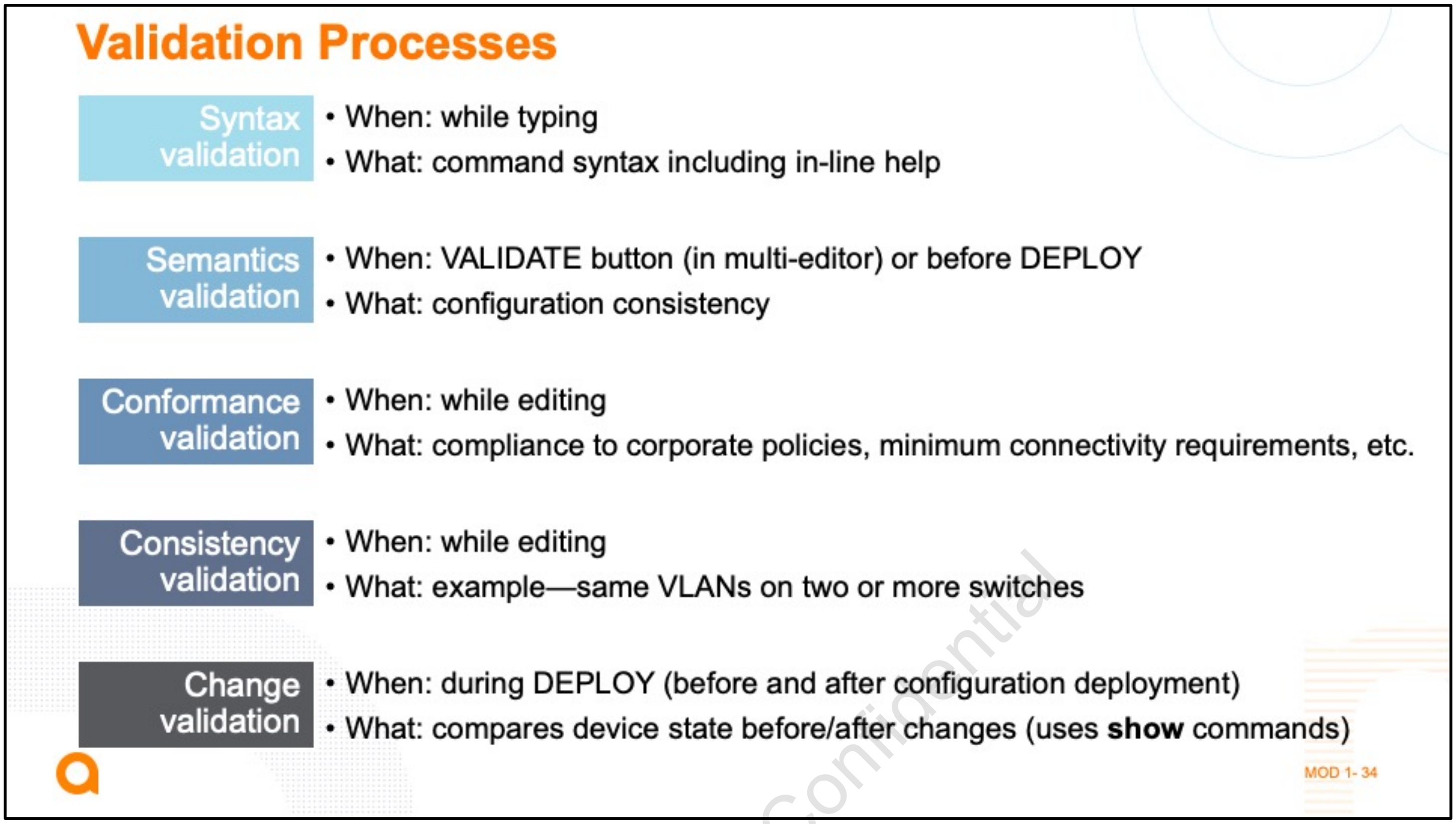
Advantages

- Scales by off-loading the process
- Uses existing config validation process in the switch
- Can check configuration against state

There are several advantages to offloading the process of validation at the device level.

As shown in the figure, NetEdit analyzes the configuration in the context of the device's state using the existing configuration validation process in the switch and creates a temporary copy in the Current State DB (CSDB) to perform the analysis. NetEdit uses the REST interface to send the configuration to the device and to receive the result in JSON format.

Aruba Training-Confidential



Validation Processes

Syntax validation	• When: while typing • What: command syntax including in-line help
Semantics validation	• When: VALIDATE button (in multi-editor) or before DEPLOY • What: configuration consistency
Conformance validation	• When: while editing • What: compliance to corporate policies, minimum connectivity requirements, etc.
Consistency validation	• When: while editing • What: example—same VLANs on two or more switches
Change validation	• When: during DEPLOY (before and after configuration deployment) • What: compares device state before/after changes (uses **show** commands)

MOD 1- 34

Using Validate, you can continuously monitor and ensure conformance to corporate policies. You can perform validation checks on all configuration changes, including those made directly on the switch (i.e., outside of NetEdit). For example, you can verify that all management IP addresses are on the management subnetwork, or that all routers running OSPF are logging adjacency changes.

You can easily customize validation tests can easily be customized and extended based on your existing knowledge of configuration commands.

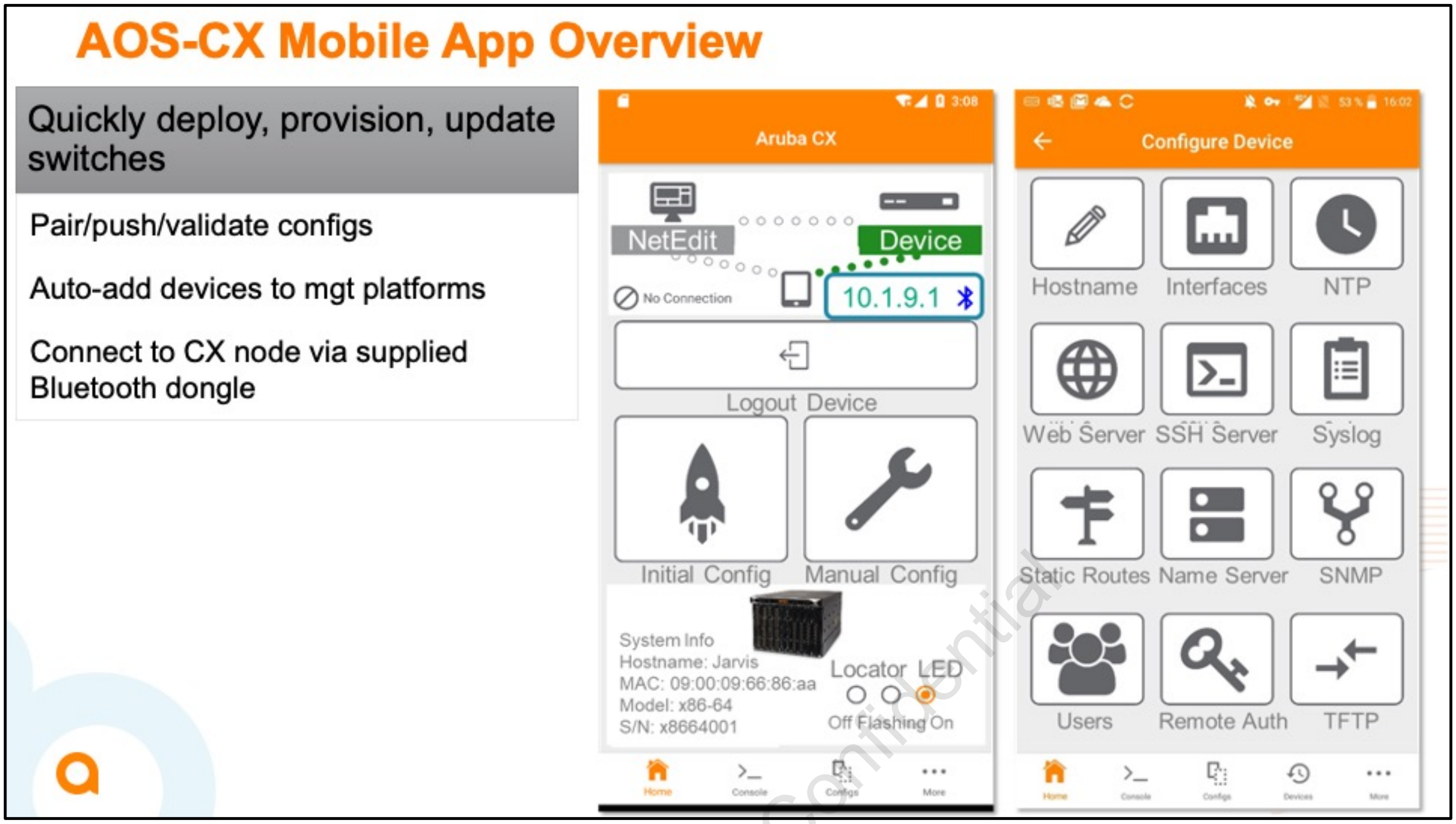
Aruba Training-Confidential

AOS-CX Mobile App

aruba
a Hewlett Packard Enterprise company

Aruba Training-Confidential

Installing switches into an enterprise network just got a whole lot easier. Let’s explore the Aruba CX Mobile App.



AOS-CX Mobile App Overview

Quickly deploy, provision, update switches

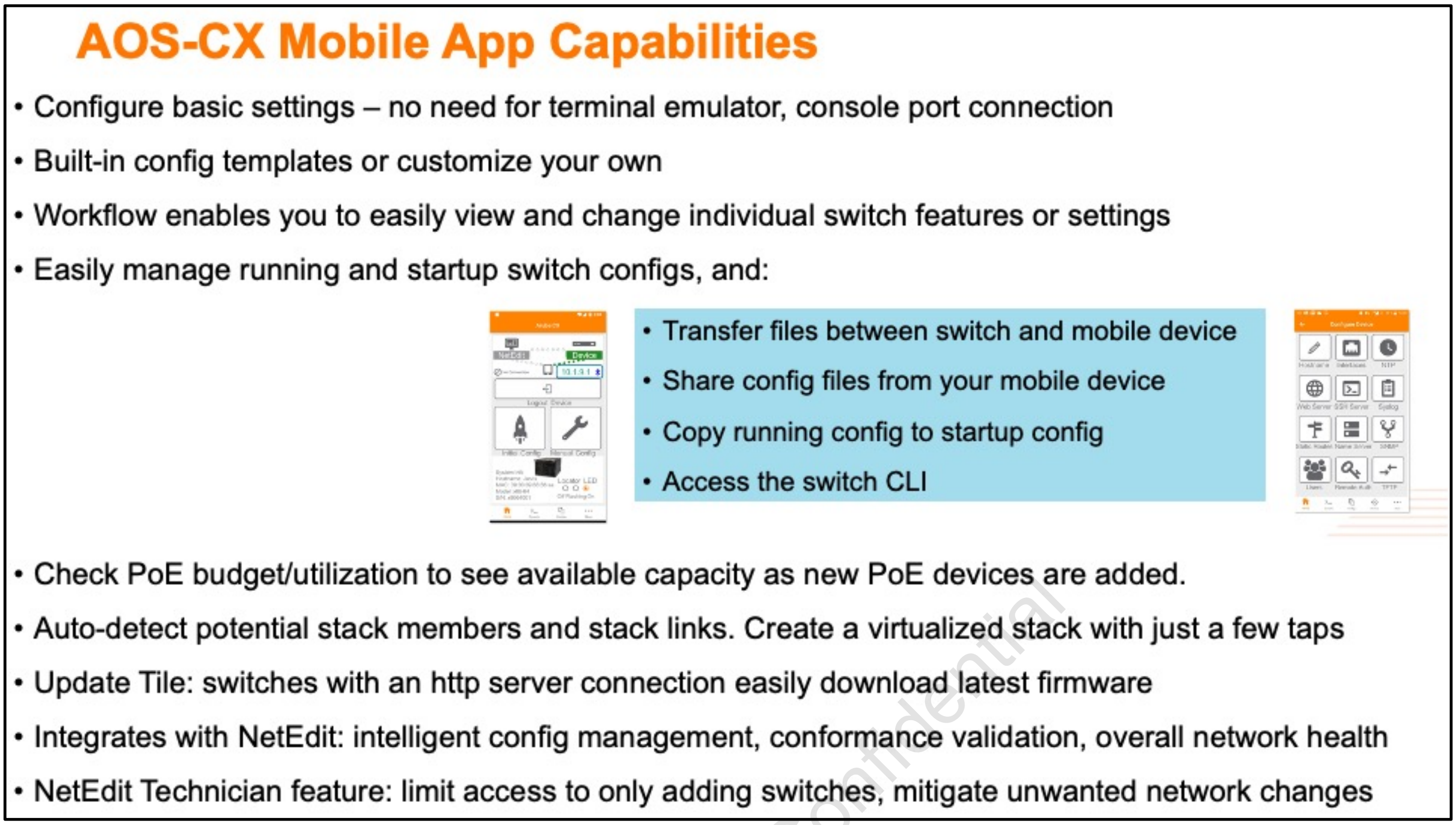
Pair/push/validate configs

Auto-add devices to mgt platforms

Connect to CX node via supplied Bluetooth dongle

The CX mobile app is the go-to application to provision a CX device. Aruba CX mobile app connects to CX node using supplied Bluetooth dongle. Connect your iOS or Android device to Aruba AOS-CX switches via Bluetooth or Wi-Fi for configuration and management

Aruba Training-Confidential

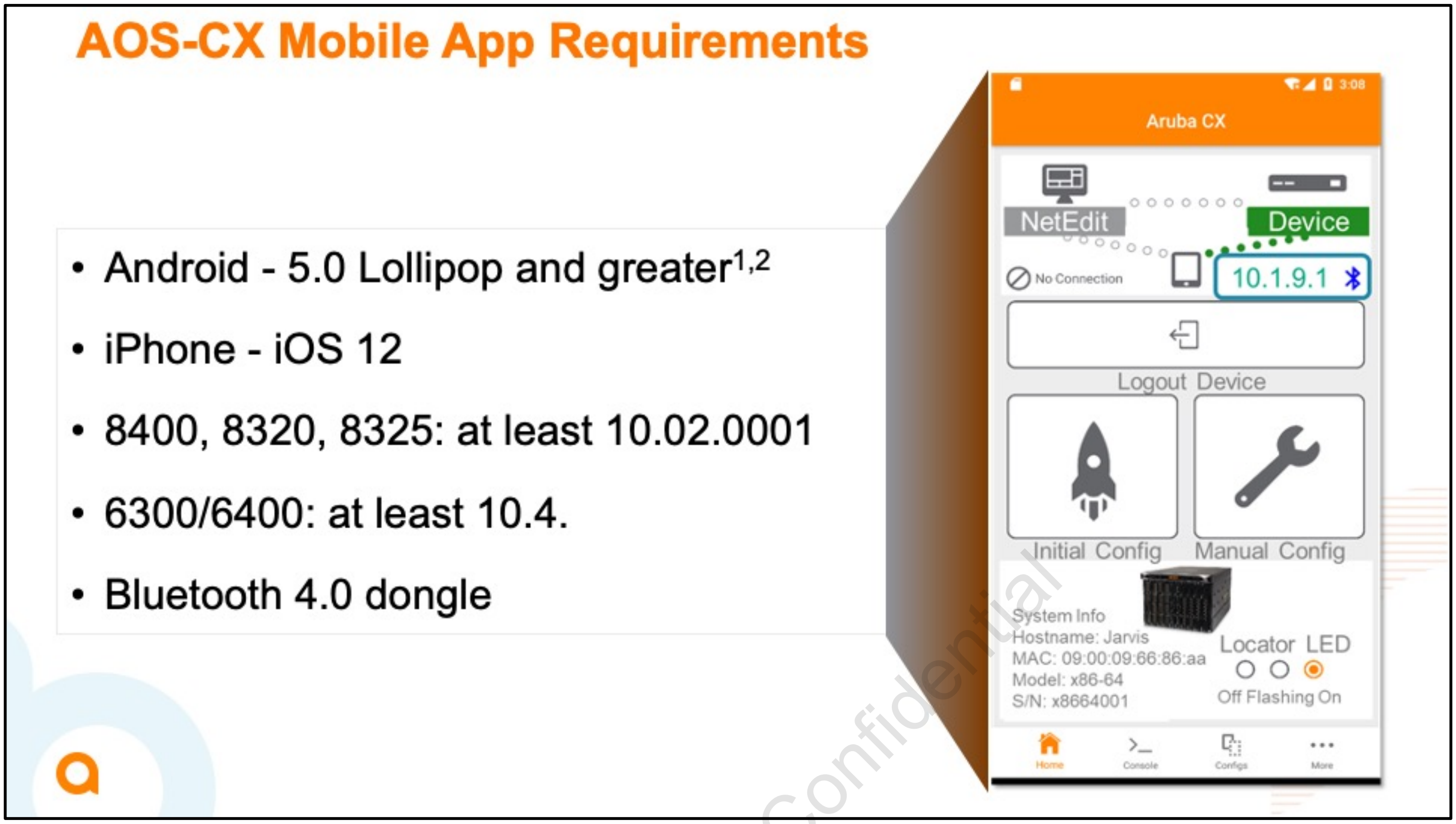


AOS-CX Mobile App Capabilities

- Configure basic settings – no need for terminal emulator, console port connection
- Built-in config templates or customize your own
- Workflow enables you to easily view and change individual switch features or settings
- Easily manage running and startup switch configs, and:
  - Transfer files between switch and mobile device
  - Share config files from your mobile device
  - Copy running config to startup config
  - Access the switch CLI
- Check PoE budget/utilization to see available capacity as new PoE devices are added.
- Auto-detect potential stack members and stack links. Create a virtualized stack with just a few taps
- Update Tile: switches with an http server connection easily download latest firmware
- Integrates with NetEdit: intelligent config management, conformance validation, overall network health
- NetEdit Technician feature: limit access to only adding switches, mitigate unwanted network changes

Please take a moment to review the mobile app capabilities.

Aruba Training-Confidential



AOS-CX Mobile App Requirements

- Android - 5.0 Lollipop and greater[1,2]
- iPhone - iOS 12
- 8400, 8320, 8325: at least 10.02.0001
- 6300/6400: at least 10.4.
- Bluetooth 4.0 dongle

Aruba Training-Confidential

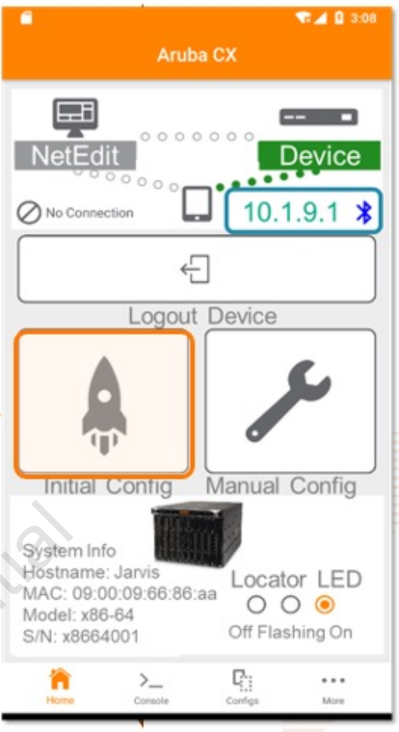
The figure summarizes minimum supported requirements for the AOS-CX Mobile AppSome Android phones don’t support simultaneously.

[1] BlueTooth IP stack and Wi-Fi IP stack. Example: OnePlus3, Sony and etc. With such devices, the NetEdit device import is not possible as part of the Initial Config upload process.

[2] On Android, Internet Sharing has to be enabled while pairing the AOS-CX switch with the phone (otherwise device will not get an IP address over Bluetooth).

AOS-CX Mobile App Initial Config

- REST API calls to AOS-CX and NetEdit server
- DryRun API validates config before sending the actual running configuration.
- Access to Template Selection, ability to deploy initial config without access to NetEdit server



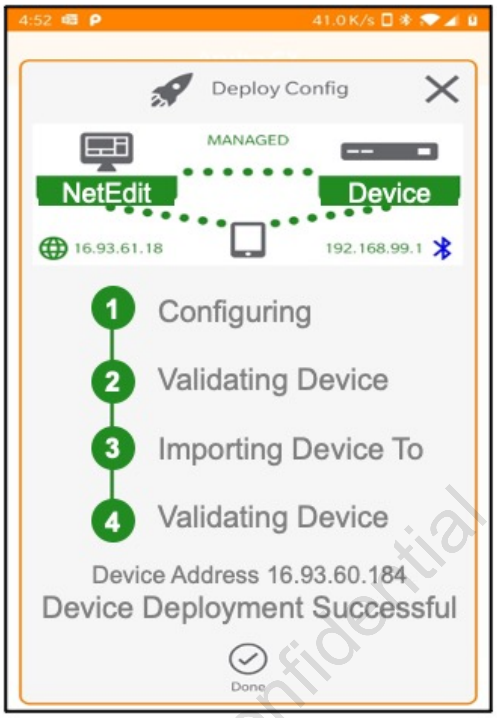
System Info
Hostname: Jarvis
MAC: 09:00:09:66:86:aa
Model: x86-64
S/N: x8664001

Initial Config Menu actions use REST API call to AOS-CX and to NetEdit server. CX mobile app uses the DryRun API to validate configuration before configuration deployment to the actual running configuration (like NetEdit). This provides access to the Template Selection and the capability to deploy initial configuration even without having access to NetEdit server.

Aruba Training-Confidential

AOS-CX Mobile App Version 2.0 Updates

- RBAC
- Firmware updates
- Stacking Automation
- PoE Utilization



1 Configuring
2 Validating Device
3 Importing Device To
4 Validating Device

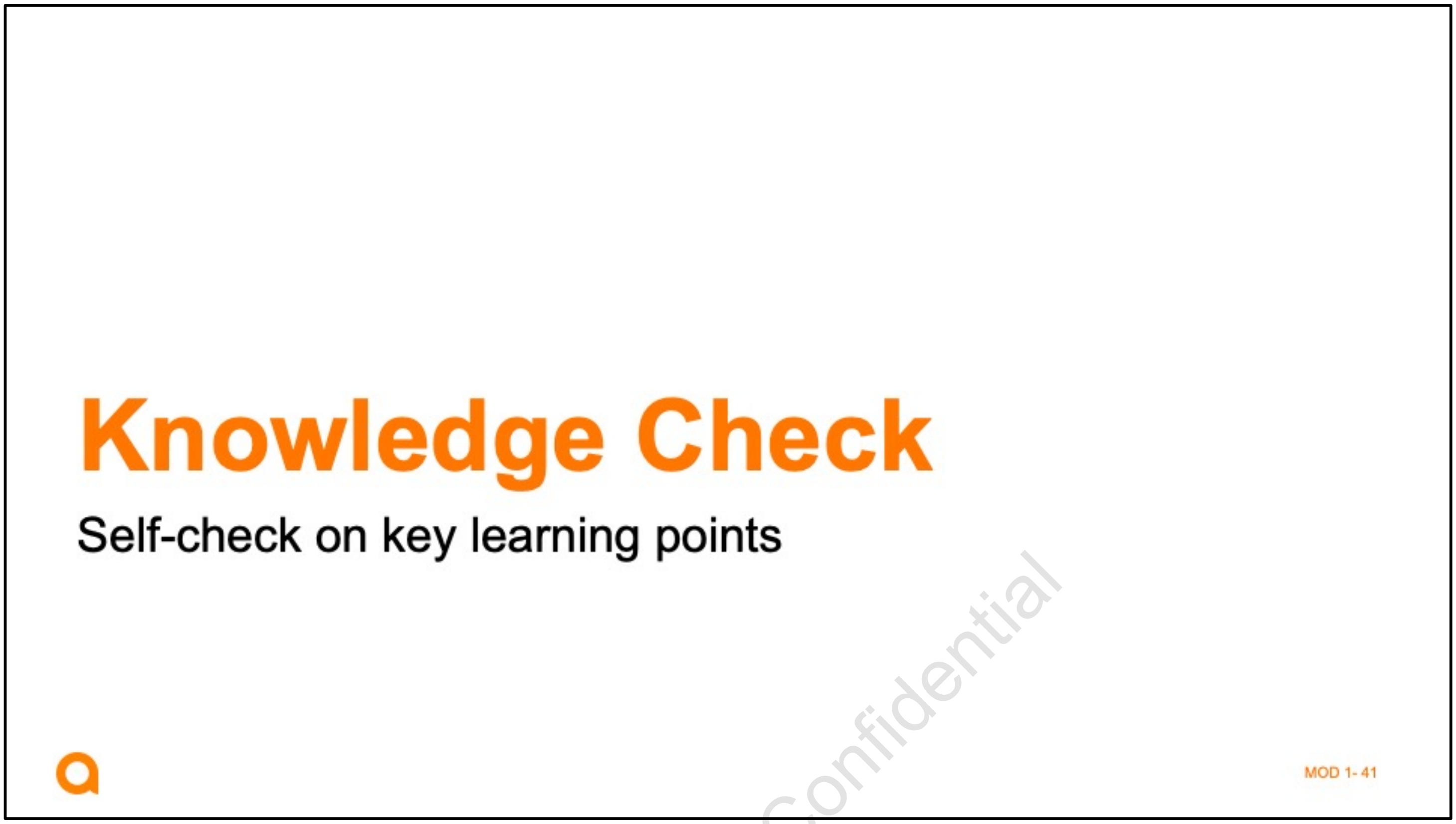
Device Address 16.93.60.184
Device Deployment Successful

MOD 1- 40

The CX mobile app is the go-to application to provision a CX device. Version 2.0 allows you to Stack switches, see PoE utilization, and upgrade devices via HTTPS.

Note: See the Aruba CX Mobile App User Guide for more information about using the app to manage/monitor a switch(es).

Aruba Training-Confidential

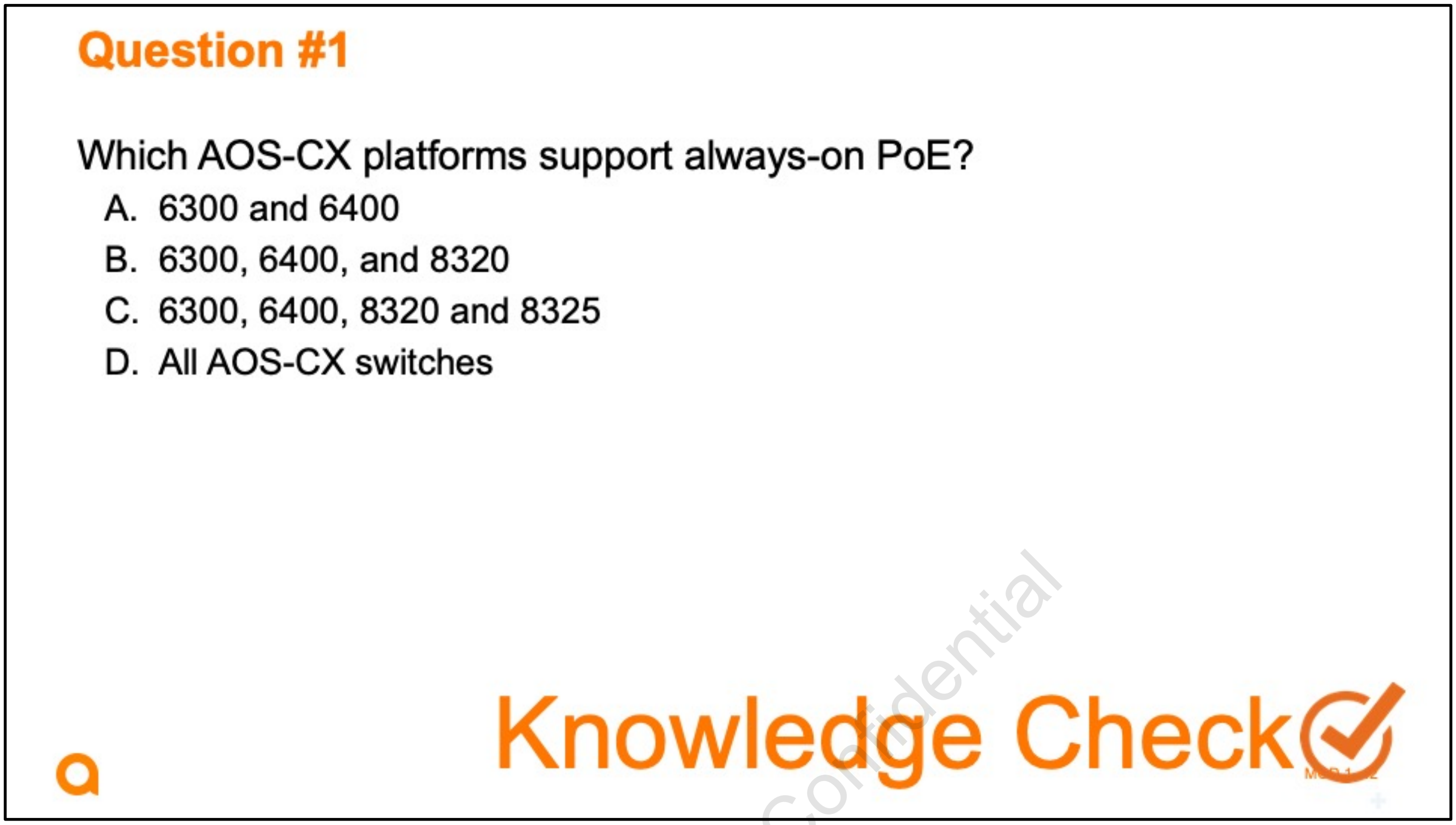


Knowledge Check

Self-check on key learning points

MOD 1- 41

Aruba Training-Confidential



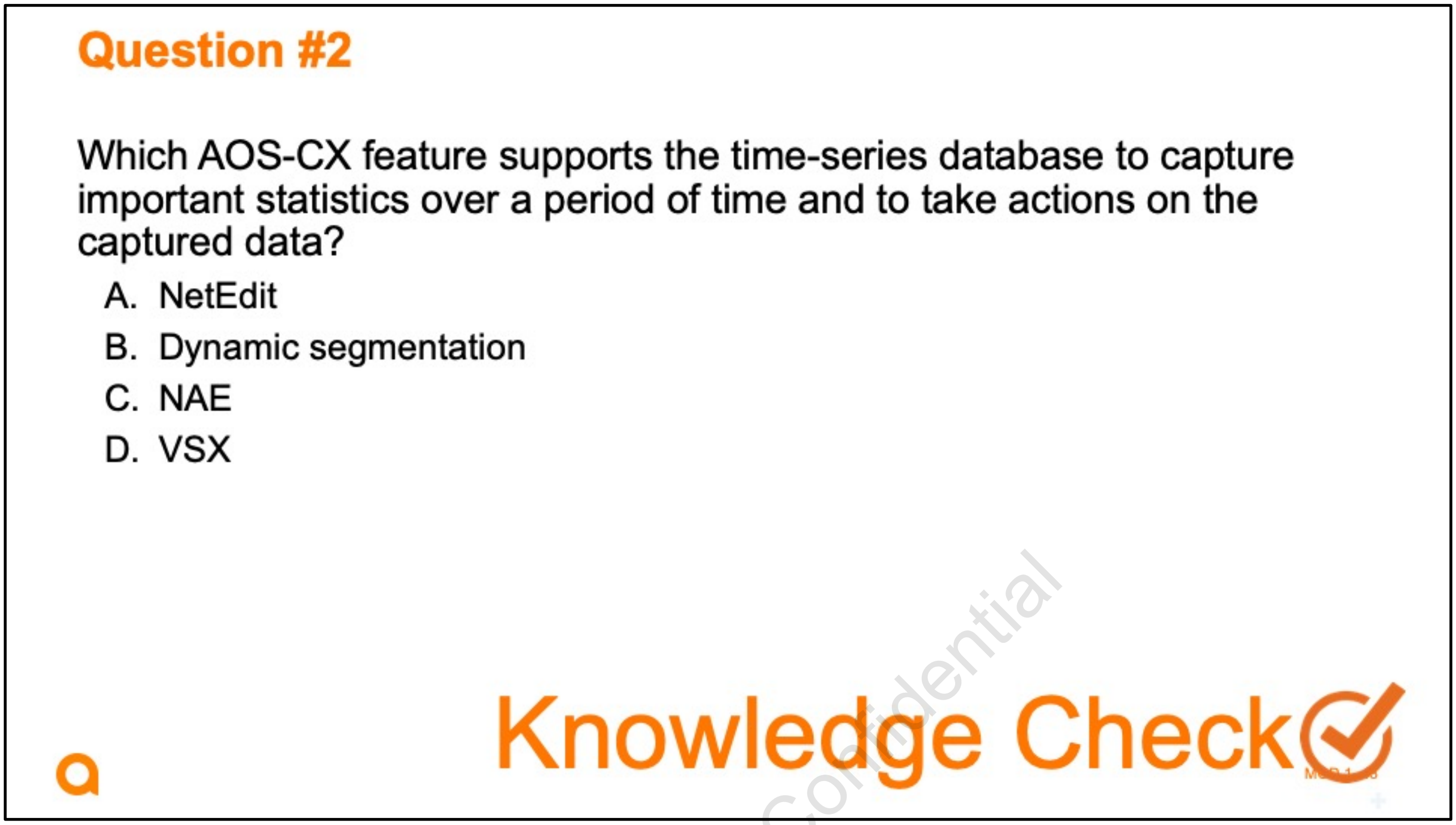
Question #1

Which AOS-CX platforms support always-on PoE?

A. 6300 and 6400
B. 6300, 6400, and 8320
C. 6300, 6400, 8320 and 8325
D. All AOS-CX switches

Knowledge Check

Aruba Training-Confidential



Question #2

Which AOS-CX feature supports the time-series database to capture important statistics over a period of time and to take actions on the captured data?

A. NetEdit

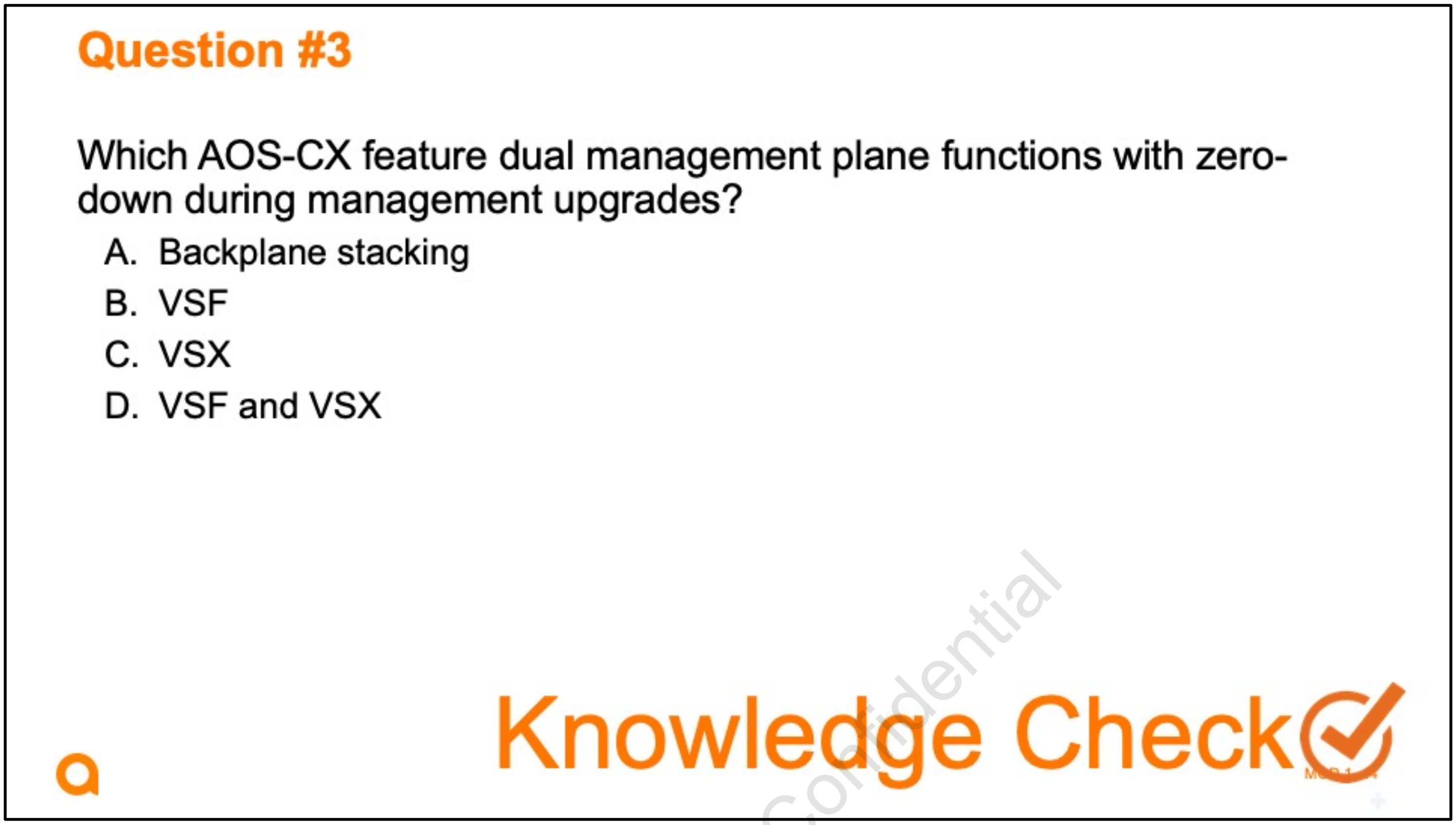
B. Dynamic segmentation

C. NAE

D. VSX

Knowledge Check

Aruba Training-Confidential



Question #3

Which AOS-CX feature dual management plane functions with zero-down during management upgrades?

A. Backplane stacking

B. VSF

C. VSX

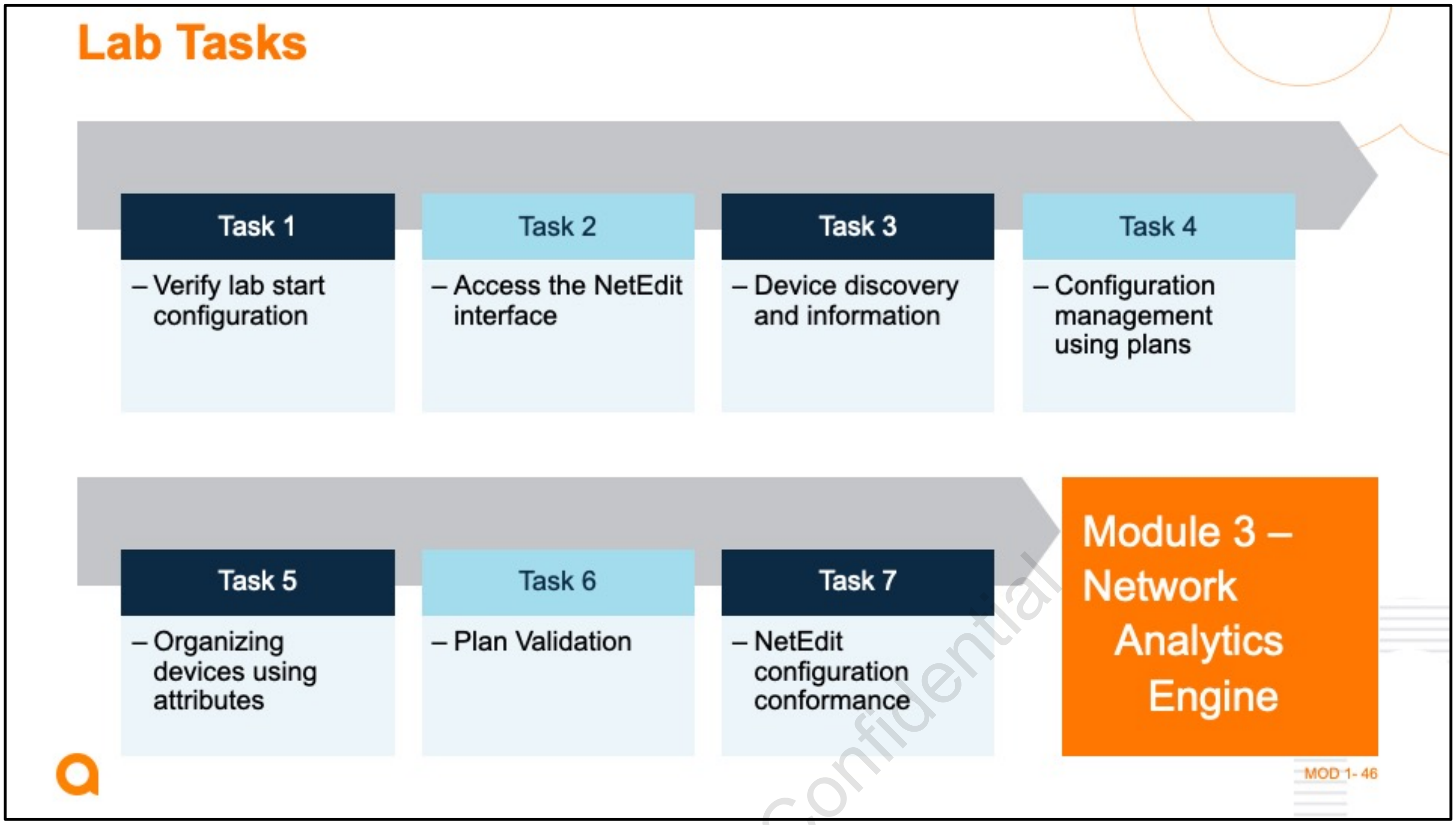
D. VSF and VSX

Knowledge Check✓

Aruba Training-Confidential

Lab Activity

Aruba Training-Co



Lab Tasks

Task 1

– Verify lab start configuration

Task 2

– Access the NetEdit interface

Task 3

– Device discovery and information

Task 4

– Configuration management using plans

Task 5

– Organizing devices using attributes

Task 6

– Plan Validation

Task 7

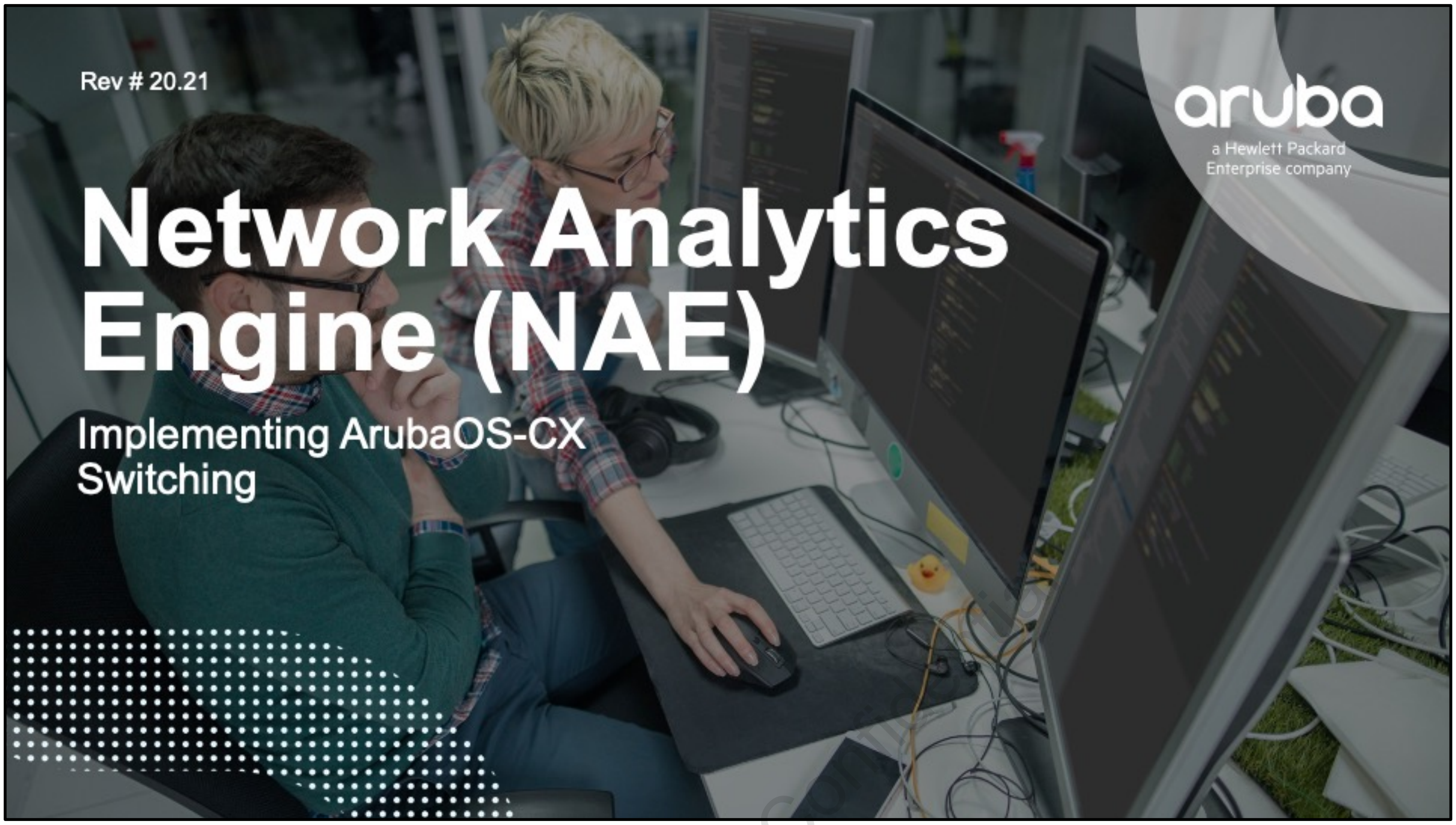
– NetEdit configuration conformance

Module 3 – Network Analytics Engine

MOD 1- 46

The figure provides a brief review of lab tasks. Please see your lab guide for details.

Aruba Training-Confidential



Rev # 20.21

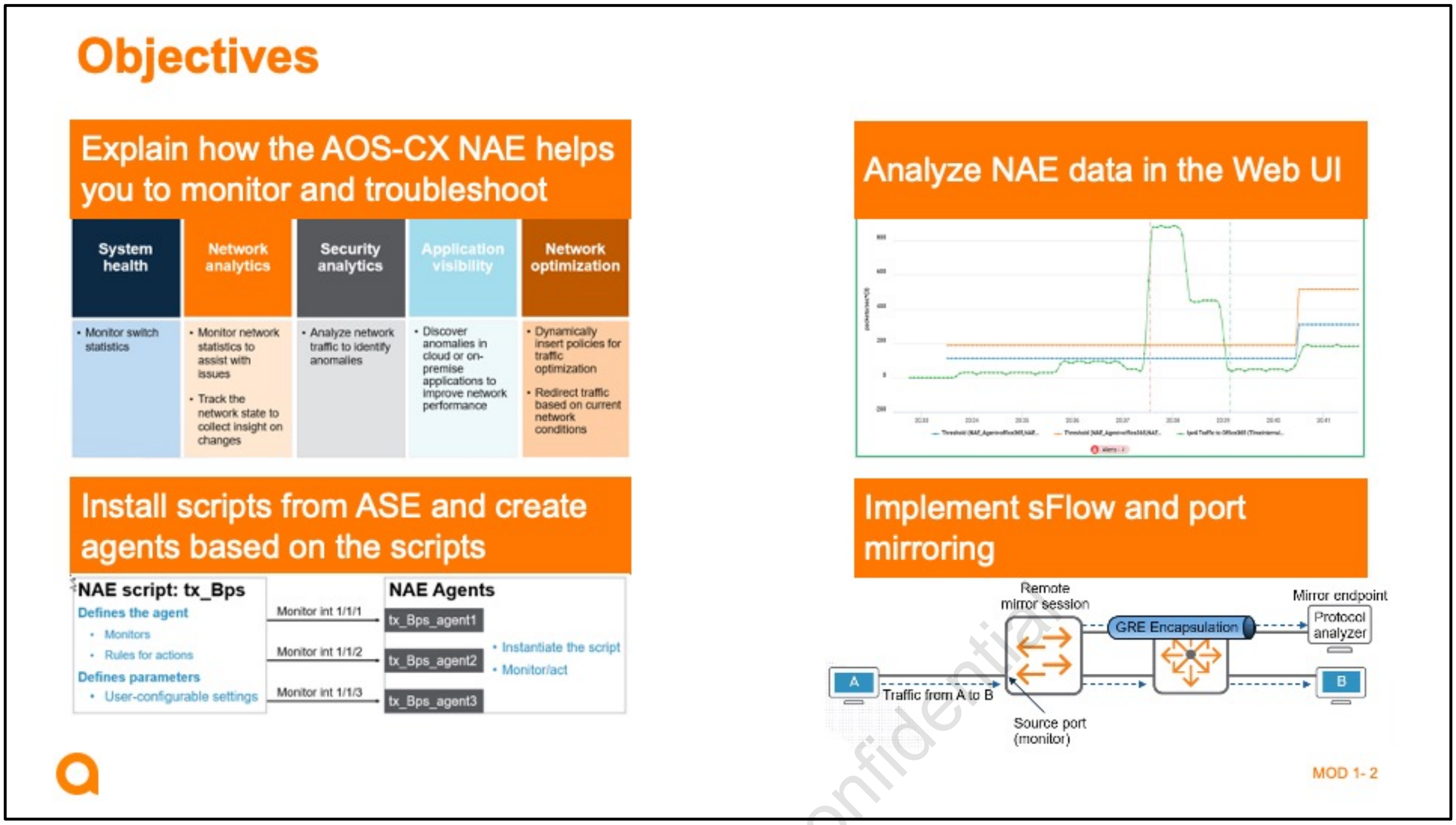
Network Analytics Engine (NAE)

Implementing ArubaOS-CX Switching

aruba
a Hewlett Packard Enterprise company

Hi team, and welcome back to the switching course. Let's talk about NAE.

Aruba Training-Confidential



Objectives

Explain how the AOS-CX NAE helps you to monitor and troubleshoot

System health	Network analytics	Security analytics	Application visibility	Network optimization
• Monitor switch statistics	• Monitor network statistics to assist with issues • Track the network state to collect insight on changes	• Analyze network traffic to identify anomalies	• Discover anomalies in cloud or on-premise applications to improve network performance	• Dynamically insert policies for traffic optimization • Redirect traffic based on current network conditions

Analyze NAE data in the Web UI

Install scripts from ASE and create agents based on the scripts

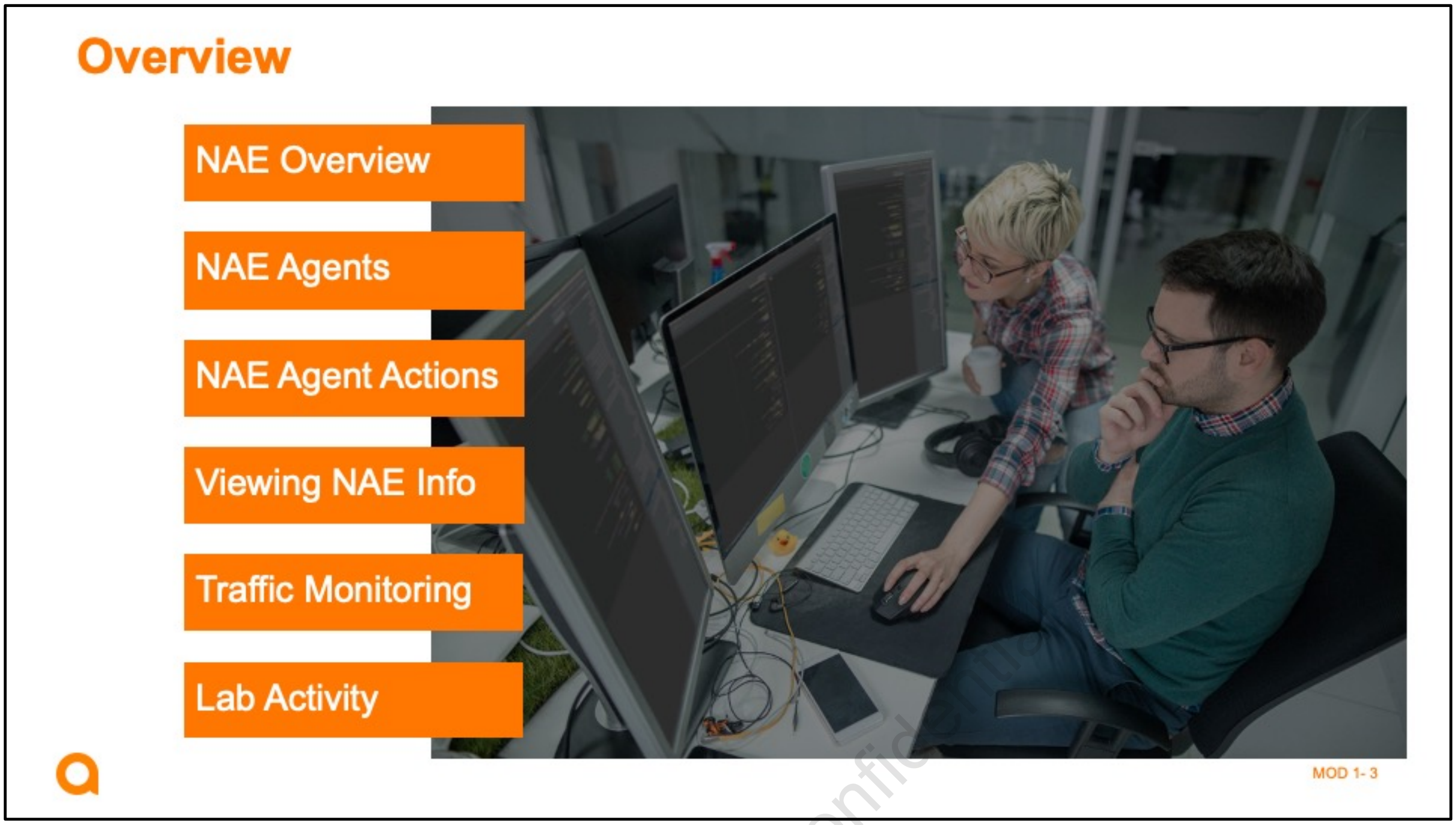
Implement sFlow and port mirroring

MOD 1- 2

Aruba Training-Confidential

After completing this module, you should be able to:

- Explain how the AOS-CX NAE helps you to monitor and troubleshoot
- Analyze NAE data in the Web UI
- Install scripts from ASE and create agents based on the scripts
- Implement sFlow and port mirroring



Overview

NAE Overview

NAE Agents

NAE Agent Actions

Viewing NAE Info

Traffic Monitoring

Lab Activity

MOD 1- 3

Aruba Training-Confidential

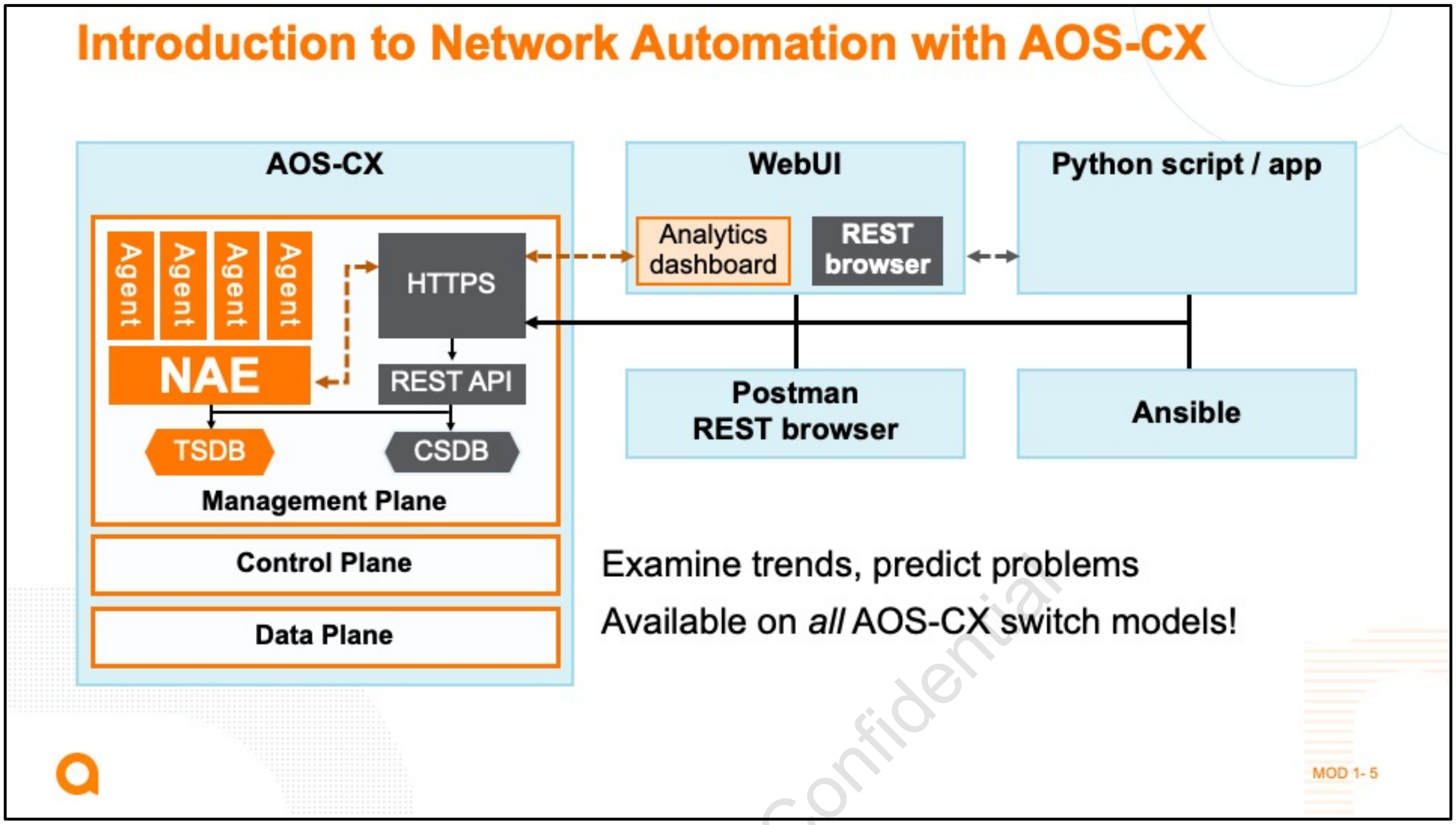
This module teaches you all about the AOS-CX Network Analytics Engine (NAE). By the time that you have completed this module, you will be able to explain how the AOS-CX NAE helps you to monitor and troubleshoot. You will be able to analyze NAE data in the Web UI. You will be able to install scripts from the Aruba Solutions Exchange (ASE) and create agents based on the scripts.

Expand the section titles for quick preview.

NAE Overview

Let’s get acquainted with NAE.

Aruba Training-Confidential



Introduction to Network Automation with AOS-CX

Examine trends, predict problems

Available on *all* AOS-CX switch models!

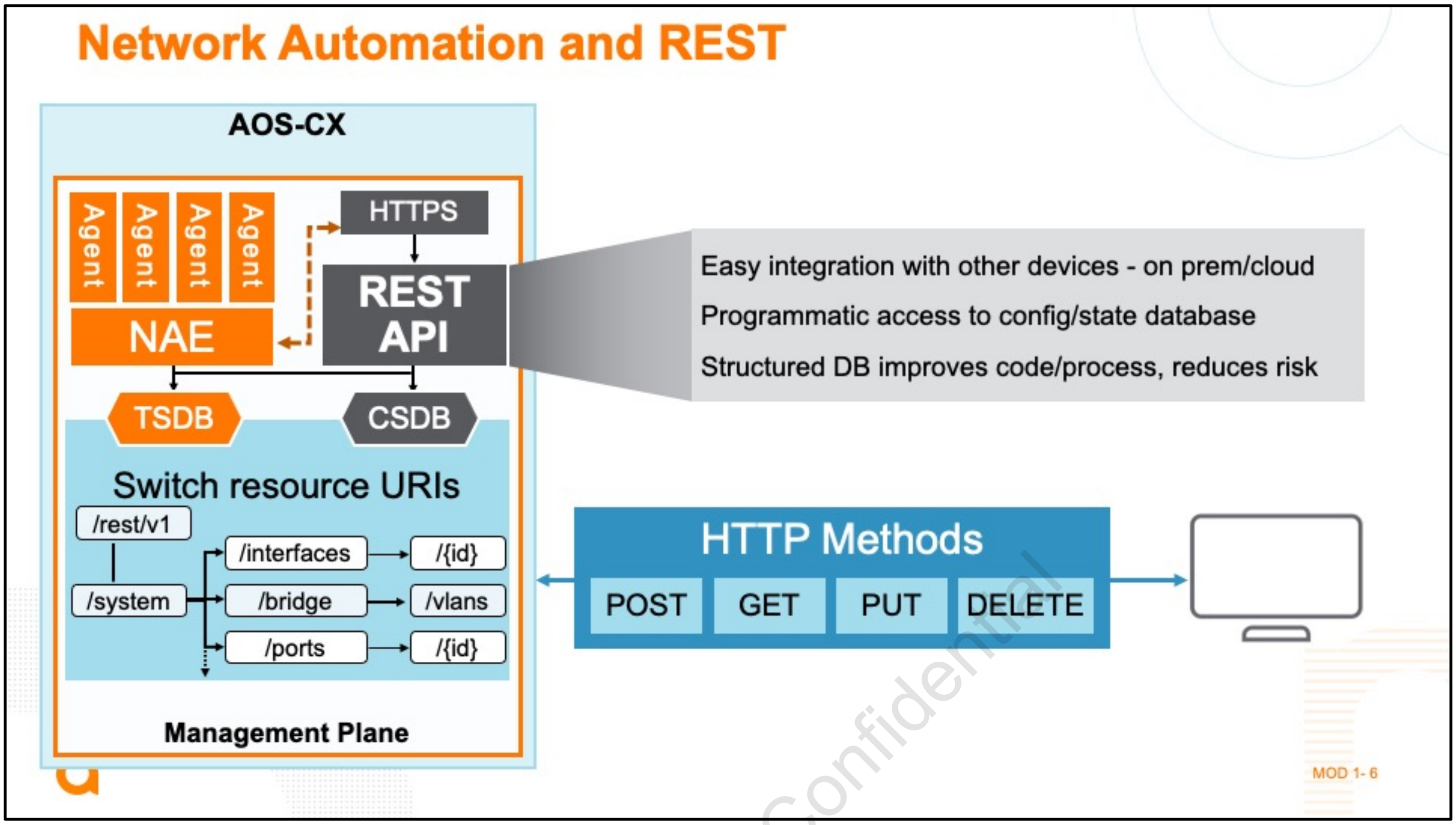
MOD 1- 5

Aruba Training-Confidential

This NAE engine is integrated with the AOS-CX system configuration and time series databases, enabling you to examine historical trends and predict future problems due to scale, security, and performance bottlenecks. With that information, you can create software modules that automatically detect issues, take appropriate actions, and reduce the time spent on manual tasks.

NAE uses python-based agents, developed by Aruba and the Aruba community, to automate network monitor and troubleshoot functions. Agents monitor the state, settings, and statistics for system components in the current state database (CSDB) – perhaps the number of bytes transmitted on an interface or the switch's operating temperature. The agent then writes ongoing results to the time series database (TSDB) for ongoing analysis.

You can interact with the NAE through the Web UI's analytics dashboard. You can also automate the NAE through the REST API using tools such as Postman, Ansible, or another choice. You can create or obtain Python scripts for new agents and load the agents on the switch using the Web UI or the REST API.



Network Automation and REST

Easy integration with other devices - on prem/cloud

Programmatic access to config/state database

Structured DB improves code/process, reduces risk

MOD 1- 6

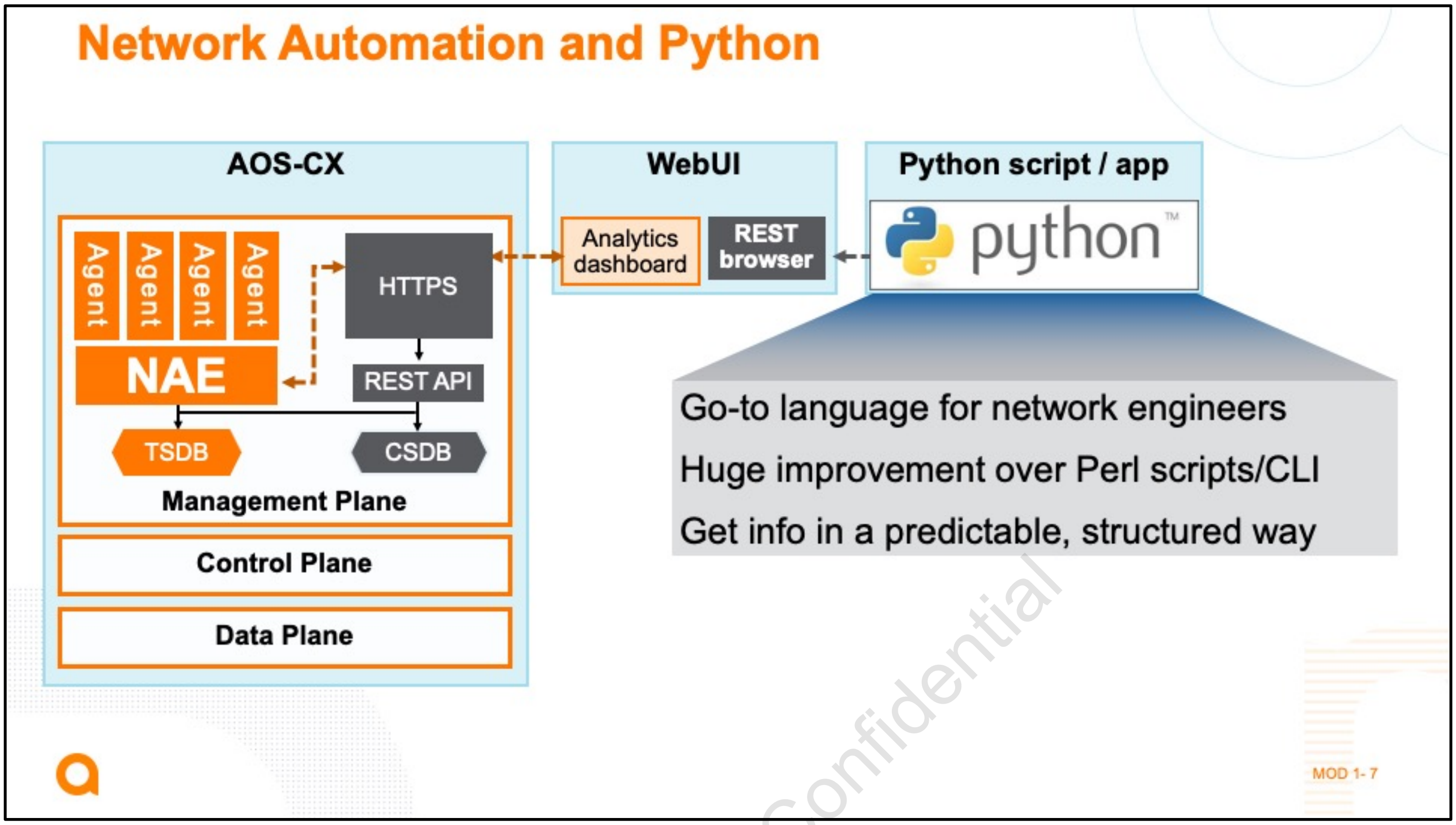
Aruba Training-Confidential

Switches running the AOS-CX software are fully programmable with a REpresentational State Transfer (REST) API, allowing easy integration with other devices both on premises and in the cloud. This programmability—combined with the Aruba Network Analytics Engine—accelerates network administrator understanding of and response to network issues.

The AOS-CX REST API enables programmatic access to the AOS-CX configuration and state database at the heart of the switch. By using a structured model, changes to the content and formatting of the CLI output do not affect the programs you write. And because the configuration is stored in a structured database instead of a text file, rolling back changes is easier than ever, thus dramatically reducing a risk of downtime and performance issues.

The AOS-CX REST API is a web service that performs operations on switch resources using HTTPS POST, GET, PUT, and DELETE methods.

A switch resource is indicated by its Uniform Resource Identifier (URI). A URI can be made up of several components, including the host name or IP address, port number, the path, and an optional query string. The AOS-CX operating system includes the AOS-CX REST API Reference, which is a web interface based on the Swagger UI. The AOS-CX REST API Reference provides the reference documentation for the REST API, including resources URIs, models, methods, and errors.



Network Automation and Python

AOS-CX

Agent

Agent

Agent

Agent

NAE

HTTPS

REST API

TSDB

CSDB

Management Plane

Control Plane

Data Plane

WebUI

Analytics dashboard

REST browser

Python script / app

python™

Go-to language for network engineers

Huge improvement over Perl scripts/CLI

Get info in a predictable, structured way

MOD 1- 7

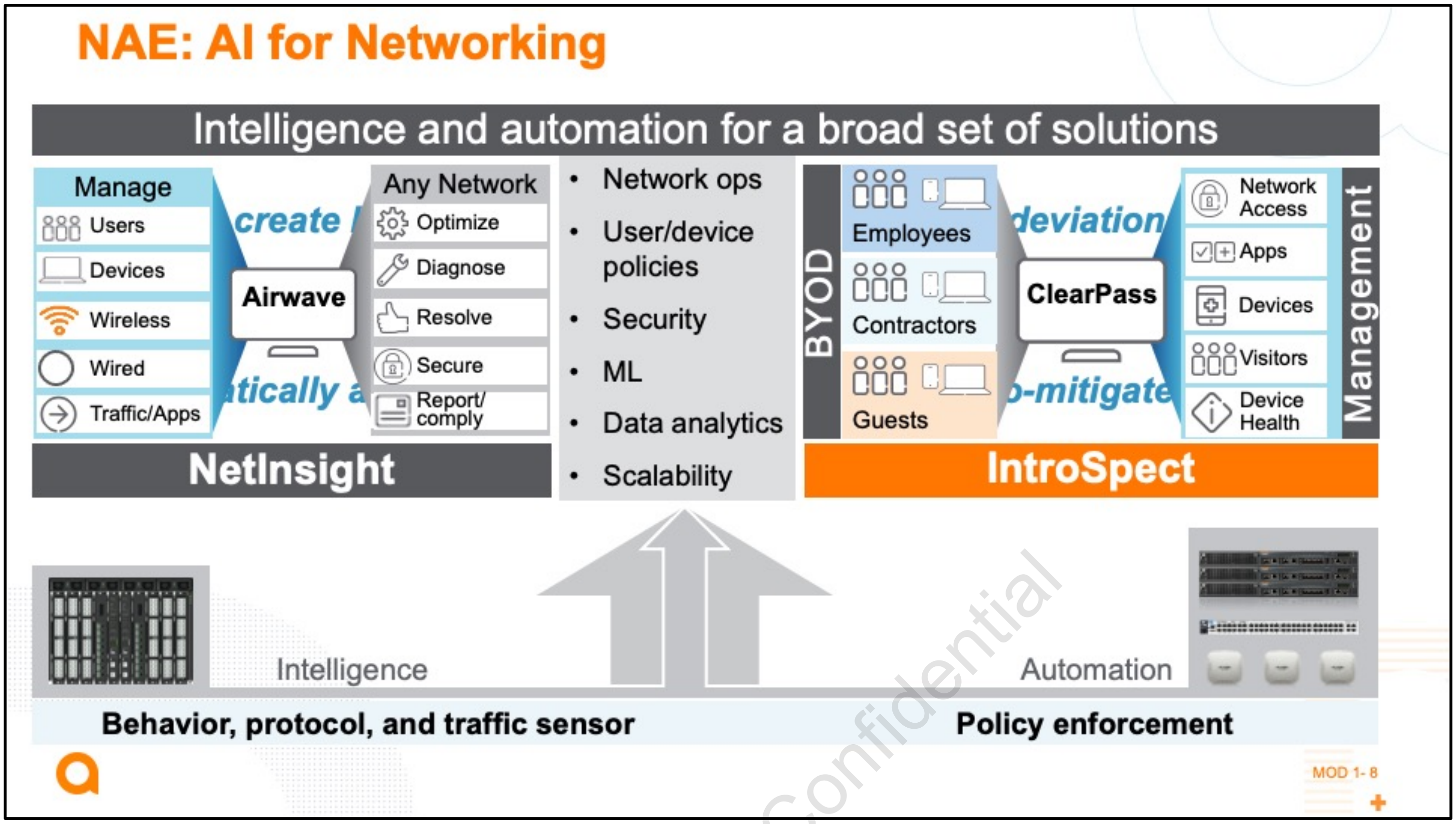
Aruba Training-Confidential

You write NAE scripts in Python, so NAE includes a built-in Python interpreter, used to validate scripts and creating agents from scripts.

Python is the go-to language for network engineers:

- It is high-level and human and machine readable.
- It is popular with an active development community.
- There are many libraries (code written for you that you can use in your programs) available.

Python and the REST API to the AOS-CX database provide powerful tools to support network automation. By using Python and the REST API, you can move far beyond CLI scripting in network automation.



NAE: AI for Networking

Intelligence and automation for a broad set of solutions

- Network ops
- User/device policies
- Security
- ML
- Data analytics
- Scalability

Behavior, protocol, and traffic sensor

Policy enforcement

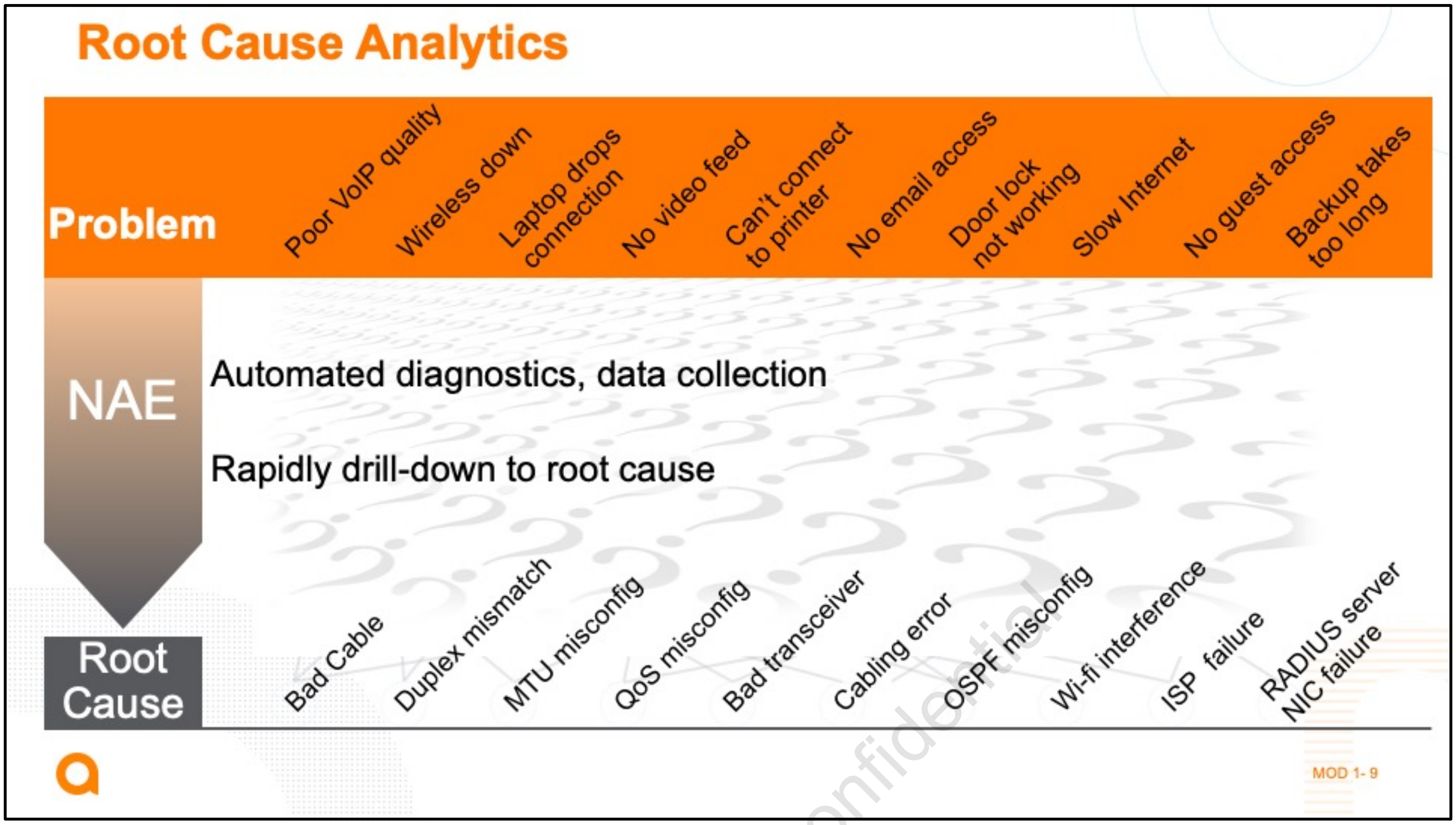
MOD 1- 8

Aruba Training-Confidential

Using NAE, intelligent AOS-CX switches provide a foundation for security, DevOps/operation automation, supportability, capacity planning, and monitoring/root cause analysis.

Agents create baselines for specific system components and monitor excessive deviations from those baselines, and other system aspects too. When pre-defined conditions occur, including excessive baseline deviations, NAE can automatically alert administrators or even mitigate the problem.

| NAE can also feed information to external management solutions using REST calls. These solutions include AirWave and ClearPass, along with Aruba | NetInsight for AI-powered analytics and Aruba IntroSpect for User and Entity Behavior Analytics (UEBA). It also integrates with third-party IT service ticketing solutions like ServiceNow and TOPdesk.



Root Cause Analytics

Automated diagnostics, data collection

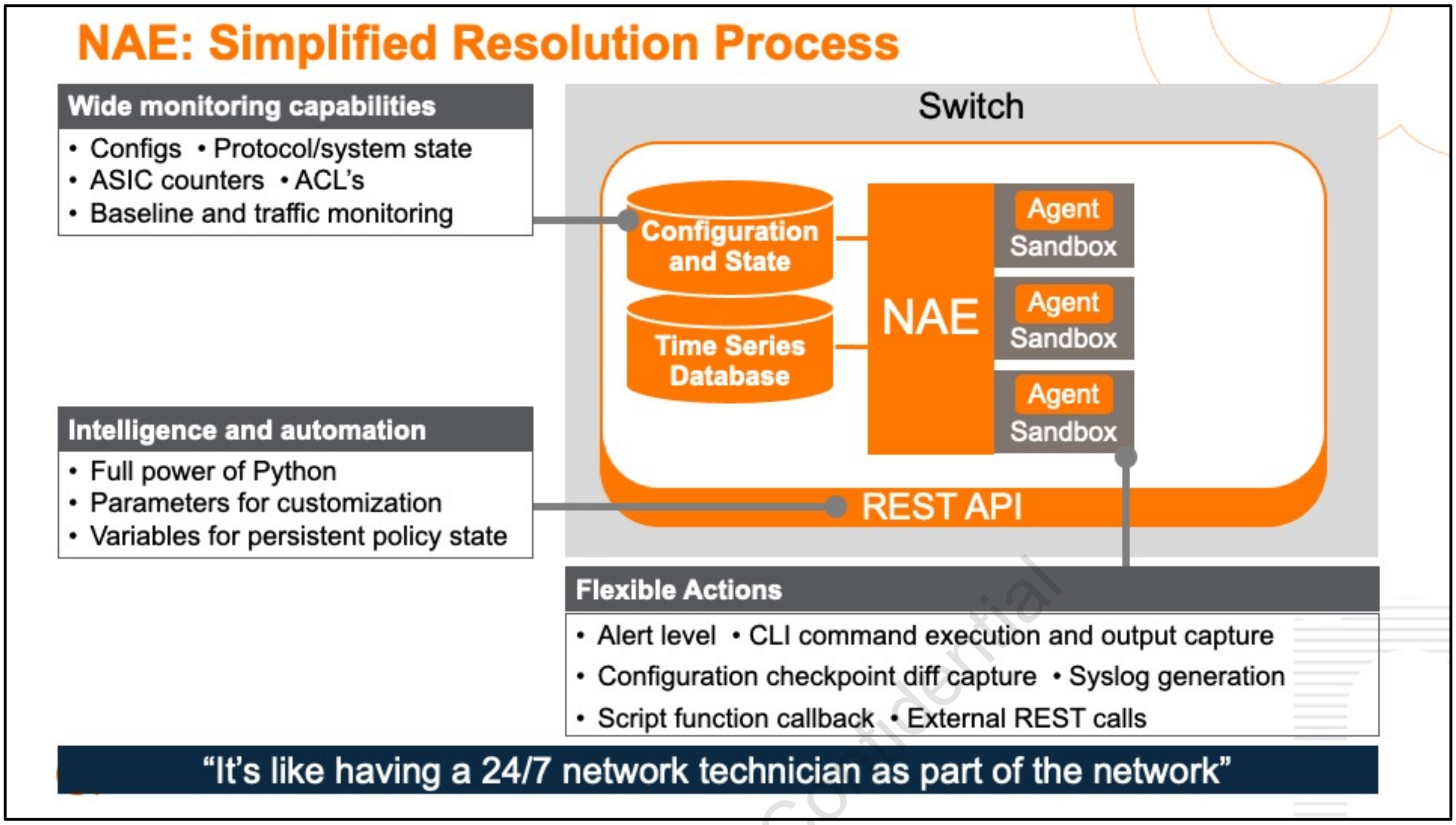
Rapidly drill-down to root cause

Aruba Training-Confidential

Modern networks face unprecedented complexities and scaling. Networks must handle more users, each with multiple devices.

A typical troubleshooting process includes many time-consuming steps. You must first become aware of a problem, perhaps reactively waiting until users complain. These user complaints can be vague, so you must assess the issue and its impact to derive a more accurate problem description. Then comes the laborious gathering of required information to formulate hypotheses, perhaps slogging through configs and log files to make and test hypothesis. Then, if appropriate, you deploy a change and check for the intended effect.

NAE can automate this process and rapidly drill down to the root cause.



NAE: Simplified Resolution Process

Wide monitoring capabilities

- Configs • Protocol/system state
- ASIC counters • ACL's
- Baseline and traffic monitoring

Intelligence and automation

- Full power of Python
- Parameters for customization
- Variables for persistent policy state

Switch

Configuration and State

Time Series Database

NAE

Agent Sandbox

Agent Sandbox

Agent Sandbox

REST API

Flexible Actions

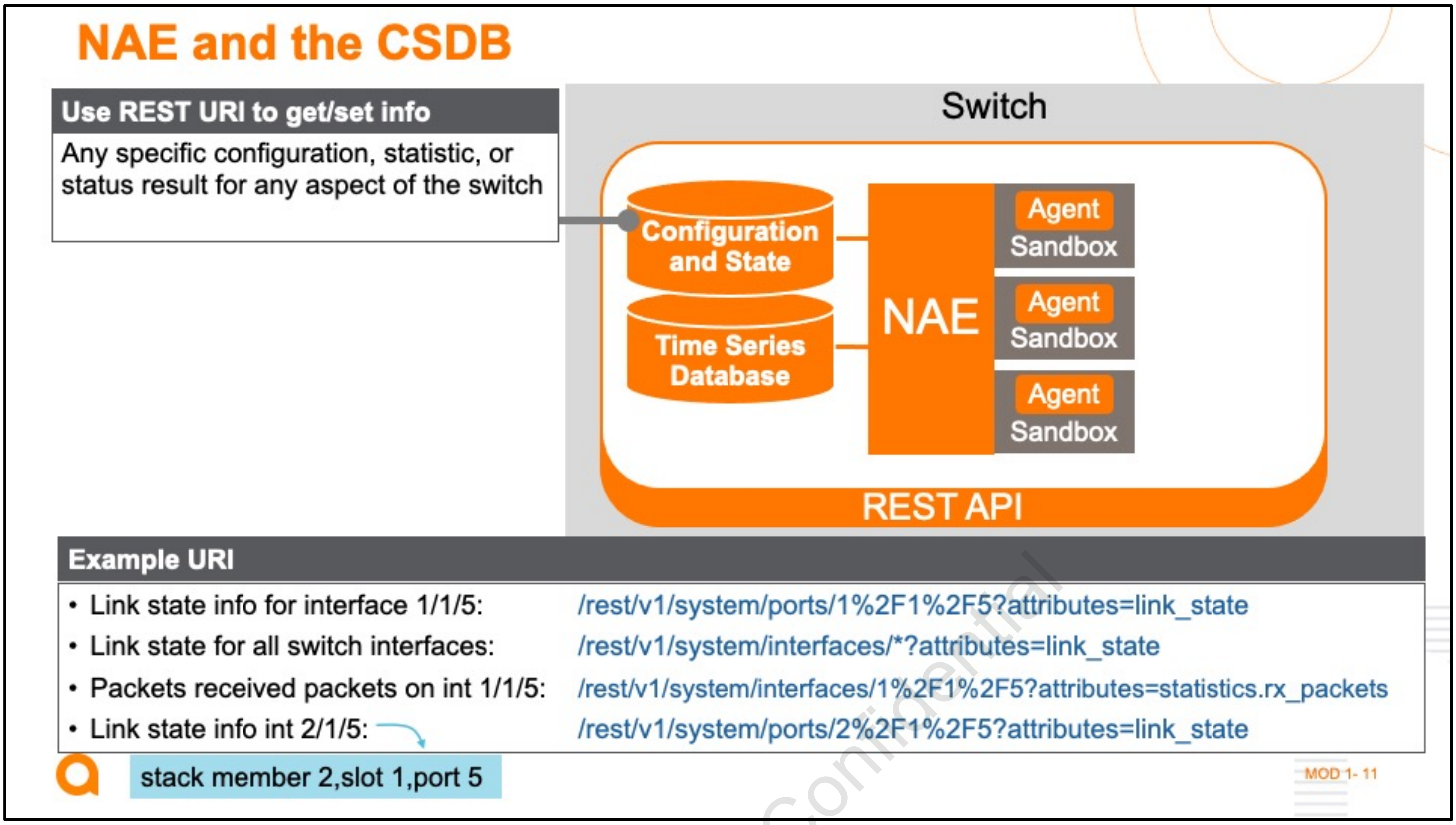
- Alert level • CLI command execution and output capture
- Configuration checkpoint diff capture • Syslog generation
- Script function callback • External REST calls

"It's like having a 24/7 network technician as part of the network"

Aruba Training-Confidential

Key to this ease of monitoring and troubleshooting is two cooperating databases. The Configuration and State DB (CSDB) monitors any value you want to see - configurations, protocol and system state, ASIC statistics, ACLs, and more. Then those statistics are mapped to the Time Series DB (TSDB).

The TSDB monitors any value you want to see in the CSDB, but correlated to a Time Series - you get a realistic, real-time time model of network behavior. NAE Agents give visual representation of values and provides automation. Every aspect of the switch configuration and state information is modeled in the AOS-CX switch database and is accessible through the REST API.



NAE and the CSDB

Use REST URI to get/set info

Any specific configuration, statistic, or status result for any aspect of the switch

Example URI

• Link state info for interface 1/1/5:	/rest/v1/system/ports/1%2F1%2F5?attributes=link_state
• Link state for all switch interfaces:	/rest/v1/system/interfaces/*?attributes=link_state
• Packets received packets on int 1/1/5:	/rest/v1/system/interfaces/1%2F1%2F5?attributes=statistics.rx_packets
• Link state info int 2/1/5:	/rest/v1/system/ports/2%2F1%2F5?attributes=link_state

stack member 2,slot 1,port 5

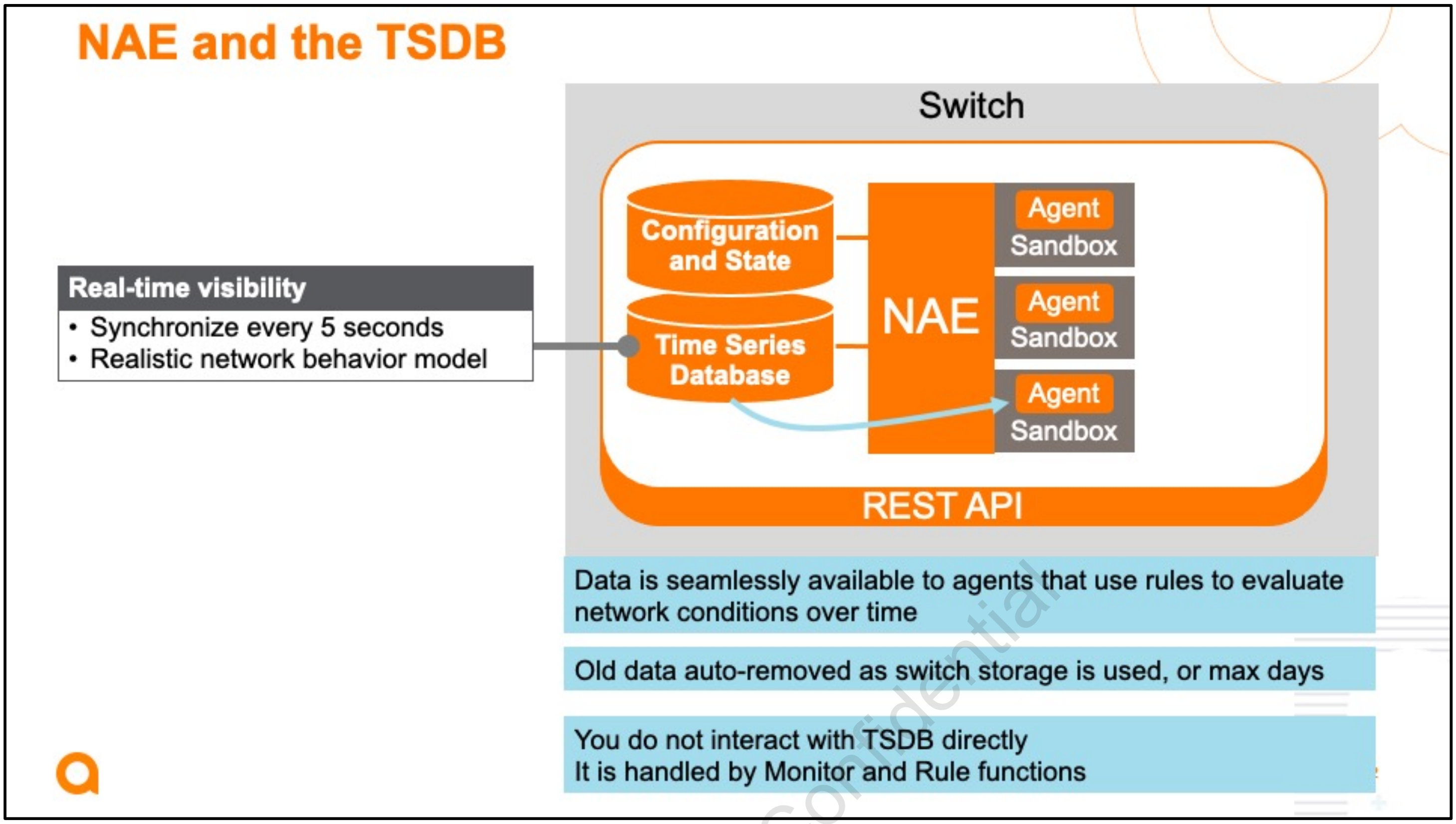
MOD 1- 11

Aruba Training-Confidential

Developers can use the REST URI to access a specific configuration, statistic, or status result for any aspect of the switch. The figure shows several example URIs.

- For link state information about interface 1/1/5:
  /rest/v1/system/ports/1%2F1%2F5?attributes=link_state
- For link state of all switch interfaces:
  /rest/v1/system/interfaces/*?attributes=link_state
- To see number of received packets on interface 1/1/5:
  /rest/v1/system/interfaces/1%2F1%2F5?attributes=statistics.rx_packets
- For link state info for interface 2/1/5 (stack member 2, slot 1, port 5):
  /rest/v1/system/ports/2%2F1%2F5?attributes=link_state

Note that switches that are members of a stack are treated as a single switch.



NAE and the TSDB

Real-time visibility

- Synchronize every 5 seconds
- Realistic network behavior model

Data is seamlessly available to agents that use rules to evaluate network conditions over time

Old data auto-removed as switch storage is used, or max days

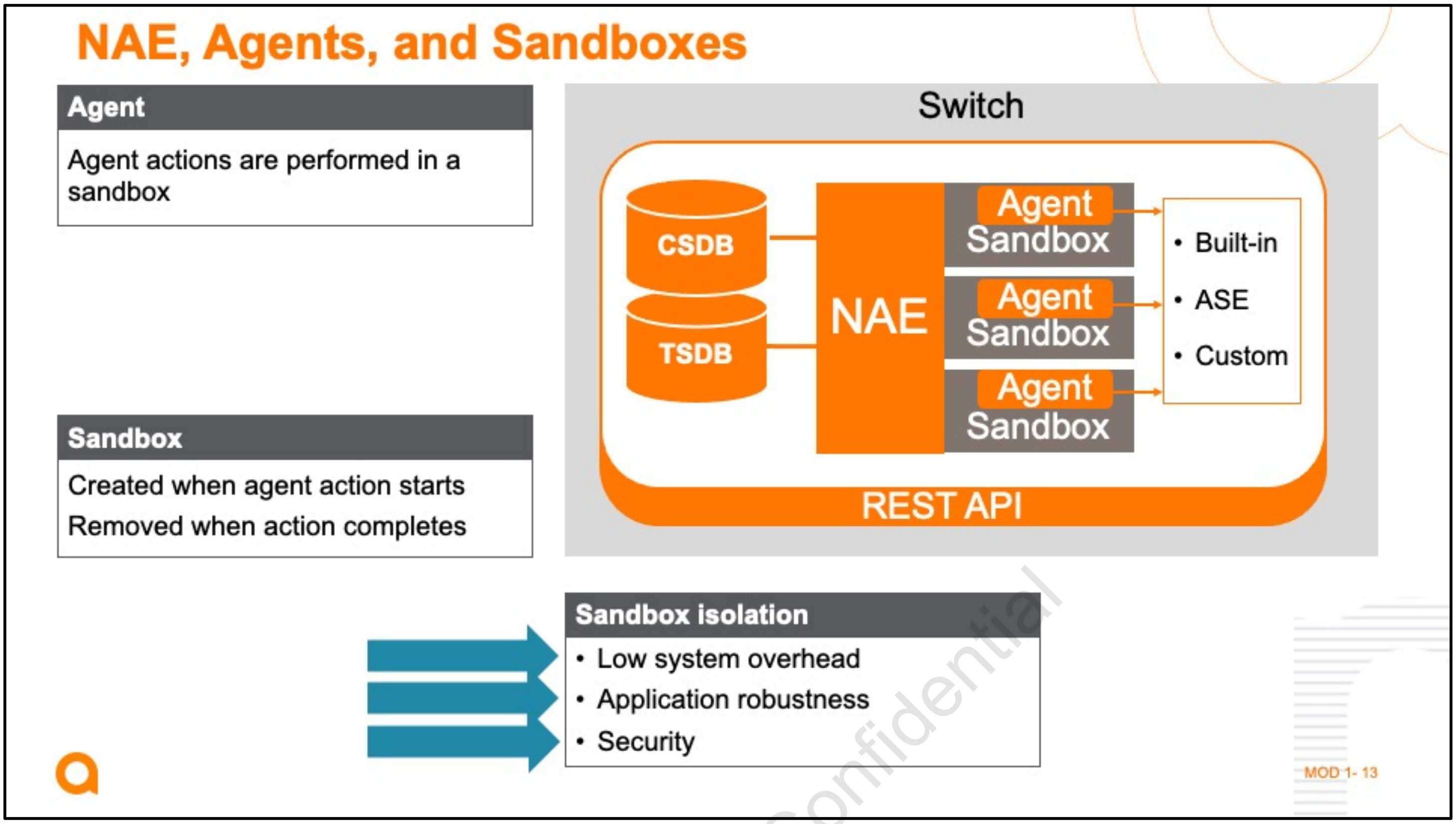
You do not interact with TSDB directly
It is handled by Monitor and Rule functions

The Aruba NAE includes a built-in time series database. Time-series data about the resources monitored by agents are automatically collected and presented in graphs in the switch Web UI. The database makes the data seamlessly available to agents that use rules that evaluate network conditions over time.

Old time-series data is removed automatically either as the storage space on the switch is used, or as the maximum number of days of data is reached. The amount of storage space consumed at any given time depends on the number of switch resources being monitored at that time. Each monitored resource creates one time series. Each time series consumes approximately 240 KB of storage for each day.

When creating a script, software developers do not interact with this database directly. Use of the database is automatically handled by the Monitor and Rule functions.

Aruba Training-Confidential



NAE, Agents, and Sandboxes

Agent

Agent actions are performed in a sandbox

Sandbox

Created when agent action starts

Removed when action completes

Switch

CSDB

TSDB

NAE

Agent Sandbox

Agent Sandbox

Agent Sandbox

- Built-in
- ASE
- Custom

REST API

Sandbox isolation

- Low system overhead
- Application robustness
- Security

MOD 1- 13

Aruba Training-Confidential

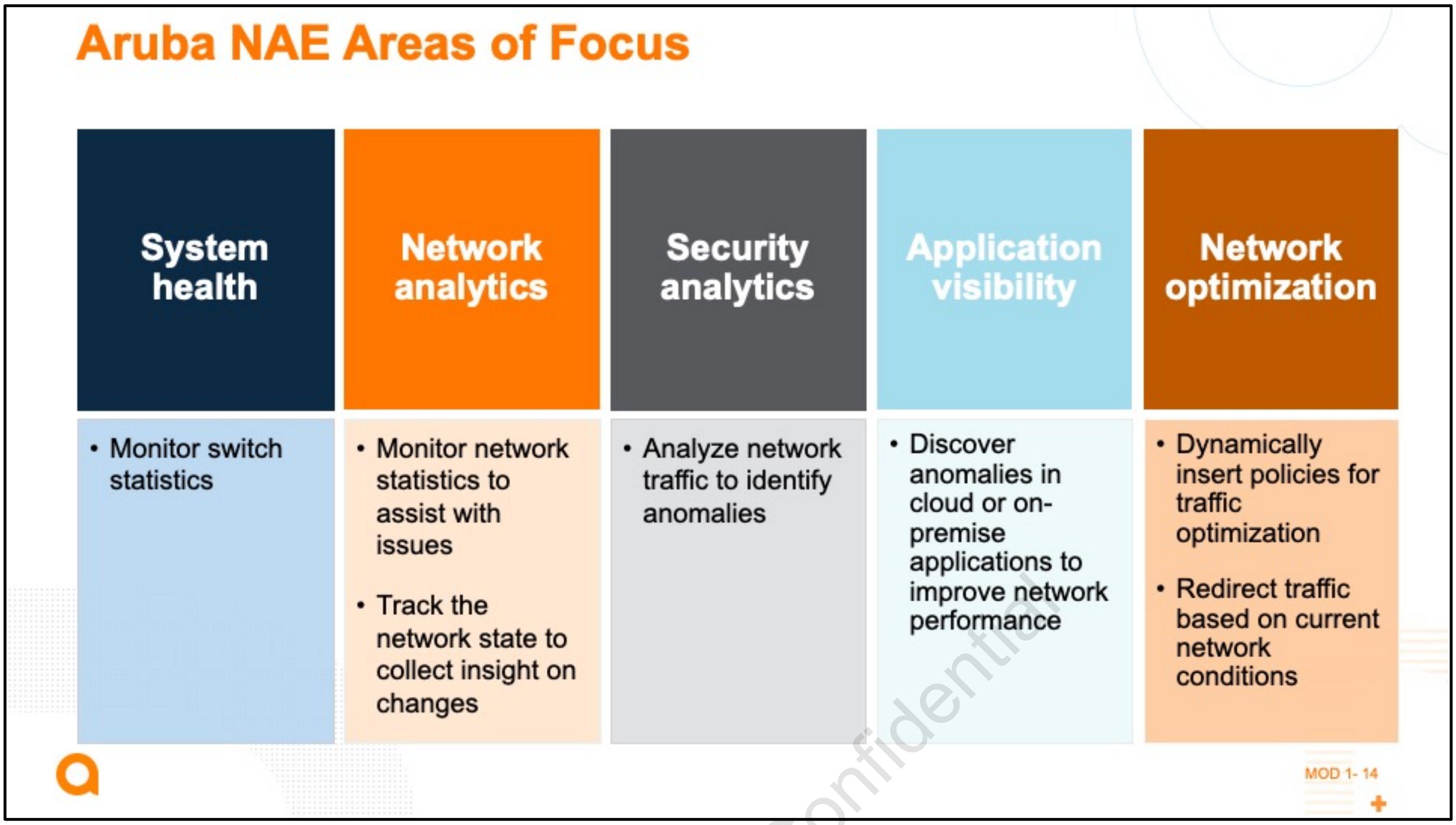
When an agent performs an action, the action is performed in a "sandbox" that is created when the action starts and removed when the action completes. The sandbox is in the default VRF, so it does not have access to the management network.

A sandbox is an isolated, tightly controlled environment in which programs can be run. Sandboxes restrict what a program can do, giving it the appropriate permissions and computing resources without allowing it access to the entire computing environment.

This design has the following benefits:

- Agents coexist and are prevented from using an excessive amount of CPU resources.
- Agents can benefit from the high-availability features of AOS-CX. During a switch failover event, the daemon that handles the sandbox can recover its state information and continue operations as before.
- Agents are prevented from accessing sensitive information—such as certificate files—in the switch operating system.

Soon you will learn about the three sources for the scripts that create NAE agents – Built-in system-created scripts, installed from the Aruba Solution Exchange (ASE) portal, and custom scripts created by you.



Aruba NAE Areas of Focus

System health	Network analytics	Security analytics	Application visibility	Network optimization
• Monitor switch statistics	• Monitor network statistics to assist with issues • Track the network state to collect insight on changes	• Analyze network traffic to identify anomalies	• Discover anomalies in cloud or on-premise applications to improve network performance	• Dynamically insert policies for traffic optimization • Redirect traffic based on current network conditions

MOD 1- 14

Aruba Training-Confidential

| With AOS-CX, Aruba's got a system that's designed to add new capabilities; and with NAE, Aruba goes beyond troubleshooting to system health, where you can monitor system statistics for root cause assistance on system issues. By using NAE, additional details can be gathered for system health issues like CPU utilization.

| Network analytics allows you to monitor network statistics to assist with Layer-1, Layer-2 and Layer-3 issues, track the state of OSPF neighbors, STP links, and etcetera, and collect insight information on network changes.

| Security analytics is especially interesting, as you can develop agents to analyze network traffic passing through the core and learn about anomalies like east-west traffic flows (perhaps traffic patterns that shouldn't be taking place, like camera going out to a different network segment in the IoT VRF).

| Application visibility allows you can leverage counters to match traffic going to applications in the cloud or data center to discover anomalies in application performance, cloud or on-prem.

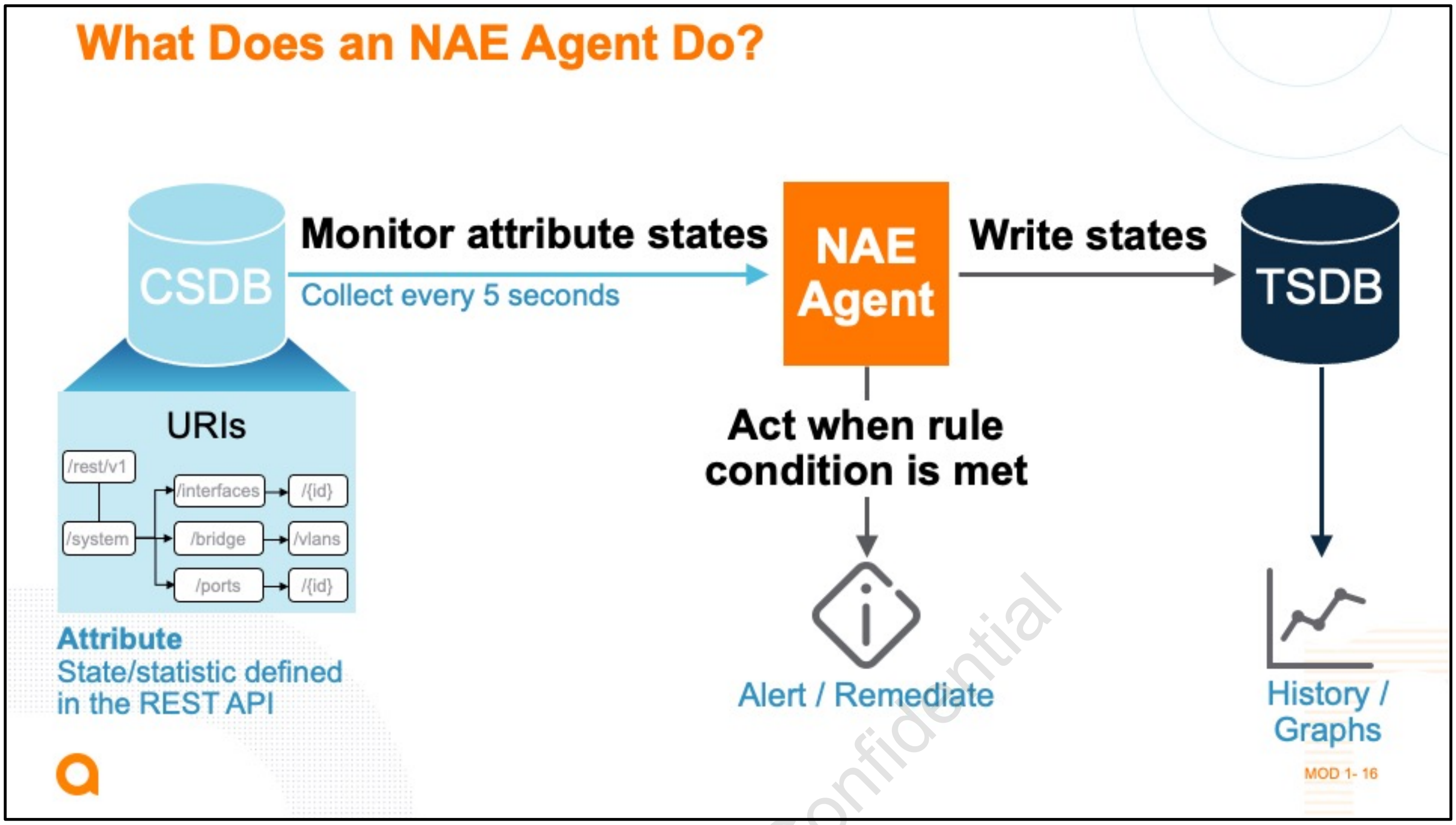
| Network optimization allows you to use knowledge of the network and it's performance to dynamically insert policies for optimization and take action such as Redirecting destination routes based on link status.

NAE Agents

MOD 1- 15

This section introduces agents and scripts used to implement NAE.

Aruba Training-Confidential



What Does an NAE Agent Do?

CSDB

Monitor attribute states

Collect every 5 seconds

NAE Agent

Write states

TSDB

URIs

/rest/v1

/system

/interfaces

/bridge

/ports

/{id}

/vlans

/{id}

Attribute

State/statistic defined in the REST API

Act when rule condition is met

Alert / Remediate

History / Graphs

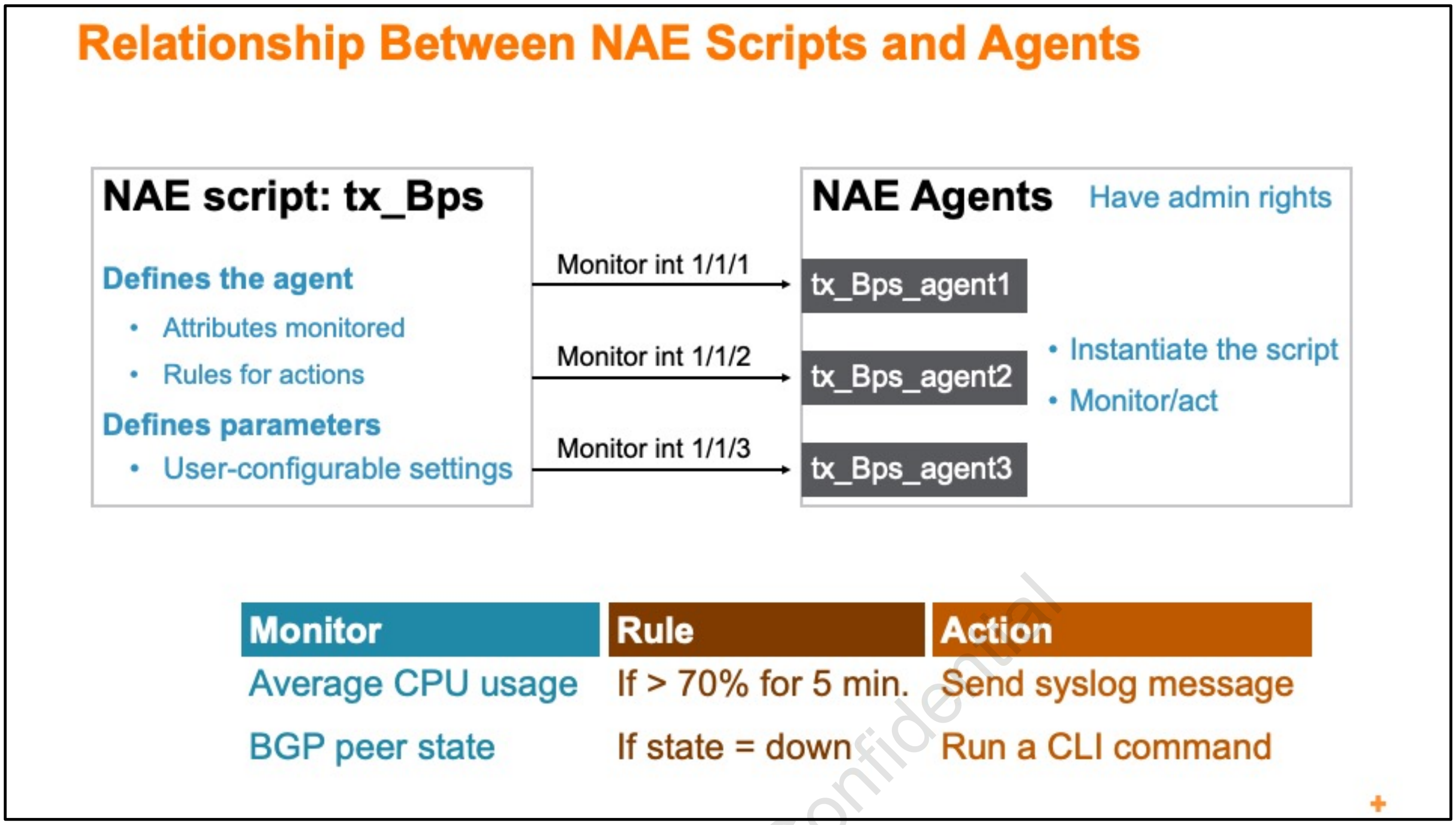
MOD 1- 16

Aruba Training-Confidential

An NAE agent is the fundamental component of the AOS-CX NAE. Every agent includes one or more monitors, which track data about a specific attribute on the switch in the current state database. An attribute is any state or statistic defined within the AOS-CX REST API. For example, it could be the number of transmitted bytes on a particular interface. Or it could be CPU utilization on a management module. This gives you great flexibility in terms of collecting the data that is relevant for your environment.

The agent collects the data every five seconds and writes the data to the time-series database on the switch's hard disk. This creates an ongoing record of the resource's state. You can see a graph of the monitored resource state over time in the AOS-CX Web UI.

If programmed with the appropriate rules, the NAE agent can take action when a particular condition is met. For example, if the agent is monitoring CPU utilization, it can set an alert when the utilization rises above a specific percentage.



Relationship Between NAE Scripts and Agents

NAE script: tx_Bps

Defines the agent

- Attributes monitored
- Rules for actions

Defines parameters

- User-configurable settings

NAE Agents Have admin rights

- Instantiate the script
- Monitor/act

Monitor	Rule	Action
Average CPU usage	If > 70% for 5 min.	Send syslog message
BGP peer state	If state = down	Run a CLI command

Aruba Training-Confidential

Each NAE agent is defined by a Python script that is loaded on the AOS-CX switch. The script defines the agent, including which attributes it monitors and rules for agent actions. The script might also define parameters - user configurable settings. For example, instead of specifying that an agent monitors transmitted bytes per second (Bps) on a specific interface, the script might specify that the agent monitors transmitted Bps on a variable interface parameter. Then users can choose the target interface when they create agents. Parameters can also define values for rule conditions, and much more – it depends on how you write the script.

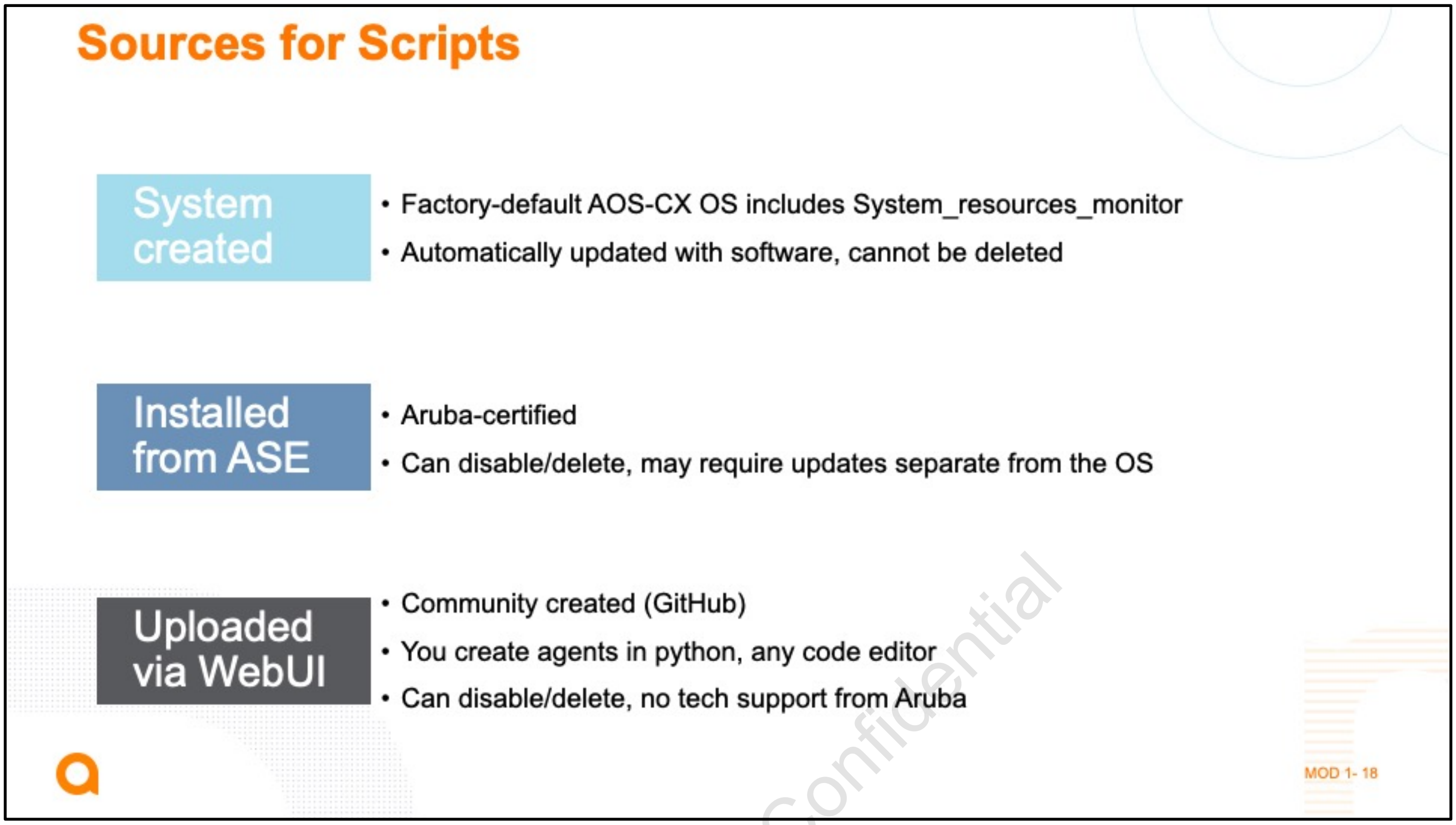
Note: An Aruba Network Analytics Engine agent is a specifically-configured executable instance of an NAE script on a switch. When the agent is enabled, it performs the tasks defined by the script. Agents have administrator rights.

An NAE script has no effect until it has an associated agent, which instantiates the script and starts monitoring the attribute and taking actions. To create an agent, you define parameter values. You can create different agents from the same script using different parameters, as shown in the figure. Perhaps script parameter indicates the interface to monitor, and you create multiple agents, each monitoring a different interface.

Other examples of tasks a script can define include the following:

- | Monitor average CPU usage and sending a system log message when the CPU usage is greater than 70% for 5 minutes.
- | Monitor the connection state of a particular BGP neighbor and execute a CLI command when the state transitions from UP to DOWN.

Some parameters that are integers—such as a CPU utilization threshold—can be changed for a given agent after that agent has been created. You can change such parameters easily through the Web UI— no programming skill is required.



Sources for Scripts

System created
- Factory-default AOS-CX OS includes System_resources_monitor
- Automatically updated with software, cannot be deleted

Installed from ASE
- Aruba-certified
- Can disable/delete, may require updates separate from the OS

Uploaded via WebUI
- Community created (GitHub)
- You create agents in python, any code editor
- Can disable/delete, no tech support from Aruba

MOD 1- 18

Aruba Training-Confidential

There are three sources for an agent script, as shown in the figure. Please click on each source to learn more.

System created, built-in scripts

Built-in scripts and agents are installed on the switch before the switch is shipped from the factory. All scripts and agents include origin information – “system” for the built-in ones, marked System Created in the Web UI.

Built-in scripts and agents cannot be deleted. You can enable, disable, and change the configuration of built-in agents:

- Built-in scripts are not displayed
- Built-in agents are displayed only if one of more parameters has been changed and saved.
- If you create an additional agent from a built-in script, that agent is considered a user-created agent, which can be deleted.

The current software release includes a single built-in script and agent that monitors several system resources:

- Built-in script: system_resource_monitor
- Built-in agent: system_resource_monitor.default

Aruba-certified scripts

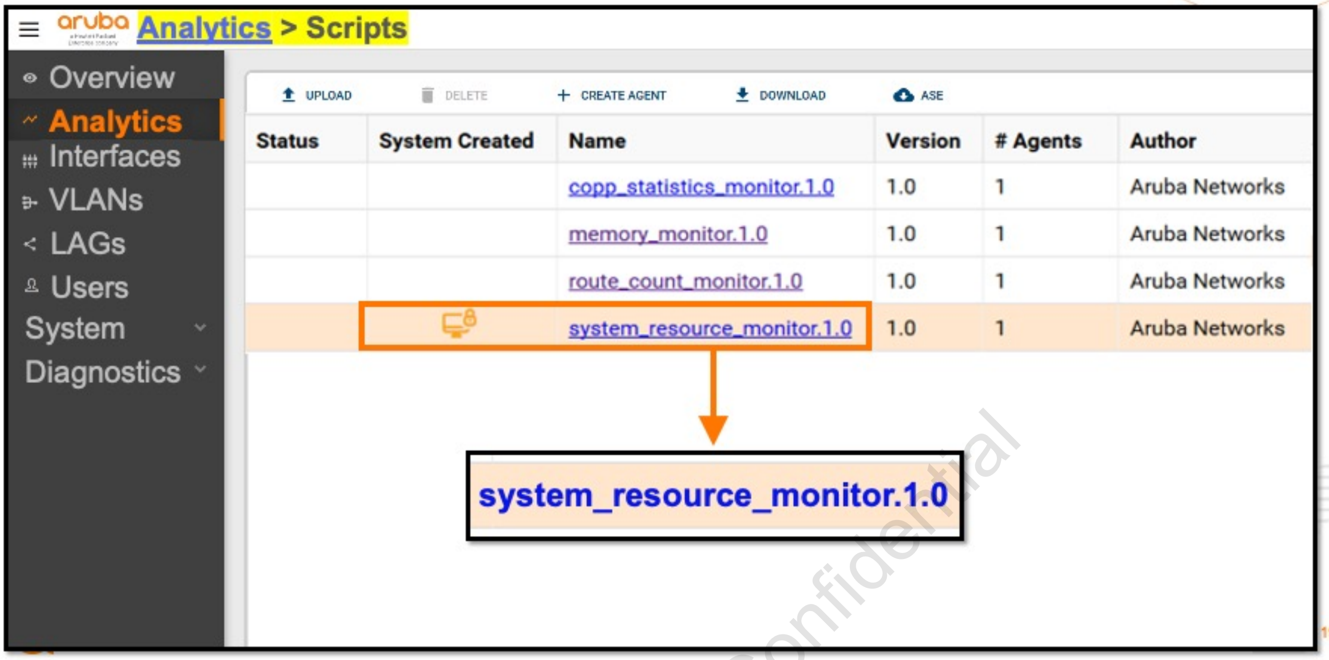
You can also obtain agent scripts from community sources. Scripts listed as Aruba-certified are written and validated by Aruba to work as advertised. Community members can also create and upload scripts - use at your own risk. Always evaluate a script before loading it on the switch to ensure that the actions executed are as expected. You can disable and remove these agents after loading them as desired. Before doing a software upgrade, you may need to remove an old version of the agent. After the upgrade, upload a new one.

Create your own scripts

Finally, you can create agents on your own with Python script. These agents are managed like the ones that you install from a community source. You can disable, remove and update them separately from the software.

Aruba Training-Confidential

System-Created Agent



Status	System Created	Name	Version	# Agents	Author
		copp_statistics_monitor.1.0	1.0	1	Aruba Networks
		memory_monitor.1.0	1.0	1	Aruba Networks
		route_count_monitor.1.0	1.0	1	Aruba Networks
		system_resource_monitor.1.0	1.0	1	Aruba Networks

system_resource_monitor.1.0

Aruba Training-Confidential

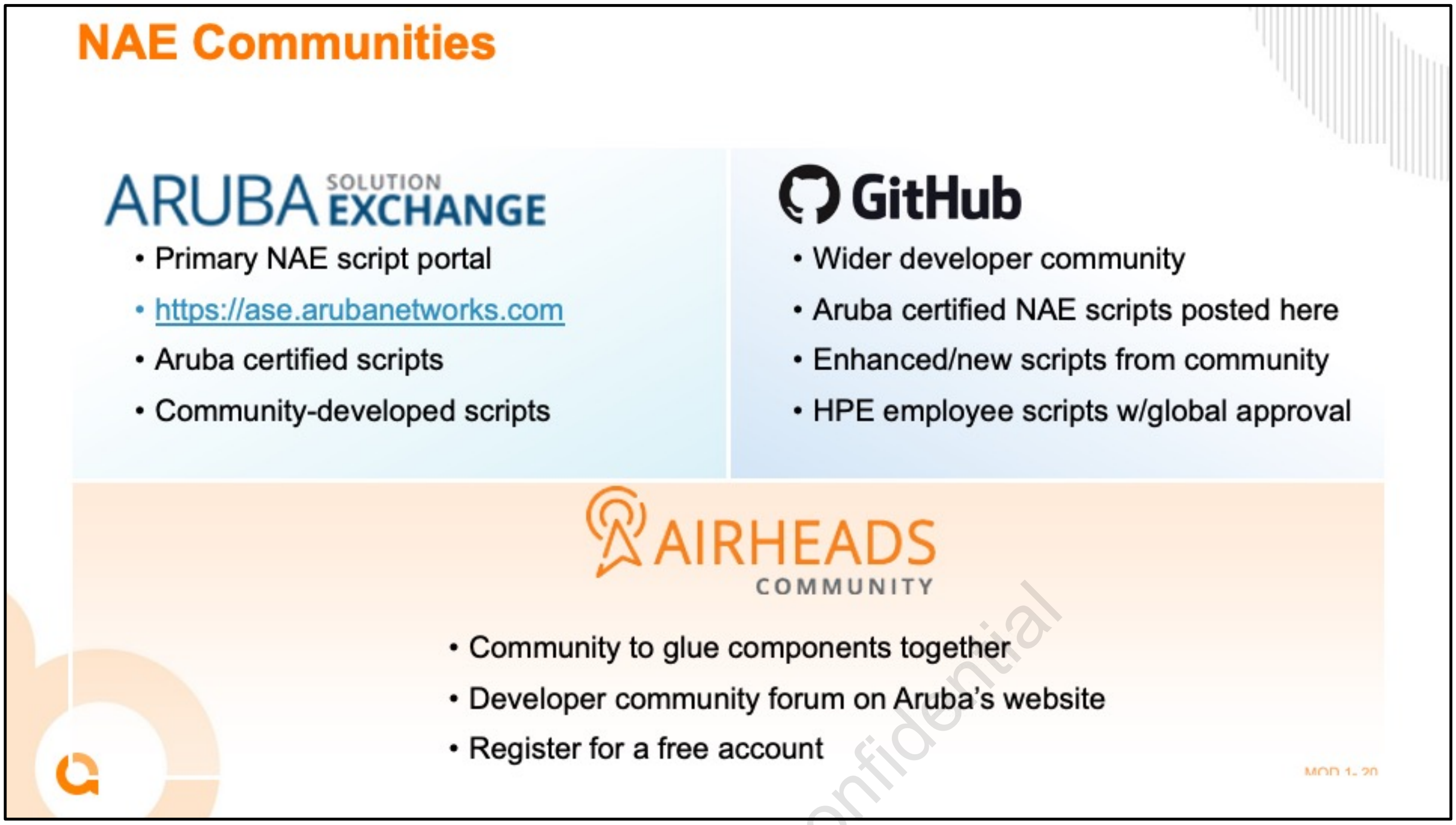
Use the AOS-CX Web UI to access information for NAE agents, scripts, and alerts. This includes time-series data graphs and other information generated by the enabled agents.

Choose **Analytics** in the UI navigation pane to see the Analytics Dashboard. You can display time-series graph panels for up to nine agents on the dashboard. However, many more agents can be enabled on a switch.

From the Analytics Dashboard, you can open Analytics detail pages. Analytics detail pages allow you to enable, disable, create, delete, or edit agents, upload scripts, and view detailed data about monitored resources. Administrator rights are required for actions that modify an agent.

As shown in the figure, the Script Management page shows all scripts loaded on the switch and the agents associated with them. Look in the System Created column to find the system created script, as highlighted in the figure

AOS-CX supports a single system created script, which has an associated agent by default. This agent monitors CPU and memory utilization over short, medium, or long time periods. It triggers alerts and log messages when utilization passes a particular threshold. The alert severity depends on whether the high usage continues over a short, medium, or long period.



NAE Communities

ARUBA SOLUTION EXCHANGE

- Primary NAE script portal
- https://ase.arubanetworks.com
- Aruba certified scripts
- Community-developed scripts

GitHub

- Wider developer community
- Aruba certified NAE scripts posted here
- Enhanced/new scripts from community
- HPE employee scripts w/global approval

AIRHEADS COMMUNITY

- Community to glue components together
- Developer community forum on Aruba's website
- Register for a free account

MOD 1- 20

Aruba Training-Confidential

The figure summarizes key aspect of the three NAE communities. Click on each one to explore further.

Aruba Solution Exchange

The primary source for additional NAE scripts is the Aruba Solutions Exchange (ASE), at https://ase.arubanetworks.com/. You can browse agents and search for particular agent types. To view or download the agent script you must log in with your Aruba SSO system user account. You can also access ASE directly through the AOS-CX switch UI, assuming good Internet connectivity – no log in required

GitHub

GitHub is a wide-ranging developer community with custom scripts for many different applications. Aruba will post all Aruba-certified NAE scripts here as well as to ASE. The GitHub community can enhance these scripts and create new scripts for sharing within the community. HPE employees can also post NAE scripts to GitHub provided that they pass global approval, which consists of approval from the Open Source Review Board (OSRB) and the legal department.

GitHub members can also create and upload scripts. These are not Aruba supported - use at your own risk. Always evaluate a script before loading it on the switch to validate expected operation. AOS-CX scripts on GitHub can be found here: https://github.com/aruba/aos-cx-python.

Airheads

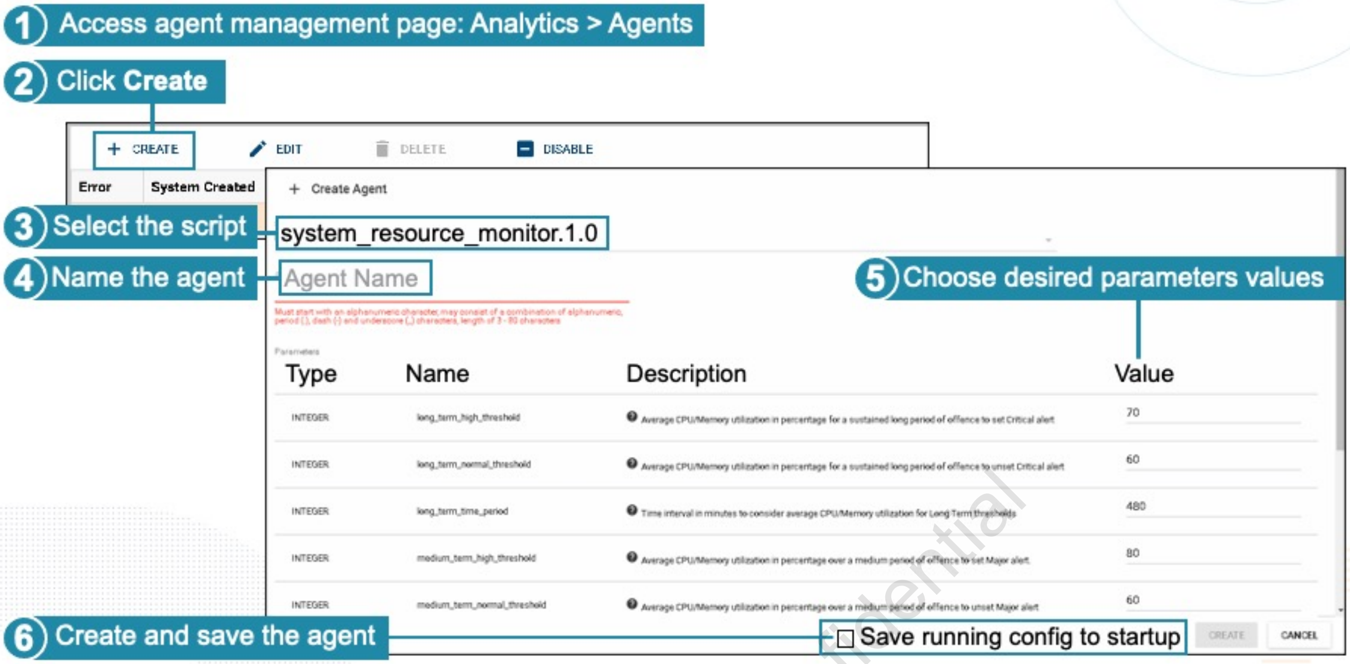
TheAirheads community is a place for members or participants to search for information, read and post about topics of interest, and learn from each other. Unregistered guests can browse or search the community for information. Registered members can post messages or comments, track discussions, and get email notifications on posting activity and other community actions.

The Airheads community is the glue that holds together all NAE solution components. Members can share their expertise and tips and point each other toward relevant agents. The Developer Community forum under the technology category has discussion forums about automation and NAE questions and support.

Aruba is committed to expanding the NAE value with extensive research and development (R&D) investment in building NAE scripts and supporting the community. The Airheads Developer Community group is specific to APIs, programming, and automation. This is the recommended place to post NAE-related questions. Go to https://community.a rubanetworks.com/t5/Developer-Community/bd-p/DeveloperComm unity. See the Airheads community at http://community.ar ubanetworks.com/.

Aruba Training-Confidential

Creating and Editing Agents For Existing Scripts



1) Access agent management page: Analytics > Agents

2) Click Create

3) Select the script: system_resource_monitor.1.0

4) Name the agent: Agent Name

5) Choose desired parameters values

6) Create and save the agent

+ CREATE
EDIT
DELETE
DISABLE
Error
System Created
+ Create Agent
Must start with an alphanumeric character, may consist of a combination of alphanumeric, period (.), dash (-) and underscore (_) characters, length of 3 - 80 characters
Parameters
□ Save running config to startup
CREATE
CANCEL

Type	Name	Description	Value
INTEGER	long_term_high_threshold	Average CPU/Memory utilization in percentage for a sustained long period of offence to set Critical alert	70
INTEGER	long_term_normal_threshold	Average CPU/Memory utilization in percentage for a sustained long period of offence to unset Critical alert	60
INTEGER	long_term_time_period	Time interval in minutes to consider average CPU/Memory utilization for Long Term thresholds	480
INTEGER	medium_term_high_threshold	Average CPU/Memory utilization in percentage over a medium period of offence to set Major alert	80
INTEGER	medium_term_normal_threshold	Average CPU/Memory utilization in percentage over a medium period of offence to unset Major alert	60

MOD 1- 21

Aruba Training-Confidential

When you install a new script, it must have an associated agent. To create an agent based on a script, access the Agent Management page – just select Agents in the Analytics Dashboard. Then select Create. You can also select Create Agent in the Script Management page.

A window appears where you can choose the script on which to base the agent. Define the agent name, customize the parameters, and save.

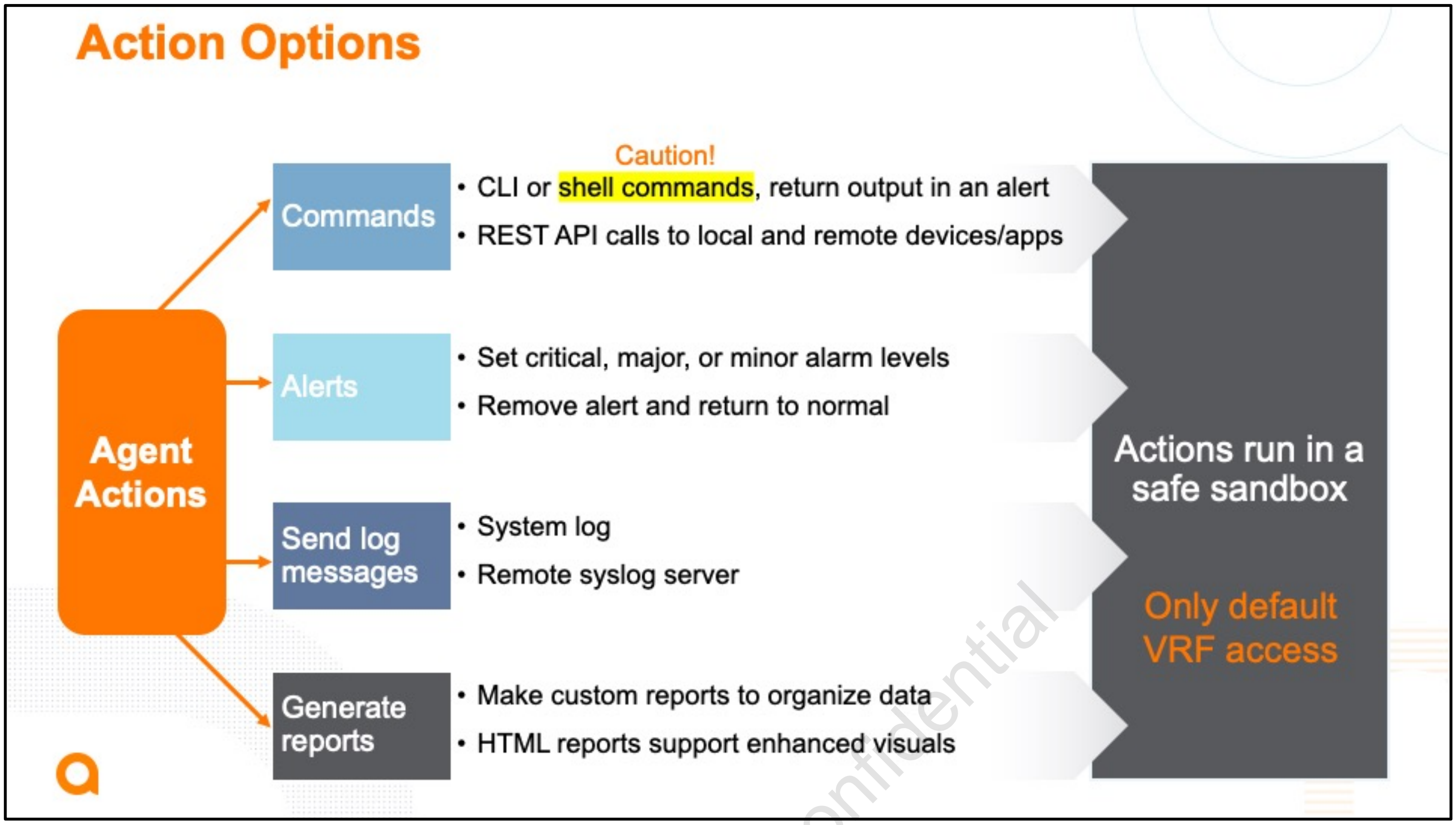
You can also select an agent in the Agent Management page and select Edit to adjust the parameters on an existing agent.

NAE Agent Actions

MOD 1- 22

Now that you've learned how to obtain agents, you can examine some of the actions that they can take in more detail.

Aruba Training-Confidential



Action Options

Agent Actions

Commands

Caution!

- CLI or shell commands, return output in an alert
- REST API calls to local and remote devices/apps

Alerts

- Set critical, major, or minor alarm levels
- Remove alert and return to normal

Send log messages

- System log
- Remote syslog server

Generate reports

- Make custom reports to organize data
- HTML reports support enhanced visuals

Actions run in a safe sandbox

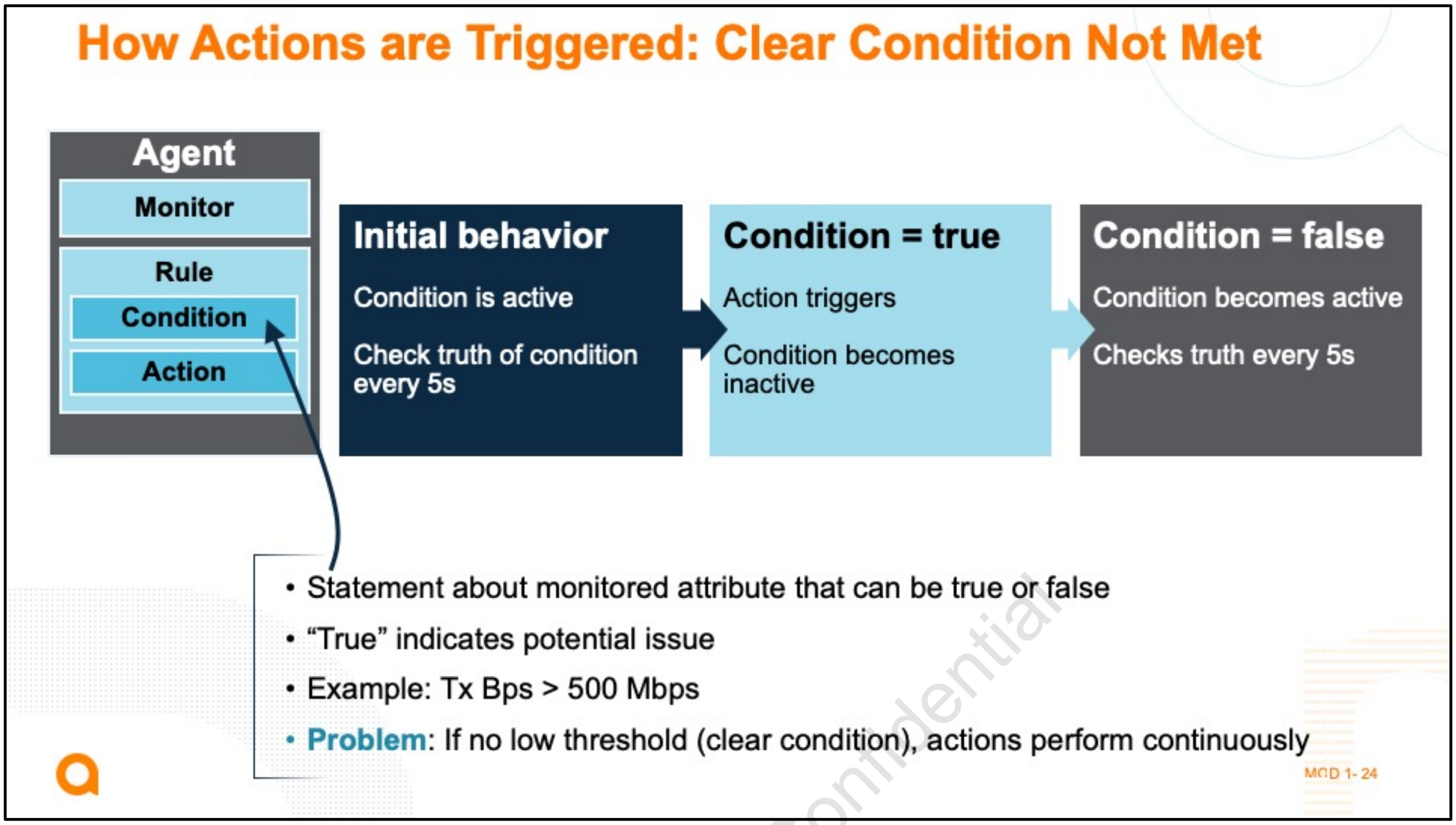
Only default VRF access

Aruba Training-Confidential

Agents can execute CLI commands or shell commands. Be very careful using shell commands in an agent! Mistakes can damage switch functionality and render it unusable. Agents can make REST API calls to itself or remote devices and applications. This allows for interaction with the all Aruba applications and devices, which have their own REST interfaces. NAE can use REST call to integrate with third party applications, including ticket generation systems like ServiceNow and TOPdesk.

An action can also send log messages to a system log or remote Syslog server, indicating information about the condition that has occurred. Finally, an action can generate a custom report with multiple pieces of relevant information about the condition.

For more complex scenarios, a rule can combine multiple actions. Note that AOS-CX uses a safe sandbox environment to execute actions, which prevents agents from using excessive CPU or obtaining access to sensitive files such as certificates. The sandbox has access to the default VRF, not the management one. If actions need access to the Internet to execute properly, make sure that the default VRF has such access.



How Actions are Triggered: Clear Condition Not Met

Initial behavior

Condition is active

Check truth of condition every 5s

Condition = true

Action triggers

Condition becomes inactive

Condition = false

Condition becomes active

Checks truth every 5s

- Statement about monitored attribute that can be true or false
- “True” indicates potential issue
- Example: Tx Bps > 500 Mbps
- **Problem**: If no low threshold (clear condition), actions perform continuously

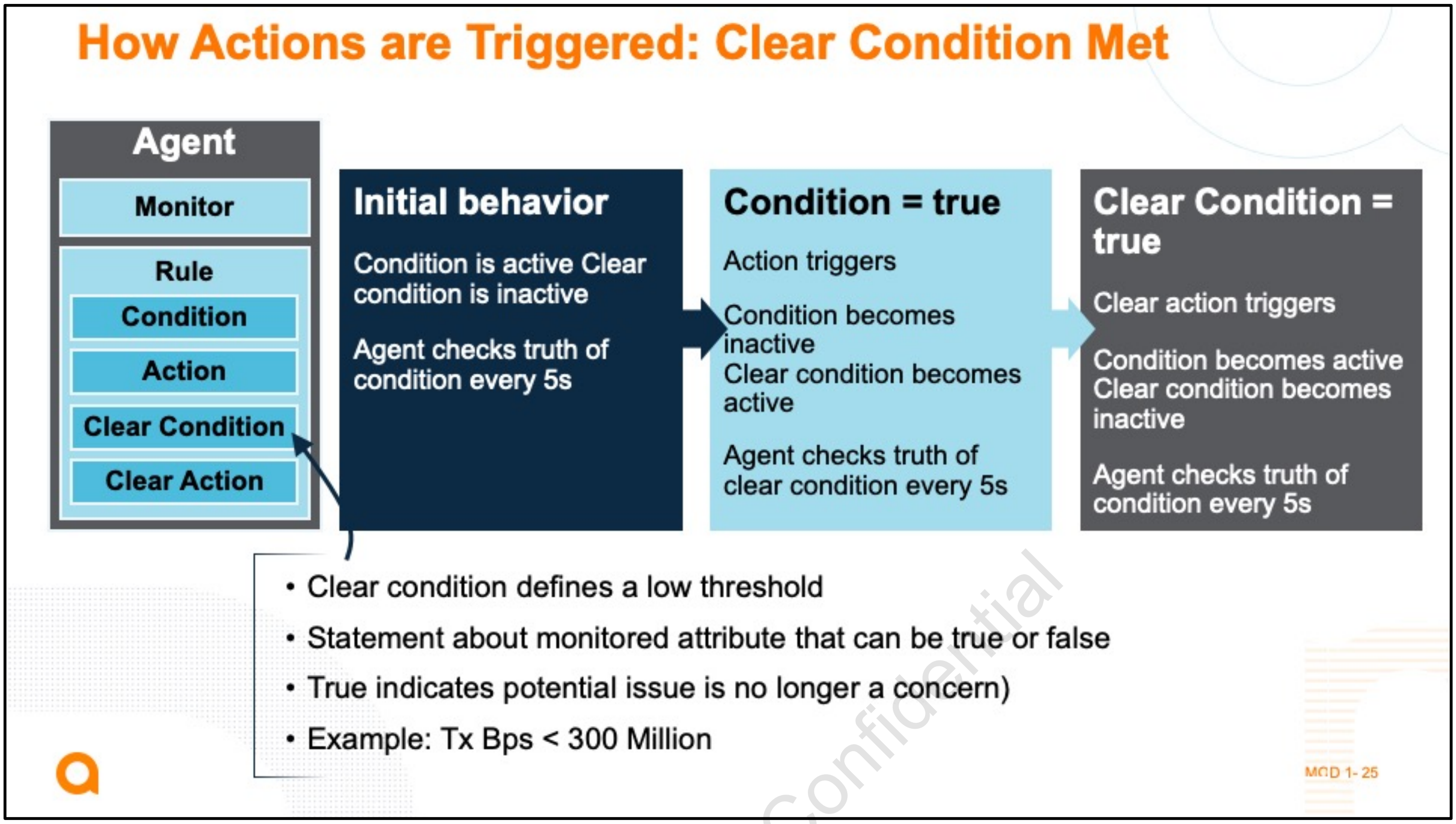
MOD 1- 24

Aruba Training-Confidential

Sometimes you only want an agent to track information in the time-series database and produce a graph. But you may also want the agent to take action based on information obtained by the monitor. You define these actions with one or more rules. A rule must have a condition, which defines a particular state for the monitored value. At any given moment, the state is either true or false. Generally, the state being true indicates a potential issue that you should know about. Suppose that an agent monitors transmitted Bps on an Ethernet interface. A rule condition might be that the monitored value is greater than 500 million Bps.

The agent checks the state against the condition every 5 seconds. The rule's action executes when the condition changes from false to true. At that point, the condition becomes inactive, and the agent does not re-execute actions if the condition remains true. This prevents constant alerts for the same ongoing issue. If the condition becomes false again (for example, the Tx Bps drops lower), the condition becomes active again, and the agent once again starts monitoring it.

In some cases, a condition might toggle frequently between true and false. For example, the condition is more than 500 million Bps transmitted on an interface. The interface might transmit 600 million Bps and change the condition from false to true. Then it transmits 490 million Bps and transitions from true to false. Then it transmits 510 million Bps and triggers the action again.



How Actions are Triggered: Clear Condition Met

Initial behavior

Condition is active Clear condition is inactive

Agent checks truth of condition every 5s

Condition = true

Action triggers

Condition becomes inactive
Clear condition becomes active

Agent checks truth of clear condition every 5s

Clear Condition = true

Clear action triggers

Condition becomes active
Clear condition becomes inactive

Agent checks truth of condition every 5s

- Clear condition defines a low threshold
- Statement about monitored attribute that can be true or false
- True indicates potential issue is no longer a concern)
- Example: Tx Bps < 300 Million

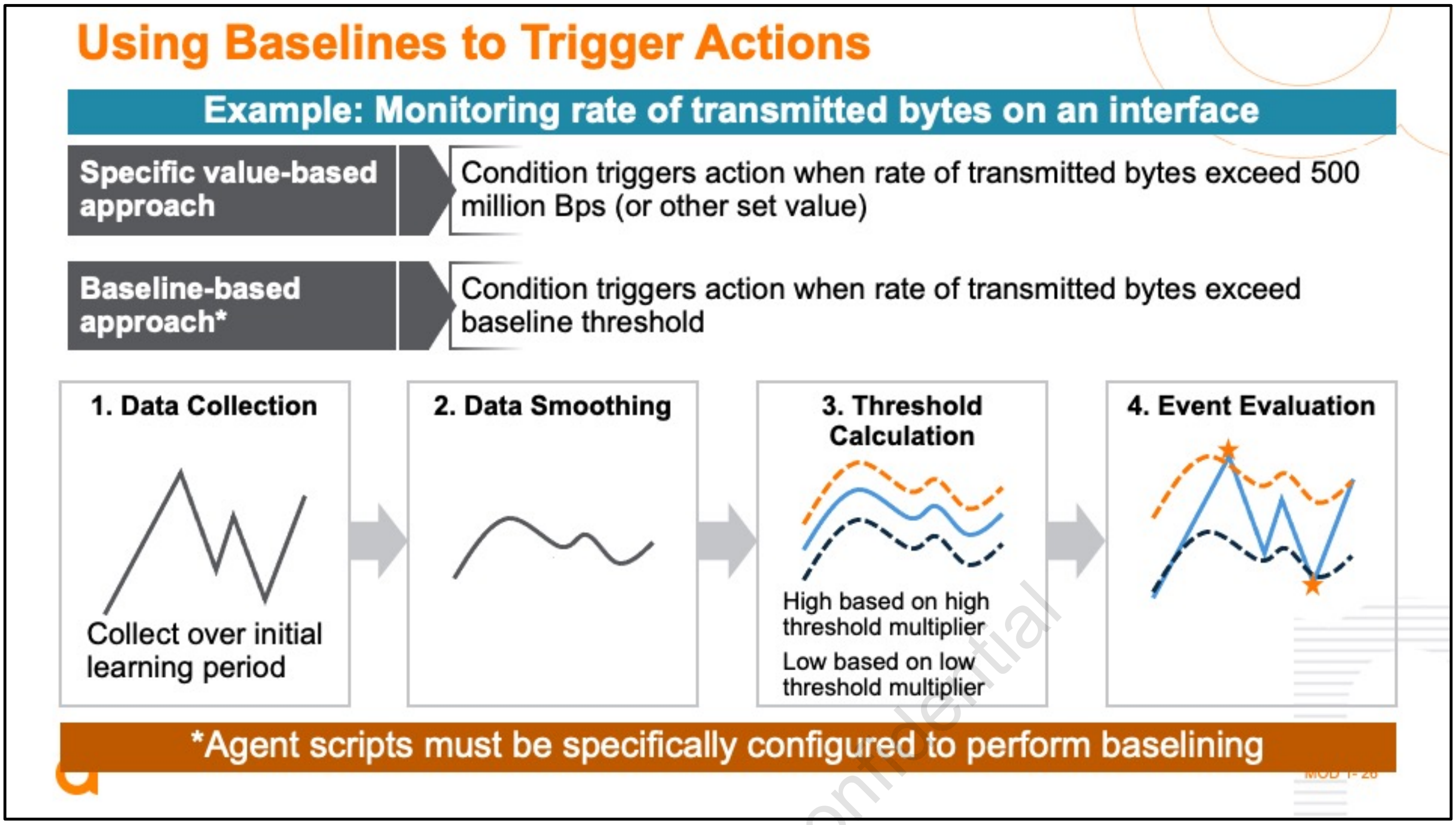
MOD 1- 25

To avoid that “continuous action” problem, you can define a clear condition, which must be met to free the main condition. The “clear” condition is a low threshold that the monitored attribute must fall below before the action is triggered again. The clear condition becomes active when the condition first transitions from false to true, and the agent begins checking the state of the clear condition every five seconds.

Even if the condition becomes false again, the clear condition remains active, and the condition inactive, as long as the clear condition is also false. Only when the clear condition becomes true does the condition transition to active. For example, the clear condition could be transmission of less than 300 million Bps. Now the rate must fall below that (transition to true) before the original condition can become active again. You can optionally define a clear action, which also executes when the clear condition changes from false to true.

This example shows just one rule, but agents can customize their behavior to nuanced conditions by using multiple rules, each with a different condition and action. For example, a condition in one rule might indicates a less severe problem and a condition in another rule might indicate a more severe problem.

Aruba Training-Confidential



Using Baselines to Trigger Actions

Example: Monitoring rate of transmitted bytes on an interface

Specific value-based approach → Condition triggers action when rate of transmitted bytes exceed 500 million Bps (or other set value)

Baseline-based approach* → Condition triggers action when rate of transmitted bytes exceed baseline threshold

1. Data Collection

Collect over initial learning period

2. Data Smoothing

3. Threshold Calculation

High based on high threshold multiplier

Low based on low threshold multiplier

4. Event Evaluation

*Agent scripts must be specifically configured to perform baselining

Aruba Training-Confidential

Often, it can be difficult to determine the best value to set for a condition. For example, you are trying to define a condition that indicates excessive traffic on an Ethernet interface. Is a rate of 500 million bytes per second excessive, or is a higher or lower value more appropriate? The answer often depends on the particular network, switch, and interface.

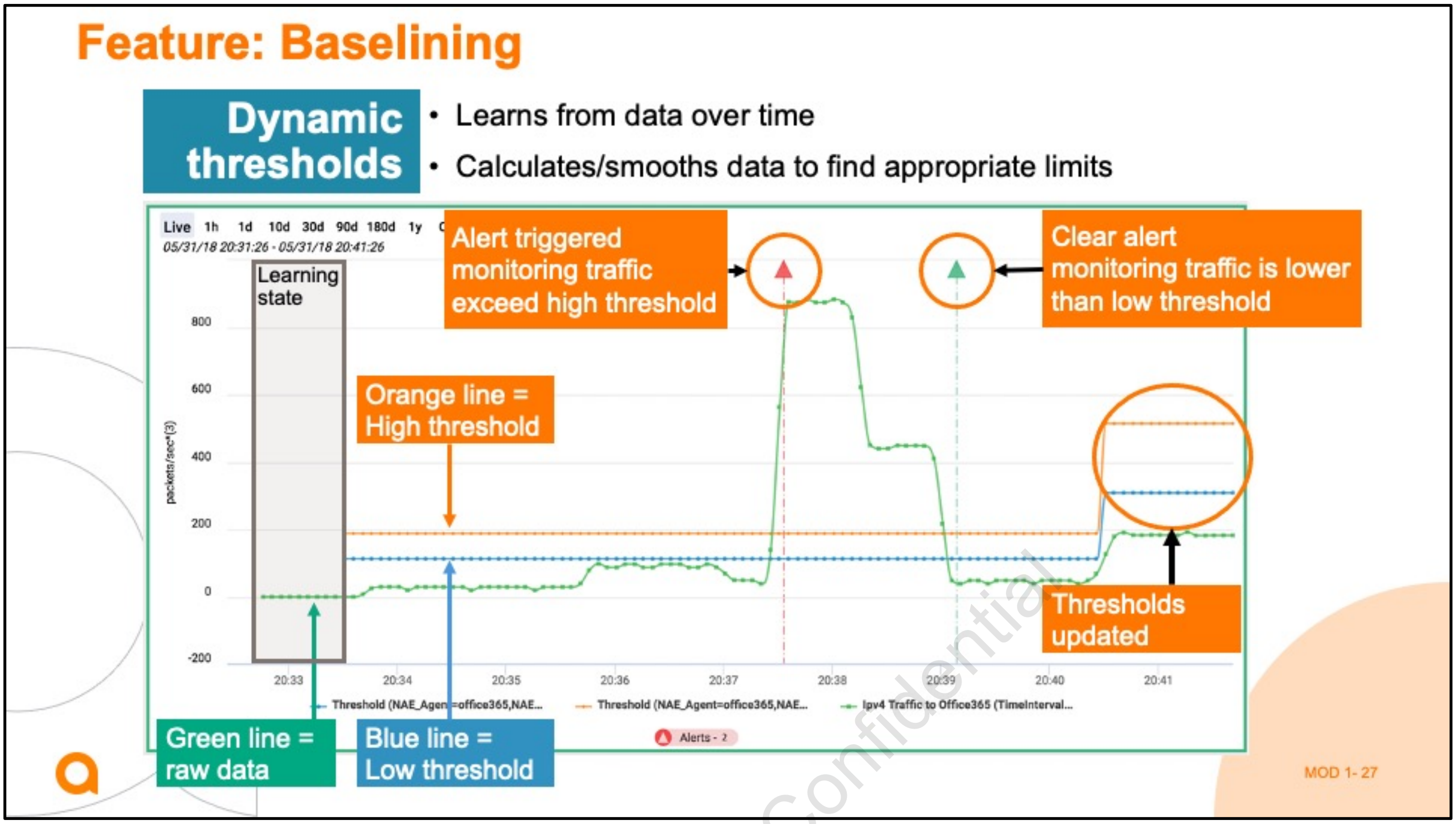
NAE can help to automate and remove the guesswork for defining conditions by using baselines. The agent can construct a baseline of normal values for the monitored resource attribute. A condition then triggers if the value deviates too much from the baseline. Understand that agents do not create baselines by default – you must specifically script the agent to create a baseline.

Agents construct a baseline by collecting data over an initial, configurable learning period. This learning period should be long enough that all representative patterns are included. For example, it should be at least a day and maybe as much as a week. The agent then uses an algorithm to smooth the data and create the baseline.

Then the agent uses a configurable threshold multiplier to determine the high threshold – when a condition changes to true and triggers an associated action. Larger multipliers allow for more deviation – smaller multipliers allow for less deviation. The low multiplier lets the agent construct a low threshold, which indicates when the condition is cleared and a clear action triggered. You could create several conditions with different multipliers that trigger different actions; the condition with the larger multiplier should have a more significant action because it indicates a greater deviation from normal.

In addition to the initial learning period, the agent also uses a continuous learning period to continue to recalibrate the baseline. This lets the agent respond to gradual changes that create a new normal. However, data from a continuous learning period that includes an anomaly that triggered a condition is discarded. This lets the agent detect the anomalies, but not use them to change the baseline.

Aruba Training-Confidential



Feature: Baselining

Dynamic thresholds

- Learns from data over time
- Calculates/smooths data to find appropriate limits

Alert triggered monitoring traffic exceed high threshold

Clear alert monitoring traffic is lower than low threshold

Learning state

Orange line = High threshold

Thresholds updated

Green line = raw data

Blue line = Low threshold

Aruba Training-Confidential

Baselining provides dynamic thresholds for NAE monitors. This is Artificial Intelligence (AI)!

In this case, the thresholds are set significantly higher because the algorithm includes all the data in its continuous learning window, which included the time in which the traffic rate was much higher than the previous threshold.

In this graph, the green line is the raw data, the orange line is the high threshold as calculated by the baseline, and the blue line is the low threshold as calculated by the baseline. The events in the timeline are as follows:

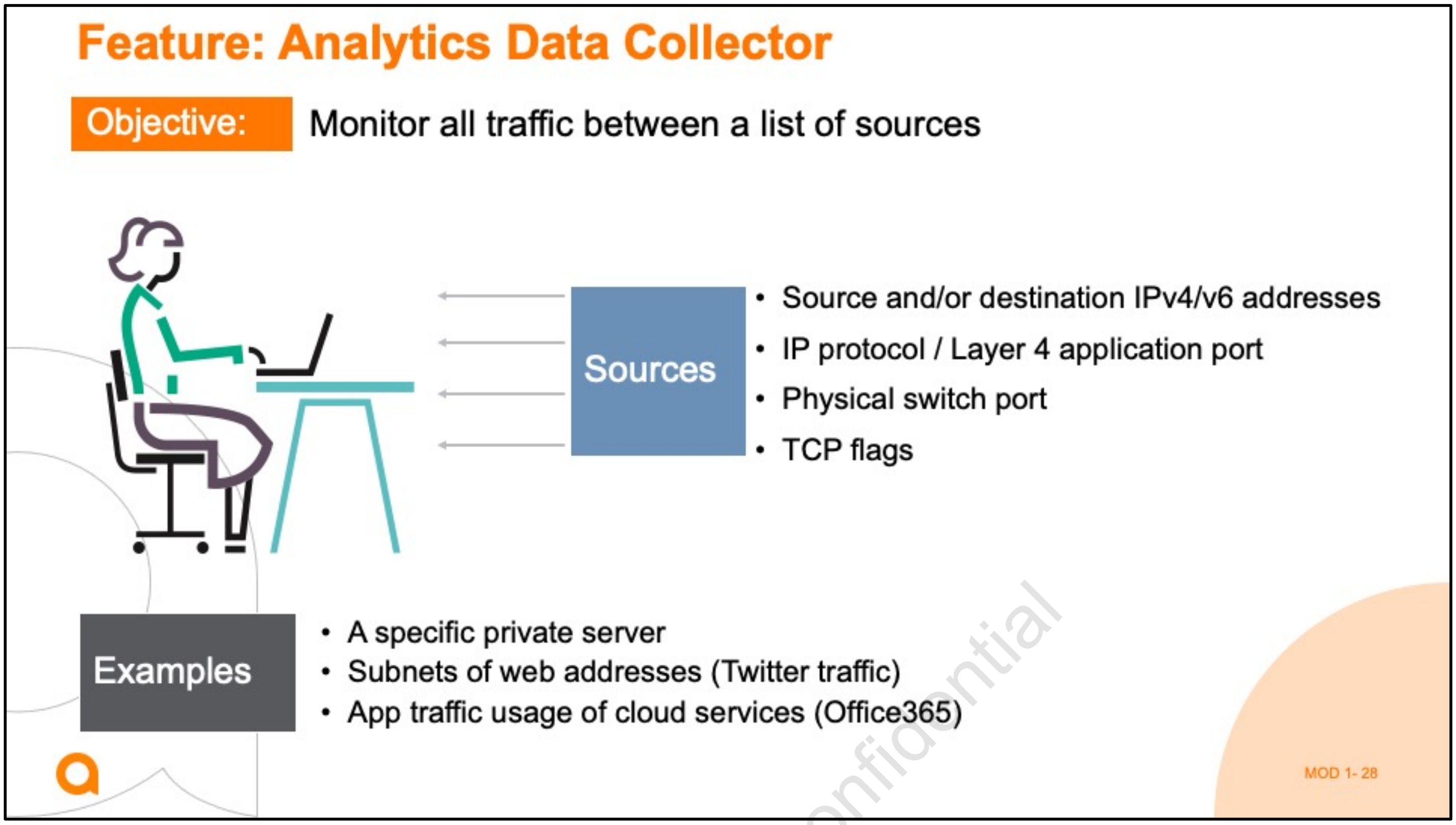
1. At 20:32:30, an agent is created and enabled. The baseline enters the learning state. Because the script did not specify default thresholds, there are no thresholds defined. In the graph, just the green line for the raw data is displayed.
2. At 20:33:30, the baseline exits the learning state and enters the active state:
    - The high threshold and the low threshold calculations are completed.
    - The graph begins the display of the orange line for the high threshold and the blue line for the low threshold.
    - The agent will generate an alert when the monitored traffic rate (in packets per second) exceeds the high threshold.
    - The agent will clear the alert when the monitored traffic rate (in packets per second)

drops below the low threshold.

3. At 20:37:33, an alert is triggered because the monitored traffic rate exceeds the high threshold.
4. At 20:39:15, the alert is cleared because the monitored traffic rate (in packets per second) is lower than the low threshold.
5. At 20:40:30, the thresholds are updated. In this case, the thresholds are set significantly higher because the algorithm includes all the data in its continuous learning window, which included the time in which the traffic rate was much higher than the previous threshold.
   If the script specified a longer initial learning time, such as one day, the calculations used to create the thresholds can include the typical fluctuations in data that can occur, resulting in more appropriate thresholds and alerts that trigger only for significant anomalies.

You can refer to the **ArubaOS-CX 10.05 Network Analytics Engine Guide** for additional reference.

Aruba Training-Confidential



Feature: Analytics Data Collector

Objective: Monitor all traffic between a list of sources

Sources

- Source and/or destination IPv4/v6 addresses
- IP protocol / Layer 4 application port
- Physical switch port
- TCP flags

Examples

- A specific private server
- Subnets of web addresses (Twitter traffic)
- App traffic usage of cloud services (Office365)

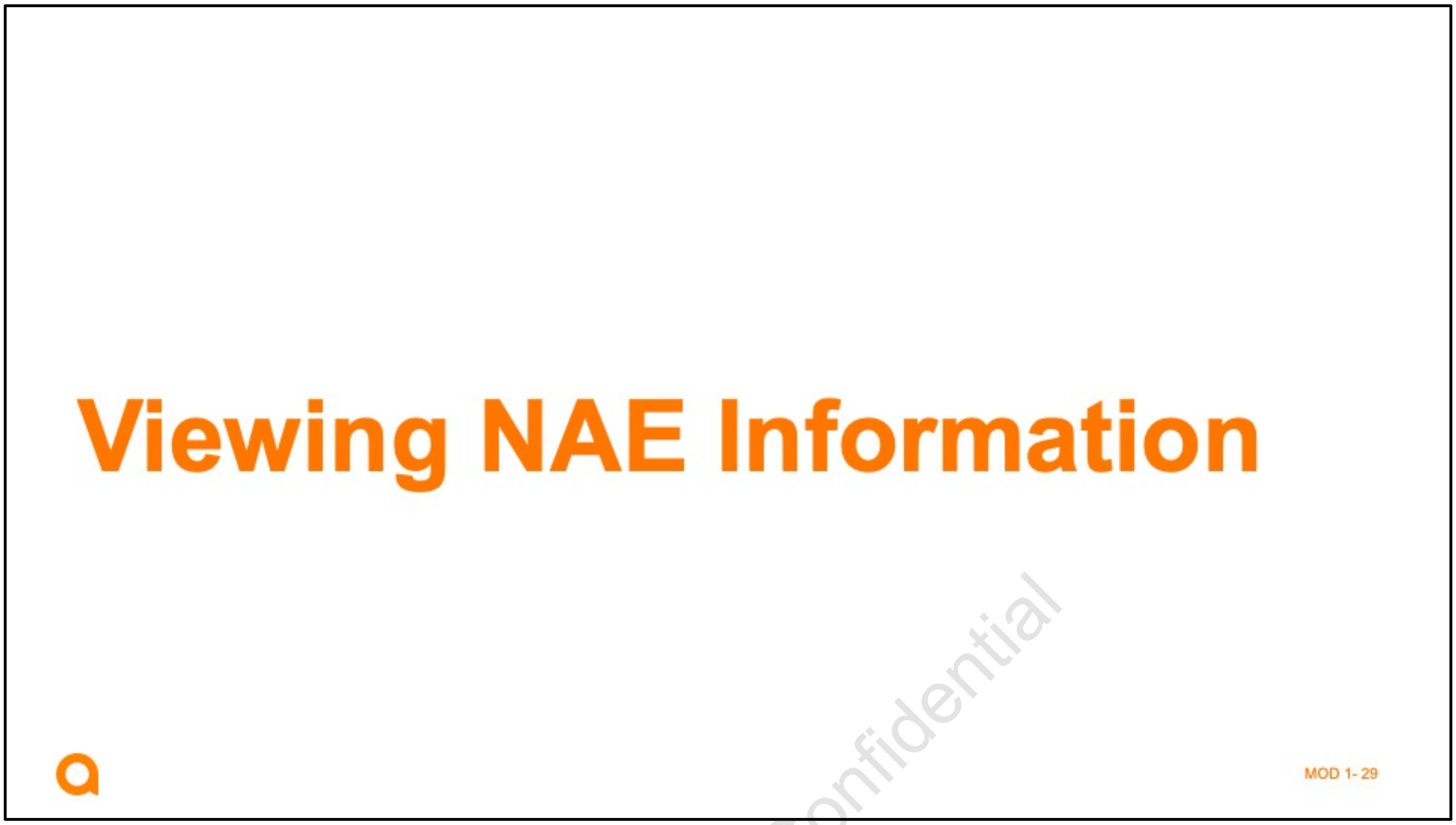
MOD 1- 28

Aruba Training-Confidential

You can set up agents with scripts that monitor traffic with specific filtering criteria, as shown in the figure.

Here are some examples:

- A specific private server
- Subnets of web addresses, such as Twitter traffic
- Application traffic usage of cloud services, such as Office365



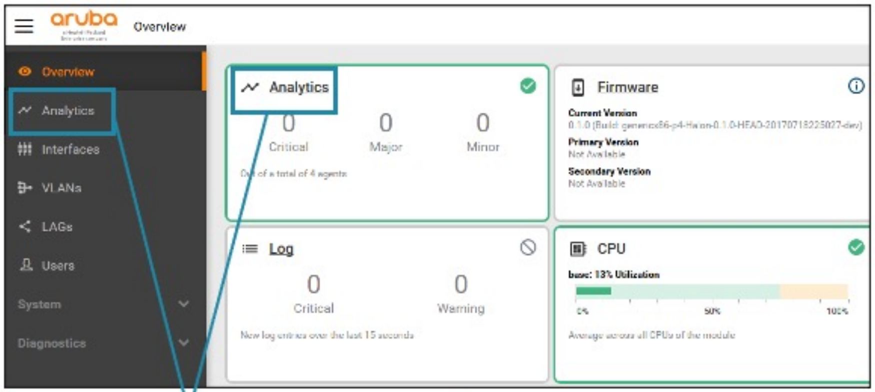
Viewing NAE Information

MOD 1- 29

You will now look at how you can view the information that NAE agents collect in the AOS-CX Web UI.

Aruba Training-Confidential

Access the Analytics Dashboard



Current Version
0.1.0 (Build: genericx86-p4-Halon-0.1.0-HEAD-20170718225027-dev)
Primary Version
Not Available
Secondary Version
Not Available

base: 13% Utilization

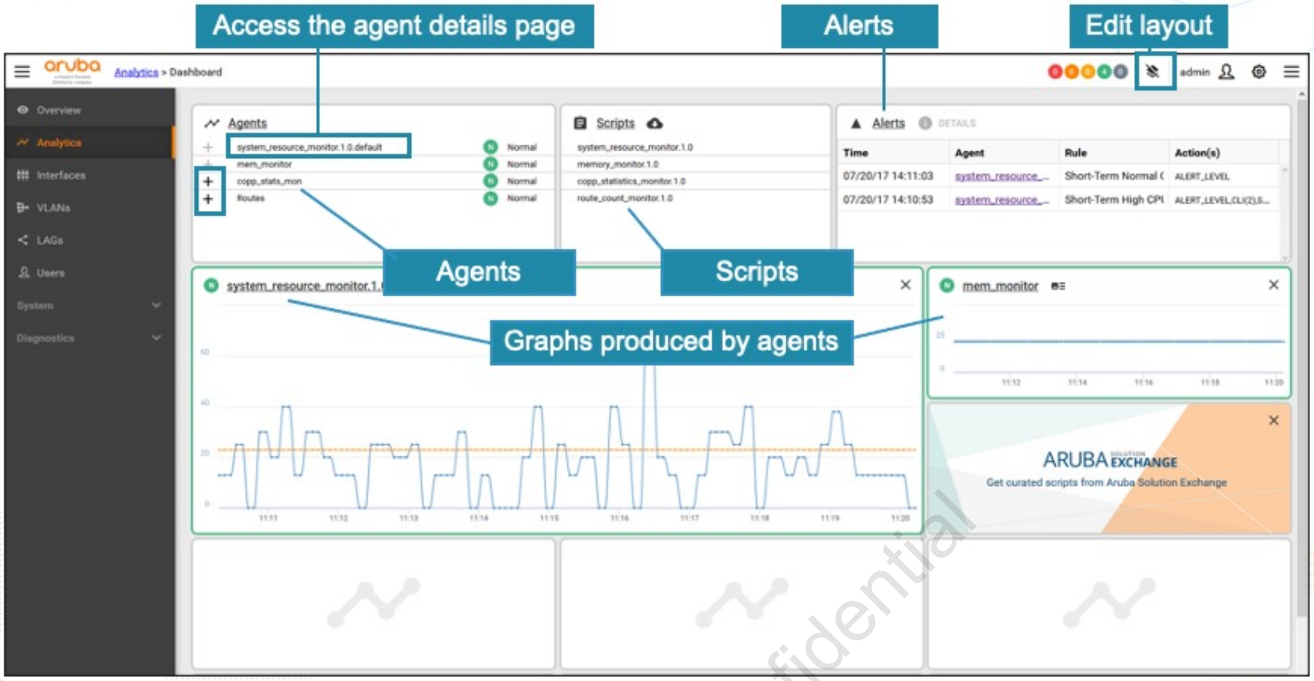
Access the Analytics dashboard

MOD 1- 30

When you enter the Web UI, you see a main dashboard. The Analytics section in this dashboard refers to the NAE functions. It shows how many critical, major, or minor alerts are active. If you want more detailed information, you should view the Analytics Dashboard. To get there, either click the Analytics section in the main dashboard or select Analytics in the sidebar.

Aruba Training-Confidential

Analytics Dashboard



Access the agent details page

Alerts

Edit layout

Agents

Scripts

Graphs produced by agents

Agents	
system_resource_monitor.1.0.default	Normal
mem_monitor	Normal
copp_stats_mon	Normal
Routes	Normal

Scripts: system_resource_monitor.1.0, memory_monitor.1.0, copp_statistics_monitor.1.0, route_count_monitor.1.0

Time	Agent	Rule	Action(s)
07/20/17 14:11:03	system_resource_...	Short-Term Normal (	ALERT_LEVEL
07/20/17 14:10:53	system_resource_...	Short-Term High CPI	ALERT_LEVEL,CLI(2),S...

system_resource_monitor.1.0
mem_monitor
ARUBA SOLUTION EXCHANGE
Get curated scripts from Aruba Solution Exchange

MOD 1- 31

Aruba Training-Confidential

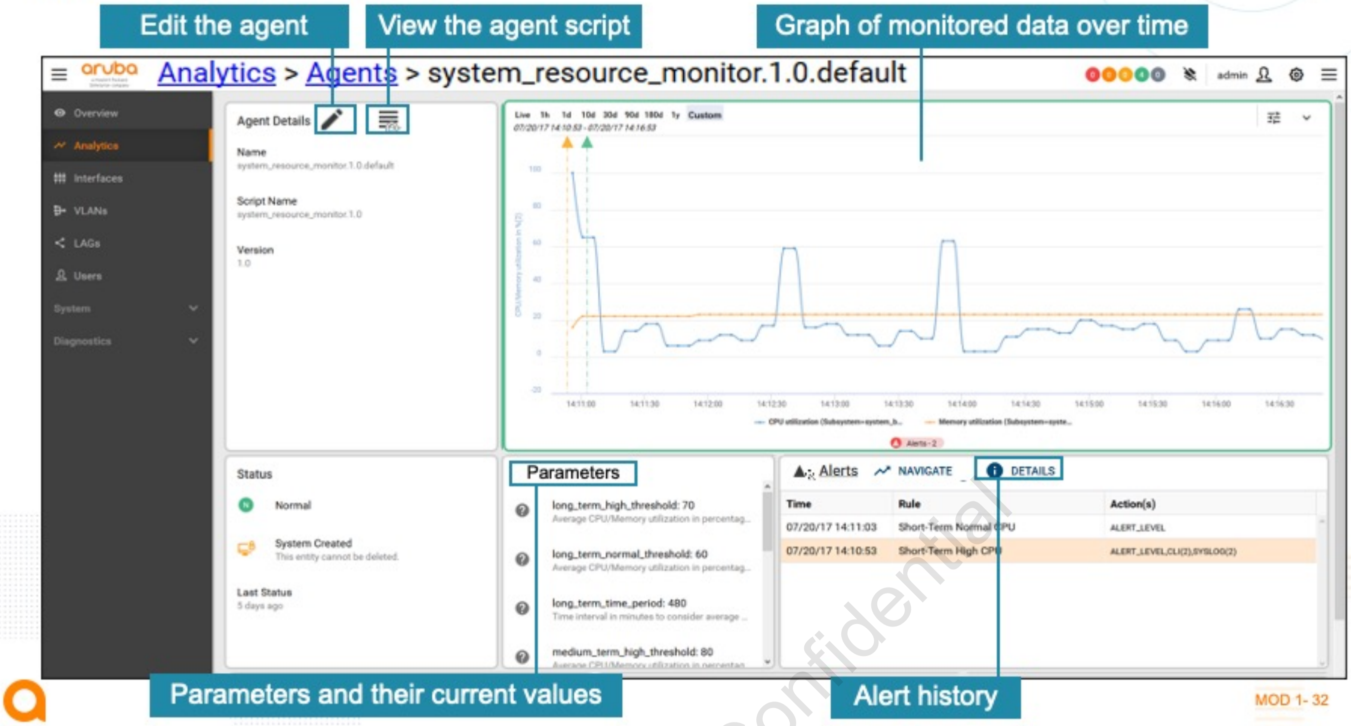
On the Analytics dashboard, the Agents section lists all of the agents on NAE and indicates whether they are operating normally.

- The dashboard also shows one graph for each of the NAE agents that has been added to it.
- You can click the plus icons in the Agents section to add the graph for another agent.
- The Scripts section shows the Python scripts associated with those agents.
- Look to the Alerts section to see more details on any minor, major, or critical alerts that the NAE agents have generate.

You can select the name of these sections to move to a page with details. For example, click the Agents link to see the Agent Management. Or you can select an individual agent to move to that agent's detail page.

You can also change the layout for the dashboard. Click this icon at the top of the bar and choose, Unlock Page Layout. Each section in the dashboard is then outlined with dashed lines. You can click a section and move it around. When you are done editing, you can click the icon again to lock the page layout.

NAE Agent Details



Edit the agent

View the agent script

Graph of monitored data over time

Parameters and their current values

Alert history

Analytics > Agents > system_resource_monitor.1.0.default

Agent Details

Name
system_resource_monitor.1.0.default

Script Name
system_resource_monitor.1.0

Version
1.0

Status

Normal

System Created
This entity cannot be deleted.

Last Status
5 days ago

Parameters

long_term_high_threshold: 70
Average CPU/Memory utilization in percentag...

long_term_normal_threshold: 60
Average CPU/Memory utilization in percentag...

long_term_time_period: 480
Time interval in minutes to consider average ...

medium_term_high_threshold: 80
Average CPU/Memory utilization in percentag...

Time	Rule	Action(s)
07/20/17 14:11:03	Short-Term Normal CPU	ALERT_LEVEL
07/20/17 14:10:53	Short-Term High CPU	ALERT_LEVEL,CLI(2),SYSLOG(2)

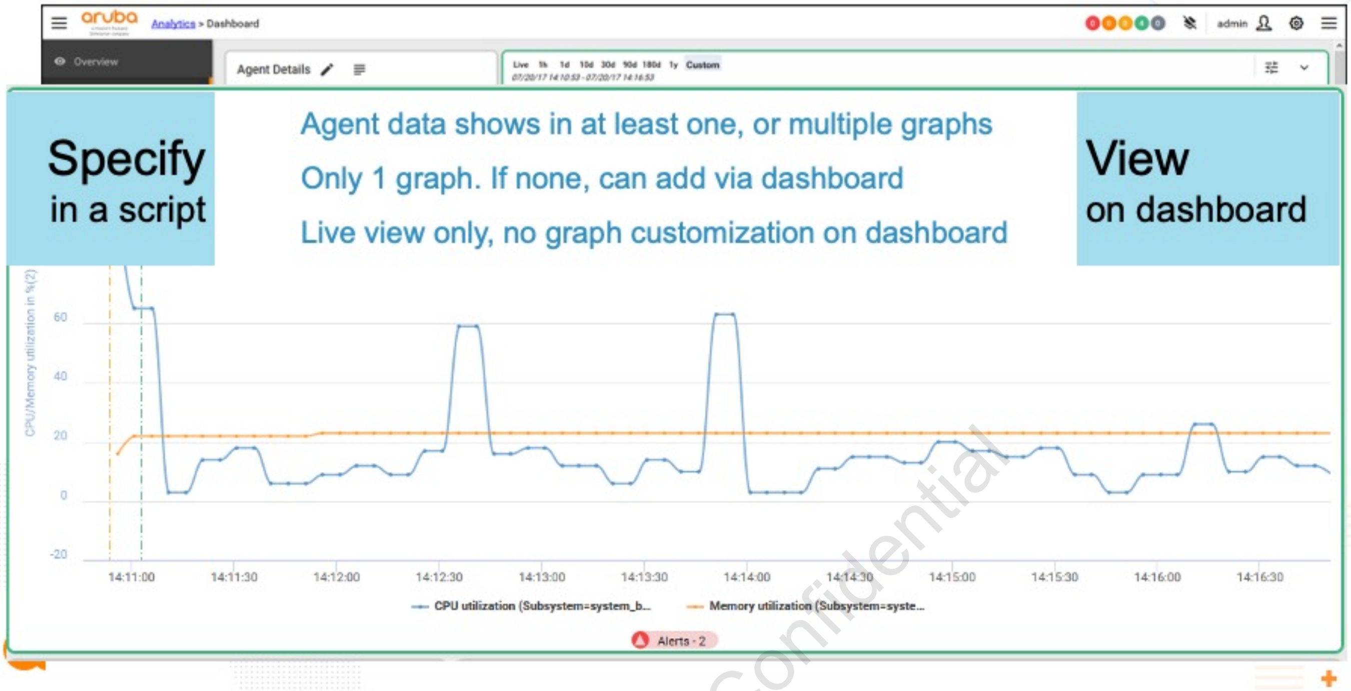
MOD 1- 32

The Agent Details page provides more detailed information about the agent, including its version and status. If the agent has multiple monitors, this page shows all of the graphs associated with those monitors. The parameters section shows agent settings that you can configure. You can also click icons to edit the agent or look at its script.

Check the alerts section to see information about the alerts this agent generates so that you can take the proper steps to resolve them. You can click a particular alert and then select Details to see a window with Alert details.

Aruba Training-Confidential

Graphs



Specify in a script

Agent data shows in at least one, or multiple graphs

Only 1 graph. If none, can add via dashboard

Live view only, no graph customization on dashboard

View on dashboard

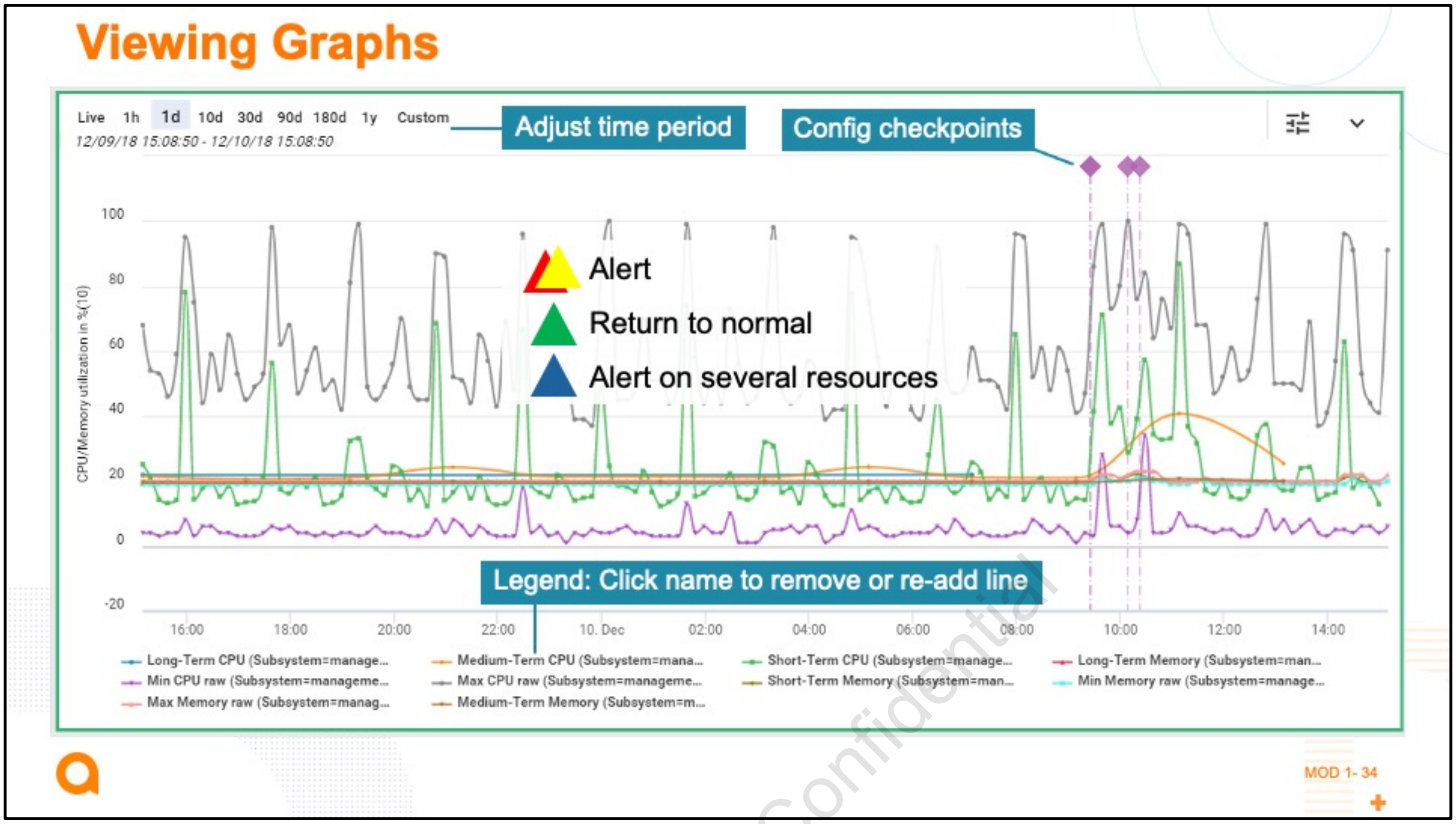
Data collected by an Analytics agent is displayed in the Web UI in one or more time series graphs.

| An agent has at least one graph, and can have multiple graphs as specified in the script. However, only one graph represents the agent on the Analytics dashboard, also specified in the script.

If the dashboard does not include a graph for an agent, you can add one from Analytics dashboard. Graphs on the Analytics dashboard represent a live view only. The graph customization toolbar is not available from the Analytics dashboard.

Aruba Training-Confidential

Viewing Graphs



Data collected by an Analytics agent is displayed in the Web UI in one or more time series graphs. An agent has at least one graph, and can have multiple graphs as specified in the script. However, only one graph represents the agent on the Analytics dashboard, also specified in the script.

If the dashboard does not include a graph for an agent, you can add one from Analytics dashboard. Graphs on the Analytics dashboard represent a live view only. The graph customization toolbar is not available from the Analytics dashboard.

The figure shows a closer view of a graph in an Agent Details page. This graph has multiple lines, one for each monitor within the agent. For example, one line shows the CPU usage averaged over short time periods, one shows CPU usage averaged over long time periods, one shows memory averaged over short time periods, and so on. The legend at the bottom of the graph shows what each line indicates. You can click any of the names in the legend to remove that line from the graph or to re-add it. In this way, you can focus on the information that you want.

Configuration checkpoints and alert indicators are overlaid on the graph. Configuration checkpoints are shown as purple diamonds. Alert indicators can include the following:

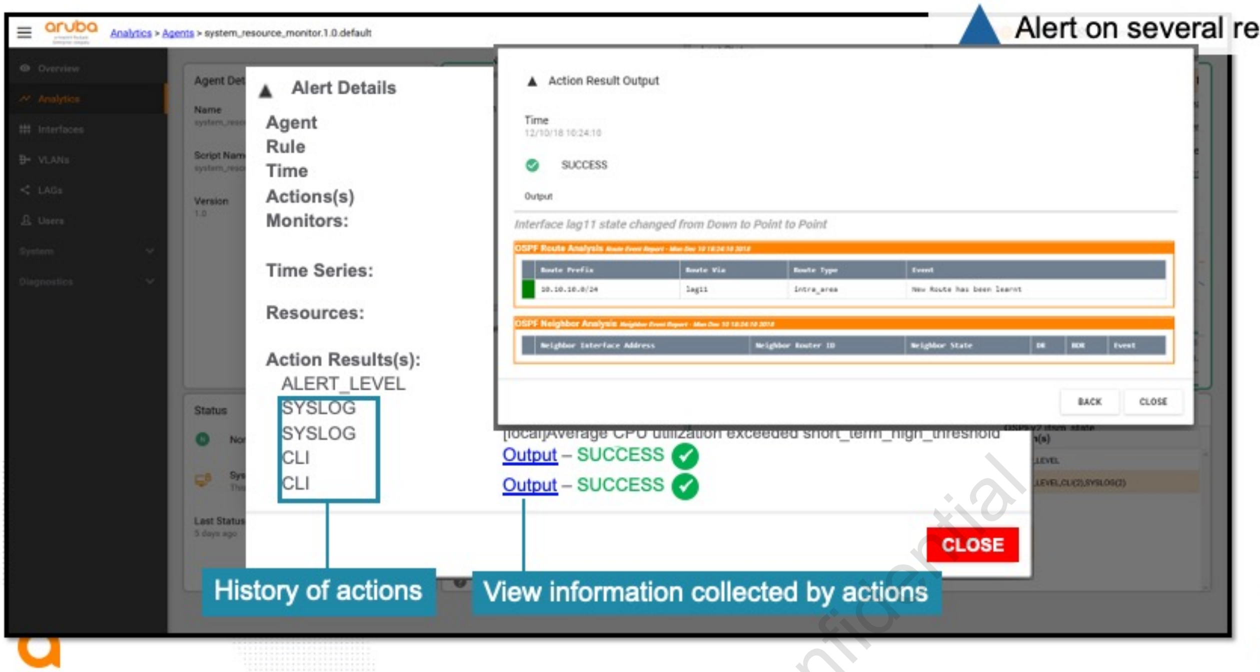
- | A red or yellow triangle for an alert
- A green triangle for a return to normal
- A blue triangle for an alert on several resources being monitored.

You can also change the graph to show different time periods such as the last hour or day up to the last year. Or you can specify a custom time period.

Aruba Training-Confidential

NAE Alert Details

▲ Alert
▲ Return to normal
▲ Alert on several resources



▲ Alert Details

Agent
Rule
Time
Actions(s)
Monitors:

Time Series:

Resources:

Action Results(s):
ALERT_LEVEL
SYSLOG
SYSLOG
CLI
CLI

[local]Average CPU utilization exceeded short_term_high_threshold
Output – SUCCESS ✓
Output – SUCCESS ✓

▲ Action Result Output

Time
12/10/18 10:24:10

✓ SUCCESS

Output

Interface lag11 state changed from Down to Point to Point

OSPF Route Analysis

	Route Prefix	Route Via	Route Type	Event
🟩	10.10.10.0/24	lag11	intra_area	New Route has been learnt

OSPF Neighbor Analysis

	Neighbor Interface Address	Neighbor Router ID	Neighbor State	DR	BDR	Event

History of actions

View information collected by actions

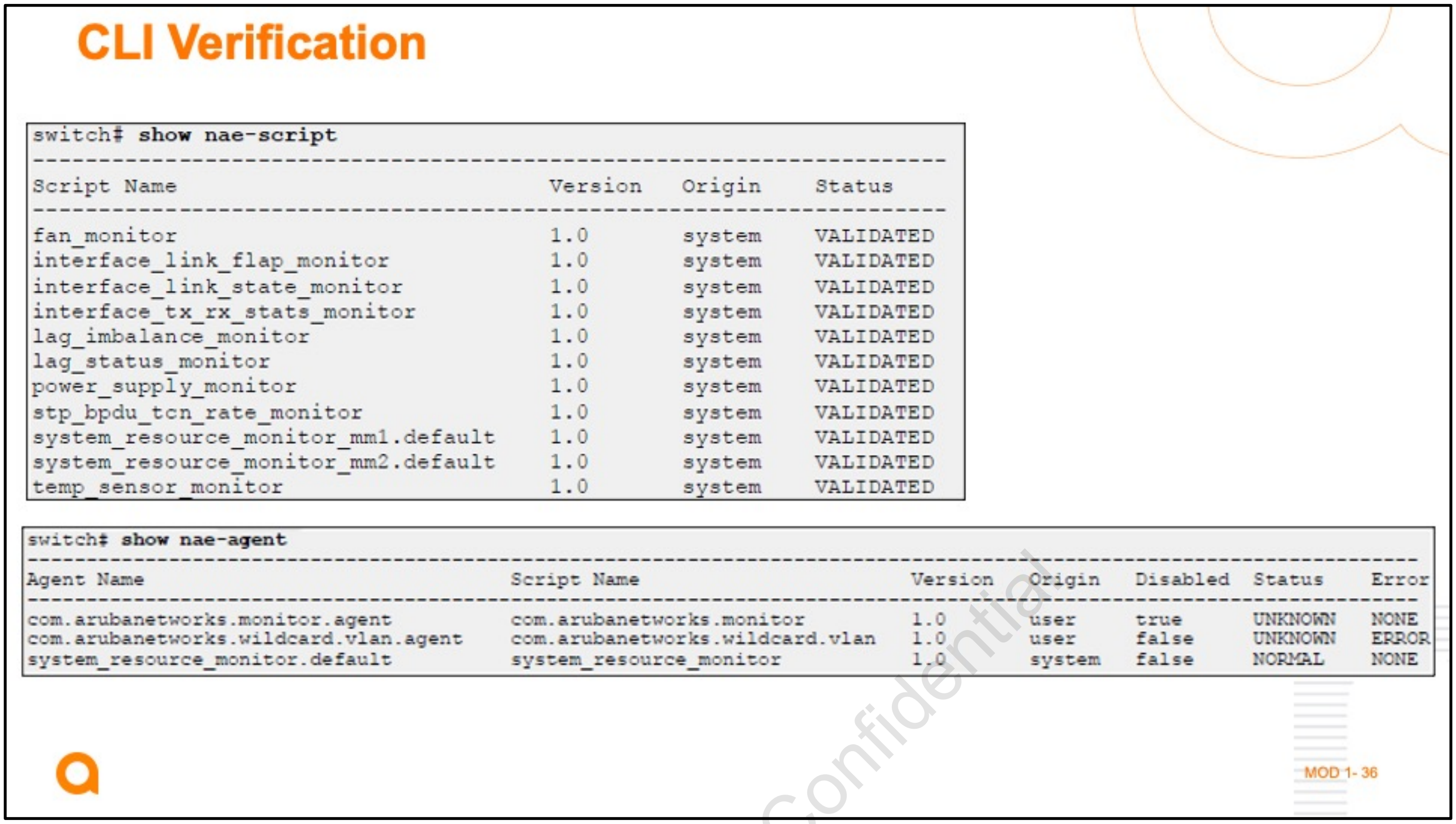
MOD 1- 35

Aruba Training-Confidential

The Alert Details window shows when and why the alert occurred, the alert level, and a history of actions taken for the condition that triggered this alert. In this example, these actions include syslog messages and CLI commands. You can also collect any information that those actions generated.

| From the Alert Details window, select Output to see the output from the CLI commands. |f the alert generated a custom report, you can click and view that report.

You can also use the REST API to obtain information about NAE and NAE agents. Refer to the documentation and reference interface for more details.



CLI Verification

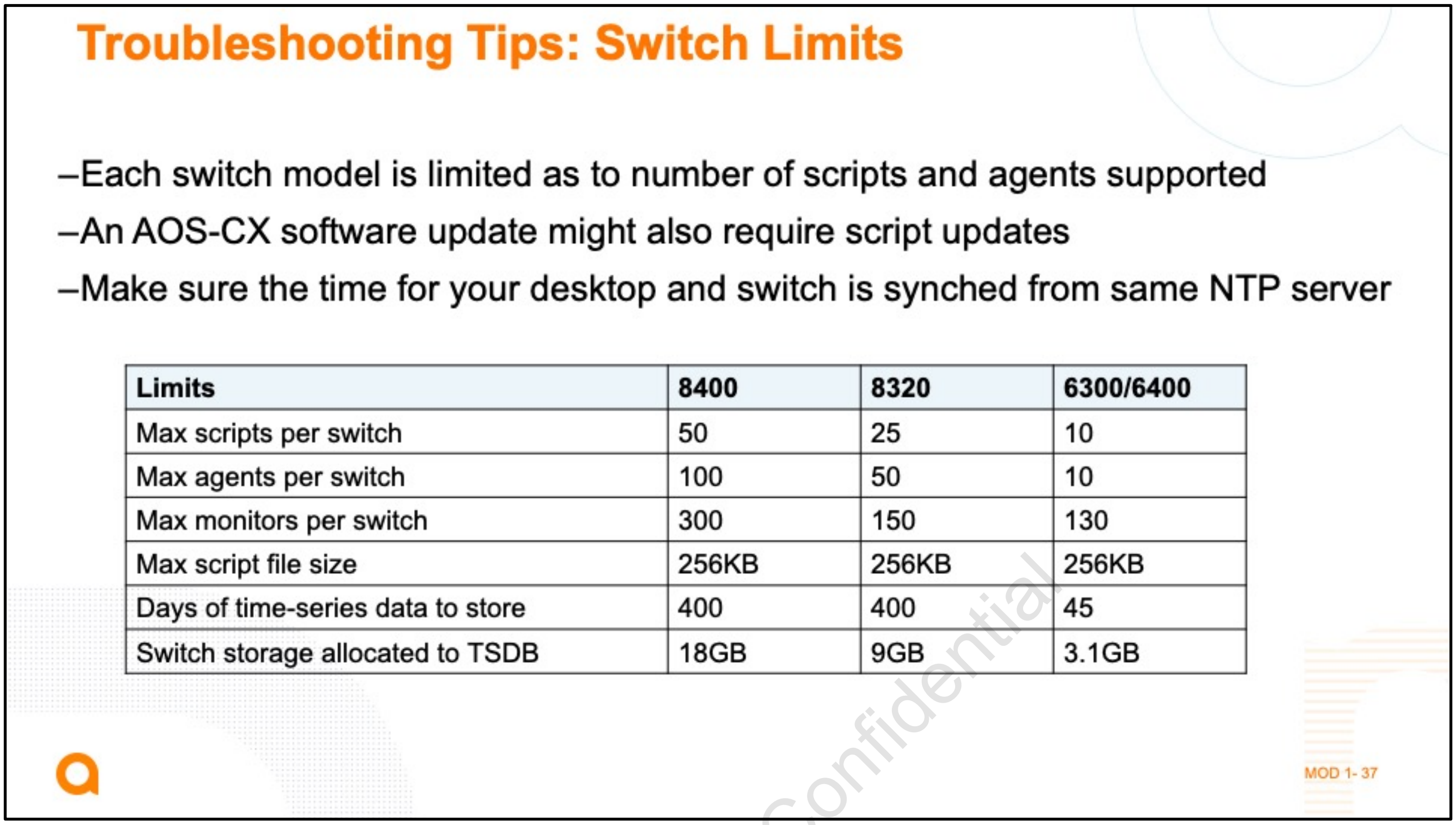
```
switch# show nae-script
---------------------------------------------------------------------
Script Name                              Version   Origin    Status
---------------------------------------------------------------------
fan_monitor                              1.0       system    VALIDATED
interface_link_flap_monitor              1.0       system    VALIDATED
interface_link_state_monitor             1.0       system    VALIDATED
interface_tx_rx_stats_monitor            1.0       system    VALIDATED
lag_imbalance_monitor                    1.0       system    VALIDATED
lag_status_monitor                       1.0       system    VALIDATED
power_supply_monitor                     1.0       system    VALIDATED
stp_bpdu_tcn_rate_monitor                1.0       system    VALIDATED
system_resource_monitor_mm1.default      1.0       system    VALIDATED
system_resource_monitor_mm2.default      1.0       system    VALIDATED
temp_sensor_monitor                      1.0       system    VALIDATED
```

```
switch# show nae-agent
------------------------------------------------------------------------------------------------------------------------
Agent Name                                 Script Name                            Version   Origin    Disabled  Status    Error
------------------------------------------------------------------------------------------------------------------------
com.arubanetworks.monitor.agent            com.arubanetworks.monitor              1.0       user      true      UNKNOWN   NONE
com.arubanetworks.wildcard.vlan.agent      com.arubanetworks.wildcard.vlan        1.0       user      false     UNKNOWN   ERROR
system_resource_monitor.default            system_resource_monitor                1.0       system    false     NORMAL    NONE
```

MOD 1- 36

Aruba Training-Confidential

You can see which scripts and agents are available from the CLI with the show nae-script and show nae-agent commands.



Troubleshooting Tips: Switch Limits

- –Each switch model is limited as to number of scripts and agents supported
- –An AOS-CX software update might also require script updates
- –Make sure the time for your desktop and switch is synched from same NTP server

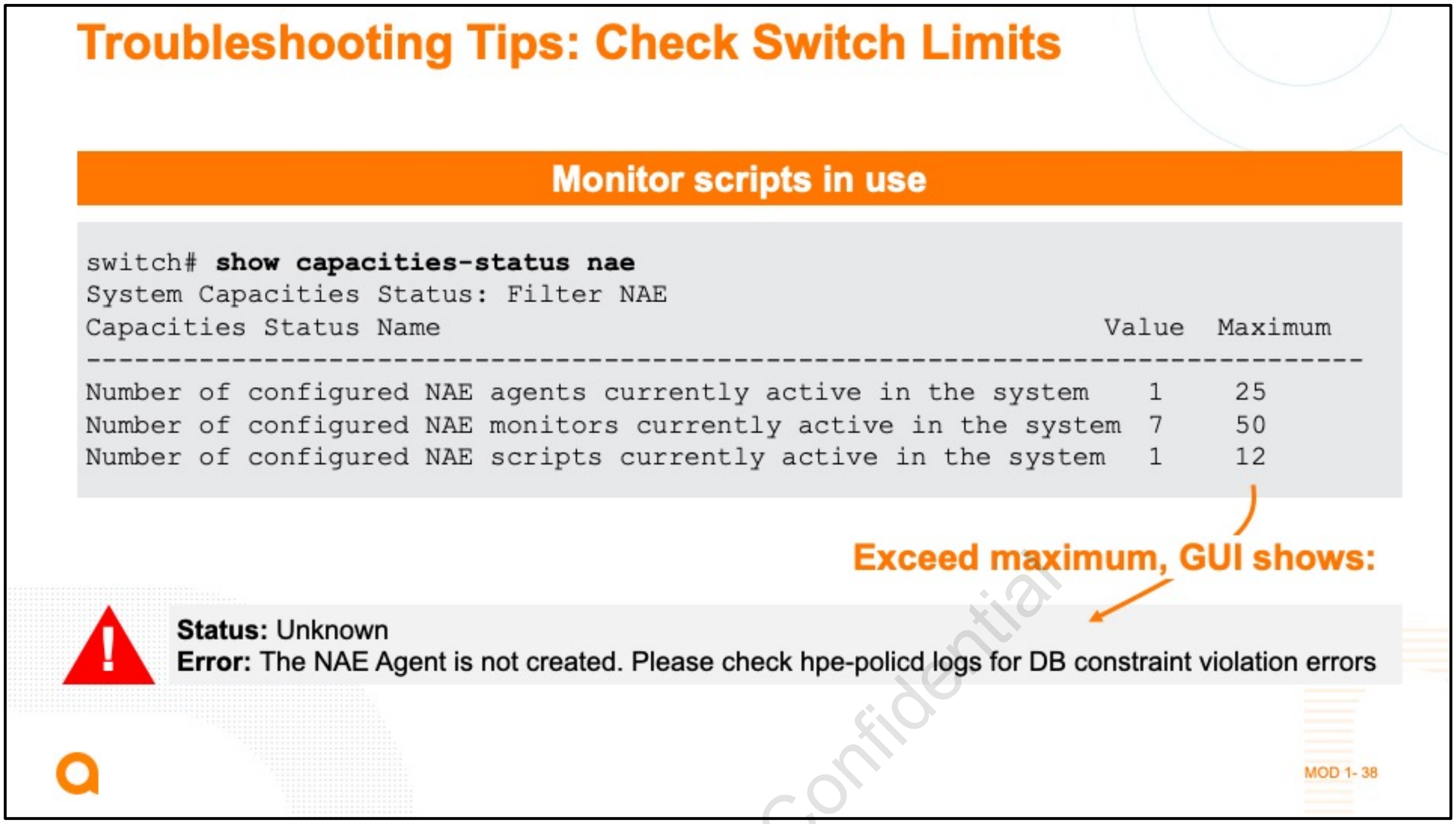
Limits	8400	8320	6300/6400
Max scripts per switch	50	25	10
Max agents per switch	100	50	10
Max monitors per switch	300	150	130
Max script file size	256KB	256KB	256KB
Days of time-series data to store	400	400	45
Switch storage allocated to TSDB	18GB	9GB	3.1GB

MOD 1- 37

Aruba NAE supported maximums for each switch model is shown in the figure

Note: The NAE limitations defined above can change based on the AOS-CX software version running on the switch. The above limitations are for AOS-CX 10.4.

Aruba Training-Confidential



Troubleshooting Tips: Check Switch Limits

Monitor scripts in use

```
switch# show capacities-status nae
System Capacities Status: Filter NAE
Capacities Status Name                                         Value  Maximum
-------------------------------------------------------------------------------
Number of configured NAE agents currently active in the system    1     25
Number of configured NAE monitors currently active in the system  7     50
Number of configured NAE scripts currently active in the system   1     12
```

Exceed maximum, GUI shows:

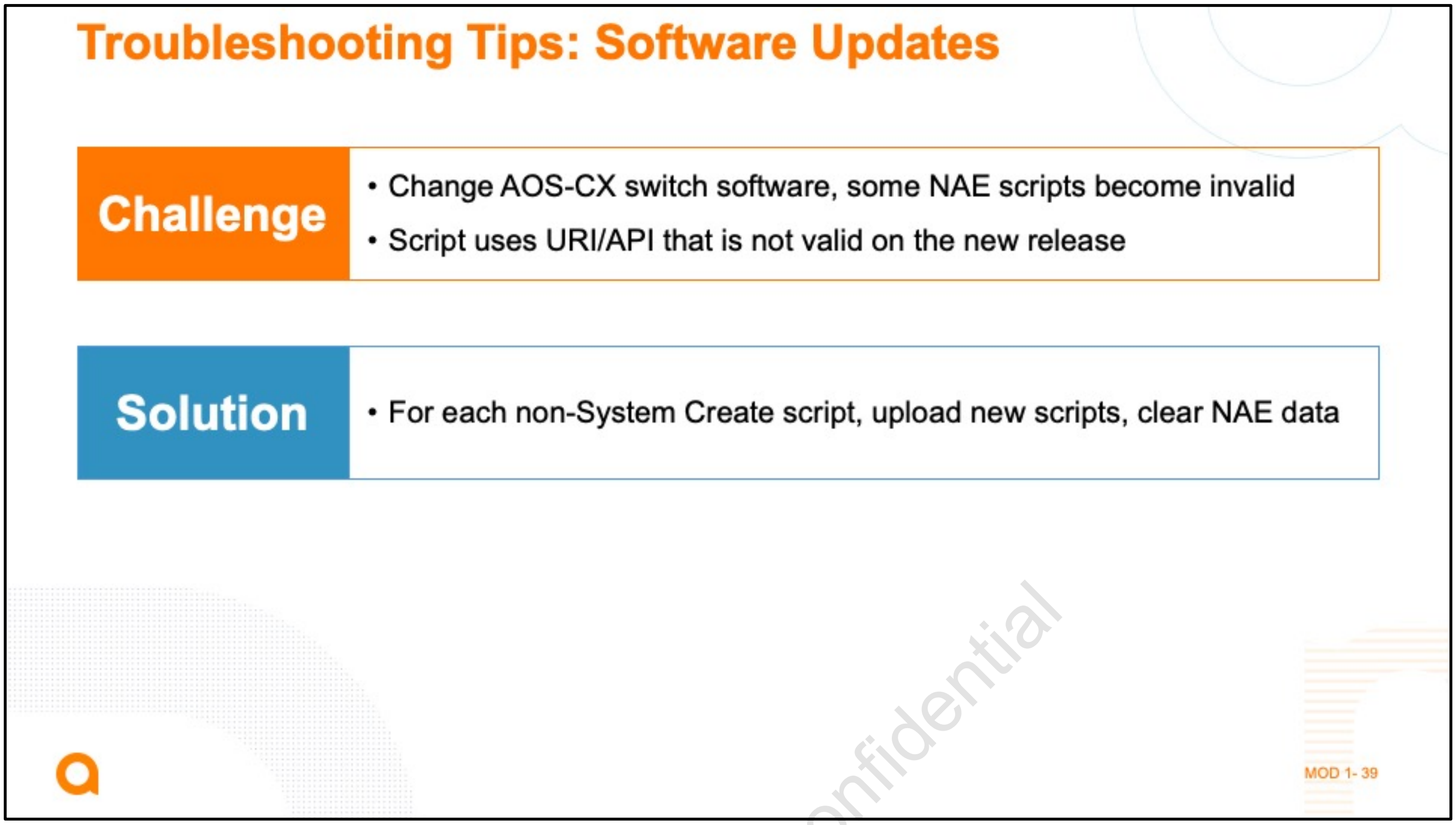
Status: Unknown
Error: The NAE Agent is not created. Please check hpe-policd logs for DB constraint violation errors

MOD 1- 38

Aruba Training-Confidential

To monitor the scripts in use, execute the **show capacities-status nae** command, as shown in the figure. Notice that it shows the current and maximum agents, monitors, and scripts supported.

Suppose you attempt to create an agent that would exceed the maximum agents supported on the switch. The agent appears in the GUI Agents panel with a red triangle error symbol and status of Unknown, with the error message as shown in the figure.



Troubleshooting Tips: Software Updates

Challenge

- Change AOS-CX switch software, some NAE scripts become invalid
- Script uses URI/API that is not valid on the new release

Solution

- For each non-System Create script, upload new scripts, clear NAE data

MOD 1- 39

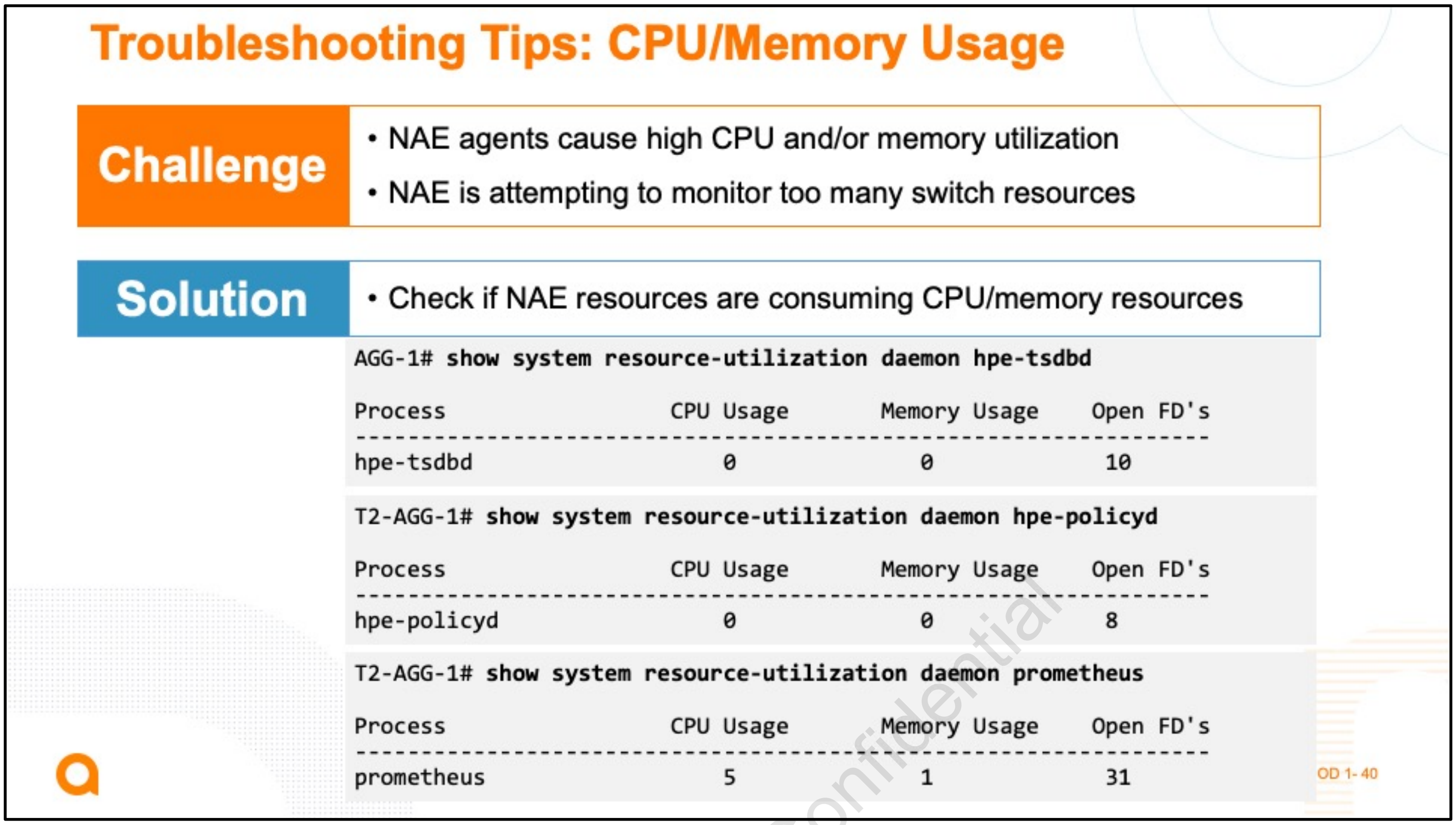
Aruba Training-Confidential

Managing NAE scripts across switch software updates

After you update or downgrade the switch AOS-CX software, some NAE scripts might become invalid. This is likely because the script uses a URI or API function that is not valid on that software release. The existing scripts might generate errors, and the NAE data might not be valid until you upload the new scripts and clear the NAE data.

1. For each script that is not marked System Created in the Web UI, locate the script version that supports the currently running software release.
   A. For Aruba-certified NAE scripts, the Aruba Solutions Exchange (ASE) includes tags that indicate the minimum and maximum supported software release.
   B. When you select the ASE download button in the Web UI, the Web UI displays only the Aruba-certified NAE scripts that are supported on the software release running on the switch.
   C. For scripts that are not Aruba-certified NAE scripts downloaded from the ASE, see the information provided by the script author about which script version is supported.
2. Follow the instructions for updating a script. The steps to update a script depend on the switch software release version that was running when the script was installed.
3. On the Analytics Dashboard, locate and close any time series graph panels for the default agent from the previous software release.

The default agent is created from the built-in script. The name of the agent is based on the name of the built-in script. Built-in scripts and agents have an origin of "system" and are marked System Created in the Web UI. Built-in scripts and their agents are updated automatically. However, sometimes the time series graph panel for the previous default agent is not closed during the update. After the software update, instead of graphed data, such panels show an error message beginning with: “Agent data not found”.



Troubleshooting Tips: CPU/Memory Usage

Challenge

- NAE agents cause high CPU and/or memory utilization
- NAE is attempting to monitor too many switch resources

Solution

- Check if NAE resources are consuming CPU/memory resources

```
AGG-1# show system resource-utilization daemon hpe-tsdbd

Process                  CPU Usage       Memory Usage    Open FD's
-------------------------------------------------------------------
hpe-tsdbd                    0               0              10
```

```
T2-AGG-1# show system resource-utilization daemon hpe-policyd

Process                  CPU Usage       Memory Usage    Open FD's
-------------------------------------------------------------------
hpe-policyd                  0               0              8
```

```
T2-AGG-1# show system resource-utilization daemon prometheus

Process                  CPU Usage       Memory Usage    Open FD's
-------------------------------------------------------------------
prometheus                   5               1              31
```

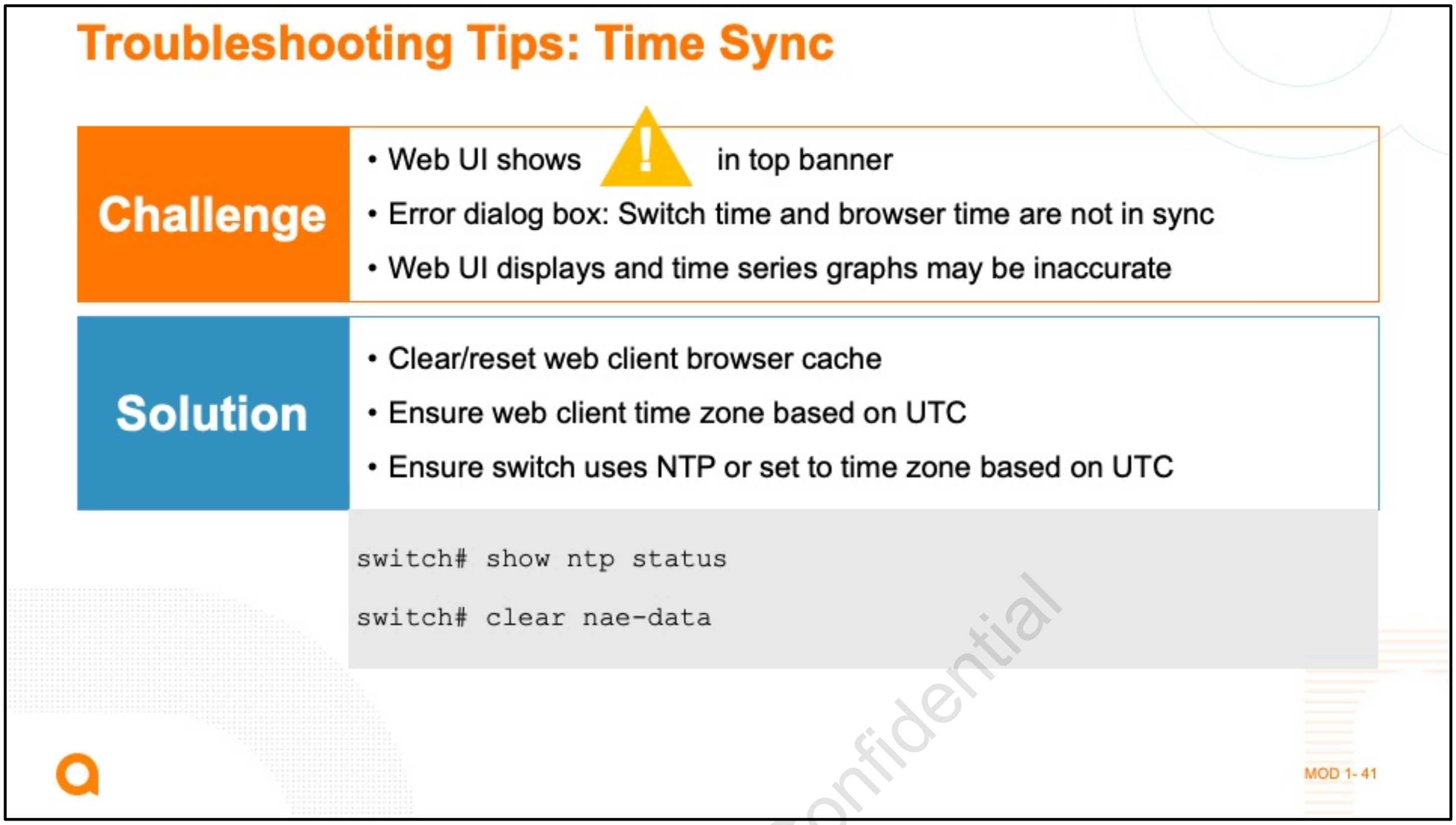
OD 1- 40

Aruba Training-Confidential

Suppose that a switch with NAE agents installed is experiencing high CPU usage, high memory usage, or both, and overall performance is affected. The cause is that the NAE is attempting to monitor too many switch resources. For example, a script uses wildcard characters in a URI to monitor all interfaces, ACLs, or VLANs, and the switch has hundreds or thousands of those items.

To validate this theory, verify that the resources associated with the NAE are consuming the memory and CPU resources by entering the following commands in the switch CLI:

- **show system resource-utilization daemon hpe-tsdbd**
- **show system resource-utilization daemon hpe-policyd**
- **show system resource-utilization daemon Prometheus**



Troubleshooting Tips: Time Sync

Challenge

- Web UI shows ⚠ in top banner
- Error dialog box: Switch time and browser time are not in sync
- Web UI displays and time series graphs may be inaccurate

Solution

- Clear/reset web client browser cache
- Ensure web client time zone based on UTC
- Ensure switch uses NTP or set to time zone based on UTC

```
switch# show ntp status

switch# clear nae-data
```

MOD 1- 41

Aruba Training-Confidential

Switch time and browser time are not in sync

The Web UI displays a yellow caution triangle in the top banner and an error dialog box with the following title: *Switch time and browser time are not in sync*.

The content of the dialog box indicates how many seconds the switch time is ahead of or behind the browser time, and states that the information displayed in the Web UI might not be accurate. In addition, time series graphs on the Analytics page might be missing expected data, might have data shown with inaccurate times, or might display incorrect data.

Things you can do to fix this issue:

- Try clearing or resetting the web client browser cache.
- Ensure that the web client from which you are viewing the Web UI is set to a time zone based on UTC. For example, if your workstation is set to Eastern Standard Time (EST), and you want to use Pacific Standard Time (PST), change the time by setting the time zone instead of by manually resetting the time.
- Ensure that the switch is set to use NTP or to a time zone based on UTC time. NTP synchronizes the time of day among a set of distributed time servers and clients to correlate events when receiving system logs and other time-specific events from multiple network devices. All NTP communications use Coordinated Universal Time (UTC). To show the NTP status, use the **show ntp status** command. After you configure the switch, clear the NAE data by entering the **clear nae-data** command from the manager context.
- If the switch is set to use NTP and there has been a significant clock change, clear the NAE data by using the **clear nae-data** command.

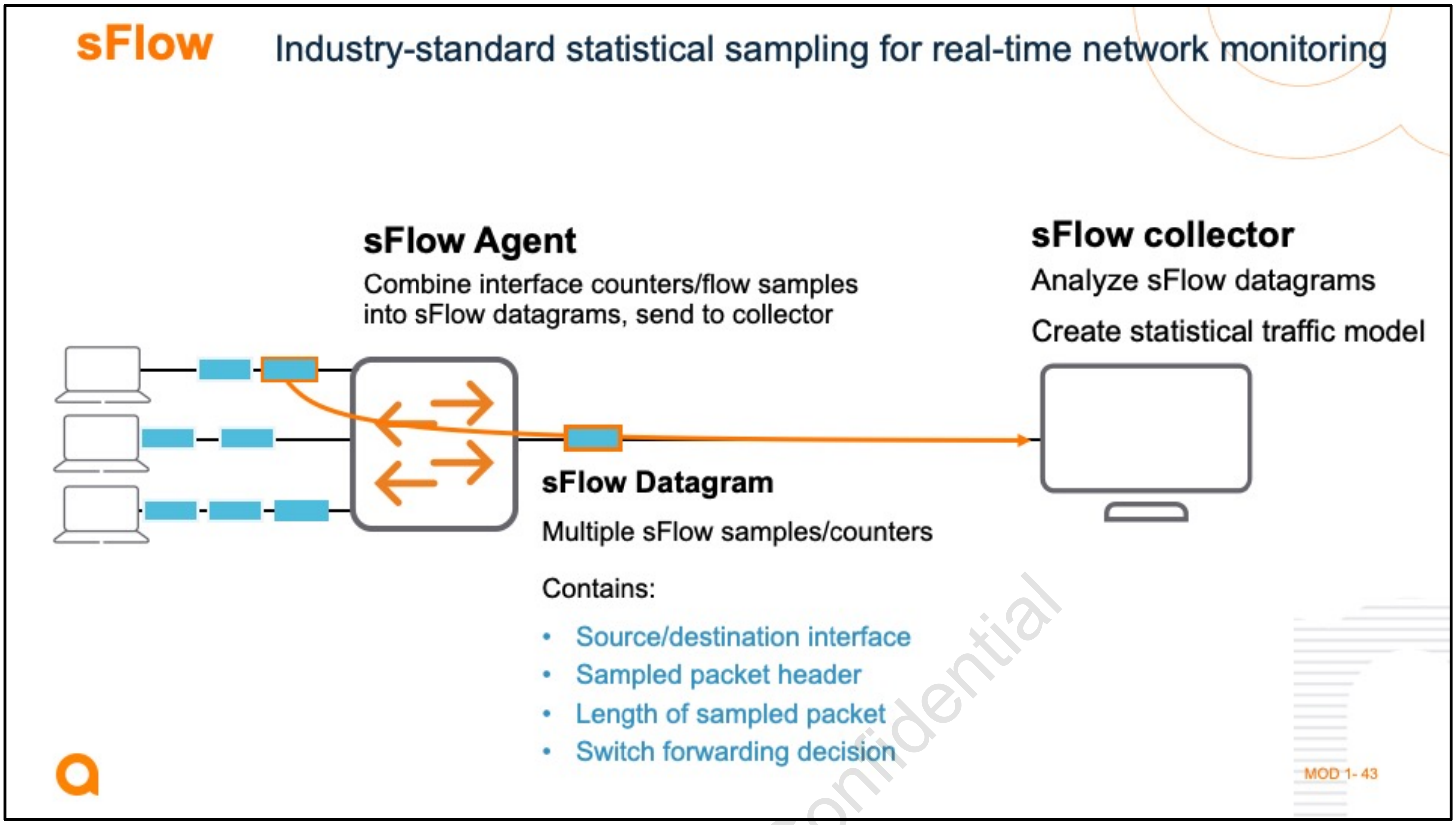
Note: See the *Network Analytics Engine Guide* for more details on using and troubleshooting the NAE feature.

Basic Traffic Monitoring

MOD 1- 42

You will now learn about a few more ways to monitor traffic on AOS-CX switches.

Aruba Training-Confidential



sFlow Industry-standard statistical sampling for real-time network monitoring

sFlow Agent

Combine interface counters/flow samples into sFlow datagrams, send to collector

sFlow collector

Analyze sFlow datagrams

Create statistical traffic model

sFlow Datagram

Multiple sFlow samples/counters

Contains:

- Source/destination interface
- Sampled packet header
- Length of sampled packet
- Switch forwarding decision

MOD 1- 43

Aruba Training-Confidential

AOS-CX switches support sFlow, short for “sampled Flow”. It is an industry-standard technology that uses statistical sampling to enable network monitoring. sFlow relies on two components:

- sFlow agents
- sFlow collector (which is also sometimes called a receiver)

sFlow agents run on network devices—such as switches, routers, and access points—gathering information about the packets transmitted on these devices and sending that information to the sFlow collector for analysis. Specifically, agents can send traffic samples and polling counters.

For traffic samples, the sFlow standard defines all the packets that a network device receives on one interface and are forwarded to another interface as a “flow.” Using a statistically accurate algorithm, the sFlow agent examines on average one of every *n*th packet, where *n* is the number of packets. The samples are taken “on average” because sFlow employs some randomness to avoid sampling packets at precise intervals that might coincide with certain traffic patterns.

An sFlow agent packages sampled information into small datagrams, which include Layer 2 through 7 information, as shown in the figure.

An sFlow agent can also use counter polling to gather traffic statistics. It polls data sources at specified intervals, adds this information to the sFlow datagrams, and sends them to a specified sFlow collector. On the AOS-CX switches, the data sources are ports (Ethernet interfaces).

The agents can include information from several packets into a single datagram. Because traffic samples are compact and agents sample only a small percentage of traffic, sFlow does not require a large amount of network bandwidth or processing power on network devices.

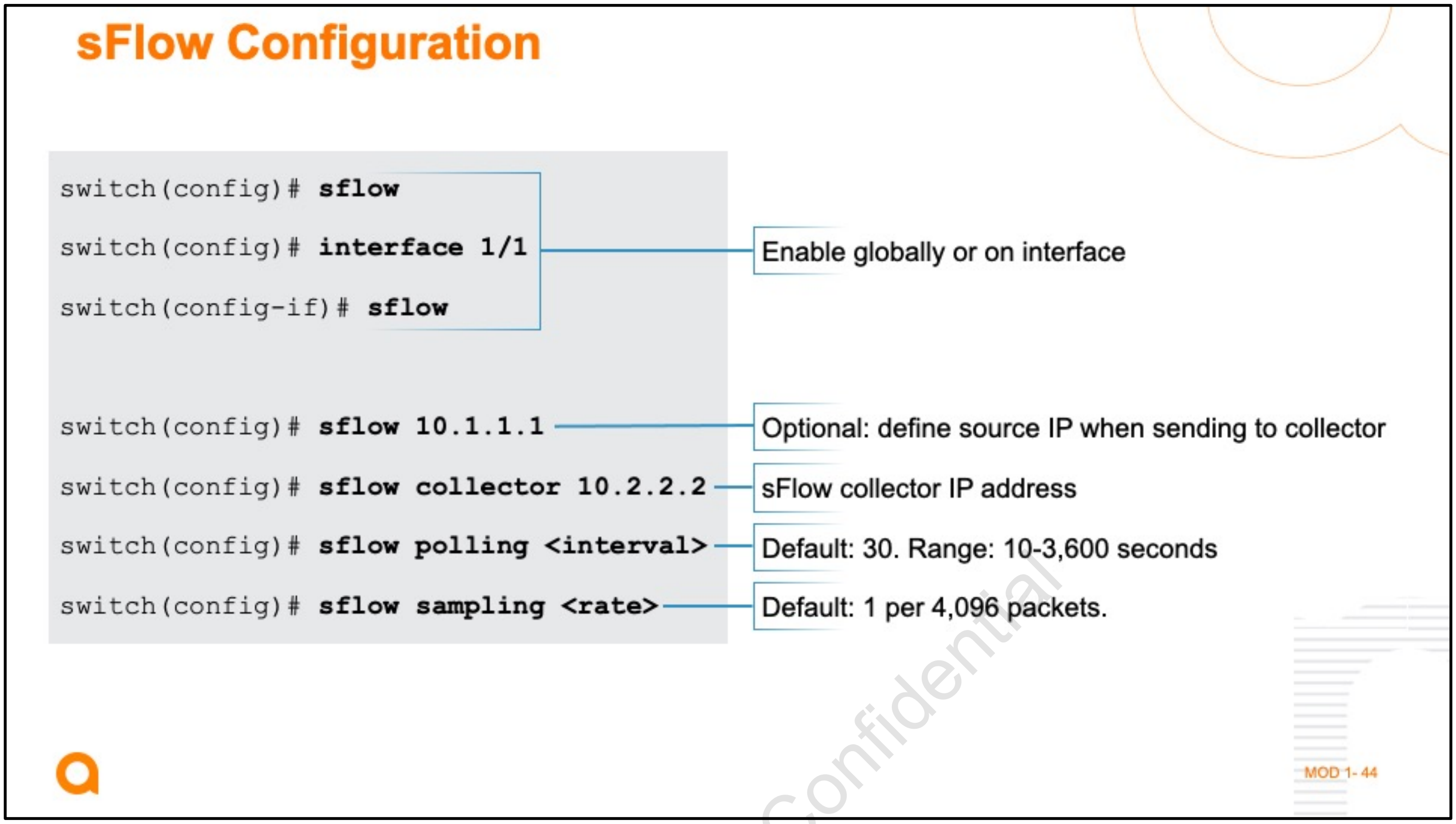
To forward the datagram to the sFlow collector, the sFlow agent appends a UDP header and an IP header to the datagram and encapsulates the packet in an Ethernet frame.

AOS-CX switches implement sFlow in hardware, as is typical, for quick, efficient operations.

The sFlow collector analyzes sFlow datagrams from each agent on wireless and wired networks. The collector then creates a statistical model of network traffic that can be used for network traffic management, baselining, security auditing, troubleshooting, and more.

Note: For more information about sFlow, visit http://sflow.org.

Aruba Training Confidential



sFlow Configuration

```
switch(config)# sflow
switch(config)# interface 1/1
switch(config-if)# sflow

switch(config)# sflow 10.1.1.1
switch(config)# sflow collector 10.2.2.2
switch(config)# sflow polling <interval>
switch(config)# sflow sampling <rate>
```

Enable globally or on interface

Optional: define source IP when sending to collector

sFlow collector IP address

Default: 30. Range: 10-3,600 seconds

Default: 1 per 4,096 packets.

MOD 1- 44

Aruba Training-Confidential

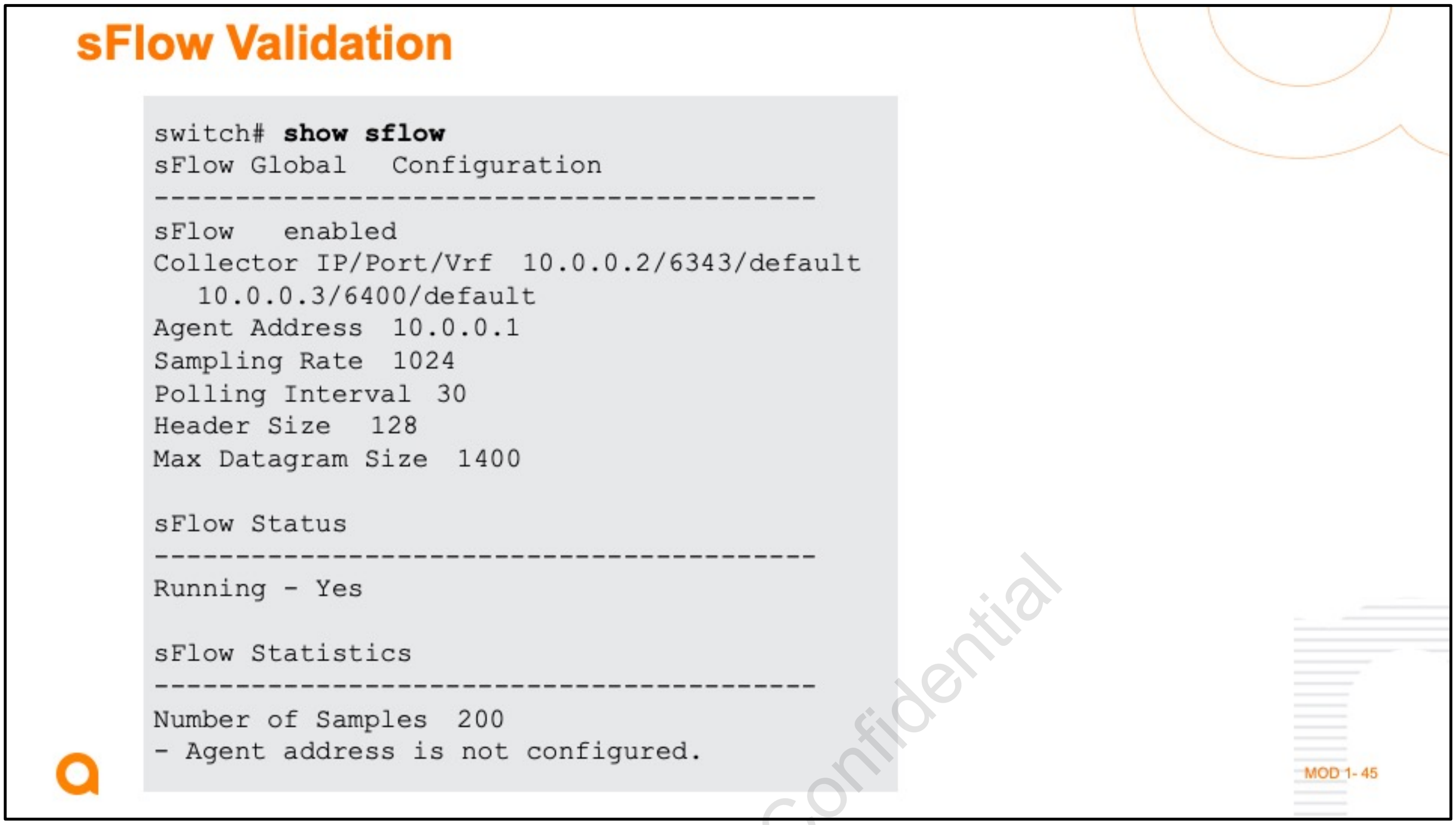
The figure shows how to configure sFlow on AOS-CX switches

The optional **sflow** <local-IP-address> command lets you define the source address on the switch to be used when sending sFlow data to a receiver/collector.

You must configure the IP address of the sFlow collector. As shown below, you can enter an optional UDP port number and/or VR name. If you do not include the port in the command, the AOS-CX switch will use the default UDP destination port, 6343.

Switch(config)# **sflow collector** <IP-address> [**udp** <port-num>] [**vrf** <VRF-name>]

You can also specify polling and packet sampling rate, as shown in the figure. The default sampling rate is 1 in every 4,096 packets for a flow. A warning message is displayed when the sampling rate is set to less than 4096 and proceeds only after user confirmation.



sFlow Validation

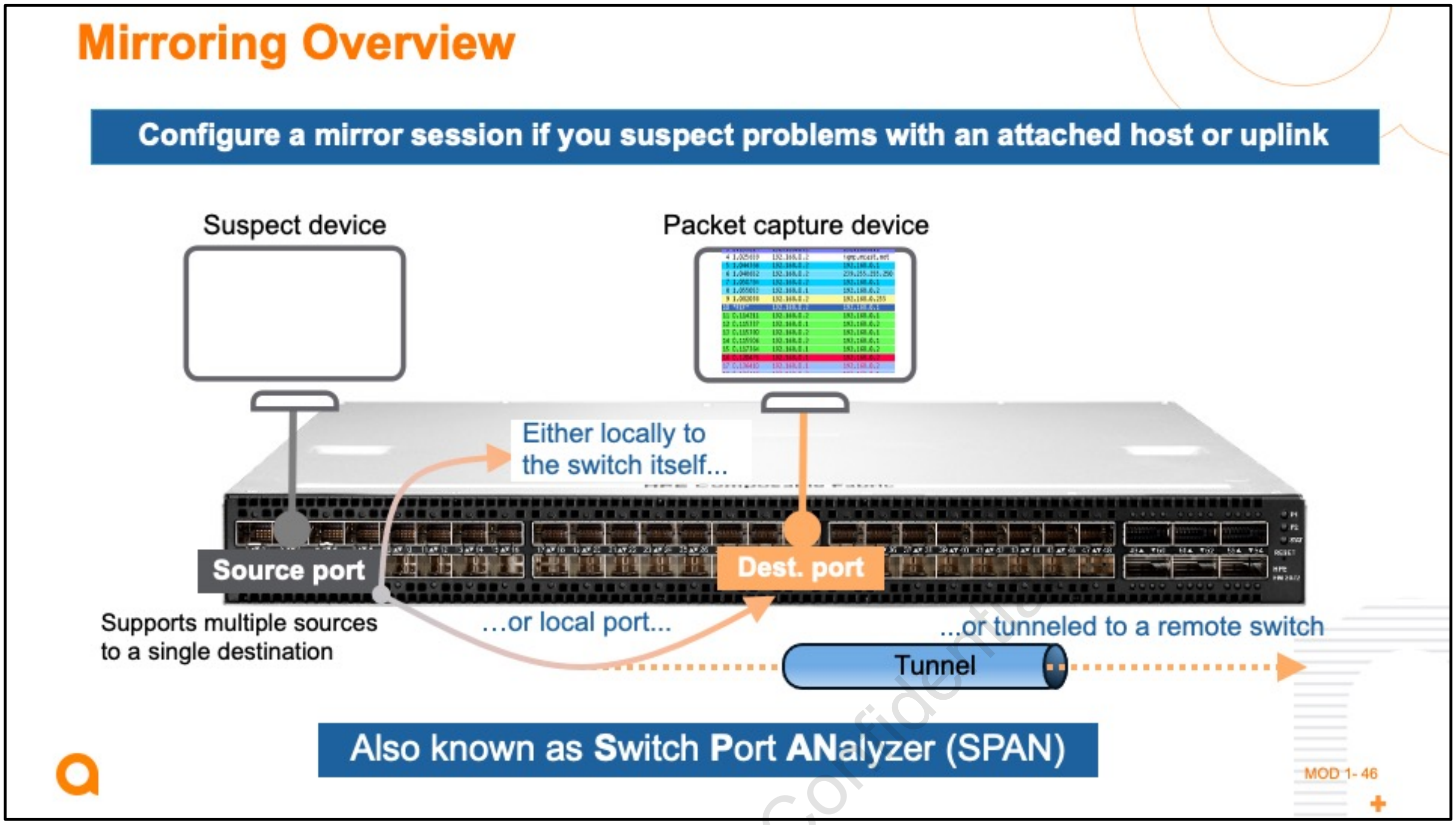
```
switch# show sflow
sFlow Global  Configuration
-----------------------------------------
sFlow   enabled
Collector IP/Port/Vrf  10.0.0.2/6343/default
   10.0.0.3/6400/default
Agent Address  10.0.0.1
Sampling Rate  1024
Polling Interval  30
Header Size   128
Max Datagram Size  1400

sFlow Status
-----------------------------------------
Running - Yes

sFlow Statistics
-----------------------------------------
Number of Samples  200
- Agent address is not configured.
```

MOD 1- 45

Aruba Training-Confidential

As shown in the figure, use **show sflow** to verify your configuration.

Note: sFlow and NAE are two different tools that you can use on your AOS-CX switch. sFlow can be used to share information easily with third-party monitoring products, like H3C IMC's Network Traffic Analyzer (NTA) module, which supports protocols like sFlow (open standard), NetFlow (Cisco-proprietary), and NetStream (H3C proprietary). NAE information is RESTful based and thus can be shared with a multitude of systems that support a RESTful API infrastructure. You can use either one of these, or both of these, depending on the type of external solution you need to share information with.



Mirroring Overview

Configure a mirror session if you suspect problems with an attached host or uplink

Suspect device

Packet capture device

Either locally to the switch itself...

Source port

Dest. port

Supports multiple sources to a single destination

...or local port...

...or tunneled to a remote switch

Tunnel

Also known as **S**witch **P**ort **AN**alyzer (SPAN)

MOD 1- 46

Aruba Training-Confidential

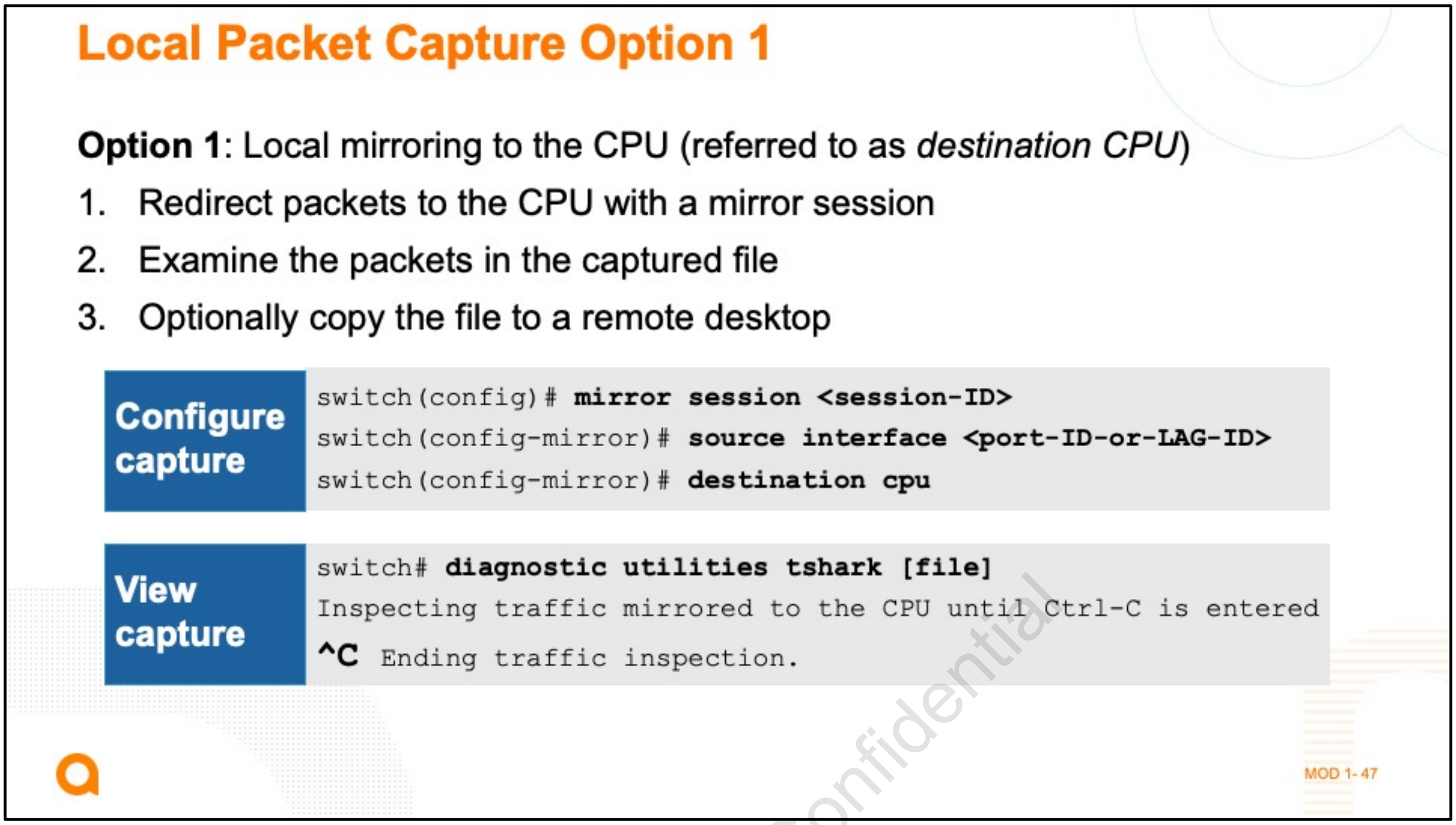
Packet captures can facilitate network troubleshooting. To facilitate this, you configure an appropriate switch or switches for port mirroring.

| You are telling the switch that traffic for some source port should be copied or mirrored to some destination. This could be:

- local storage on the switch itself
- A local port on the switch, to which you have attached some capture device – perhaps a Linux or Windows host running the popular Wireshark utility or similar. You can then use the packet capture utilities on your capture device for traffic analysis.
- Tunneled to some remote switch, where a capture device is attached.

| This collection of interfaces and settings is called a mirror session, which can have multiple sources, but only a single destination. Be careful! Multiple sources can overwhelm a single destination. For example, if you configure a multiple 10G source ports to a single 10G destination port, the destination may become saturated, causing packet loss.

| Mirroring is sometimes referred to as SPAN, for Switch Port ANalyzer. ArubaOS-CX v10.02 and higher support the ERSPAN feature, in which you send mirrored traffic over a tunnel to a remote switch.



Local Packet Capture Option 1

Option 1: Local mirroring to the CPU (referred to as *destination CPU*)

1. Redirect packets to the CPU with a mirror session
2. Examine the packets in the captured file
3. Optionally copy the file to a remote desktop

Configure capture

```
switch(config)# mirror session <session-ID>
switch(config-mirror)# source interface <port-ID-or-LAG-ID>
switch(config-mirror)# destination cpu
```

View capture

```
switch# diagnostic utilities tshark [file]
Inspecting traffic mirrored to the CPU until Ctrl-C is entered
^C Ending traffic inspection.
```

MOD 1- 47

Aruba Training-Confidential

There are two options for looking at packets locally on the AOS-CX switches via the CLI - Local mirroring and tcpdump

The option 1 configuration is shown in the figure.

The source interface can be a physical port or a LAG. You can configure from one to four mirror sessions to use the CPU (the switch itself). Use the **show mirror** command to verify your configuration.

By default, you don't see any information, even though the traffic being captured is "seen" by the CPU. To start examining the traffic, use the command **diagnostic utilities tshark [file]**, as shown in the figure. Type Ctrl-C to stop viewing the capture.

If you specify the file parameter, the switch captures packets from a mirror-to-CPU session, and saves the most recent 32MB to PCAP file which can then be copied and analyzed. A common option to analyze packet captures is the free WireShark utility. When capturing a mirror-to-CPU session to a file, packets will not be dumped to the console. Otherwise you are looking at a summary of the captured packets.

If you chose the file option, the file is stored in flash. To delete it, execute this:

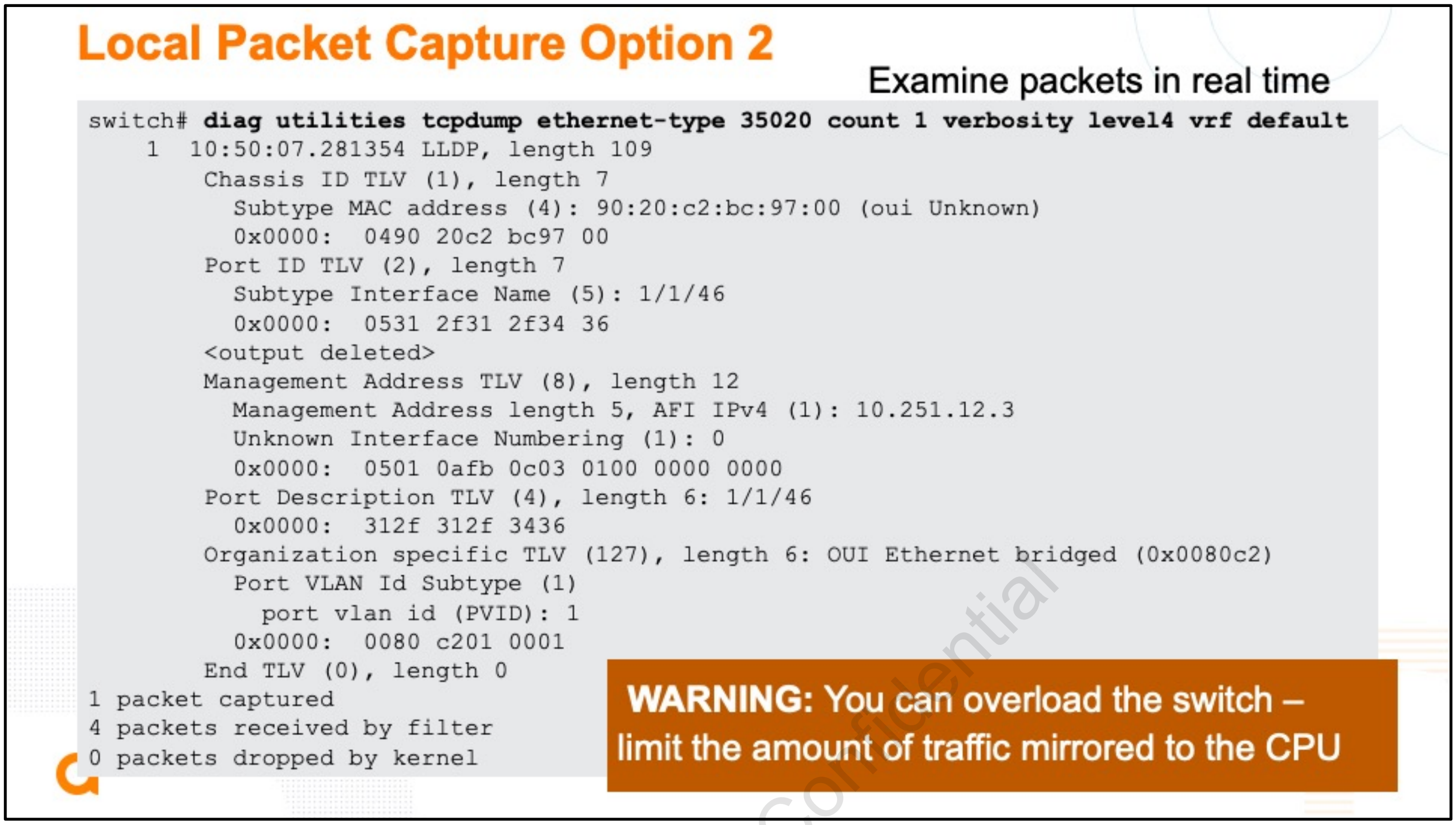
switch# **diagnostic utilities tshark delete**

To copy the file from the switch to a remote computer, use this command:

switch# **copy tshark-pcap <remote-URL> [vrf <VRF-name>]**

You can use TFTPor SFTP for the copy. Note: AOS-CX only supports one packet capture file. If you start a new diagnostic session, the old file is overwritten by a new one.

Aruba Training-Confidential



Local Packet Capture Option 2

Examine packets in real time

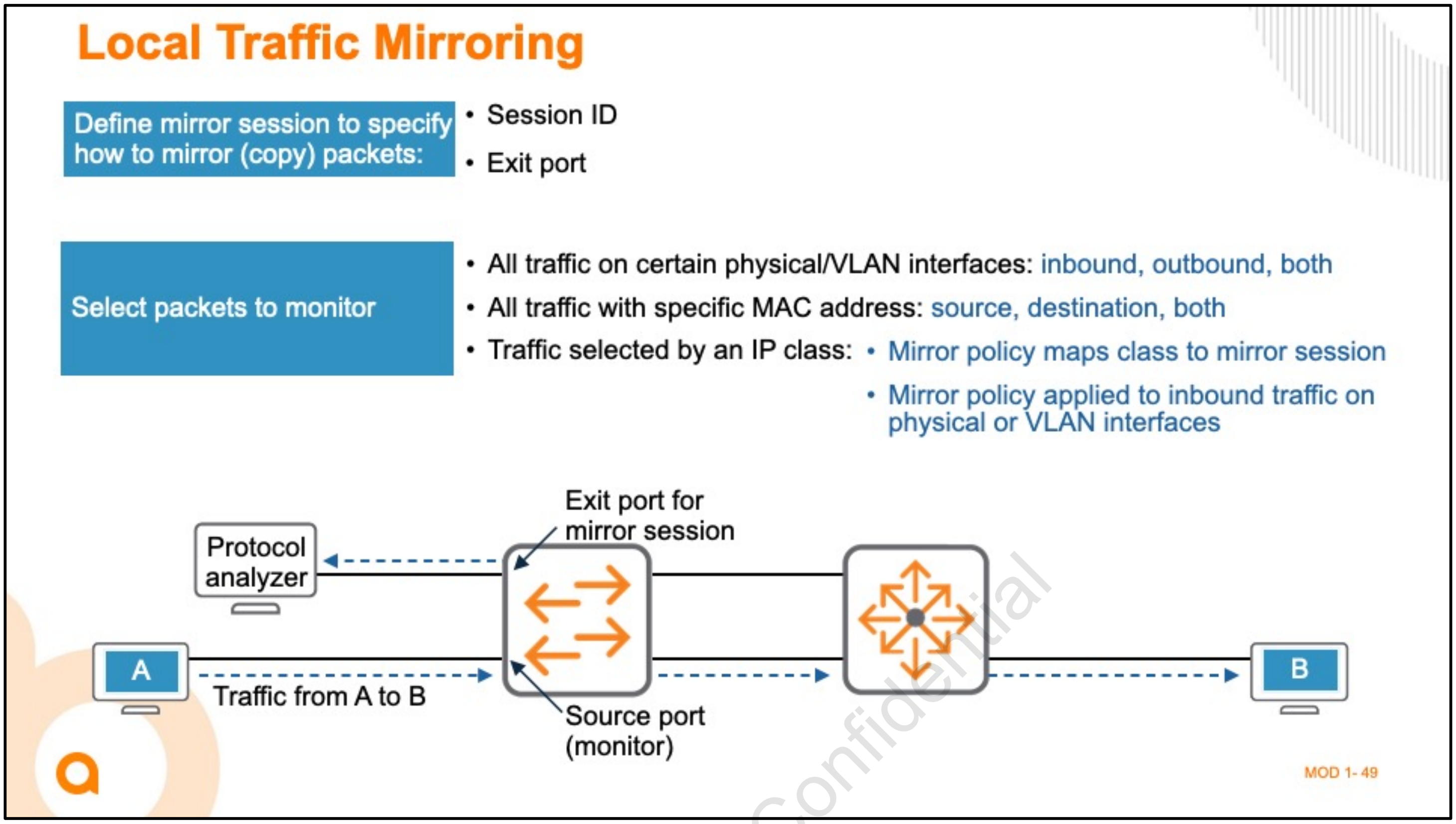
```
switch# diag utilities tcpdump ethernet-type 35020 count 1 verbosity level4 vrf default
    1  10:50:07.281354 LLDP, length 109
        Chassis ID TLV (1), length 7
          Subtype MAC address (4): 90:20:c2:bc:97:00 (oui Unknown)
          0x0000:  0490 20c2 bc97 00
        Port ID TLV (2), length 7
          Subtype Interface Name (5): 1/1/46
          0x0000:  0531 2f31 2f34 36
        <output deleted>
        Management Address TLV (8), length 12
          Management Address length 5, AFI IPv4 (1): 10.251.12.3
          Unknown Interface Numbering (1): 0
          0x0000:  0501 0afb 0c03 0100 0000 0000
        Port Description TLV (4), length 6: 1/1/46
          0x0000:  312f 312f 3436
        Organization specific TLV (127), length 6: OUI Ethernet bridged (0x0080c2)
          Port VLAN Id Subtype (1)
            port vlan id (PVID): 1
          0x0000:  0080 c201 0001
        End TLV (0), length 0
1 packet captured
4 packets received by filter
0 packets dropped by kernel
```

WARNING: You can overload the switch – limit the amount of traffic mirrored to the CPU

Another option is to look at a summary of the packets, in real-time, using Linux's tcpdump utility. Note that you do not have access to the Linux CLI, but the AOS-CX switch utilizes this utility from the AOS-CX CLI. This is accomplished using the **diag utilities tcpdump** command. Here's an example:

Warning: Any packet capturing, without the proper filtering, and displaying it to the CLI of the switch can be very CPU-intensive and affect the operation of the switch. Care should be taken when using the diag utilities tcpdump command.

Aruba Training-Confidential



Local Traffic Mirroring

Define mirror session to specify how to mirror (copy) packets:

- Session ID
- Exit port

Select packets to monitor

- All traffic on certain physical/VLAN interfaces: inbound, outbound, both
- All traffic with specific MAC address: source, destination, both
- Traffic selected by an IP class:
  - Mirror policy maps class to mirror session
  - Mirror policy applied to inbound traffic on physical or VLAN interfaces

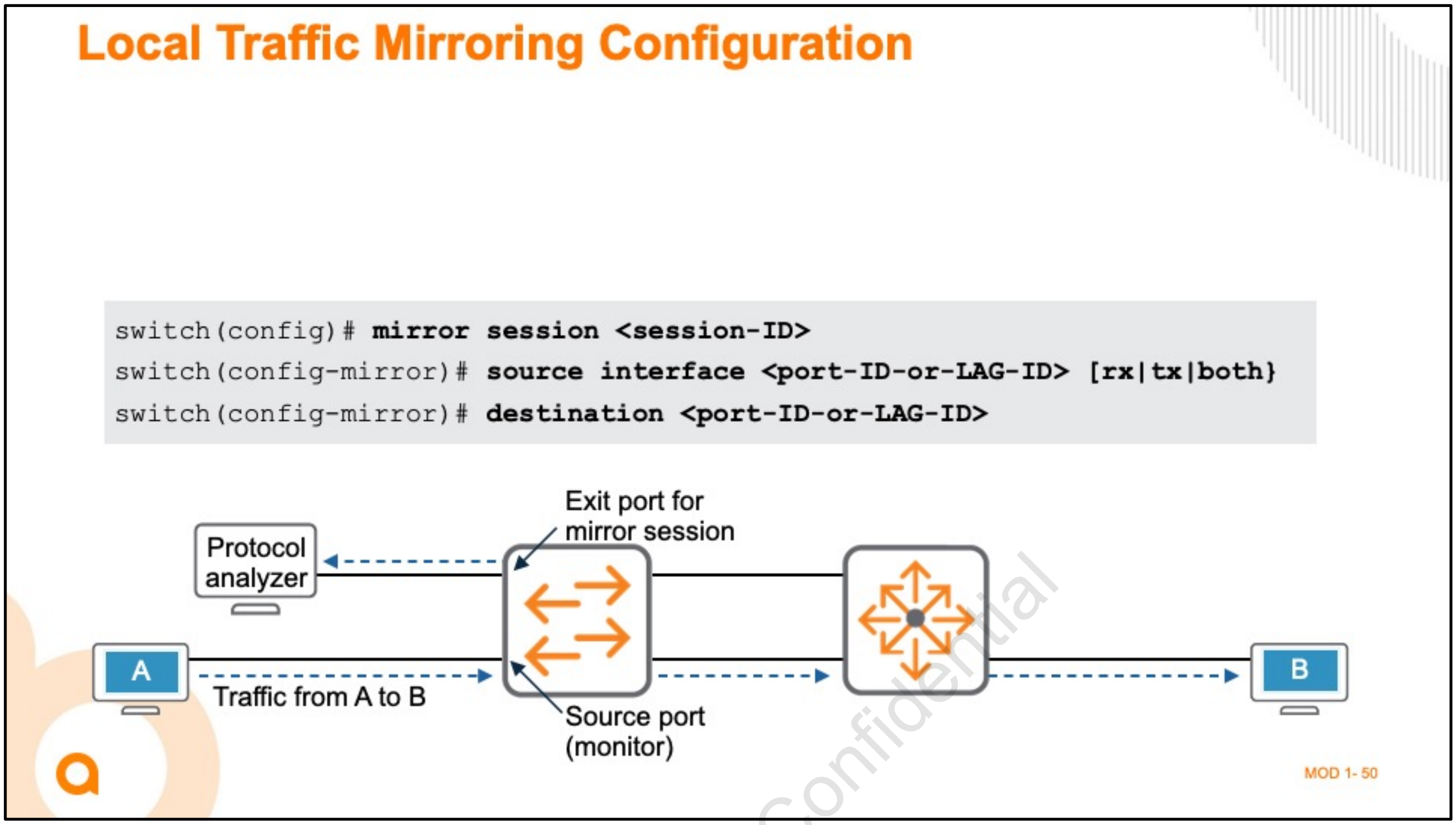
MOD 1- 49

Traffic mirroring allows a switch to copy frames that arrive or are sent from a particular port or VLAN and forward them to another port. You can use port mirroring to monitor network traffic to detect threats, troubleshoot problems, or manage the network. For example, you might want to send certain traffic to a security appliance such as an Intrusion Detection System (IDS) device, which can examine the traffic and detect possible threats. Or you may want to send traffic to a network protocol analyzer, which you can use to examine the traffic and troubleshoot a network problem.

Aruba Training-Confidential

The simplest form of traffic mirroring is local mirroring, in which the switch mirroring the packets sends the copies out a local port.

The mirroring feature (when mirroring received traffic) and the sFlow sampling feature both require the receive (rx) capability of a given port. If both features are configured and enabled to use the receive capability on the same port, only one of the features can perform its task. This interaction does not affect transmit (tx) mirroring because sFlow does not use the transmit (tx) capability of a port.



Local Traffic Mirroring Configuration

```
switch(config)# mirror session <session-ID>
switch(config-mirror)# source interface <port-ID-or-LAG-ID> [rx|tx|both}
switch(config-mirror)# destination <port-ID-or-LAG-ID>
```

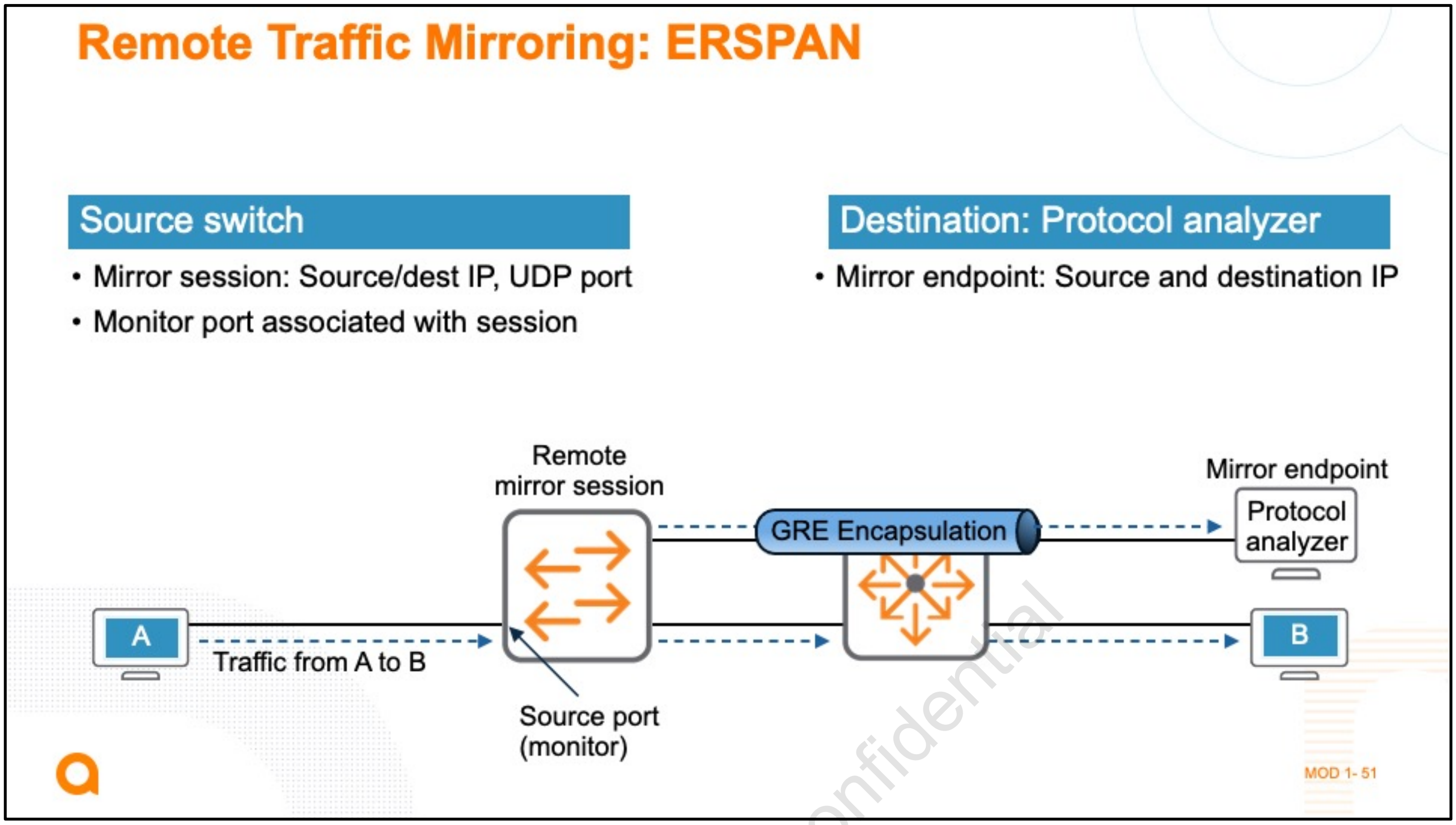
MOD 1- 50

Aruba Training-Confidential

The figure shows how to configure local traffic mirroring on AOS-CX switches. You first enter the mirror command, followed by the session number, and then specify the exit port in the interface context. The source and exit port can be a single physical interface or a link aggregation for local mirroring.

You can configure one or more mirror sessions to use the same exit port. Or you can specify different exit ports in different sessions to send mirrored traffic to different destination device or on different links to the same destination device. In the latter case, you'd want to make sure to set up different sources for each mirroring session. You can create up to four mirror sessions (1-4).

Use the **show mirror** command to verify your configuration.



Remote Traffic Mirroring: ERSPAN

Source switch

- Mirror session: Source/dest IP, UDP port
- Monitor port associated with session

Destination: Protocol analyzer

- Mirror endpoint: Source and destination IP

Aruba Training-Confidential

ERSPAN is an acronym that stands for *encapsulated remote switched port analyzer*. ERSPAN mirrors traffic on one or more source ports and forwards the mirrored traffic to a destination on a remote device. The traffic is encapsulated in generic routing encapsulation (GRE) and is, therefore, routable across a layer 3 network between the source switch and the destination device, like a packet sniffer, e.g. Wireshark.

AOS-CX switches also support remote port mirroring, in which the switch forwards the mirrored packets to a destination device. The switch achieves this by encapsulating the mirrored packets in a GRE header that uses the remote switch's IP address as the destination. The remote, or destination, switch is configured to decapsulate traffic from this GRE tunnel and to forward the traffic out an exit port (see the previous slide on setting up local mirroring).

This command specifies the tunnel to where all mirrored traffic for the session will be transmitted. Only one tunnel destination is allowed per session.

Switch(config)# **mirror session** <session-ID> {**rx** | **tx** | **both**}

Switch(config-mirror)# **source interface** <port-ID-or-LAG-ID>

Switch(config-mirror)# **destination tunnel** <destination-IP-address> **source** <source-IP-address>
[**dscp** <dscp-code>] [**vrf** <VRF-name>]

You may configure multiple mirror sessions with the same source/destination IP address pair, however, only one of those sessions sharing the same source/destination IP address pair can be enabled at a given time.

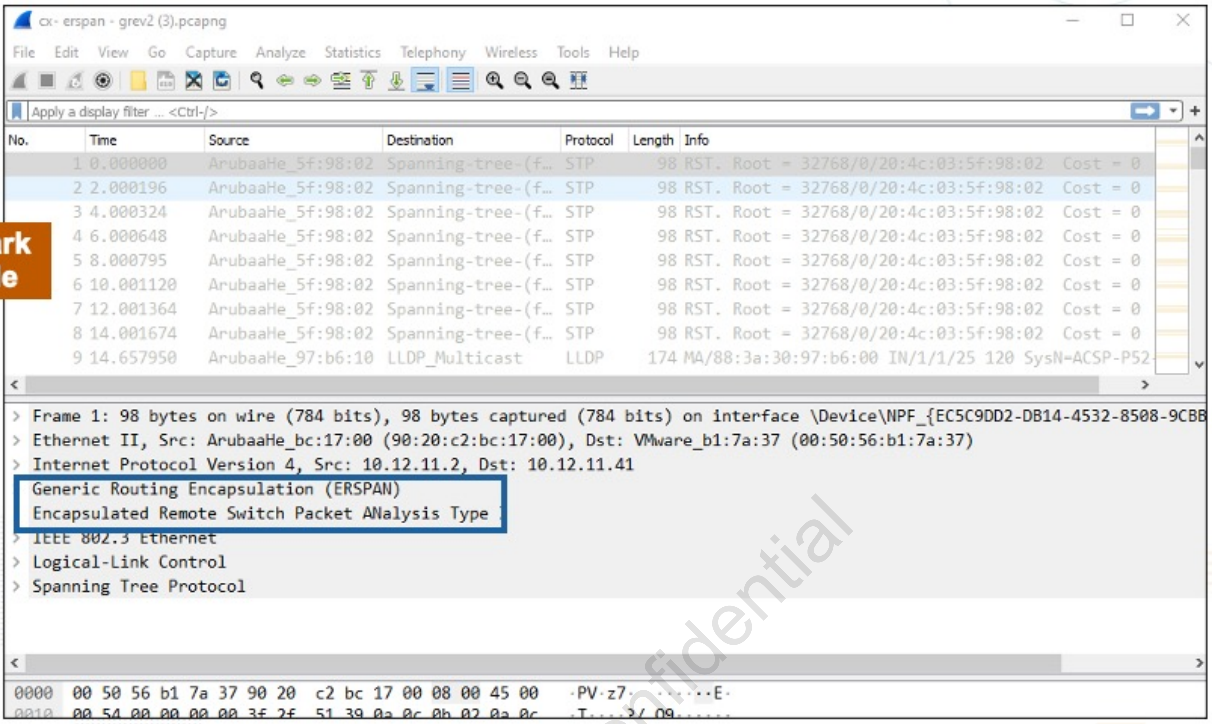
ERSPAN is not supported leaving the switch by the OOB port. If VRF management is configured for an ERSPAN session, the session will be in "mirror_err_tunnel_ oob_port_not_supp orted" operation status. ERSPAN is not supported leaving the switch encapsulated within another tunnel (e.g. GRE IPv4). When the path to the destination IP address will leave via a tunnel, the session will be in "tunnel_route_resol ution_not_populate d" operation status.

Note: AOS switches support mirroring to other AOS switches. AOS-CX switches, however, do not support this feature. Instead, the remote mirroring must be to a device that supports it, like Wireshark.

Aruba Training-Confidential

ERSPAN Output

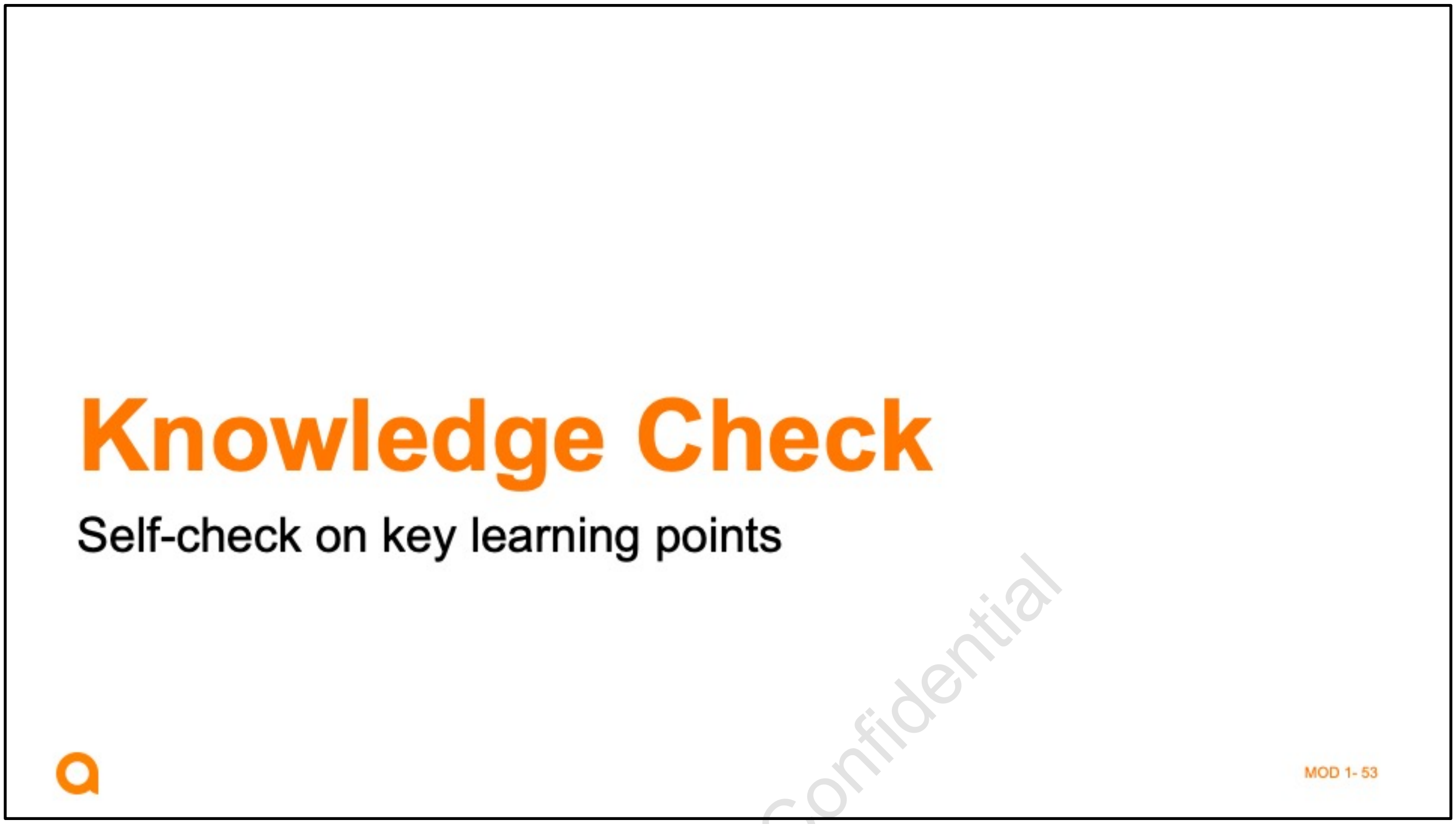
Wireshark example


cx- erspan - grev2 (3).pcapng
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>

```
No.  Time       Source             Destination          Protocol Length Info
   1 0.000000   ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   2 2.000196   ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   3 4.000324   ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   4 6.000648   ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   5 8.000795   ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   6 10.001120  ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   7 12.001364  ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   8 14.001674  ArubaaHe_5f:98:02  Spanning-tree-(f…    STP       98 RST. Root = 32768/0/20:4c:03:5f:98:02  Cost = 0
   9 14.657950  ArubaaHe_97:b6:10  LLDP_Multicast       LLDP     174 MA/88:3a:30:97:b6:00 IN/1/1/25 120 SysN=ACSP-P52

> Frame 1: 98 bytes on wire (784 bits), 98 bytes captured (784 bits) on interface \Device\NPF_{EC5C9DD2-DB14-4532-8508-9CBB
> Ethernet II, Src: ArubaaHe_bc:17:00 (90:20:c2:bc:17:00), Dst: VMware_b1:7a:37 (00:50:56:b1:7a:37)
> Internet Protocol Version 4, Src: 10.12.11.2, Dst: 10.12.11.41
  Generic Routing Encapsulation (ERSPAN)
  Encapsulated Remote Switch Packet ANalysis Type
> IEEE 802.3 Ethernet
> Logical-Link Control
> Spanning Tree Protocol

0000  00 50 56 b1 7a 37 90 20  c2 bc 17 00 08 00 45 00   ·PV·z7·  ······E·
0010  00 54 00 00 00 00 3f 2f  51 39 0a 0c 0b 02 0a 0c   ·T····?/ Q9······
```

Here's an example where a remote mirroring session was set up on an AOS-CX switch to forward packets to Wireshark on a remote desktop. As you can see from the highlighted part of the graphic, the remote desktop is receiving the encapsulated (tunneled traffic) denoted by the ERSPAN packets.

Aruba Training-Confidential

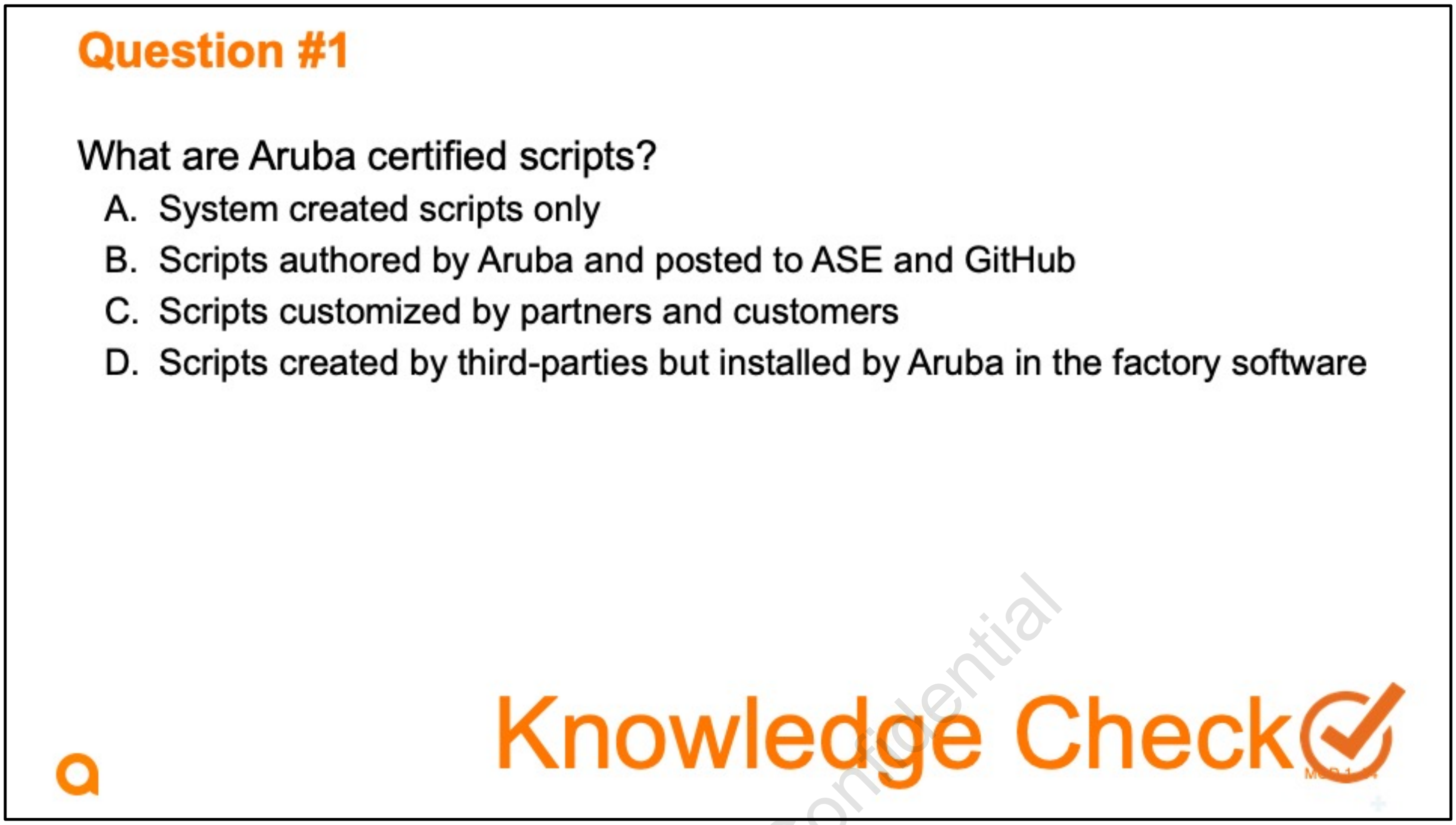


Knowledge Check

Self-check on key learning points

MOD 1- 53

Aruba Training-Confidential



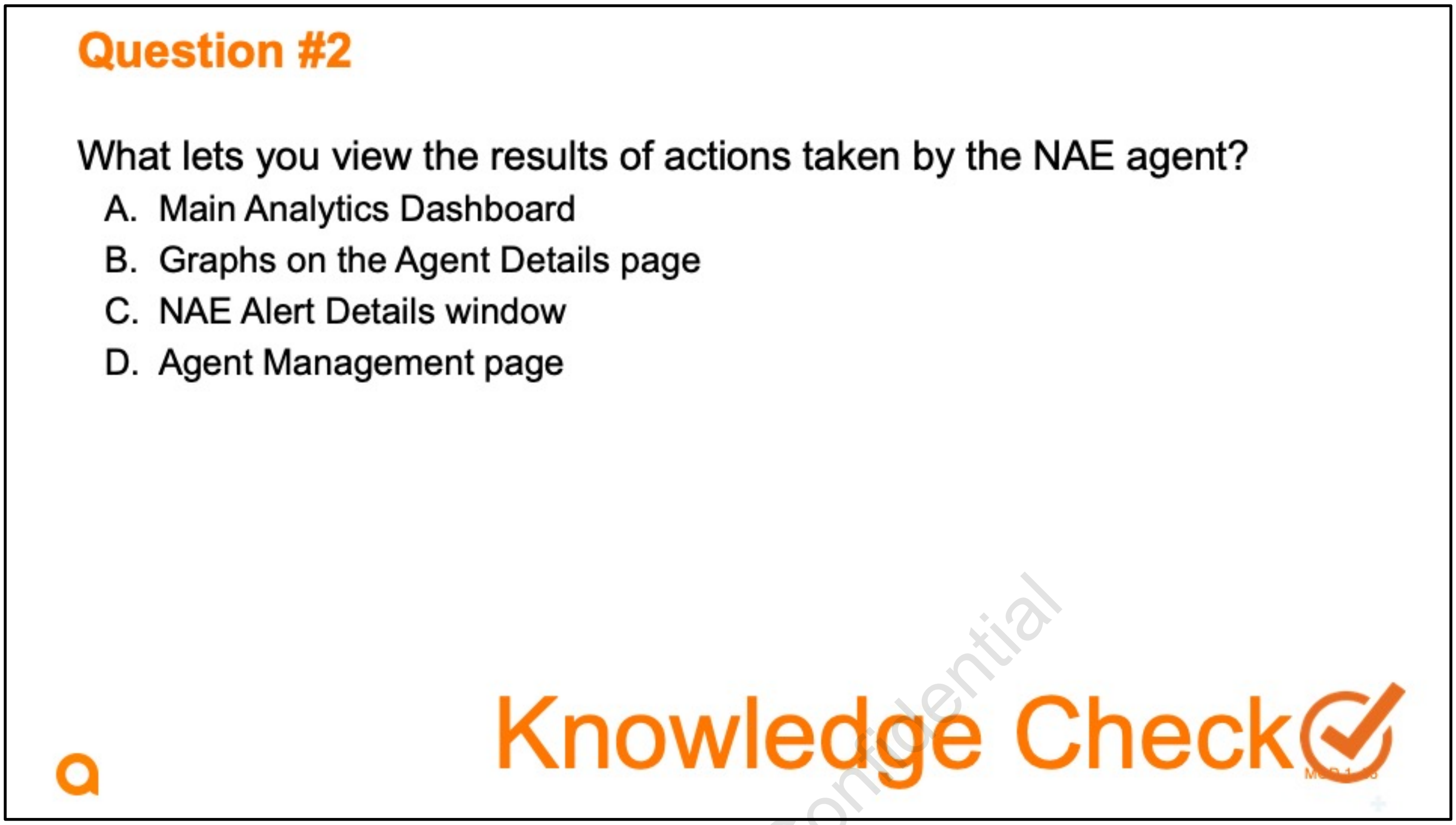
Question #1

What are Aruba certified scripts?

A. System created scripts only
B. Scripts authored by Aruba and posted to ASE and GitHub
C. Scripts customized by partners and customers
D. Scripts created by third-parties but installed by Aruba in the factory software

Knowledge Check

Aruba Training-Confidential



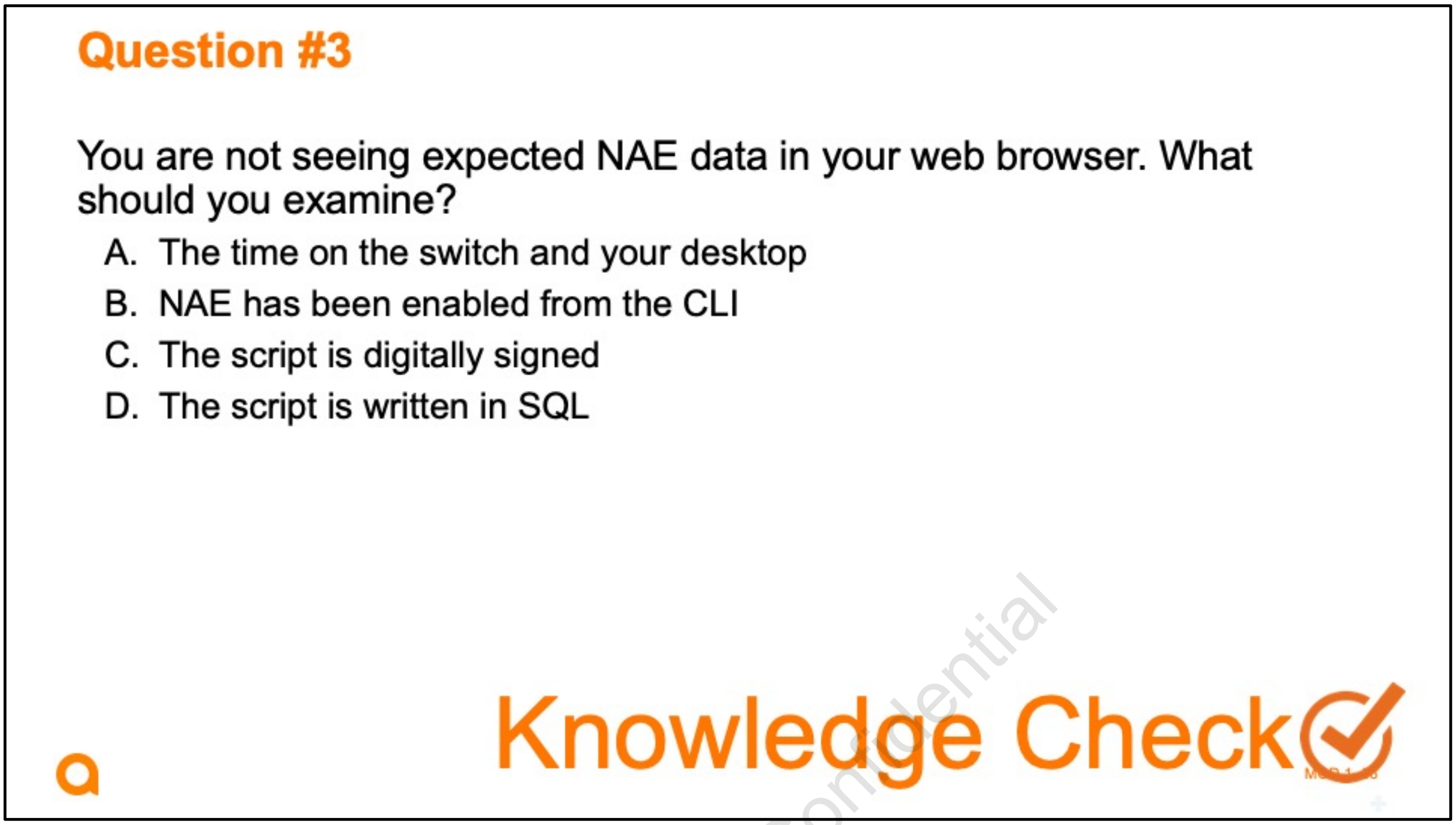
Question #2

What lets you view the results of actions taken by the NAE agent?

A. Main Analytics Dashboard
B. Graphs on the Agent Details page
C. NAE Alert Details window
D. Agent Management page

Knowledge Check

Aruba Training-Confidential



Question #3

You are not seeing expected NAE data in your web browser. What should you examine?

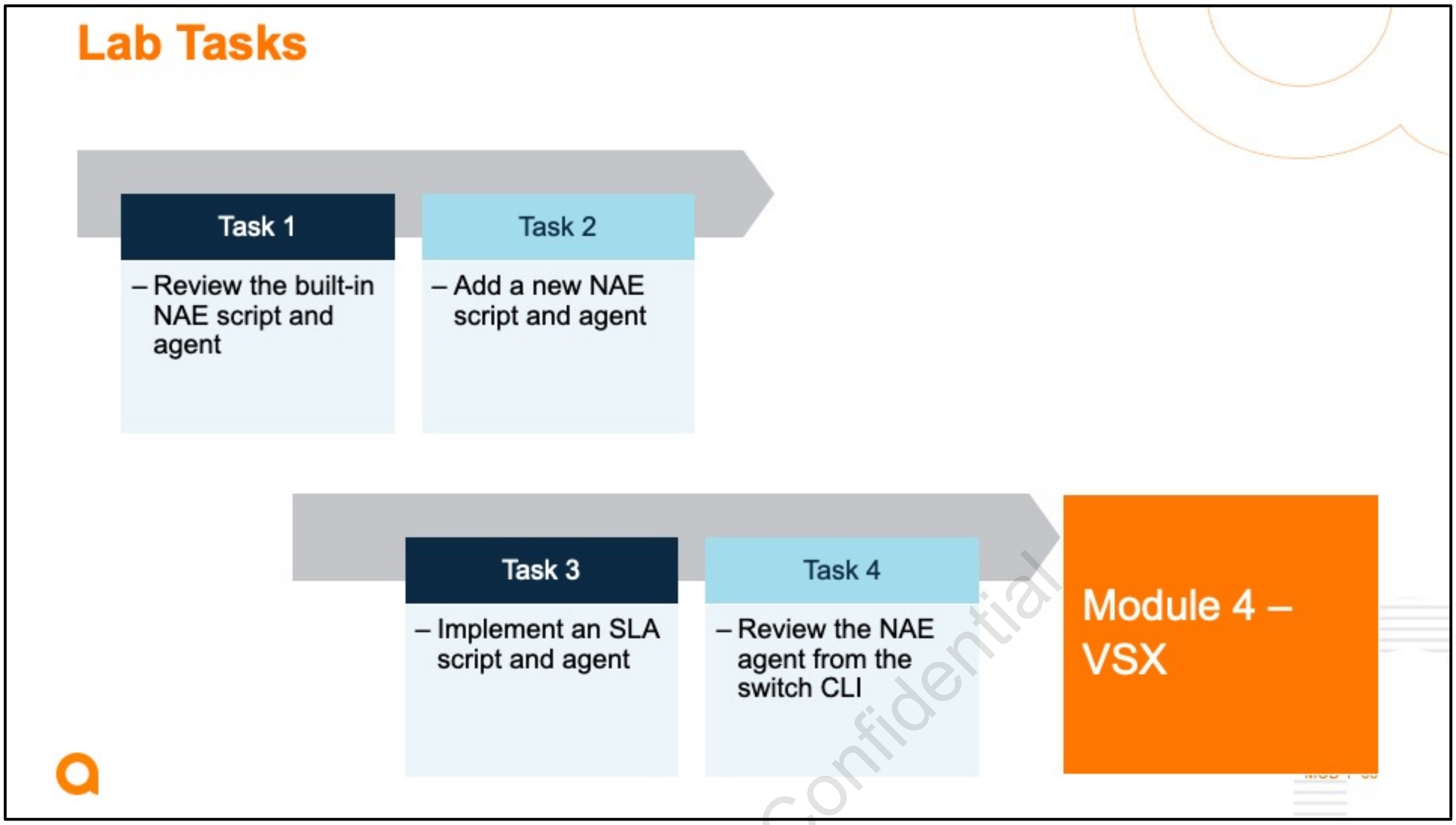
A. The time on the switch and your desktop
B. NAE has been enabled from the CLI
C. The script is digitally signed
D. The script is written in SQL

Knowledge Check

Aruba Training-Confidential

Lab Activity

Aruba Training-O



Lab Tasks

Task 1

– Review the built-in NAE script and agent

Task 2

– Add a new NAE script and agent

Task 3

– Implement an SLA script and agent

Task 4

– Review the NAE agent from the switch CLI

Module 4 – VSX

Aruba Training-Confidential

The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with Module 4 – VSX.

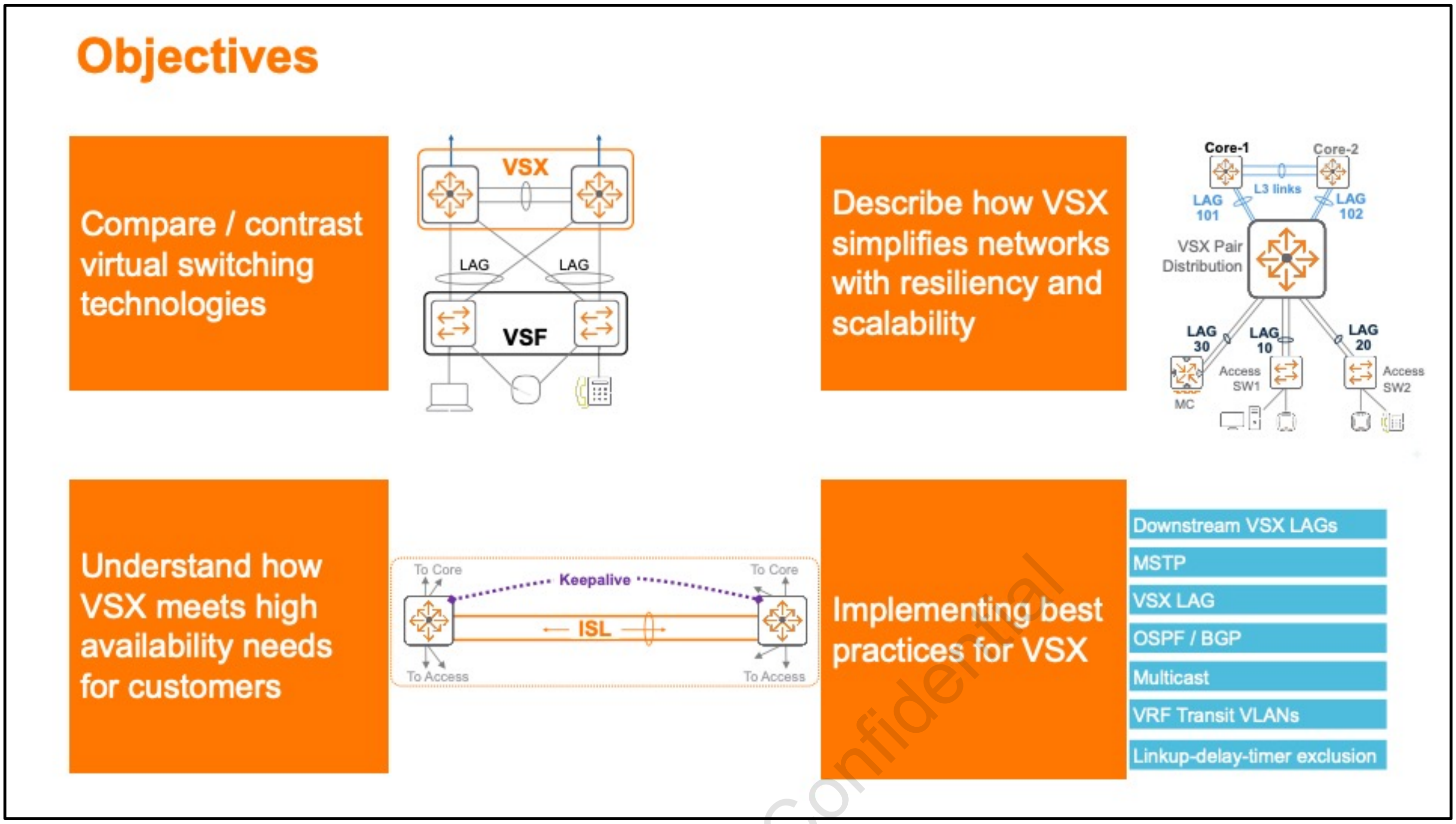
Rev # 20.21

VSX

Implementing ArubaOS-CX Switching

Hi everyone, and welcome back. This module is about the Virtual Switching eXtension (VSX) feature.

Aruba Training-Conf



Objectives

Compare / contrast virtual switching technologies

Describe how VSX simplifies networks with resiliency and scalability

Understand how VSX meets high availability needs for customers

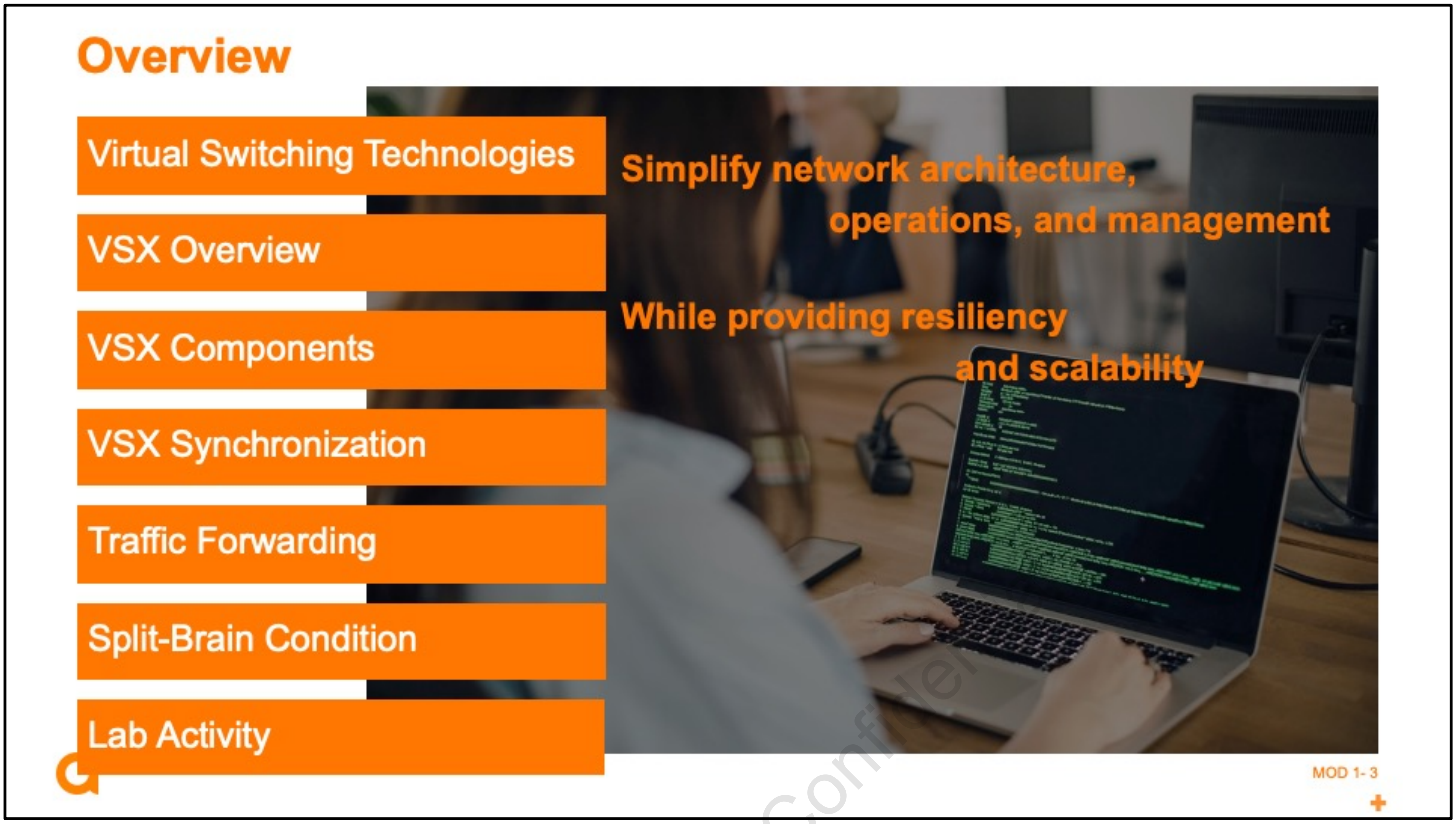
Implementing best practices for VSX

- Downstream VSX LAGs
- MSTP
- VSX LAG
- OSPF / BGP
- Multicast
- VRF Transit VLANs
- Linkup-delay-timer exclusion

Aruba Training-Confidential

After completing this module, you should be able to:

- Compare and contrast virtual switching technologies
- Understand how Aruba’s VSX technology helps you simplify your network architecture, operations, and management while providing a resiliency and scalability.
- Understand how VSX meets high availability needs for customers.
- Implementing best practices for VSX



Overview

- Virtual Switching Technologies
- VSX Overview
- VSX Components
- VSX Synchronization
- Traffic Forwarding
- Split-Brain Condition
- Lab Activity

Simplify network architecture, operations, and management

While providing resiliency and scalability

MOD 1- 3

This module focuses on high availability technologies that help you simplify your network architecture, operations, and management while providing a resiliency and scalability. You'll learn how to use Virtual Switching Extension (VSX) and the benefits of using VSX to virtualization the aggregation/core of the campus network. VSX provides advanced capabilities beyond what was previously available with the VSX-LAG feature. You'll learn how VSX meets high availability needs of the customers to provide redundancy, resiliency, performance, and simplicity.

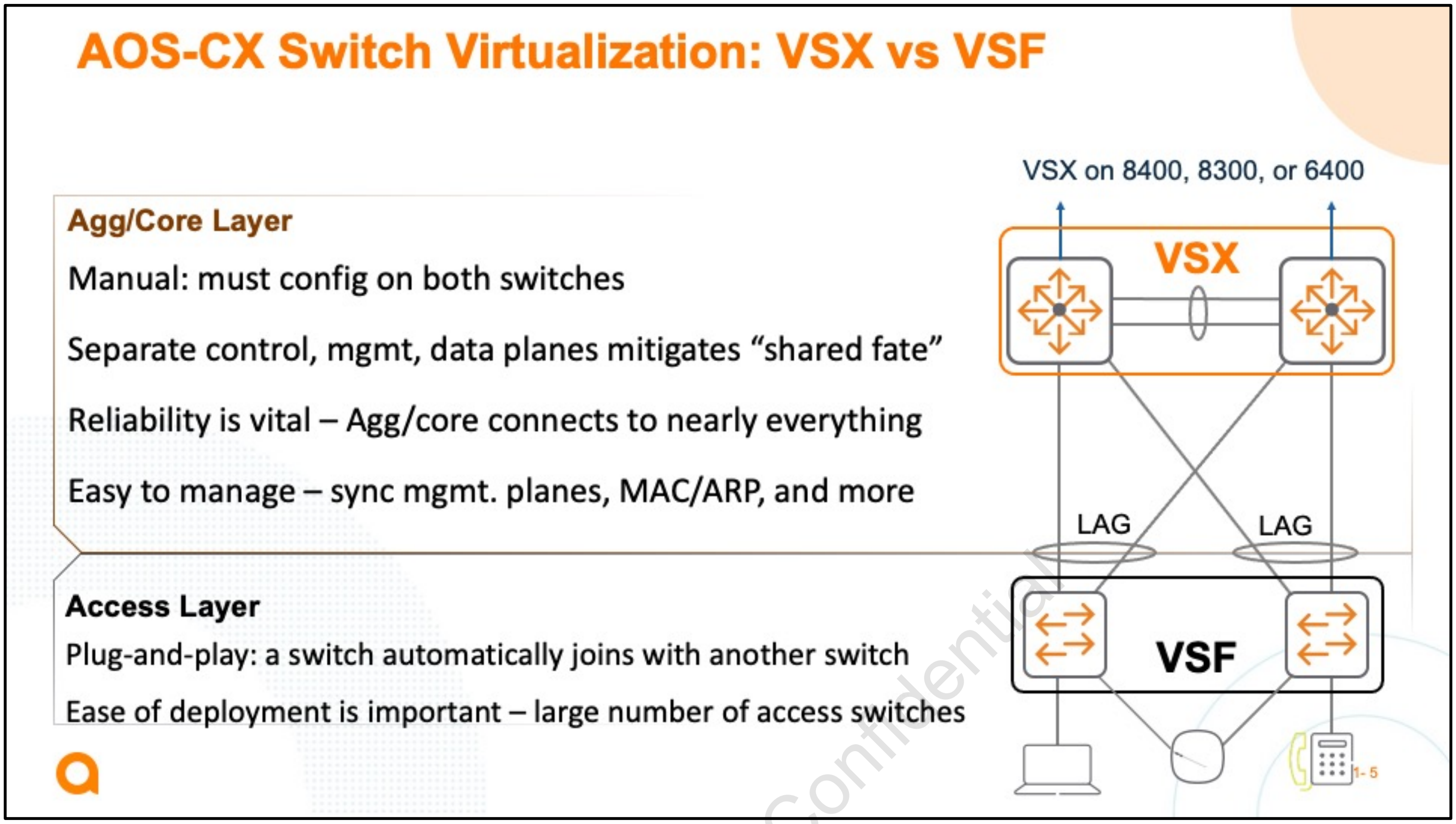
Aruba Training-Confidential

Virtual Switching Technologies

aruba
a Hewlett Packard Enterprise company

This section introduces the virtual switching technologies that Aruba has designed for the ArubaOS and AOS-CX switches.

Aruba Training-Confidential



AOS-CX Switch Virtualization: VSX vs VSF

Agg/Core Layer

Manual: must config on both switches

Separate control, mgmt, data planes mitigates “shared fate”

Reliability is vital – Agg/core connects to nearly everything

Easy to manage – sync mgmt. planes, MAC/ARP, and more

Access Layer

Plug-and-play: a switch automatically joins with another switch

Ease of deployment is important – large number of access switches

Aruba Training-Confidential

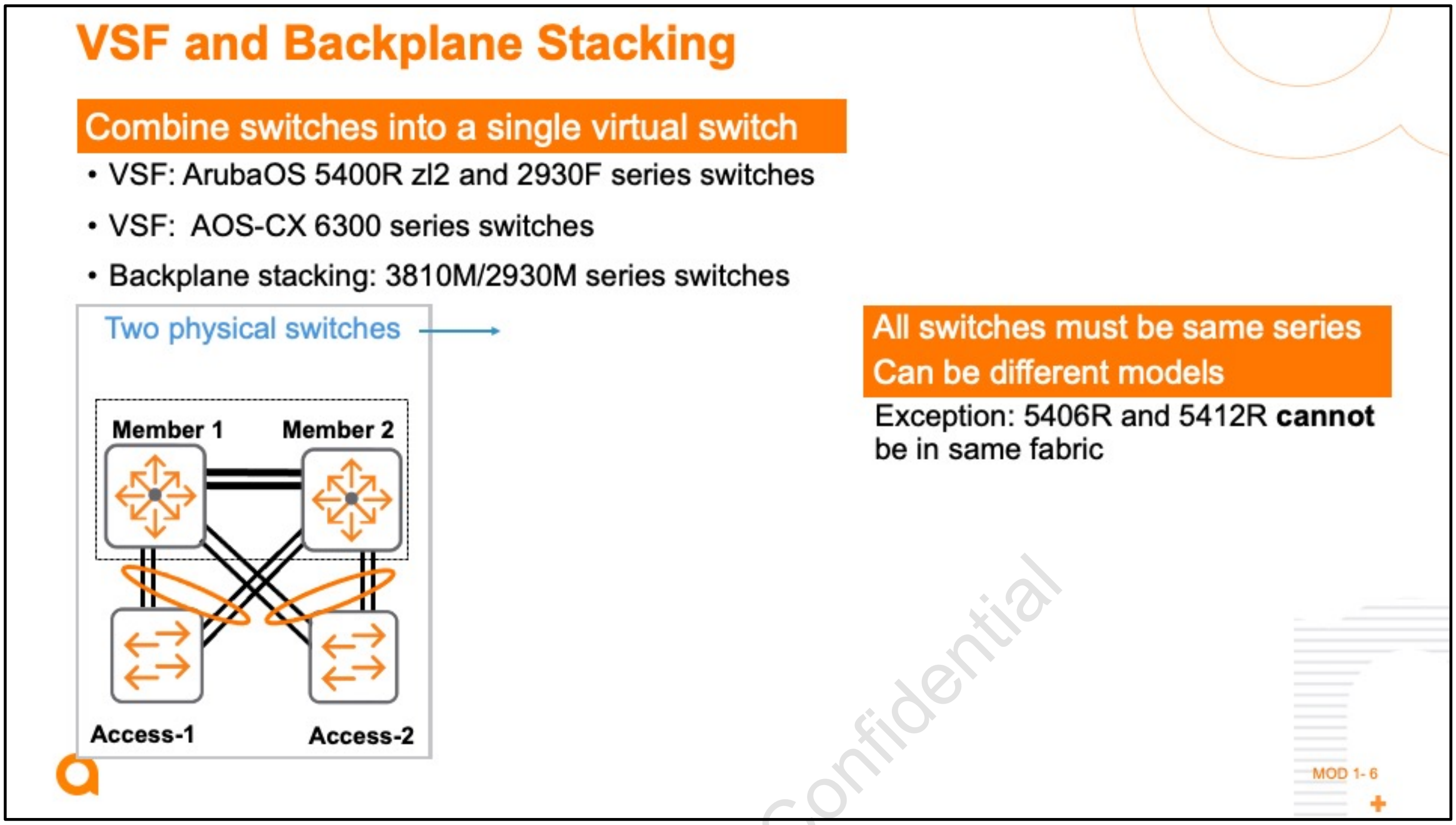
VSX and VSF are similar to each other, and even to earlier virtualization techniques like Virtual Router Redundancy Protocol (VRRP) – multiple devices are perceived as a single device. This improves resiliency while reduces configuration and management overhead. But there are distinct differences between VSX and VSF, as summarized in the figure.

You use **VSX** at the campus network aggregation and core layers. You use **VSF** primarily at the access layer. The differences in features and function are intentional – because priorities differ between access and core.

Configuring VSF and backplane stacking is a plug-and-play affair - a switch, at default settings, automatically unites with another switch. This ease of deployment is especially important at the access layer – you may have dozens, or hundreds of access switches.

You must manually configure VSX on both switches. VSX maintains separate control plane and management planes, and much of the data plane. This mitigates the “shared fate” nature of common planes, for improved uptimes – you get high availability and redundancy. This is especially important when aggregating all those access layer switches. And it remains easy to manage, because you can synchronize the management planes, and optionally, other features - MAC and ARP addresses, and more.

Let’s zoom in on VSF.



VSF and Backplane Stacking

Combine switches into a single virtual switch

- VSF: ArubaOS 5400R zl2 and 2930F series switches
- VSF: AOS-CX 6300 series switches
- Backplane stacking: 3810M/2930M series switches

All switches must be same series
Can be different models

Exception: 5406R and 5412R **cannot** be in same fabric

MOD 1- 6

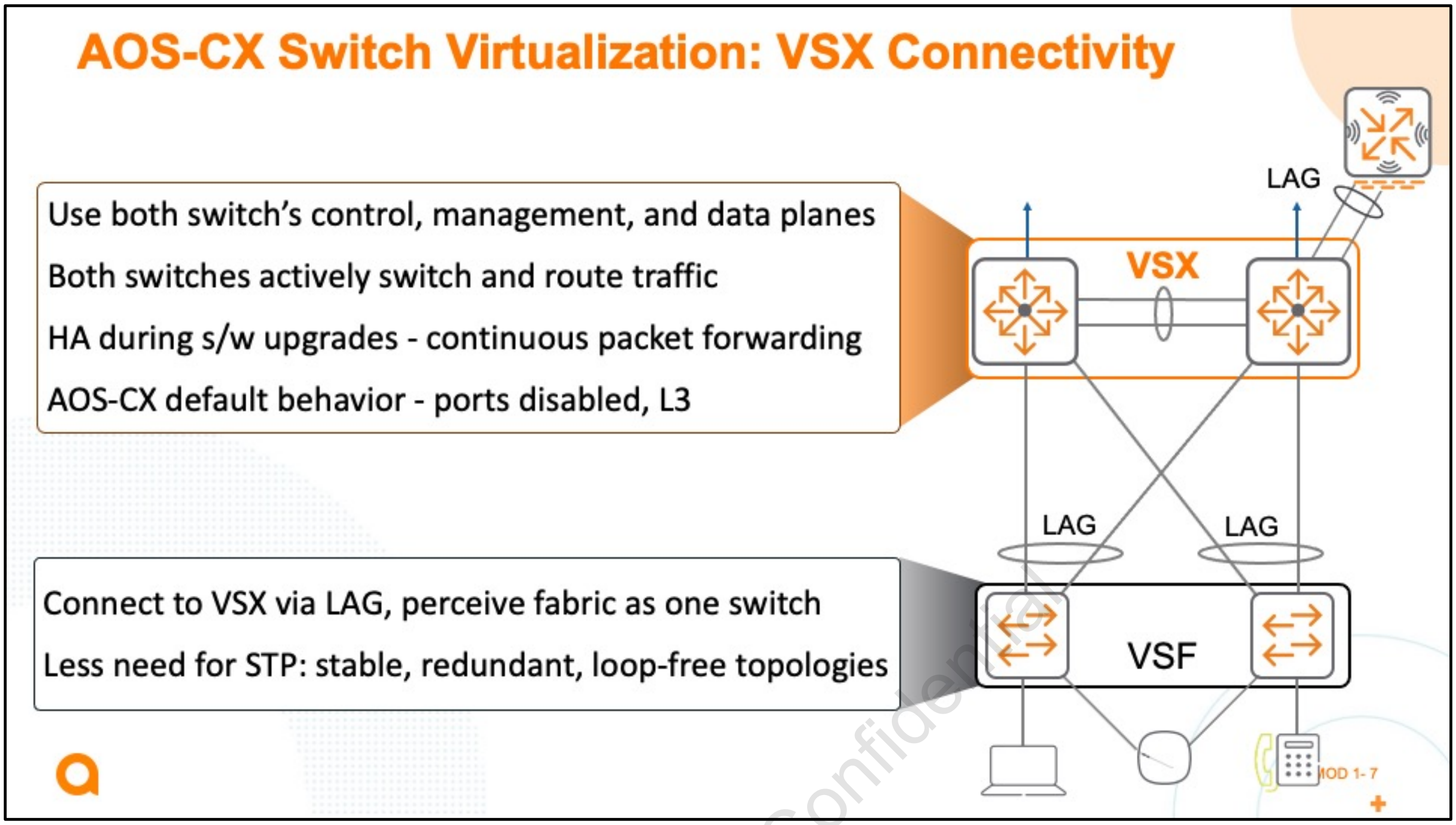
Virtual Switching Framework (VSF) and backplane stacking are Aruba technologies that combine multiple switches into a single virtual switch. You manage the combined switches as a single entity. You also treat switches as a single entity from the viewpoint of designing Layer 2 and 3 protocols. This provides significant advantages for a wired network design.

| ArubaOS 5400R zl2 and 2930F series switches and AOS-CX 6300 series switches support VSF. ArubaOS 3810M and 2930M switches support backplane stacking.

A VSF fabric must consist of switches that are the same series, such as two Aruba 5406R switches or two Aruba 2930F switches. For the 2930F series, you can combine different models in the same fabric. For example, you can combine a 2930F-48G PoE+ switch with a 2930F-48G switch.

| You cannot combine 5406R and 5412R switches in the same VSF fabric. The two switches can support different interface modules, although Aruba recommends that you install the same modules in both switches for consistency.

Aruba Training-Confidential



AOS-CX Switch Virtualization: VSX Connectivity

Use both switch's control, management, and data planes

Both switches actively switch and route traffic

HA during s/w upgrades - continuous packet forwarding

AOS-CX default behavior - ports disabled, L3

Connect to VSX via LAG, perceive fabric as one switch

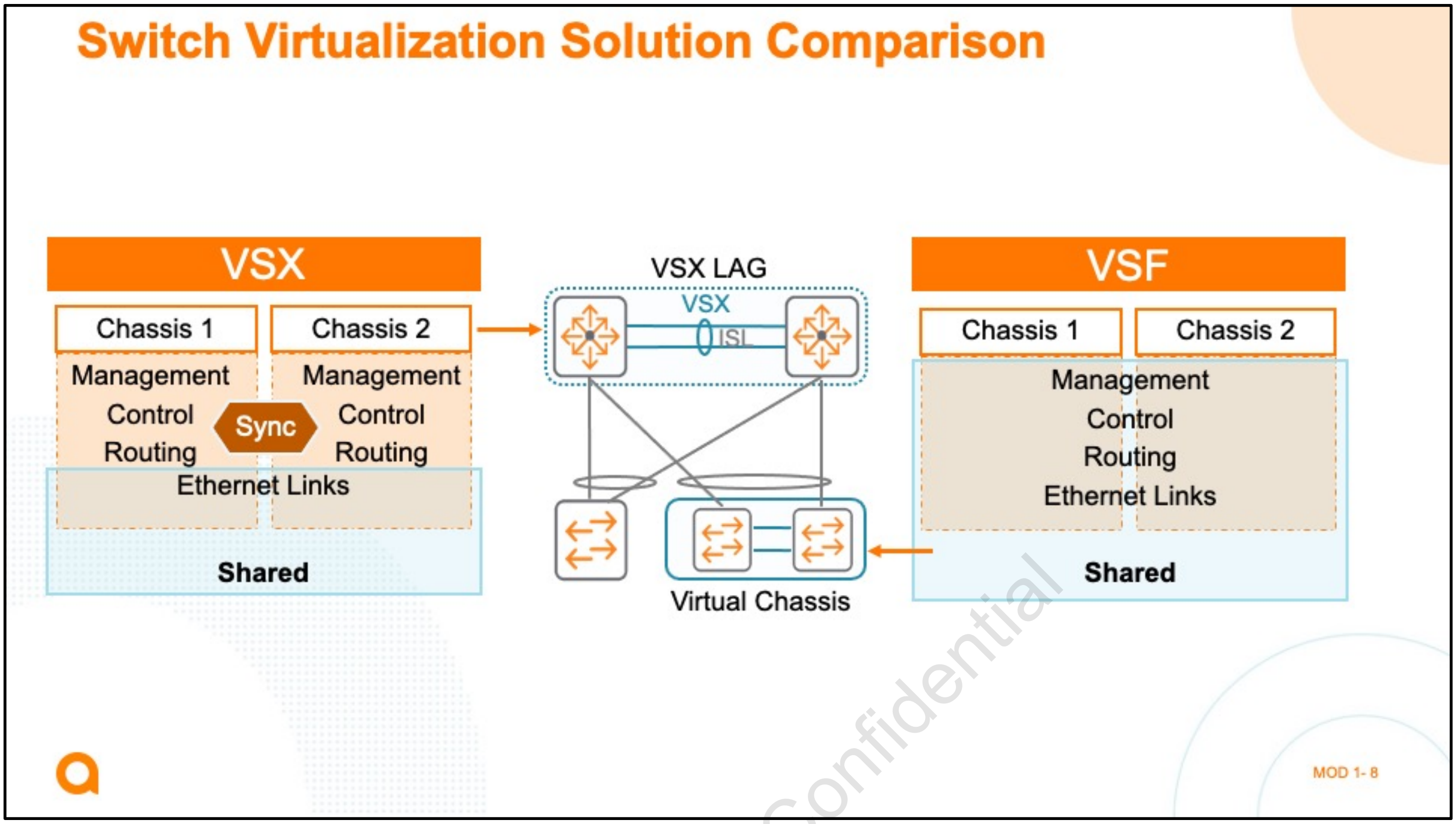
Less need for STP: stable, redundant, loop-free topologies

Aruba Training-Confidential

The VSX fabric leverages both switch's control, management, and data planes, while appearing to other devices as a single switch. They support an active-active path for Layer 2 traffic, and for Layer 3 unicast and multicast traffic. In other words, both switches actively switch and route traffic.

VSX maintains the AOS-CX default behavior - ports are disabled and operate at Layer 3. Finally, VSX delivers high availability during software upgrades, with near zero downtime and continuous packet forwarding.

| Access switches and mobility controllers connect to core switches via LAG connections. If configured for LACP, they perceive the VSX fabric as a single switch. This mitigates reliance on STP to ensure stable, redundant, loop-free topologies.



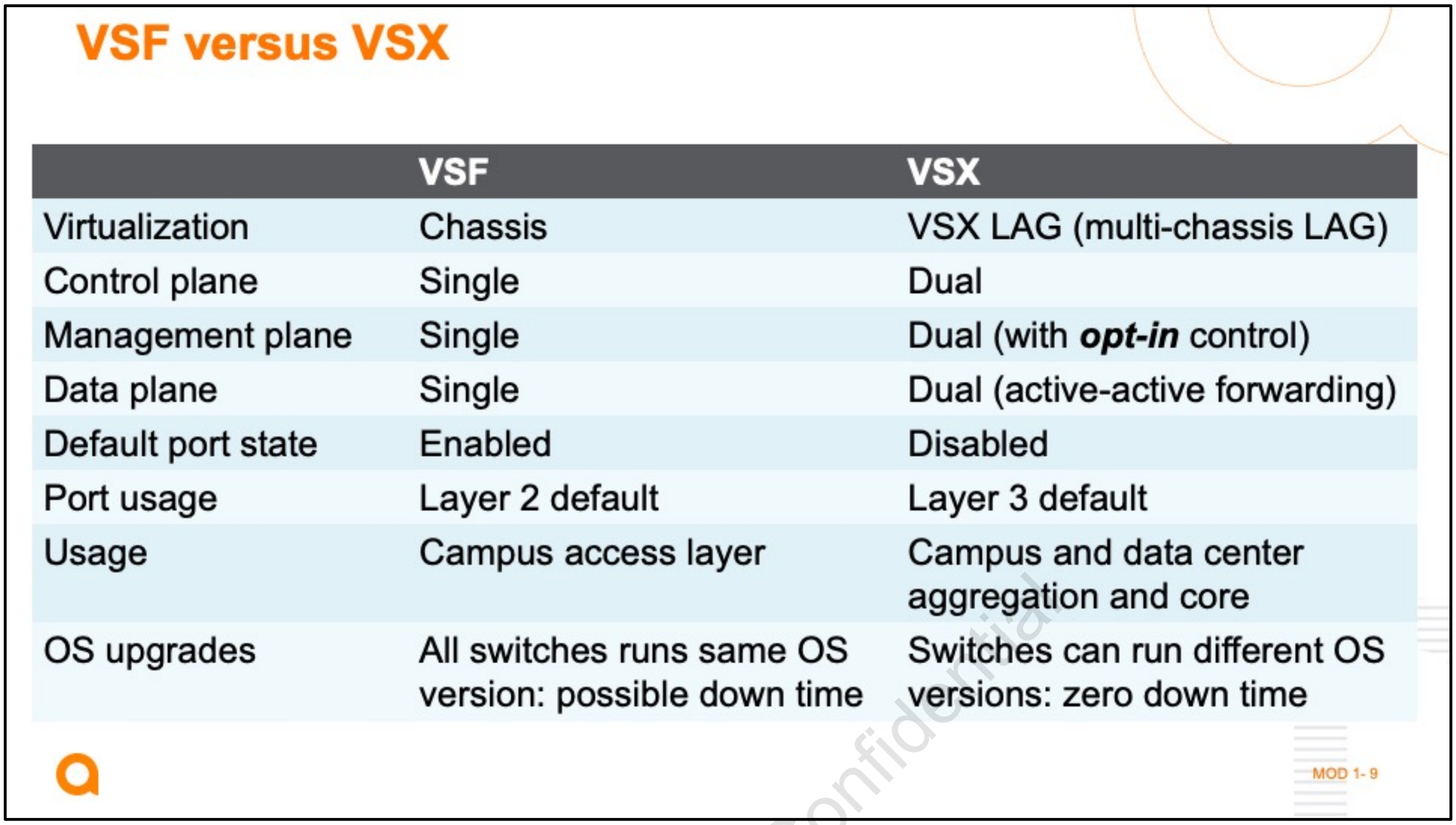
Switch Virtualization Solution Comparison

VSX virtualizes the control plane of two switches to function as one device at Layer 2 and as independent devices at Layer 3. From a datapath perspective, each device does an independent forwarding lookup to decide how to handle the traffic. Some of the forwarding databases, such as the MAC forwarding database and neighbor tables, are synchronized between the two devices using a proprietary VSX control plane. Some of the forwarding databases are built independently by each switch. Configuration synchronization is one aspect of this VSX solution where the primary switch configuration is synced to the secondary switch.

VSF allows multiple ArubaOS switches of the same model to act as a single virtual device. VSF helps to provide fast failover, scalability, manageability, and high availability.

A VSF fabric has one active management plane and one active control plane, which the management module in the commander switch provides. The VSF commander proxies the control plane to the standby member. Both members of the fabric participate in the forwarding plane. Interface modules from both members are combined as if over one large switch connected by the VSF link.

Aruba Training-Confidential



VSF versus VSX

	VSF	VSX
Virtualization	Chassis	VSX LAG (multi-chassis LAG)
Control plane	Single	Dual
Management plane	Single	Dual (with ***opt-in*** control)
Data plane	Single	Dual (active-active forwarding)
Default port state	Enabled	Disabled
Port usage	Layer 2 default	Layer 3 default
Usage	Campus access layer	Campus and data center aggregation and core
OS upgrades	All switches runs same OS version: possible down time	Switches can run different OS versions: zero down time

MOD 1- 9

The figure summarizes the differences between VSF and VSX, as a study and review aid.

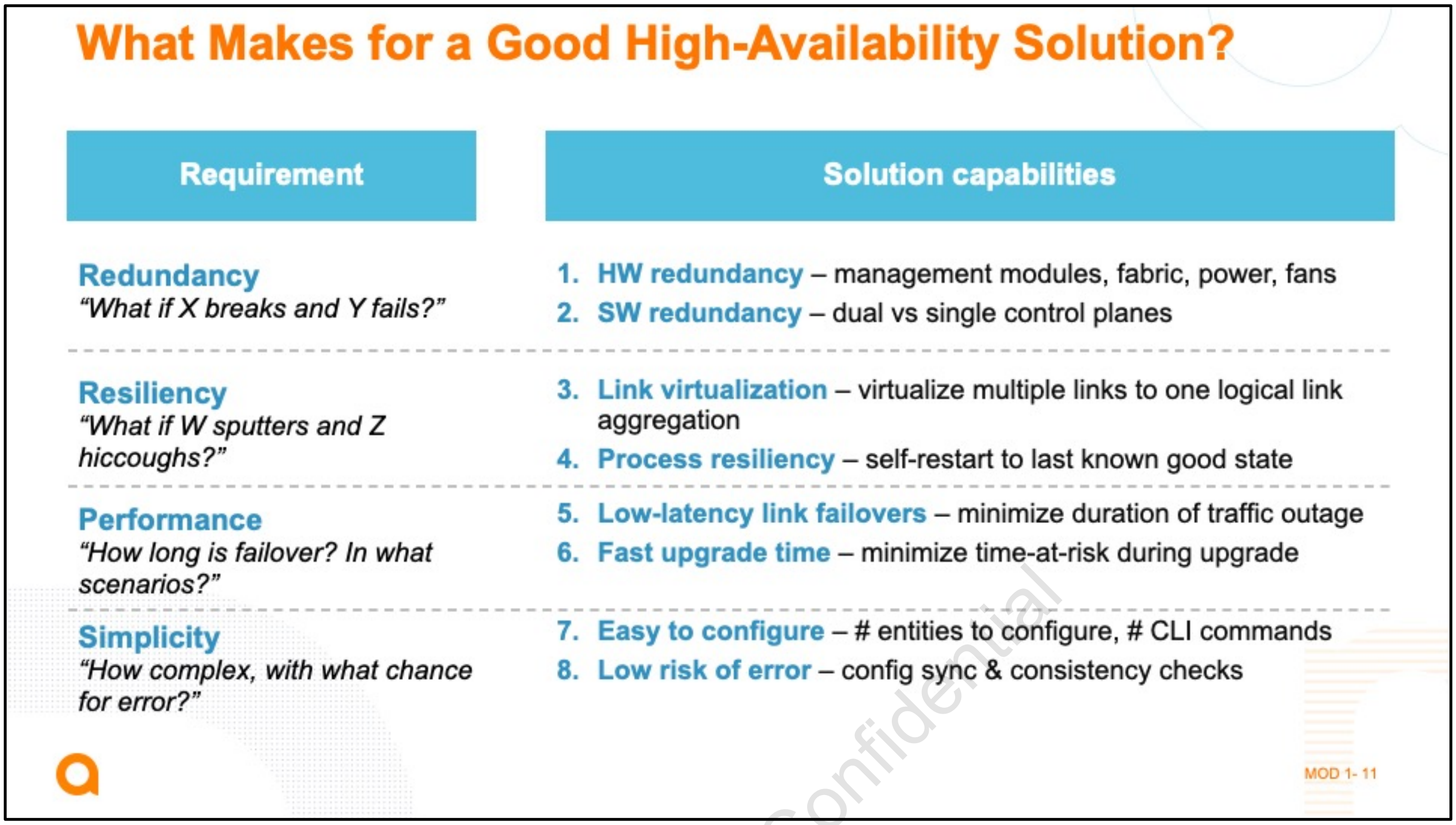
Aruba Training-Confidential

VSX Overview

MOD 1- 10

This section provides a brief overview of AOS-CX’s VSX technology.

Aruba Training-Confidential



What Makes for a Good High-Availability Solution?

Requirement	Solution capabilities
Redundancy *"What if X breaks and Y fails?"*	1. **HW redundancy** – management modules, fabric, power, fans 2. **SW redundancy** – dual vs single control planes
Resiliency *"What if W sputters and Z hiccoughs?"*	3. **Link virtualization** – virtualize multiple links to one logical link aggregation 4. **Process resiliency** – self-restart to last known good state
Performance *"How long is failover? In what scenarios?"*	5. **Low-latency link failovers** – minimize duration of traffic outage 6. **Fast upgrade time** – minimize time-at-risk during upgrade
Simplicity *"How complex, with what chance for error?"*	7. **Easy to configure** – # entities to configure, # CLI commands 8. **Low risk of error** – config sync & consistency checks

MOD 1- 11

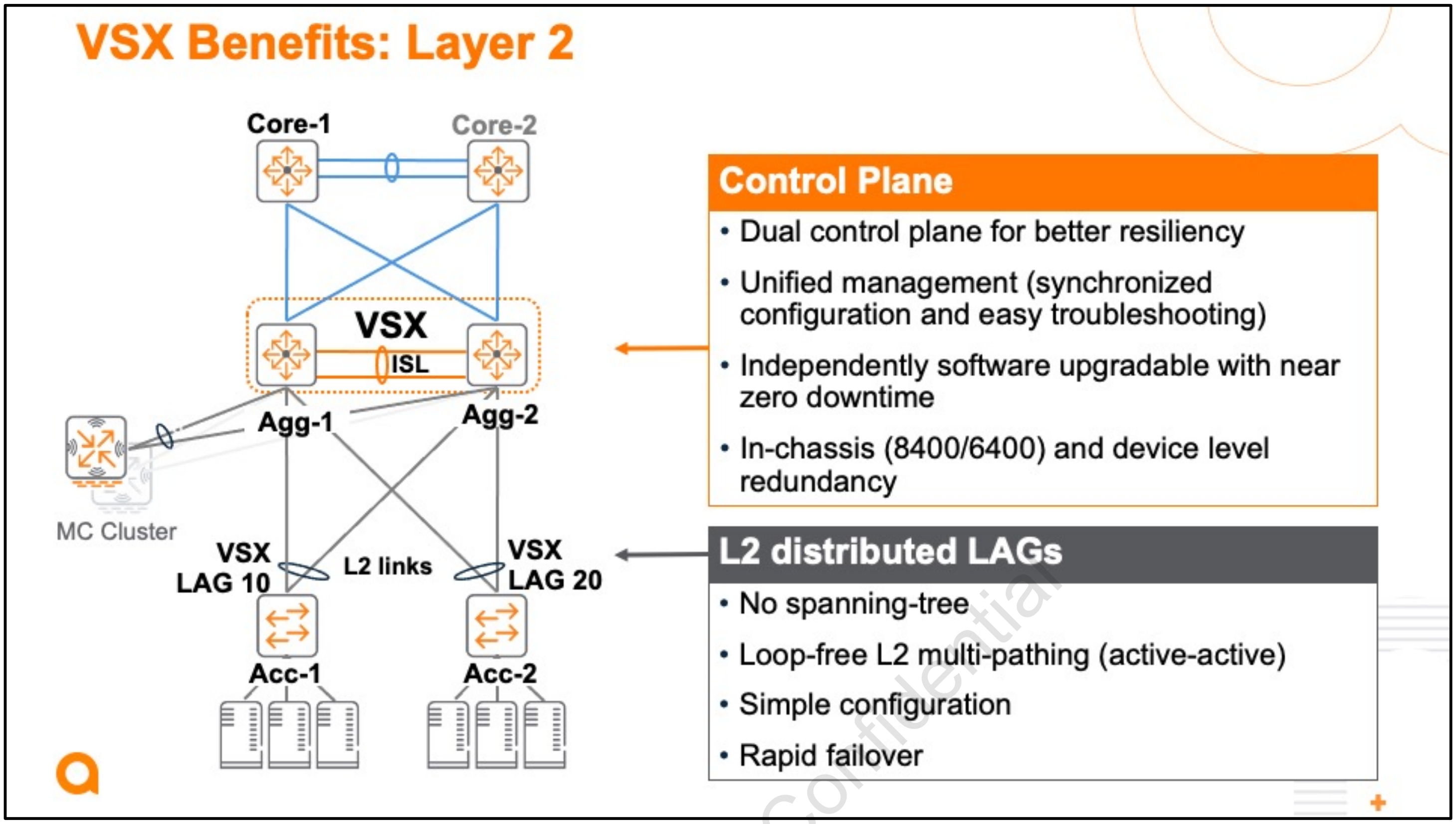
Aruba Training-Confidential

AOS-CX switches, particularly when deployed with VSX, meet customers' high availability needs by providing redundancy, resiliency, performance, and simplicity. Redundancy consists of having at least two of every required component so that the system can continue functioning if one component fails. AOS-CX switches support redundancy at the hardware level with support for multiple management modules, fabric modules, power modules, and fans.

At the software level, VSX maintains two control planes, one in each switch, to ensure redundancy. Resiliency builds on redundancy by allowing the system to take advantage of redundant components. AOS-CX switches in a VSX fabric support link virtualization. This feature enables multiple redundant links, including ones on different members of the VSX fabric, to operate as one logical link, called a link aggregation (LAG). AOS-CX switches also support process resiliency; if a process encounters an error, the system can self-restart to the last known good state.

A good high availability solution should also provide good performance with fast failover in all scenarios. VSX minimizes the duration of traffic outages by providing low-latency link failovers. If one link in the LAG fails, traffic instantly fails over to other links. VSX supports live upgrades so the VSX fabric remains active as one member and then the next upgrades. In addition, the AOS-CX switches have a fast upgrade time, which minimizes the time that the network is at risk, with just one up member during the upgrade process.

Finally, a highly available solution must be simple to configure to lower the chances that operator error introduces unplanned downtime. It is easy to configure VSX on AOS-CX switches with simple CLI commands. VSX also supports configuration synchronization and consistency checks between the members of the fabric to lower the risk of error.



VSX Benefits: Layer 2

Control Plane

- Dual control plane for better resiliency
- Unified management (synchronized configuration and easy troubleshooting)
- Independently software upgradable with near zero downtime
- In-chassis (8400/6400) and device level redundancy

L2 distributed LAGs

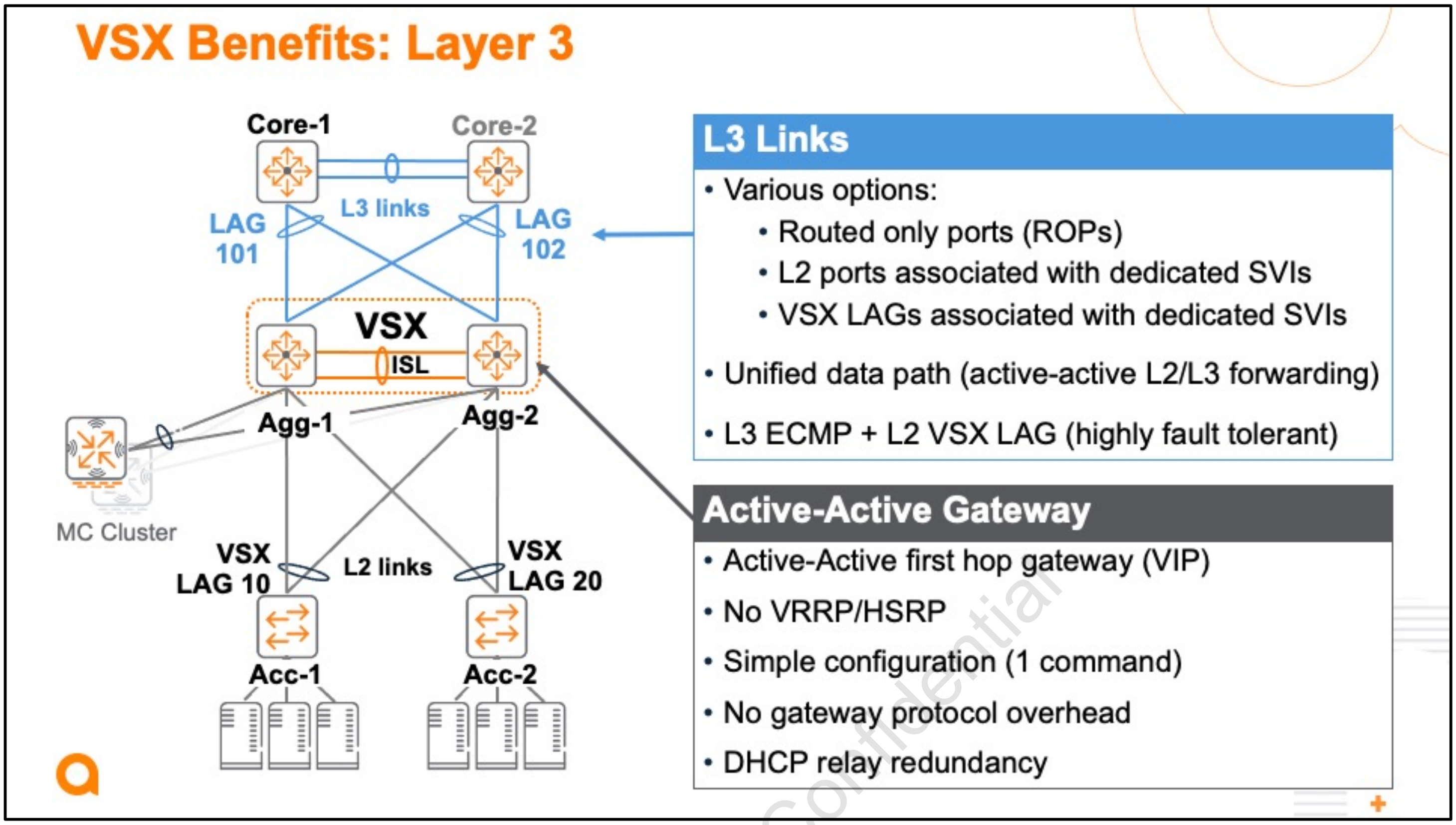
- No spanning-tree
- Loop-free L2 multi-pathing (active-active)
- Simple configuration
- Rapid failover

Here is an example of an Aruba-recommended topology in which you deploy VSX at the aggregation layer.

While you can deploy VSX at the core if the topology requires, you may prefer to keep redundant core switches entirely independent. A dual control plane provides the best resiliency and software upgrades with near zero downtime. At the same time, VSX permits unified management on the features that administrators choose to synchronize.

| The VSX pair uses link aggregations (LAGs) with links on both VSX switches to connect to access layer switches. This type of LAG is called a multi-chassis (MLAG), distributed, or VSX LAG. These VSX LAGs operate at layer 2 because the access layer switches do not route, and the aggregation layer VSX pair acts as the default gateway for the access VLANs. The layer 2 LAGs eliminate the need for Multiple Spanning Tree Protocol (MSTP) or Rapid Per-VLAN Spanning Tree (RPVST) on the redundant links. They provide a loop-free path that permits the use of all the links in the aggregation (active-active design). Configuring a LAG-based topology is simple, and if a link fails, failover occurs very quickly.

Aruba Training-Confidential



VSX Benefits: Layer 3

L3 Links

- Various options:
  - Routed only ports (ROPs)
  - L2 ports associated with dedicated SVIs
  - VSX LAGs associated with dedicated SVIs
- Unified data path (active-active L2/L3 forwarding)
- L3 ECMP + L2 VSX LAG (highly fault tolerant)

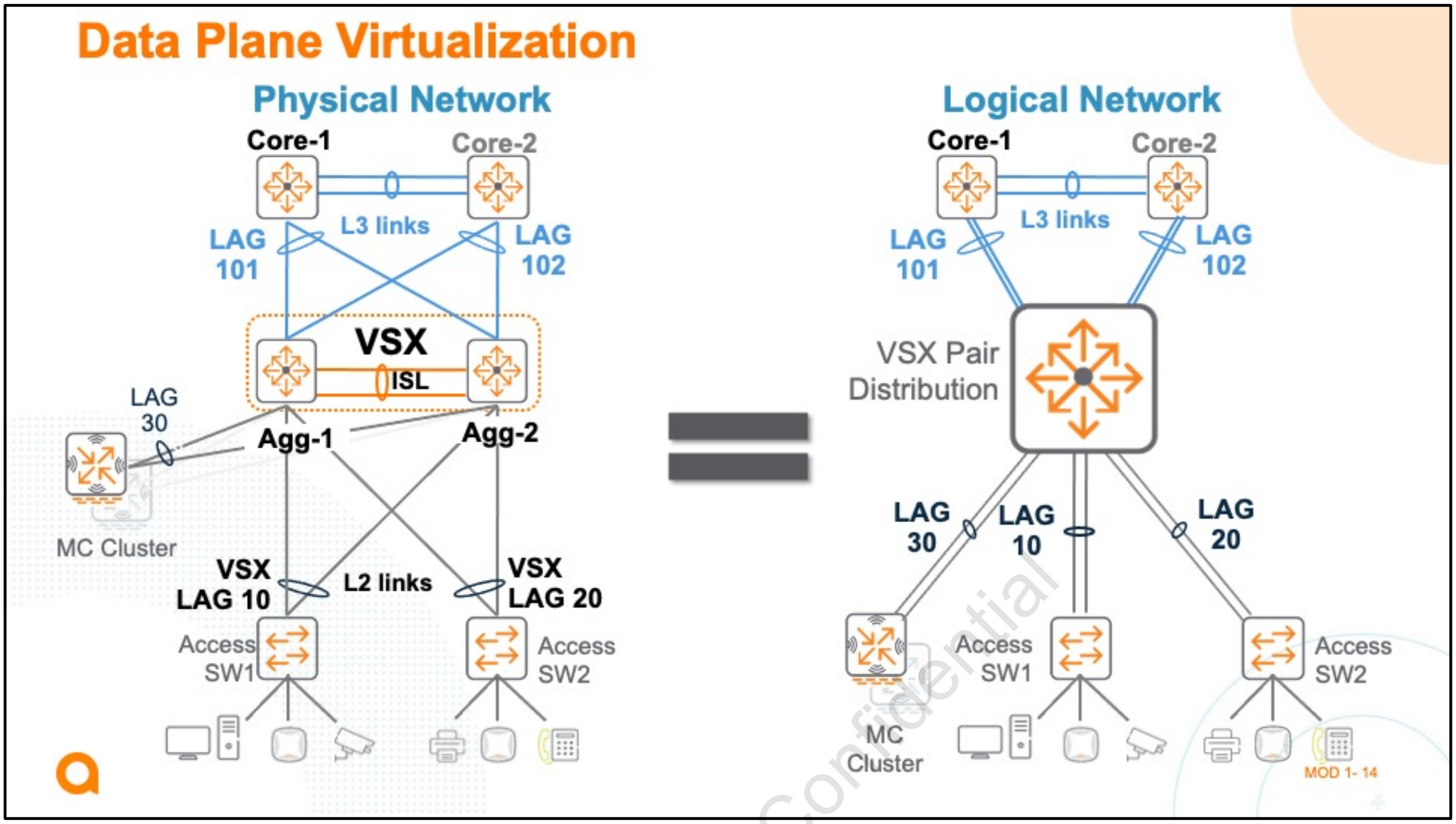
Active-Active Gateway

- Active-Active first hop gateway (VIP)
- No VRRP/HSRP
- Simple configuration (1 command)
- No gateway protocol overhead
- DHCP relay redundancy

Aruba Training-Confidential

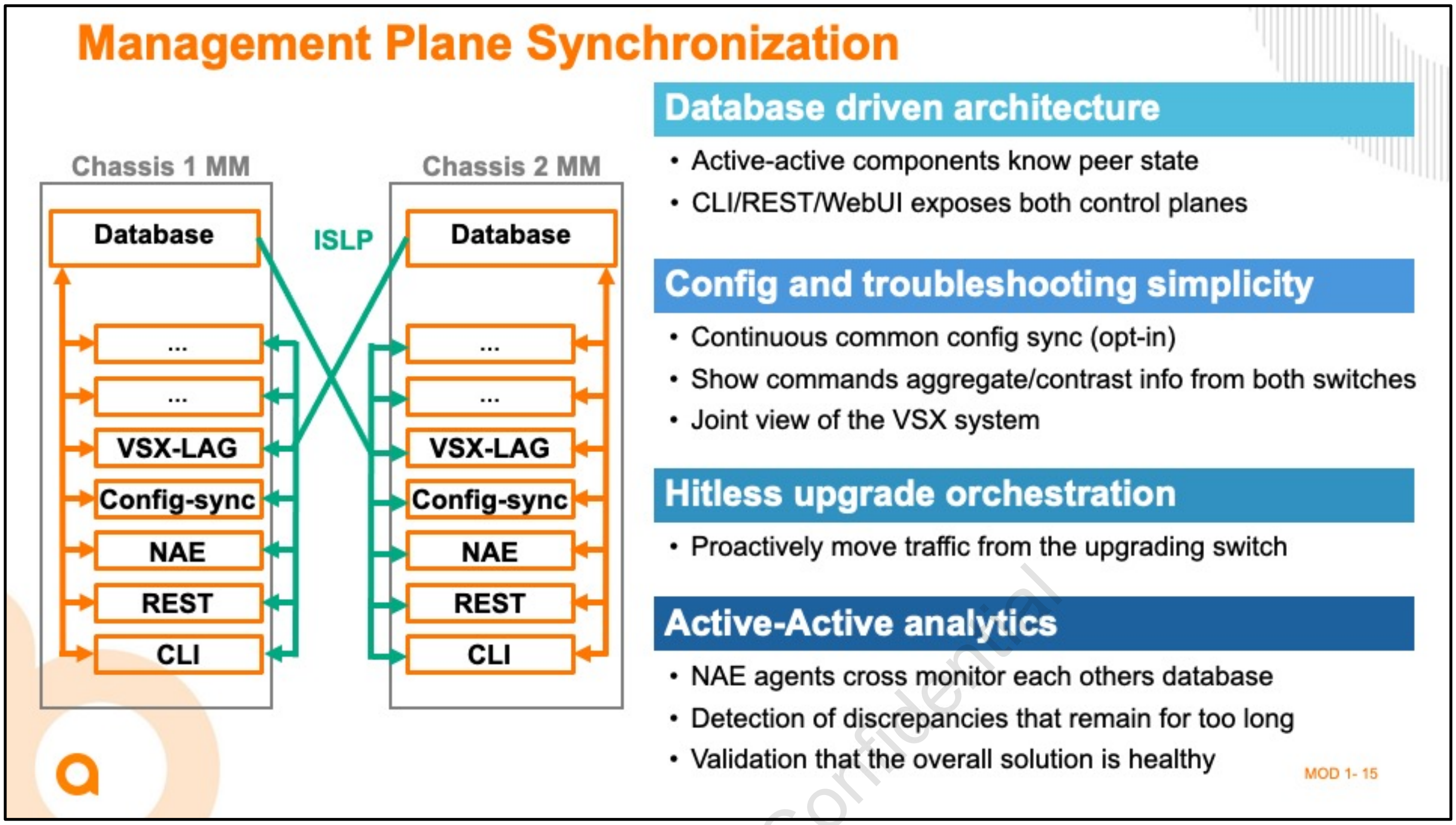
The VSX pair has layer 3 links to the core. Each VSX switch can connect independently to each core switch on a routed-only port (ROP) or on a layer 2 port associated with layer 3 VLAN interfaces (also known as Switch Virtual Interfaces – SVIs) dedicated to that port. Alternatively, both VSX switches could connect to each core switch together on a VSX LAG. This LAG would operate at layer 2 and associate with one or more VSIs. Whichever option you select, VSX can support active-active layer 3 routing through Equal Cost Multipath (ECMP) and an efficient path for unicast traffic. With the final option, the combination of layer 3 ECMP and the VSX LAG creates a highly fault-tolerant system.

| Remember that the VSX pair acts as the default gateway for the access VLANs. To do so, the pair uses the active gateway feature. This feature allows each switch in the pair to act as an active default gateway for the VLAN using a shared virtual IP address (VIP) and virtual MAC address. It eliminates the need for Virtual Router Redundancy Protocol (VRRP) or Hot Standby Router Protocol (HSRP). Simple to configure, the active gateway feature relies on VSX operations so it does not add any protocol overhead. It also supports redundancy for DHCP relay functions.



Data plane virtualization supports VSX LAGs, so other LAG-connected devices perceive the VSX switch pair as a single device. In this example, those are both downstream access layer switches and upstream core switches.

Aruba Training-Confidential



Management Plane Synchronization

Database driven architecture

- Active-active components know peer state
- CLI/REST/WebUI exposes both control planes

Config and troubleshooting simplicity

- Continuous common config sync (opt-in)
- Show commands aggregate/contrast info from both switches
- Joint view of the VSX system

Hitless upgrade orchestration

- Proactively move traffic from the upgrading switch

Active-Active analytics

- NAE agents cross monitor each others database
- Detection of discrepancies that remain for too long
- Validation that the overall solution is healthy

MOD 1- 15

Aruba Training-Confidential

Database-driven architecture

The VSX dual management plane supports synchronization, made possible partially through AOS-CX's innovative database-centric design. VSX peers can synchronize portions of their database so that each peer can actively forward traffic while also knowing the state of its peer.

VSX simplifies management functions - the CLI, REST interface, and web UI exposes both control planes to admins and business applications. In addition to simplifying management, this design enables analytics across the redundant pair.

Config and troubleshooting simplicity

You can choose which specific components to synchronize, such as an interface or VLAN. Configuration for that component is then continuously synchronized for easier configuration and fewer misconfigurations. Show commands aggregate information from several configuration pieces across both switches and compare them - quickly pick out mismatched settings. You can add the **vsx-peer** option to many commands, to view information about both the local switch and the peer – you get a simple joint view of the VSX system.

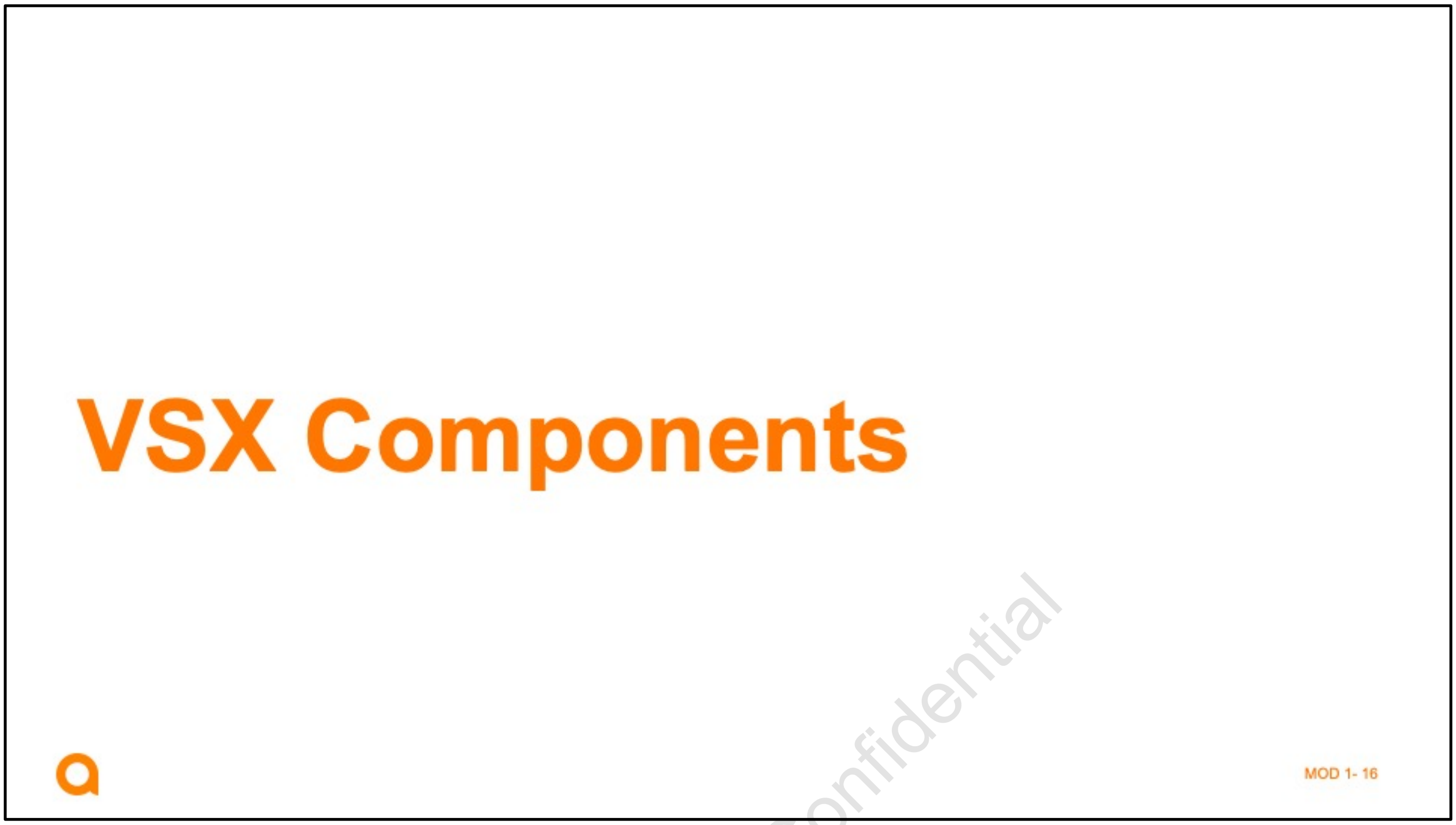
Hitless upgrade orchestration

VSX synchronization also helps to orchestrate hitless live upgrades. When one member needs to undergo an upgrade, the other member knows, and the pair can proactively move traffic away from the member that needs to be upgraded toward the other member.

Active-Active analytics

While each switch in the VSX pair maintains its own NAE agent, which stores sensor values locally, NAE is aware of VSX. The agent on each member can connect to the other member’s database and cross monitor it to detect discrepancies. For example, each NAE agent monitors the number of objects in the primary and secondary members’ databases. If synchronization is working correctly, the number of objects should be identical. An alert can trigger if, for example, there are 20 percent more objects in the primary database than the secondary database for at least five minutes. Active-active analytics helps to validate that the overall solution is healthy.

Aruba Training-Confidential

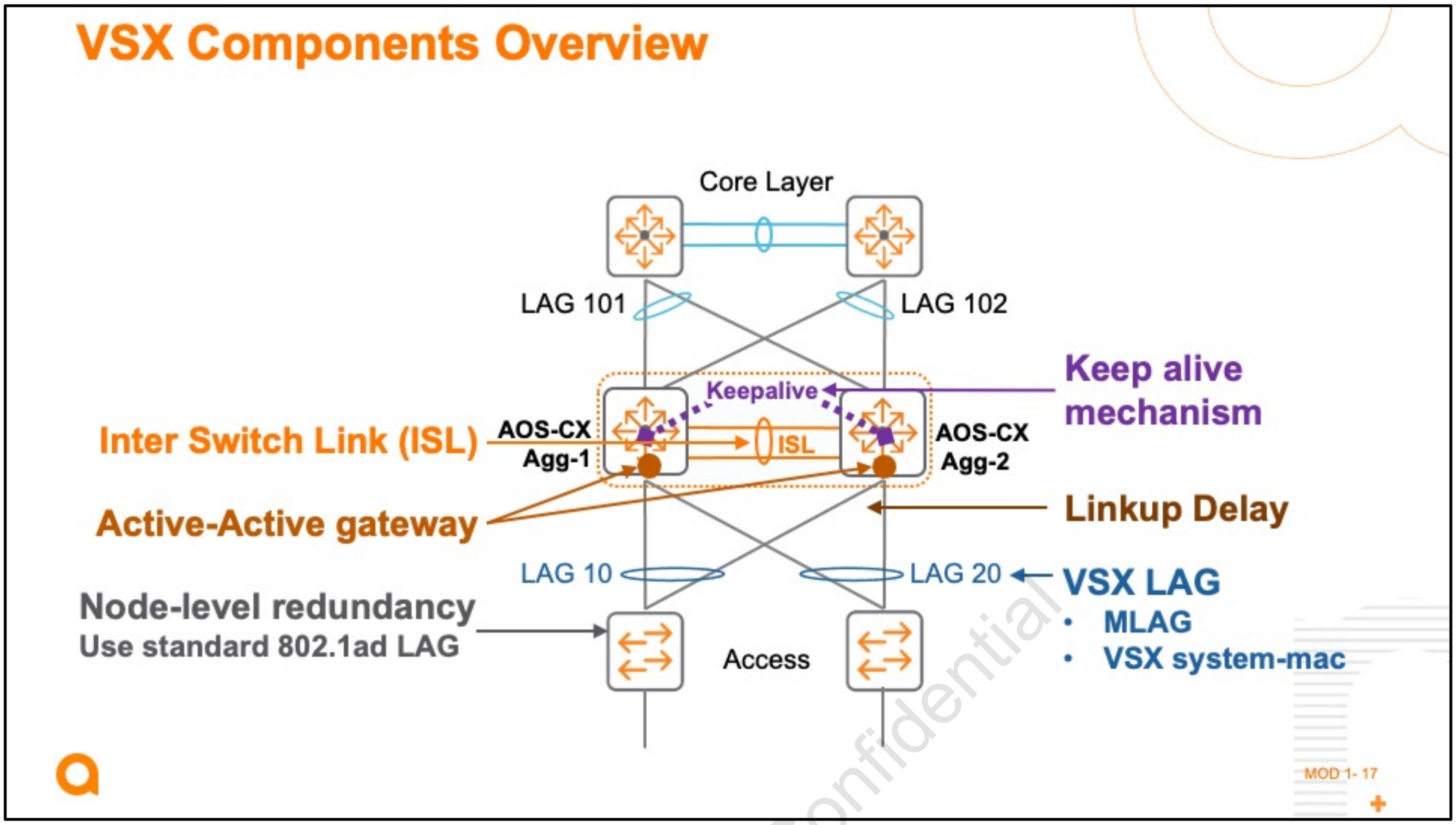


VSX Components

MOD 1- 16

This section introduces the components involved with VSX.

Aruba Training-Confidential



VSX Components Overview

Aruba Training-Confidential

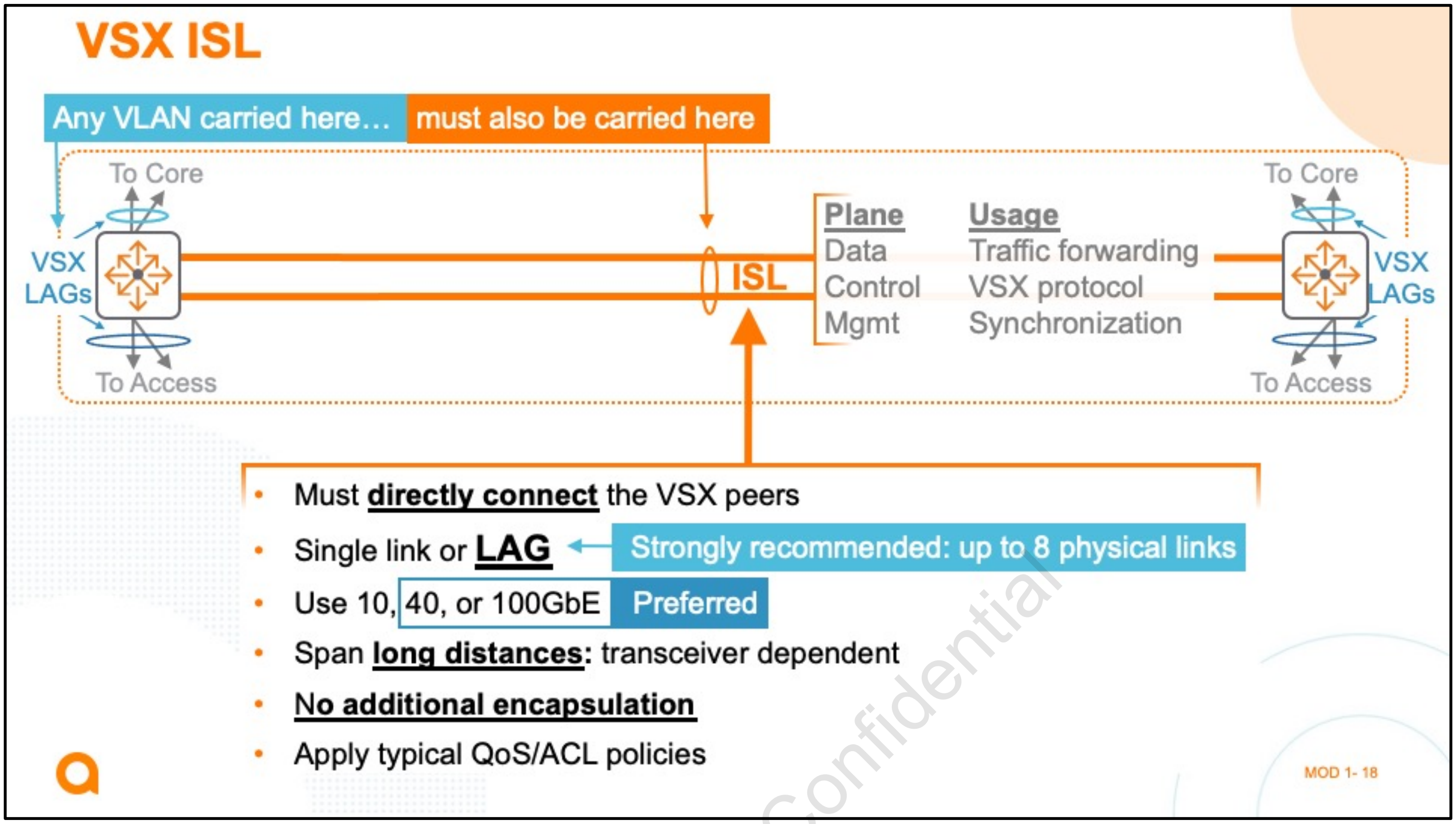
VSX supports Multi-chassis Link Aggregation (MLAG) - two or more links from two switches form a LAG that acts as a single logical interface. While the IEEE standard 802.3ad is limited to aggregating links on a single switch or device, VSX uses proprietary technology to overcome this limitation, to support link aggregation across multiple switches in the same VSX stack.

| An Inter-Switch Link (ISL) connects the two switches, and a keep-alive mechanism is used for resiliency.

| Also, VSX provides node-level network redundancy for when one switch fails. You must configure the downstream device as a standard 802.3ad LAG interface – they perceive the VSX pair as a single device. Downstream devices can be anything that supports 802.3ad Link Aggregation Control Protocol (LACP).

| In VSX, one device acts as primary and the other device as secondary, but BOTH devices forward traffic, for an Active-Active gateway solution.

| And you will soon learn about the linkup delay feature, which enhances initial forwarding readiness.



VSX ISL

Plane	Usage
Data	Traffic forwarding
Control	VSX protocol
Mgmt	Synchronization

- Must **directly connect** the VSX peers
- Single link or **LAG** ← Strongly recommended: up to 8 physical links
- Use 10, 40, or 100GbE Preferred
- Span **long distances**: transceiver dependent
- **No additional encapsulation**
- Apply typical QoS/ACL policies

MOD 1- 18

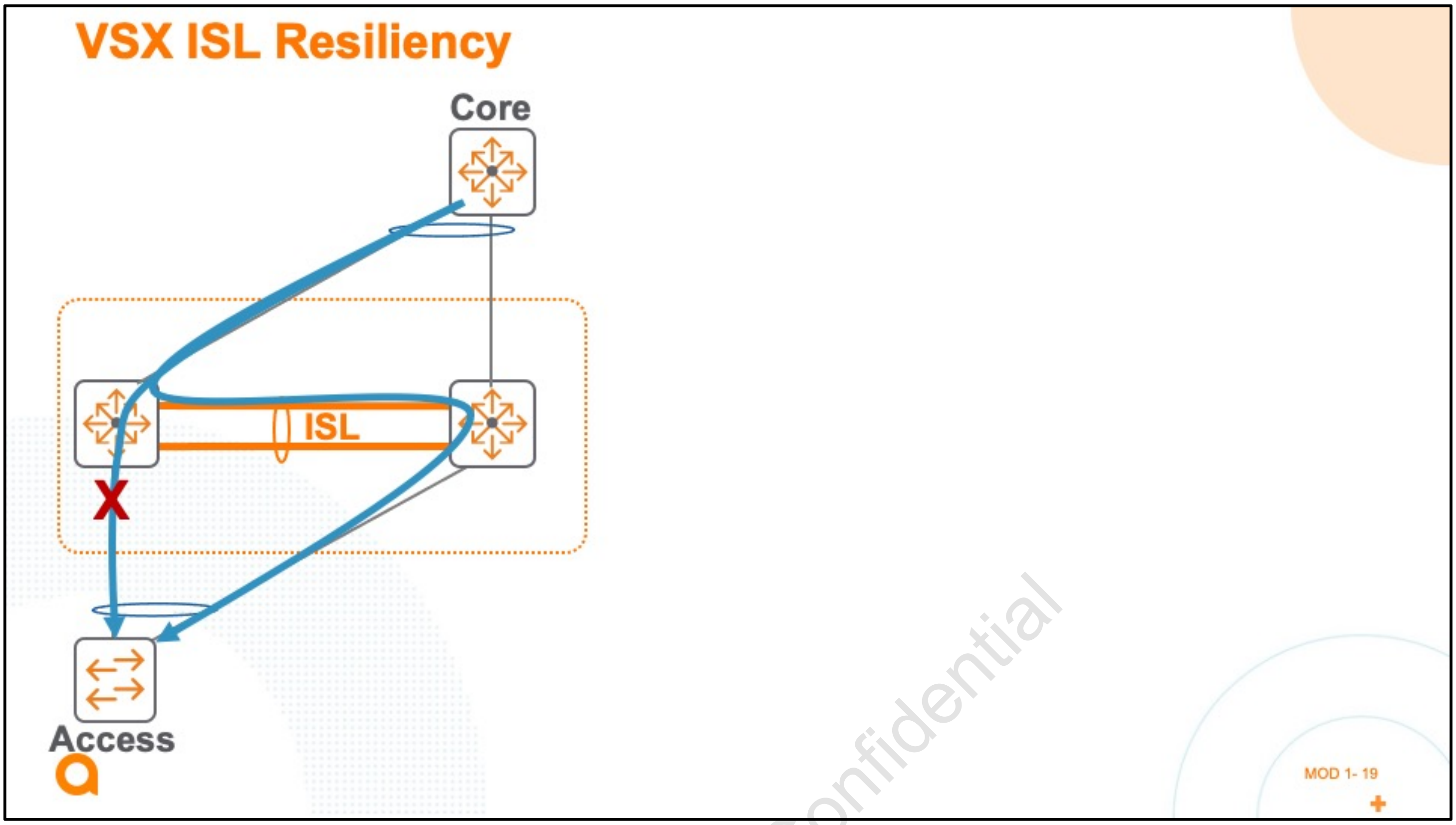
Aruba Training-Confidential

An Inter-Switch Link (ISL) is a Layer 2 interface between two VSX peer switches. You must configure each VSX switch with an ISL that directly connects to its peer VSX switch. The ISL is used at all three switch processing planes:

- Data plane: forward production traffic
- Control plane: exchange VSX protocol messages
- Management plane: synchronization

The ISL can consist of a single physical link, but Aruba strongly recommends link aggregation. You can have up to 8 physical links for resiliency, and all must be the same speed. You can use 10Gbps, but are strongly encouraged to use 40 or 100Gbps to ensure adequate bandwidth for the data path. The ISL uses front plane ports on the member switches and can use any media. If you select fiber media, the ISL can span long distances, depending on the fiber and transceiver type.

The ISL interface is a member of all VLANs by default. You can change ISL membership through the CLI, but you must ensure that any VLAN that is carried on a VSX LAG is also carried on the ISL. In the data path, traffic is forwarded natively with no additional encapsulation, unlike with VSF. Also note that you can control the ISL with QoS and ACL policies much as you do other interfaces on the switch. Of course, you must be careful that the ACL does not block required traffic. The ISL also preserves DSCP remarking. For example, you could configure a VSX LAG to re-mark traffic with a specific DSCP.



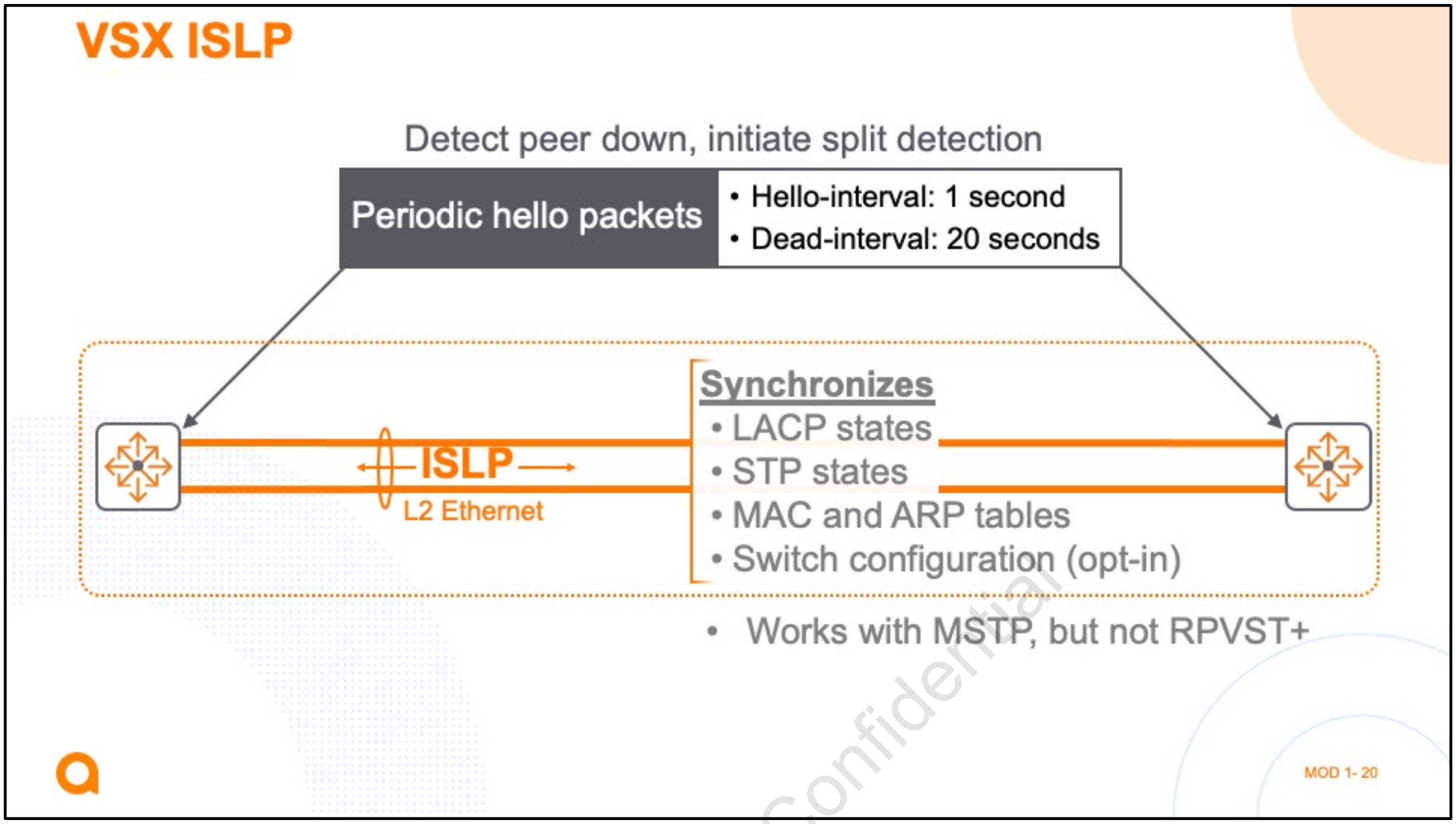
VSX ISL Resiliency

Aruba Training-Confidential

Recall that the VSX switches use LAG links to connect up to Core switches and down to Access switches.

| If a local VSX switch’s link fails, it transmits the traffic across the ISL so its peer can forward it. The packet keeps the DSCP value that it would have had if transmitted over the LAG locally. However, the ISL interfaces on the other peer must be configured to trust the DSCP settings on ingress frames.

This all works with the help of the Inter-Switch Link Protocol (ISLP).



VSX ISLP

Detect peer down, initiate split detection

Periodic hello packets

- Hello-interval: 1 second
- Dead-interval: 20 seconds

Synchronizes

- LACP states
- STP states
- MAC and ARP tables
- Switch configuration (opt-in)

- Works with MSTP, but not RPVST+

MOD 1- 20

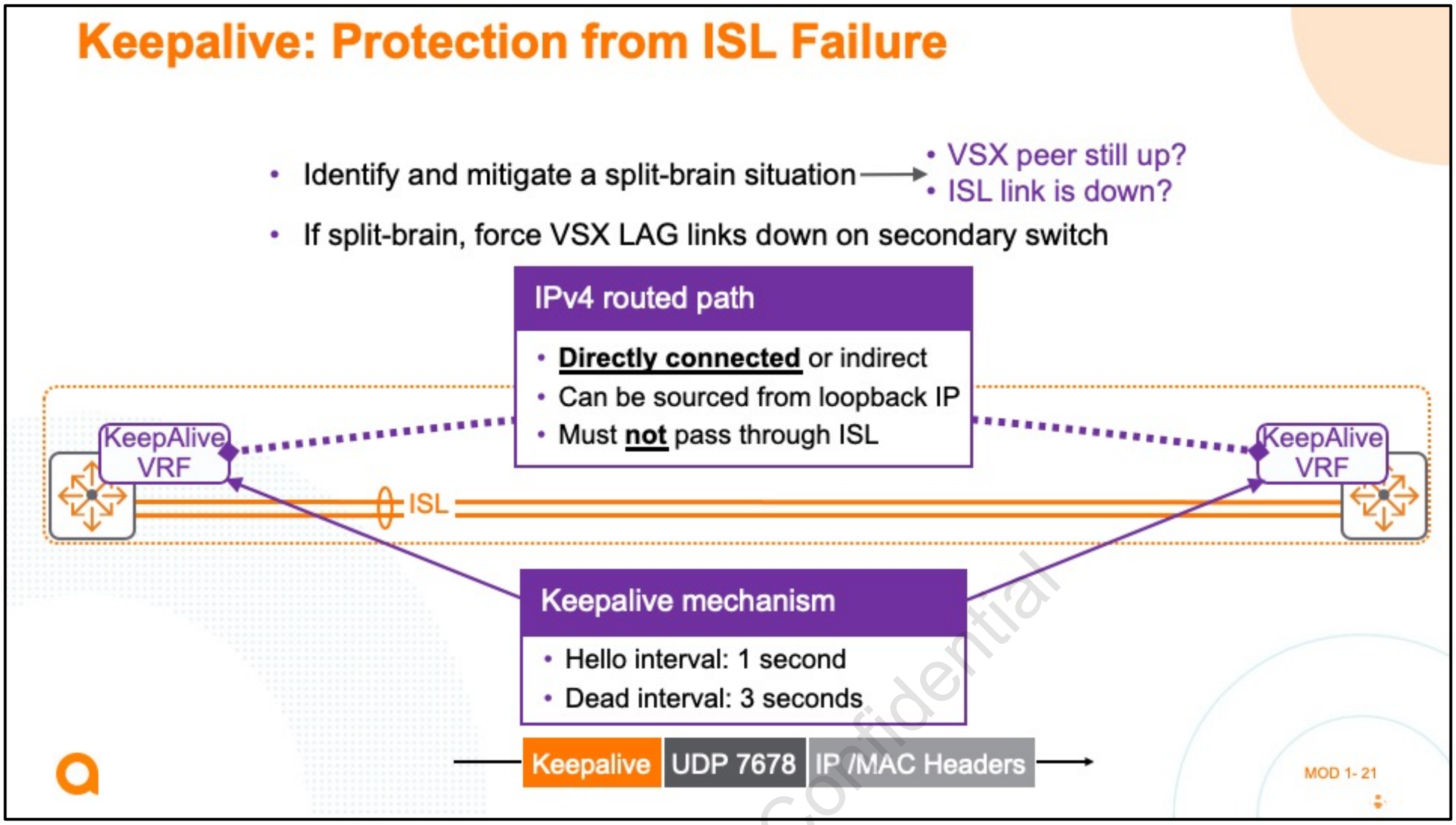
Aruba Training-Confidential

The Inter-Switch Link Protocol (ISLP) runs over the ISL. When you first configure the VSX pair, ISLP synchronizes the switches. Then it periodically synchronizes LACP states, MSTP States, MAC tables, ARP tables, and config info. VSX also supports interoperability with Multiple Spanning Tree Protocol (MSTP), but not with RPVST+.

ISLP runs at Layer 2. In other words, it is encapsulated in Ethernet frames. Since VSX switches are directly connected, there is no need for IP.

ISLP also helps peers to detect each other’s status. They periodically exchange hello packet at one second intervals by default. The dead interval is 20 seconds by default. It determines when a switch decides that the peer is no longer available and initiates split detection. While the dead interval might seem long, understand that each VSX peer maintains its own control plane. If one peer goes down, all links in VSX LAGs go down. Traffic instantly fails over to the links on the other peer, which is continues to actively forward without interruption. In most cases, peers determine that they must initiate split detection when the ISL link is down. The ISL link is considered down when all physical member links are down for the hold time, which is 0 seconds by default.

ISLP uses version control and provides backward compatibility regarding VSX synchronization capabilities. Processes used for VSX include vsxd for VSX LAGs, vsx-synchd for synchronization, and vsx-mgmtd to synchronize REST, NAE, and configuration settings across the pair.



Keepalive: Protection from ISL Failure

- Identify and mitigate a split-brain situation →
  - VSX peer still up?
  - ISL link is down?
- If split-brain, force VSX LAG links down on secondary switch

IPv4 routed path

- **Directly connected** or indirect
- Can be sourced from loopback IP
- Must **not** pass through ISL

Keepalive mechanism

- Hello interval: 1 second
- Dead interval: 3 seconds

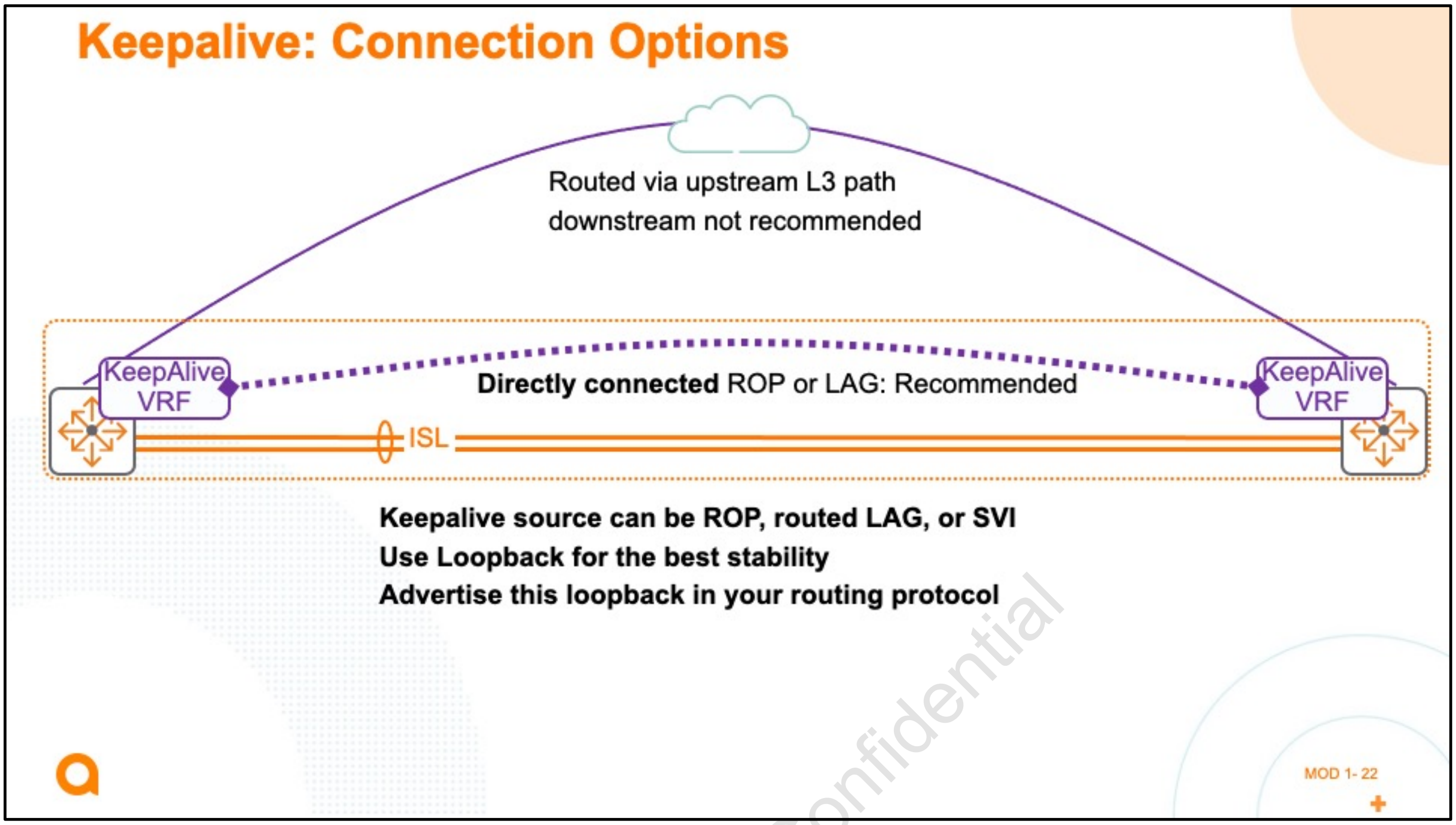
Aruba-Training-Confidential

You should configure each VSX switch with a keepalive connection to its peer. If the ISL goes down, each VSX peer uses keepalive communications to identify whether the peer is still up, but unavailable over the ISL— a so-called split brain condition. If the keepalive communications do indicate that the peer is still up, the primary VSX switch keeps its VSX LAG links up, and the secondary VSX switch forces its VSX LAG links down. The primary switch was selected when VSX established based on a user-configured role setting. If you configure the switches with the same role the lowest MAC address acts as a tie-breaker.

| Peers exchange keepalives over a routed network; the path can consist of a direct Layer 3 link or an indirect link through the upstream Layer 3 network. You can set the source for keepalive packets to a loopback interface for stability. In either case, though, make sure that keepalives do not pass through the ISL.

| Keepalive packets are UDP based. By default, they use port 7678, but the UDP port is configurable.

| By default, keepalive hello packets are sent every one second, but you can configure the interval to a value from 1 to 5 seconds. The keepalive dead-interval is 3 seconds, but has a configurable range of 2 to 20 seconds. If a device does not receive a keepalive packet from its peer within the dead interval, it treats the peer device as out-of-service and does not implement the split-brain protection mechanisms.



Keepalive: Connection Options

Routed via upstream L3 path
downstream not recommended
KeepAlive VRF
KeepAlive VRF
Directly connected ROP or LAG: Recommended
ISL

Keepalive source can be ROP, routed LAG, or SVI
Use Loopback for the best stability
Advertise this loopback in your routing protocol

MOD 1- 22

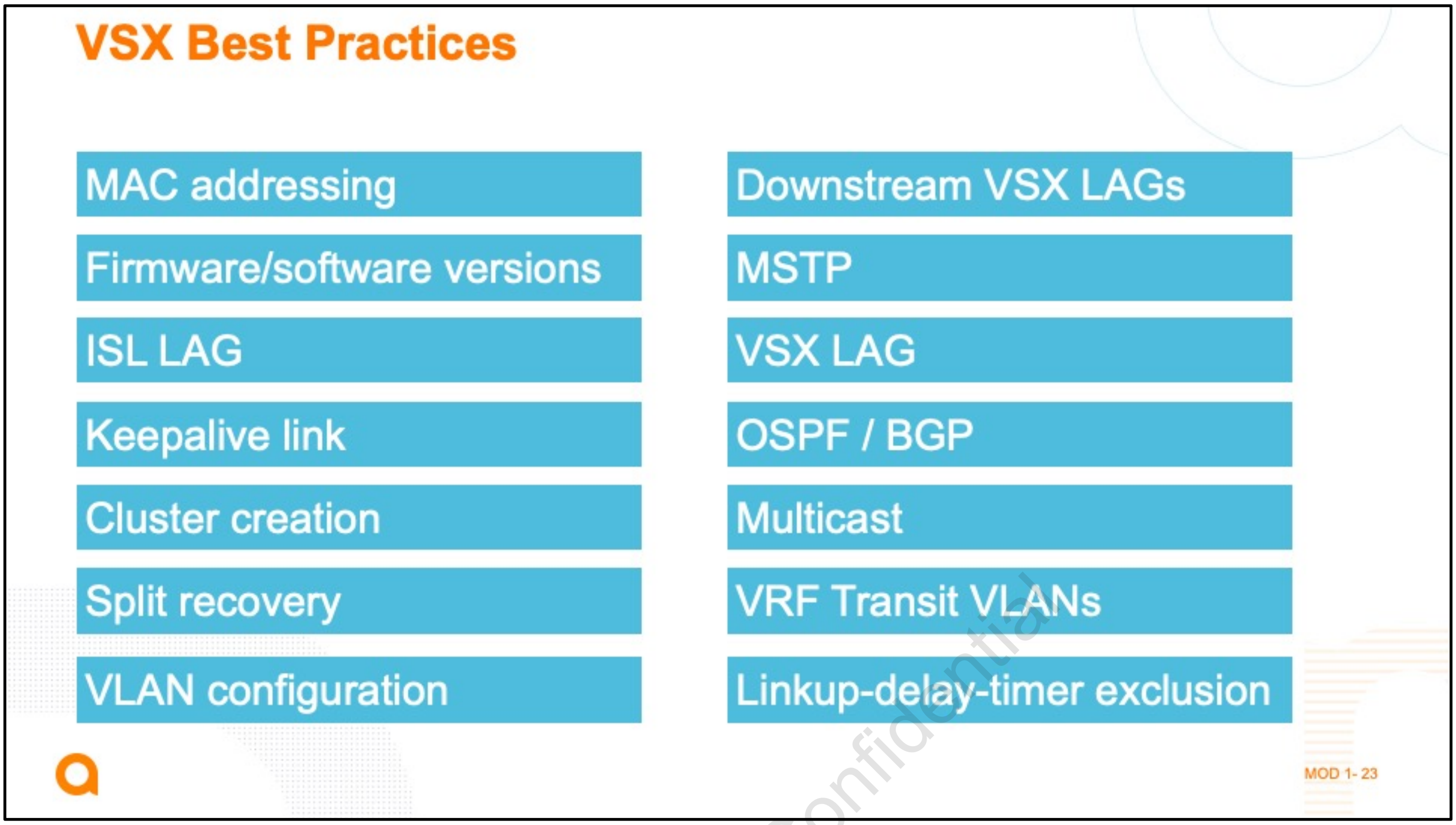
Aruba Training-Confidential

You have two options for establishing the keepalive links. The most crucial rule is that the keepalive communications must *not* pass through the ISL.

- Use a directly connected ROP on each peer or a directly connected route LAG. Aruba recommends this type of configuration.
- | Use the through routed path option. Keepalives pass through the upstream Layer 3 network; Aruba does not recommend passing the keepalives through the downstream network. The indirect option is cost-optimized because it requires fewer links.

| You can choose an ROP, routed LAG, SVI, or loopback interface for the keepalive source. A loopback interface provides the most stability because the peers can always reach each other’s loopback interfaces if even one Layer 3 path is available between each other. If you tie the keepalives to a particular physical interface or SVI, that interface might go down. Remember to advertise the source loopback interface in the dynamic routing protocol.

Note: For additional hardening, Aruba recommends assigning keepalive source interface to its *own* VRF.



VSX Best Practices

- MAC addressing
- Firmware/software versions
- ISL LAG
- Keepalive link
- Cluster creation
- Split recovery
- VLAN configuration
- Downstream VSX LAGs
- MSTP
- VSX LAG
- OSPF / BGP
- Multicast
- VRF Transit VLANs
- Linkup-delay-timer exclusion

MOD 1- 23

Aruba Training-Confidential

MAC addressing

One of the main VSX best practice is to set VSX system-mac. Do not leave it blank so the default system-mac is used. You want the VSX system-mac to be independent from the physical hardware MAC address. Thus, hardware replacement on the VSX primary switch will not affect your configuration, and so has no impact on the VSX secondary because the cluster ID remains unchanged. In other words, VSX primary hardware replacement is hitless for the VSX secondary. Otherwise the VSX secondary would have to join a new cluster ID, ID from VSX primary, and would temporarily turn-off its VSX LAG ports.

Use locally administered unicast MAC address when assigning system-mac or active-gateway virtual MAC address. There are 4 ranges reserved for private use for unicast (with second least significant bit of the first octet of the unicast address set to 1). x is any Hexadecimal value.

- x2-xx-xx-xx-xx-xx
- x6-xx-xx-xx-xx-xx
- xA-xx-xx-xx-xx-xx
- xE-xx-xx-xx-xx-xx

Here is an example proposal to ensure that unique values are used in the administrative domain, where XX is the Unique Cluster ID in the function, and Y is the Virtual MAC ID (0 to 15):

Function	System-mac	Active-gateway Virtual MAC
Access / TOR Layer	02:00:00:00:XX:00	12:00:00:00:XX:0Y
Aggregation	02:01:00:00:XX:00	12:01:00:00:XX:0Y
Core / Spine	02:02:00:00:XX:00	12:02:00:00:XX:0Y

The scope of this VMAC is purely link-local. Consequently, the same Virtual MAC address value can be used on any L3 VLAN interface (SVI).

If some servers or systems have dual-attachment to two different SVIs, and the you would like to see distinct MAC addresses for the next-hops over these separate interfaces, then 16 VMACs are available. For dual-stack IPv4 and IPv6, 16 VMACs can be used for IPv4 and the same VMACs can be used for IPv6. It is however a best practice to use only 8 VMACs for IPv4 and 8 different VMACs for IPV6.

Note: Any other allocation rules can be chosen according to administrative rules in place by the network operational team. Multicast or broadcast MAC addresses must not be used for System-mac.

Firmware/software version

Please install same version on both CX Switches that will create the VSX cluster. It is better to avoid any version mismatch during the creation of the cluster, as a warning would appear

ISL LAG

It is assumed that 2x 40G or 2x50G or 2x100G direct fibers / DACs are already interconnecting AGG-1 and AGG-2.

The best practice for ISL bandwidth is at least 2x40G (QSFP+) or 2x 50G (SFP56) or 2x100G (QSFP28). It is technically possible to use 2x10G or 2x25G; however it is recommended to plan for any uplink failure and associated impact on the bandwidth requirement for the ISL. If the uplinks from AGG-1 fail, traffic from AGG-1 is redirected to AGG-2 over the ISL before reaching the upstream layer. This is fine if there is enough bandwidth remaining for the ISL protocol and control-plane communication. It is recommended to size the ISL bandwidth to be equal to, at least, the sum of uplinks bandwidth of one VSX switch. The best practice rule is to size the

Aruba Training-Confidential

ISL bandwidth according to the failure domain target.

The best practice for ISL physical ports is to select at least two ports of the same speed (2x40G or 2x50G or 2x100G), and, in case of a chassis, to select these ports from different Line Cards.

The best practice for LAG numbering is to use the last available LAG ID (ie. 256 in AOS-CX10.4) for the ISL, so that LAG ID=1 is used for connecting the Access Switch#1 on port 1/1/1, so that LAG 2 is used to connect the second Access Switch on port 1/1/2, and so on…

Other ISL LAG Best Practices

- VLAN trunking: permit ALL VLANs, for simpler configuration. Specifying a restrictive list of VLAN IDs is entirely valid if the network admin wants more control.
- LACP timers: keep the default long timer (30s for lacp rate slow).
- Hashing algorithm: keep the default l3-src-dst (alternative being l2-src-dst).
- MTU: Configure on all devices the appropriate size to support features such as Dynamic Segmentation as well as other protocols/functions which require MTUs larger than 1500 bytes. Care should be taken to ensure that the IP path from access devices (switches or APs) can provide a MTU of at least 1564 bytes to the mobility controllers. Similarly, for datacenter server connectivity, largest MTU will ensure server jumbo frame traffic over ISL. Recommendation: Ethernet MTU = 9198 bytes.
- ACLs: Do not set any access-list on the ISL LAG to avoid designing complex and unnecessary ACL. The ISL is like a virtual data back-plane and security filtering is processed before or/and after crossing the ISL.
- QoS trust mode: Rely on the qos trust dscp that is globally configured on the Aggregation switches. If not configured globally (which is not the recommendation), qos trust dscp has to be set on the ISL LAG.

Aruba Training Confidential

Keepalive Link

The best practice for the Keepalive connection is to use a direct L3 circuit, which can be a low speed port (1G transceiver is enough, 1GBASE-T works as well) between both VSX nodes. This circuit need not be directly connected and the path can include active L2 and L3 equipment. Although this requires an additional dedicated port, it brings simplicity of configuration and operations. In the Appendix D, VSX keepalive over upstream layer 3 routing domain is documented as an alternative for those who want to protect from a fiber path cut that would impact ISL and keepalive simultaneously; or when the associated cost of a dedicated port is too high (100G). In case of a chassis (6400 or 8400), if possible, it is recommended to use a port from a different Line Card than the ones used for the ISL ports.

The best practice for Keepalive routing is to use a dedicated VRF. This is entirely optional and the default VRF can also be used, typically for the single VRF model with UDP keepalive over the upstream L3 domain. Having a dedicated VRF for Keepalive simplifies the operations and prevents any impact from routing change on the default VRF.

The best practice for Keepalive subnet is to use a /31 subnet as only 2 nodes will communicate together.

Aruba Training-Confidential

Cluster creation

Keep ISL timers (dead-interval, hello-interval, hold-time, peer-detect-interval) at default values - i.e. no specific configuration.

The best practice for role (primary or secondary) is to have a meaningful relationship with the switch hostname/identification. Example: AGG-1 is VSX primary and AGG-2 is VSX secondary.

Best practice for vsx-sync includes vsx-global. Thanks to this vsx-sync FeatureGroup parameter, the VSX management-plane will synchronize the following VSX settings: inter-switch-link hello-interval, dead-interval, hold-time, peer-detect-interval, keepalive udp-port, hello-interval, keepalive dead-interval, system-mac, split-recovery, linkup-delay-timer

VSX automatically tags the native VLAN configured on the LAG used for ISL.

Split-recovery

The best practice for VSX split-recovery is to keep the default split-recovery enabled (no configuration change). This best practice might be revisited in case of VSX and VXLAN. (VXLAN is beyond the scope of this course.)

VLAN configuration

The best practice for VLANs configuration is to configure the VLANs on the VSX primary with the vsx-sync attribute and let the VSX config-sync automatically synchronize the VLANs on the VSX secondary.

Downstream VSX-LAGs

For simplicity is assumed that connected Access Switches are already configured with uplinks - link-aggregation and trunked VLANs. The best practice for VSX LAG is to create the multi-chassis LAG interface on the VSX primary with all settings and then create the mirrored lag interface on the VSX secondary. LAG interface settings (including description) will be synchronized automatically. Only “no shut” in the lag interface context must be performed on the VSX secondary. Once the multi-chassis lag interface is created, it can be assigned to the physical port.

The best practice for allowed VLANs is to exclude the native VLAN 1 from being propagated. This is a very robust method to avoid Layer2 storm propagation due to potential loop initiated on an access switch. In case of access switch Zero-Touch-Provisioning use-case, this trunking exclusion is performed after the ZTP process.

The best practice for LAG numbering is to use LAG ID=1 for connecting the Access Switch#1 on port 1/1/1, LAG 2 used to connect a second Access Switch on port 1/1/2, and so on…

The best practice for LACP timers on the VSX LAG is to keep the default long timer (30s = lacp rate slow).

The best practice for MTU is to configure the appropriate size to support features such as Dynamic Segmentation or server jumbo frame. Ensure that the IP path from the access devices (switches or APs) can provide a MTU of at least 1564 bytes to the mobility controllers and that the server jumbo packet of 9000 bytes can be encapsulated. Flexibility should be anticipated to perform VXLAN encapsulation from the access switch (9000+50 bytes) or VXLAN encapsulation from the aggregation layer MTU+50. So the recommended Ethernet MTU is 9100 bytes for the downstream VSX LAG to the access layer and a MTU of 9000 bytes for endpoints or servers. The SVI IP MTU should match the MTU size on the aggregation layer, so the recommended IP MTU is 9100 bytes.

The best practice for hashing algorithm on the VSX LAG is to keep the default l3-src-dst (alternative being l2-src-dst). This option has an effect only if at least 2 ports per VSX node are members of the same VSX LAG.

Note: Most of the time the VSX LAG includes only two links: one link from the primary and one link from the secondary. Consequently, hashing algorithm selection has no effect on the traffic path as it is forwarded to the local port of the VSX LAG on

Aruba Training-Confidential

the switch receiving the traffic.

The **show lacp interfaces multi-chassis** command is very useful to get a complete status of the local LACP partnership as well as the VSX peer partnership details. Actor = local node, Partner = LACP neighbor (the access switch), Remote Actor = the VSX peer, Remote Partner = LACP neighbor of the VSX peer. Note that the port id of the VSX secondary is equal to 1000+ID_of_the_primary (in the example 1001). ALFNCD LACP state-flags should appear on all entries.

The best practice for the LACP fallback feature is to enable it on the VSX LAGs for the following use-cases: PXE boot, access switch ZTP, server NIC driver migration from active/standby to LACP. When applied to the VSX primary, LACP fallback is automatically synced on the VSX secondary.

Aruba Training-Confidential

MSTP

The best practice on Aggregation layer are:

- Do not use loop-protect (MSTP used instead).
- Use the default common instance 0: MST0
- Lower the spanning-tree priority to 4 to make VSX aggregation the STP root bridge (easier for support)
- Use root-guard on all downlinks to prevent any access switches from becoming Root Bridge.
- Keep the default port-type admin-network
- Let VSX secondary synchronized by vsx-sync process.

The best practice on Access layer are:

- Use loop-protect for all endpoint access ports (not configured on uplinks). Set the re-enable timer to 1hour.
- Keep the default common instance 0: MST0
- Keep the default spanning-tree priority of 8.
- All endpoint access ports are admin-edge, should not receive any BPDU (BDPU guard), should not trigger any Topology Change Notification (tcn-guard).
- Use loop-protection on all endpoint access ports as an extra-protection mechanism (in case of MSTP BPDUs are filtered by insertion of unmanaged switches which create a loop).
- Use loop-guard on all uplinks to prevent any flood due to failure of BPDU reception (fiber strand cut).

VSX LAG

- ACLs: If any ACL is used, the best practice is to have ACLs synchronized on secondary through vsx-sync. Any ACL applied on a VSX LAG on the VSX primary will get applied on the VSX secondary as well.
- QoS: QoS Marking being performed on the access layer, the aggregation switch is configure in the global context with qos trust dscp. No further configuration is needed as this was already set in previously.
- SVI (VLAN L3 interface): The best practice for SVI active-gateway is to set the active-gateway Virtual IP and Virtual MAC on the VSX primary and get the value synchronized on the VSX secondary with vsx-sync command.
    - The best practice for active-gateway VMAC is to use the same VMAC for all IPv4 SVIs. The scope of this VMAC is purely link-local. If some servers or systems have dual-attachment to two different SVIs, and the system administrator would like to see distinct MAC addresses for the next-hops over these separate interfaces, then 16 VMACs are available. For dual-stack IPv4 and IPv6, 16 VMACs can be used for IPv4 and the same VMACs can be used for IPv6. It is however a best practice to use only 8 VMACs for IPv4 and 8 different VMACs for IPV6.
- Mutlinetting: If it is used, set one VIP per secondary subnet and disable ip icmp redirect.
- IP MTU: configure all SVI IP MTUs to match the size of the L2 MTU: IP MTU recommended value = 9100. This parameter must be identical and manually set on both VSX nodes.
- DHCP relay: Configure the ip helper-address on the VSX primary and let vsx-sync configuring the same on the VSX secondary.

Aruba Training-Confidential

OSPF / BGP

It is a best practice to create a dedicated Transit VLAN between the VSX primary and the VSX secondary to exchange routes information for subnets that are not attached to both VSX nodes (ex: loopback addresses of each VSX node). This dedicated Transit VLAN provides better control and will not carry user data traffic in nominal situation or very limited in case of east-west traffic between single-attached endpoints.

There are two strategies to inject endpoint subnets into the routing table: either through OSPF or through BGP

OSPF: Most of the Campus deployments use OSPF to exchange route information for end-devices. This is simple and can scale very well with appropriate usage of areas. This is the target of this current document.

BGP: Lot of new DC deployment use BGP as a routing protocol due to the usage of

EVPN based VXLAN. Such a design is coming in the Campus as well. Also, for more complex and granular routing engineering, BGP communities and route-map can offer a level of control that OSPF can not provide. This can be exposed in a future white paper.

There are two options to inject end-user subnets into OSPF DataBase: using ospf command on the SVI (VLAN L3 interface), or redistributing the connected into OSPF with route-map control.

The best practice is to use the **ospf** command on SVI as offering a simpler configuration like for the area the subnets belongs to. This principle is selected as the OSPF best practice in the following described configuration. More details on OSPF best practices can be found on IP routing configuration guide.

The best practice for point-to-point interconnectivity subnet is to use /31 subnet.

The best practice for OSPF configuration is to use vsx-sync ospf synchronization option and have OSPF parameters automatically synced on the VSX secondary. As shown on the configuration step, very few elements must be configured on the secondary. Pay attention that to get such a benefit, the interface ID should be mirrored; i.e. if interface 1/1/49 is used for uplink on the VSX primary, it is strongly recommended to use the same ID 1/1/49 on the VSX secondary, otherwise OSPF synchronization will not synchronized the proper interface.

The best practice for OSPF cost is to have VSX primary <-> VSX secondary cost lower than Core-1 <-> Core-2 cost, as it is frequent that the ISL bandwidth is higher than the inter core devices bandwidth. In case of single-attachment subnet on one of the VSX node and non-meshed topology, the traffic from core would be sent to the VSX peer closest to the attached destination, avoiding consuming inter-core bandwidth.

Same concept applies for south-to north traffic pattern. In the below example, OSPF cost for Transit VLAN over ISL is set to 50, and 1000 for Core devices. OSPF cost is synchronized from the VSX primary to the VSX secondary, so the importance to use mirrored interface ID.

For BGP recommendations, please read the IP Routing Guide for detailed information.

Note: VSX can synchronize the full BGP configuration between the VSX primary and the VSX secondary. Most of the BGP configuration of the VSX secondary is the same than on the VSX primary. Except the configuration for iBGP peering between the VSX nodes inside the cluster, or for remote eBGP upstream peers with neighbor IP address being the physical IP address of theL3 point-to-point circuit (ex: here 1/1/50 with

Aruba Training Confidential

10.0.0.6/31 remote IP address). In such a case, the specific neighbor parameters are excluded from the VSX configuration synchronization with the following command on the VSX primary only: neighbor <IP_adress> vsx-sync-exclude.

Multicast

Please read the Multicast Guide for detailed information.

For multicast on VSX cluster, the best practice is to configure PIM Dual-DR or active/active under the PIM router command. With active-active command, the proxy-DR will also learn the multicast routes, and will allow fast recovery time if the actual DR fails.

Aruba Training-Confidential

VRF Transit VLANs

The best practice for upstream routing domain connectivity is to create one Transit VLAN per VRF per upstream VSX LAG. Consequently, according to the proposed topology, for VRF1, VLAN 101 is created and configured on upstream VSX LAG connecting CORE-1 (VSX LAG 101), VLAN 102 is created on upstream VSX LAG connecting CORE-2 (VSX LAG 102). Similarly, VLAN 201 and 202 are proposed for VRF2 and VLAN 301/302 for VRF3.

Note: for simplicity and readability, only 2 VRFs are documented: VRF1 and VRF2. VRF3 is not included in the configuration to alleviate the document. Configuration for VRF3 or any additional VRFs is similar to VRF1 and VRF2.

These Transit VLANs 101/102 and 201/202 will also serve for VSX intra-cluster routing (unlike the single VRF scenario where a dedicated Transit VLAN (VLAN 2) was created for Transit point-to-point routing between VSX nodes.

Upstream VSX LAG Configuration (VSX LAG 101/102)

In this section, for simplicity, it is assumed that the Core equipment are already configured with link-aggregation and trunked Transit VLANs.

The best practice for upstream VSX LAG is to create the multi-chassis lag interface on the VSX primary with all settings and then create the mirrored lag interface on the VSX secondary. LAG interface settings (including description) will be synchronized automatically. Only “no shut” in the lag interface has to be performed on the VSX secondary. Once multi-chassis lag interface is created it is assigned to the physical port.

The best practice for allowed VLANs is to exclude the native VLAN 1 from being propagated and to allow only the Transit VLANs corresponding to the facing Core device: i.e. VSX LAG.101 permitting VLAN 101 and 201, VSX LAG 102 permitting VLAN 102 and 202.

Aruba Training-Confidential

The best practice example for LAG numbering is to use LAG ID=101 for connecting the Core-1, LAG ID=102 for connecting Core-2. Any other numbering practice is possible as long as it does not introduce confusion with the downstream VSX LAGs.

The best practice for LACP timers on the VSX LAG is to keep the default long timer (30s for lacp rate slow). The best practice for MTU on these upstream VSX LAGs is to configure the maximum value (9198 bytes) like for ISL.

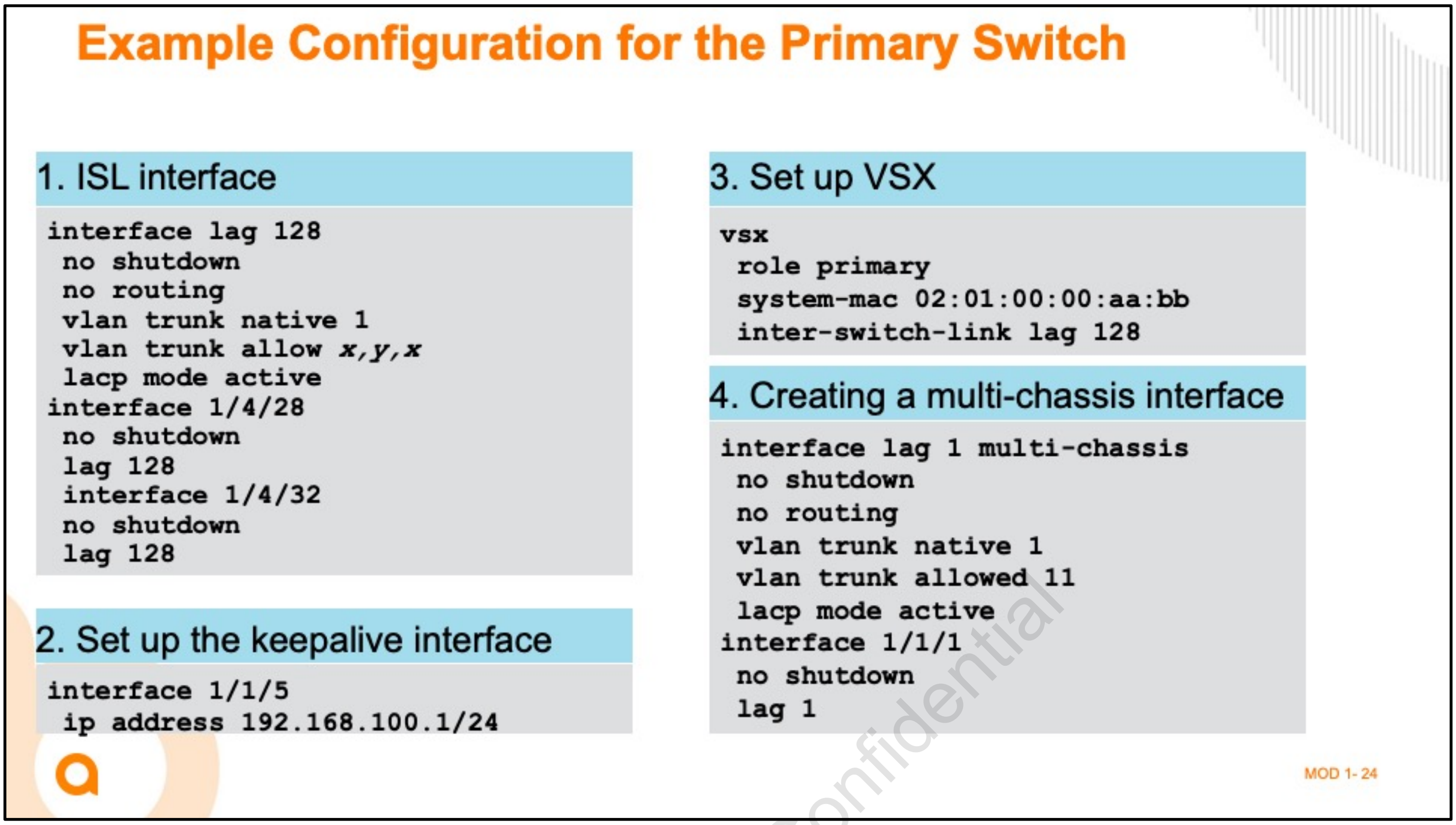
The best practice for hashing algorithm on the VSX LAG is to keep the default l3-src-dst (alternative being l2-src-dst), and would have an effect only if at least 2 ports per VSX node are members of the same VSX LAG.

VSX linkup-delay-timer exclusion

The best practice for VSX LAG exclusion for linkup-delay-timer is to exclude the upstream VSX LAGs that are used for Transit VLANs (in the topology above, VSX LAG 101 and VSX LAG 102). By excluding these upstream VSX LAGs, the routing protocols can establish peering over upstream Transit VLANs as soon as the upstream ports (to L3 Core-1 and Core-2) are UP, without waiting for the linkup-delay timer to complete. The benefit is that routes from L3 core are already learnt when ASIC is ready to forward traffic to downstream.

Note: For more information about an overview of VSX, best practices, and its configuration, see the VSX Configuration Best Practices guide. This has additional best practices based on using a multiple VRF routing model and connecting an access layer VSX cluster to an aggregation VSX cluster. You will also find useful information on maintaining and troubleshooting a VSX topology. And lastly, there are copious configuration examples included in the guide.

Aruba Training-Confidential



Example Configuration for the Primary Switch

1. ISL interface

```
interface lag 128
 no shutdown
 no routing
 vlan trunk native 1
 vlan trunk allow x,y,x
 lacp mode active
interface 1/4/28
 no shutdown
 lag 128
 interface 1/4/32
 no shutdown
 lag 128
```

2. Set up the keepalive interface

```
interface 1/1/5
 ip address 192.168.100.1/24
```

3. Set up VSX

```
vsx
 role primary
 system-mac 02:01:00:00:aa:bb
 inter-switch-link lag 128
```

4. Creating a multi-chassis interface

```
interface lag 1 multi-chassis
 no shutdown
 no routing
 vlan trunk native 1
 vlan trunk allowed 11
 lacp mode active
interface 1/1/1
 no shutdown
 lag 1
```

MOD 1- 24

Aruba Training-Confidential

This slide introduces VSX configuration.

First, set up the ISL link - a normal Layer-2 LAG. When a native VLAN is defined, as shown here, the switch automatically executes the **vlan trunk allowed all** command to ensure that the default VLAN is allowed on the trunk. In this example, LAG 128 is being used as the ISL.

The same list of VLANs that are trunked over the VSX LAGs must be configured on the primary and secondary VSX switches in the global configuration. The list of VLANs can be synced to the secondary switch if the vsx-sync command is used in the VLAN context. Also verify that the VLAN set is also permitted on the ISL on the primary and secondary VSX switches. To configure VLAN trunking on the ISL, enter the **vlan trunk allowed** [<VLAN-LIST> | all] command. If a native VLAN is defined, the switch automatically runs the vlan trunk allowed all command to ensure that the default VLAN is allowed on the trunk. To allow only specific VLANs on the trunk, enter the **vlan trunk allowed** <VLAN-LIST> command, for example: vlan trunk allowed 2,3,4

Then activate the LAG on the physical interface(s) between the two switches. Remember that to follow best practices, you should minimally use 2 links in the LAG.

Next, define the Layer-3 interface that will be used for sending keepalives between the two switches. This can be a direct link between the switches. Each switch will need unique IP addressing.

After defining the keepalive interface, you need to set up VSX. Enter the VSX context and define the role: **primary** or **secondary**. Next, specify the LAG that will be used for the ISL link. In the above example, this is LAG 128.

That is all that is necessary to initially set up VSX. For detailed information on configuring, tuning, and troubleshooting VSX, see the *Virtual Switching Extension (VSX) Guide*.

To set up a multi-chassis LAG between this VSX cluster and another device, define a new LAG, but specify the **multi-chassis** parameter: this parameter tells the VSX cluster that there is at least one link on each switch connecting to a neighboring device, like a switch.

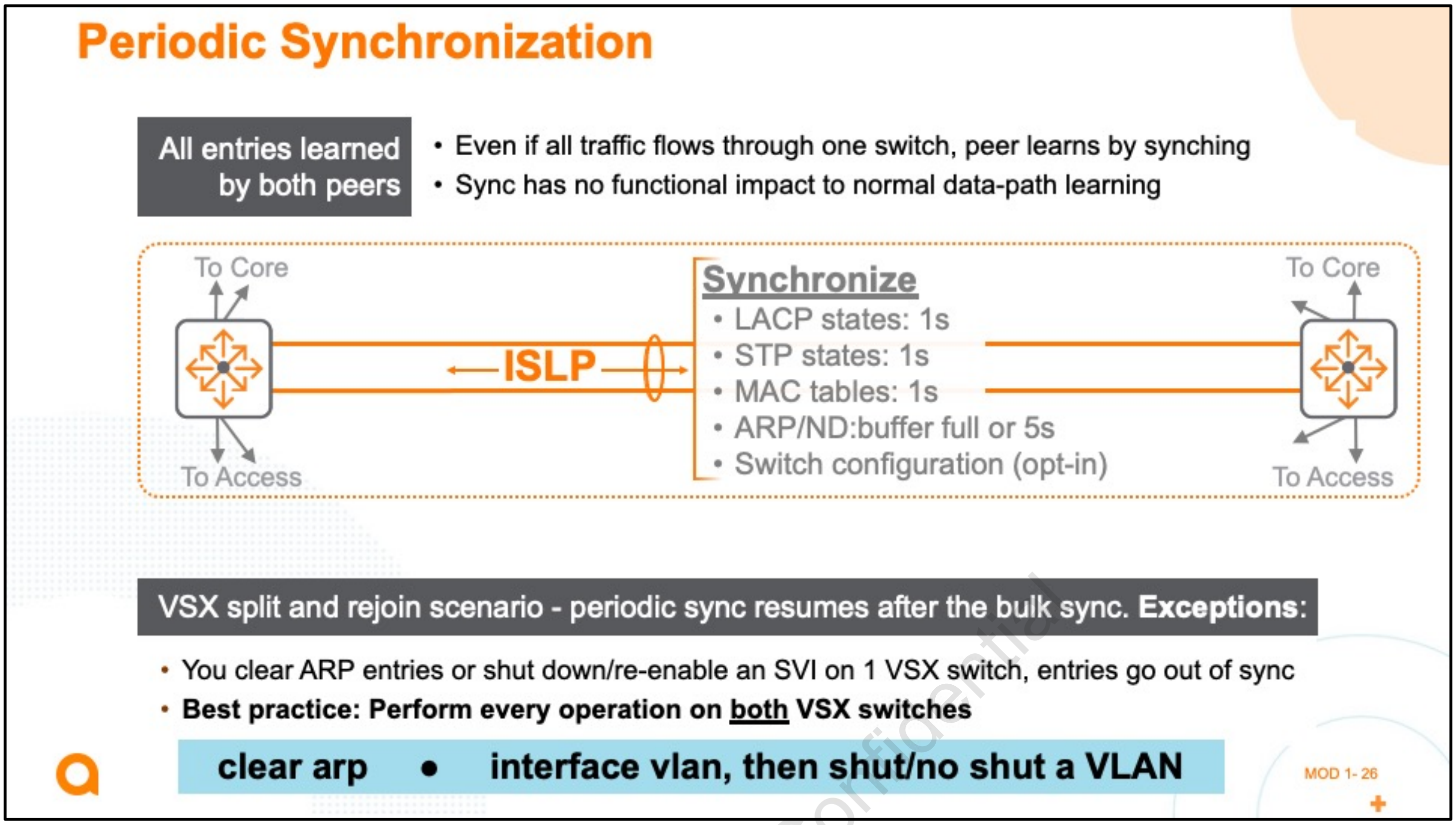
Aruba Training-Confidential

VSX Synchronization

MOD 1- 25

This section introduces what information is synchronized between a VSX pair, as well as the synchronization process itself.

Aruba Training-Confidential



Periodic Synchronization

All entries learned by both peers

- Even if all traffic flows through one switch, peer learns by synching
- Sync has no functional impact to normal data-path learning

Synchronize
- LACP states: 1s
- STP states: 1s
- MAC tables: 1s
- ARP/ND:buffer full or 5s
- Switch configuration (opt-in)

VSX split and rejoin scenario - periodic sync resumes after the bulk sync. **Exceptions:**

- You clear ARP entries or shut down/re-enable an SVI on 1 VSX switch, entries go out of sync
- **Best practice: Perform every operation on both VSX switches**

clear arp • interface vlan, then shut/no shut a VLAN

MOD 1- 26

Every second, each VSX switch synchronizes its learned MAC addresses with its VSX peer using ISLP. Each VSX switch also buffers and sends its new neighbor ARP/ND entries to the VSX peer every time its buffer is full (64 entries) or when it hits a timeout of 5 seconds.

| In a VSX scenario where all traffic from the core to the access layer flows through one VSX switch only, the other VSX switch still learns the necessary ARP/ND entries from its VSX peer. VSX has no functional impact on the VSX switches' normal datapath based learning.

| If a VSX split occurs and the switches then rejoin together, the switches perform a bulk sync; after that, periodic syncs resume.

Inter-VRF Route Leaking (IVRL) leaks neighbor entries from a source VRF to a destination VRF. This process creates two neighbor entries, one learned through the datapath in the source VRF and one that was induced in the destination VRF. The VSX switches do not sync the induced entries either during the initial ARP sync or the periodic syncs. Instead, each switch learns the ARP entries in the source VRF (either through the normal datapath process or through the ARP sync). The local IVRL process then induces the entries into the destination VRF on each VSX switch individually.

Keep one caveat about ARP synching in mind. If an admin clears ARP entries or enables and disables a VLAN interface on one VSX switch, but not on the other, ARP entries will go out of

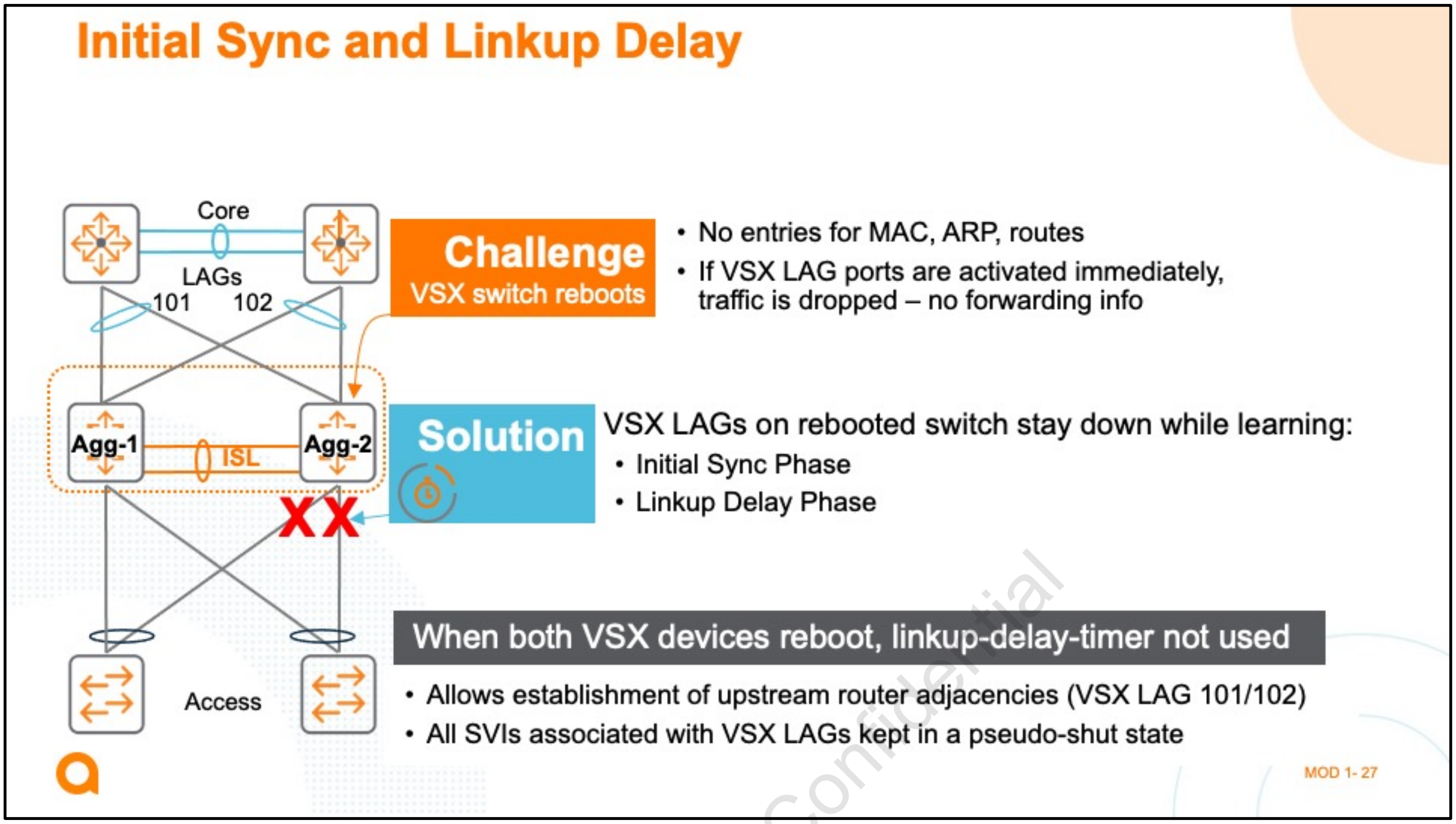
Aruba Training-Confidential

sync between the switches. The admin must perform such operations on both VSX switches so that both devices will re-learn entries and sync up.

| So, if you enter one or more of the following commands on one VSX switch, but not on the other VSX switch, the ARP entries on both switches will become unsynchronized.

- Clear ARP
- Interface VLAN
- Shutdown for a VLAN
- No shutdown for a VLAN

Aruba Training-Confidential



Initial Sync and Linkup Delay

Challenge
VSX switch reboots

- No entries for MAC, ARP, routes
- If VSX LAG ports are activated immediately, traffic is dropped – no forwarding info

Solution

VSX LAGs on rebooted switch stay down while learning:

- Initial Sync Phase
- Linkup Delay Phase

When both VSX devices reboot, linkup-delay-timer not used

- Allows establishment of upstream router adjacencies (VSX LAG 101/102)
- All SVIs associated with VSX LAGs kept in a pseudo-shut state

MOD 1- 27

Aruba Training-Confidential

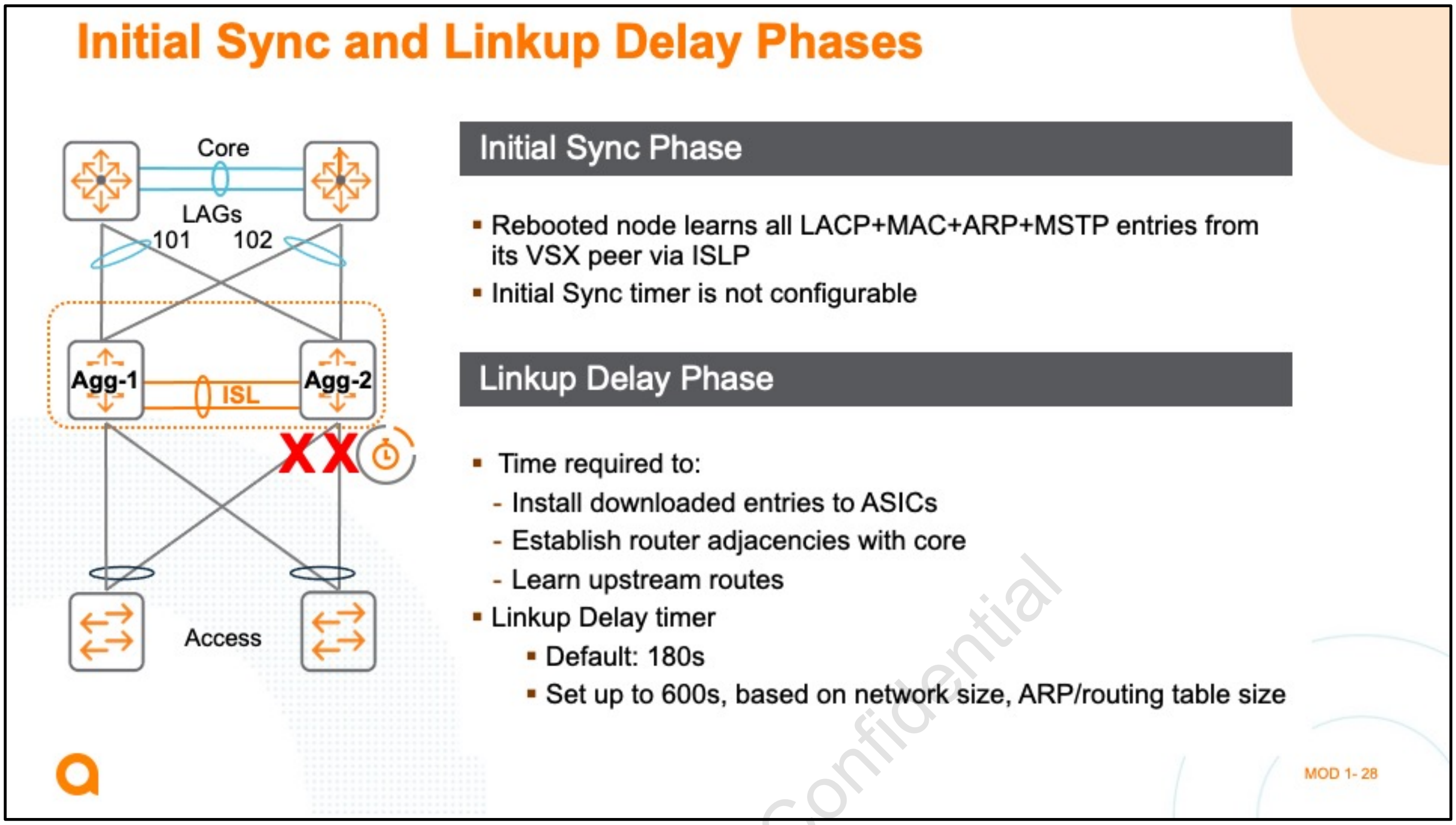
After a VSX switch reboots, it has no entries for ARP, MAC, and routes. If downstream VSX LAG ports activate before all this information is relearned, traffic drops. To avoid a traffic drop, VSX LAGs on the rebooted switch stay down until the restoration of LACP, MAC, ARP databases, and MSTP states, if MSTP is used.

The learning process for the VSX LAGs has two phases:

- **Initial sync phase**: The LACP states, MAC address table, ARP table, and potentially MSTP states download from the forwarding switch to the freshly rebooted switch.
- **Link-up delay phase**: The downloaded entries are installed into the ASIC. Router adjacencies with core nodes and learned upstream routes are also established. You can configure the link-up delay phase with the *linkup-delay-timer <DELAY-TIMER>* command. The default value is 180 seconds. Set the link-up delay timer to the maximum value of 600 seconds for a network with many MAC addresses, a large ARP table, or a large routing table.

When both VSX switches reboot, the link-up delay timer is not used because both switches must relearn the LACP states, MAC address table, and ARP table. To get upstream router adjacencies established during the link-up delay, you must exclude the upstream LAGs (LAG101, say) from the scope of the link-up delay. Run the *linkup-deal-timer exclude lag-list <LAG-LIST>* for identifying the LAGs for exclusion. Until linkup delay timer expires, all SVIs associated with VSX LAGs are kept in a pseudo-shut state.

Let’s get more detail on this two-phase learning process.



Initial Sync and Linkup Delay Phases

Initial Sync Phase

- Rebooted node learns all LACP+MAC+ARP+MSTP entries from its VSX peer via ISLP
- Initial Sync timer is not configurable

Linkup Delay Phase

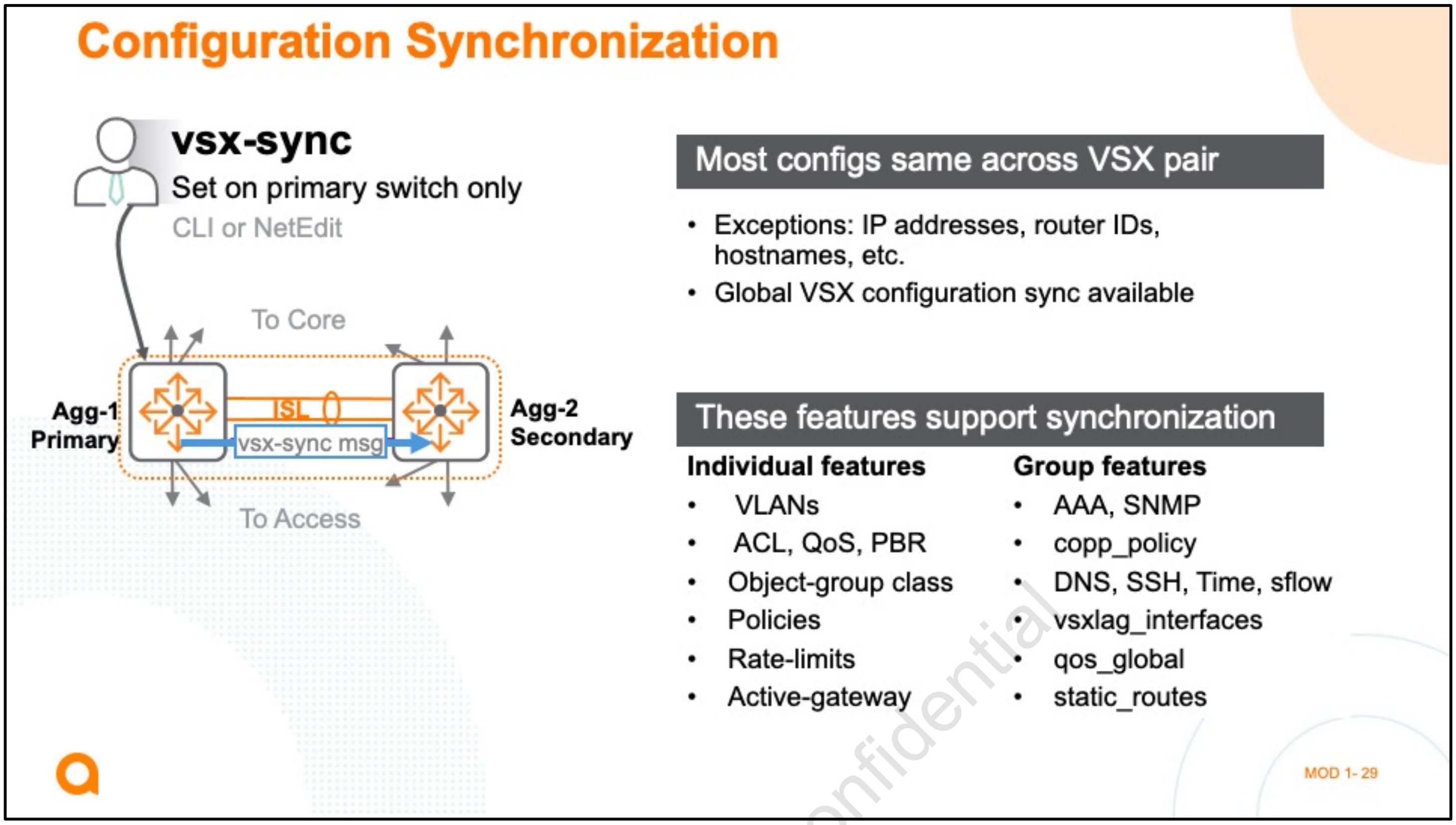
- Time required to:
  - Install downloaded entries to ASICs
  - Establish router adjacencies with core
  - Learn upstream routes
- Linkup Delay timer
  - Default: 180s
  - Set up to 600s, based on network size, ARP/routing table size

The figure summarizes the two phases of Linkup Delay.

The Initial Sync phase is the download phase, in which the rebooted node learns all the LACP, MAC, and ARP database entries from its VSX peer through ISLP. The Initial Sync timer is not configurable. It is the required time to download database information from the peer.

The Linkup Delay phase should last for as long as is required to install the downloaded entries to the ASIC, to establish router adjacencies with core devices, and to learn upstream routes. During this phase, links are forced down. The Linkup Delay timer's default value is 180 seconds. Depending on the network size, ARP table size, and routing table size, the timer may need to be set to a higher value - the maximum is 600 seconds.

Aruba Training-Confidential



Configuration Synchronization

Most configs same across VSX pair

- Exceptions: IP addresses, router IDs, hostnames, etc.
- Global VSX configuration sync available

These features support synchronization

Individual features

- VLANs
- ACL, QoS, PBR
- Object-group class
- Policies
- Rate-limits
- Active-gateway

Group features

- AAA, SNMP
- copp_policy
- DNS, SSH, Time, sflow
- vsxlag_interfaces
- qos_global
- static_routes

MOD 1- 29

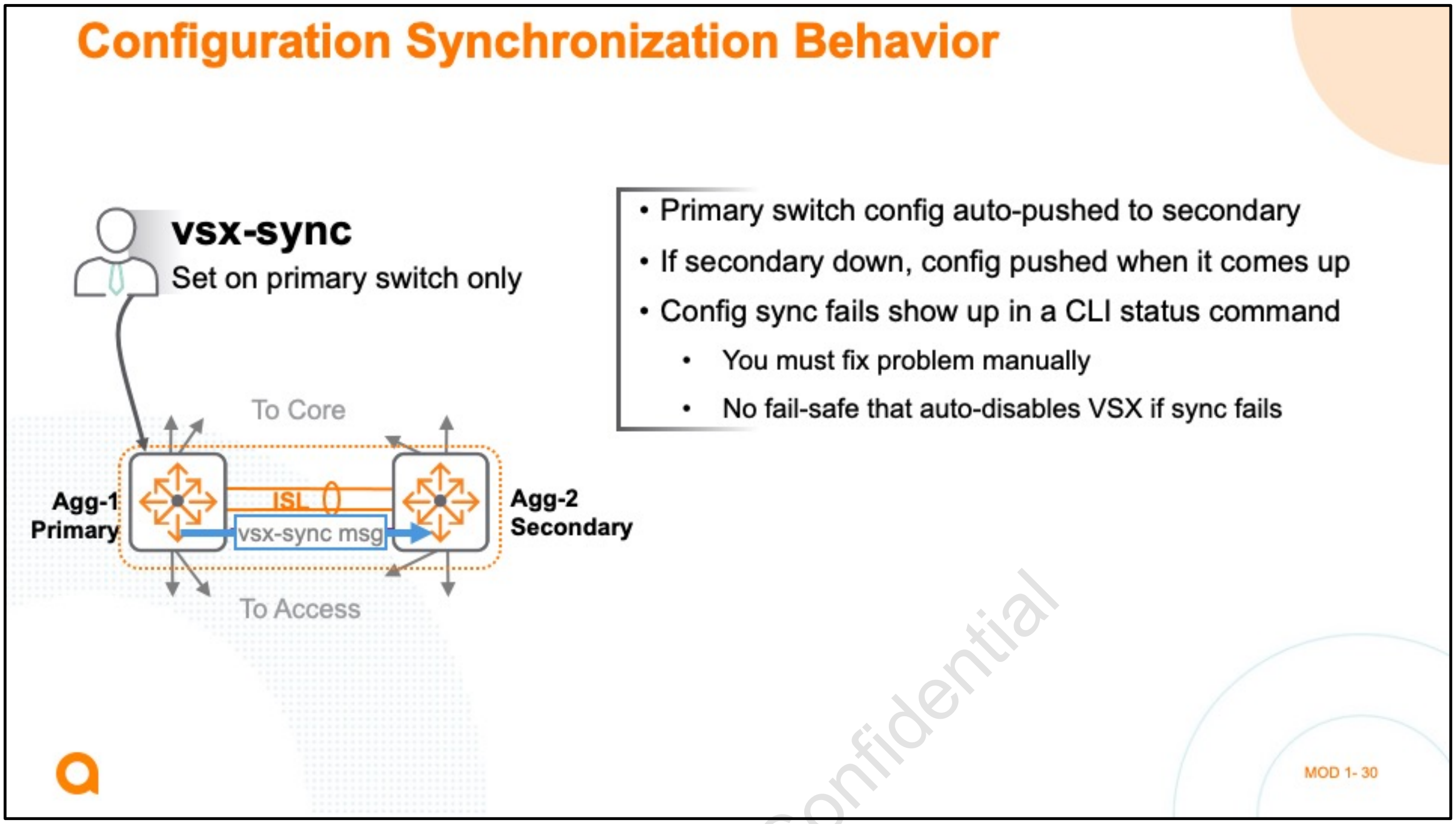
Aruba Training-Confidential

Dataplane sync and config sync are two different things. Dataplane sync was covered earlier. Let's dive into config-sync (vsx-sync).

You can use the **vsx-sync** parameter from the CLI or NetEdit to ensure that the configurations are "synced" between the VSX peers.

The two VSX switches should have nearly the same configuration with a few exceptions. These exceptions include host IP addresses on Layer 3 interfaces, router IDs, hostnames, and a few other settings. Configuration synchronization makes it simpler to ensure identical configs. By default, VSX configuration synchronization is enabled along with VSX. However, enabling synchronization globally has no effect until you select individual configuration items on which to enable synchronization. The **vsx-sync** attribute enables synchronization on an item. You can disable configuration synchronization globally, in which case you can no longer apply the **vsx-sync** attribute to any configuration items.

The figure shows individual and group features supported by vsx-sync.



Configuration Synchronization Behavior

- Primary switch config auto-pushed to secondary
- If secondary down, config pushed when it comes up
- Config sync fails show up in a CLI status command
  - You must fix problem manually
  - No fail-safe that auto-disables VSX if sync fails

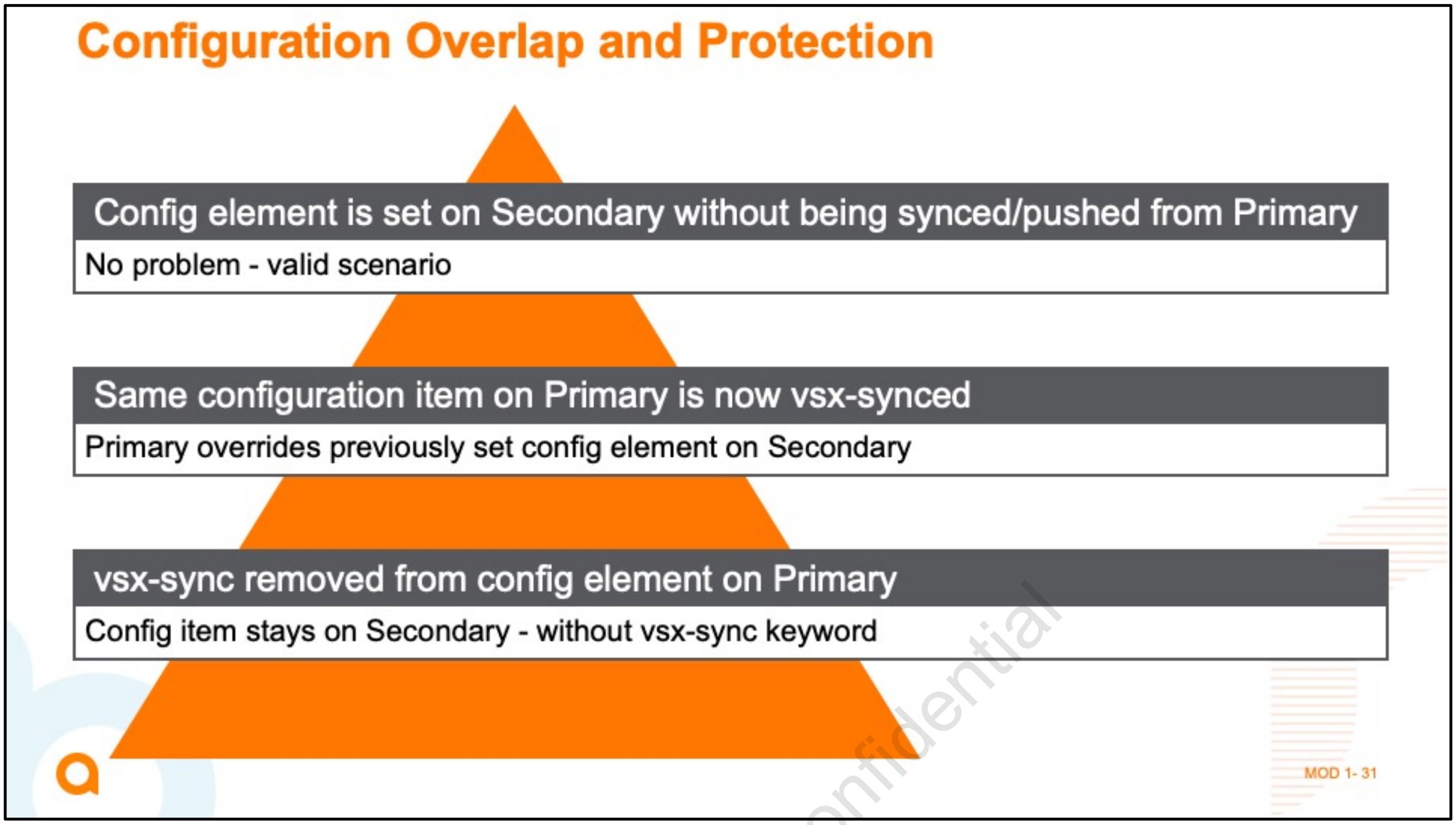
MOD 1- 30

You must set the **vsx-sync** attribute on the desired items from the VSX primary switch. The primary switch the automatically pushes this config to the secondary switch. If the secondary switch is not available when you configure the primary, the configuration is pushed to the secondary switch after the switch comes up. If a certain config sync operation fails, the CLI status commands will indicate the failure. You must manually fix the inconsistency. There is no fail-safe mechanism that disables VSX if configuration synchronization fails.

You can configure **vsx-sync** on a secondary switch if config-sync is disabled or if the ISL is down.

The **vsx-synch** attribute uses this “opt-in” model due to customers’ and partners’ preference that they be able to control exactly which features are synched.

Aruba Training-Confidential



Configuration Overlap and Protection

Config element is set on Secondary without being synced/pushed from Primary

No problem - valid scenario

Same configuration item on Primary is now vsx-synced

Primary overrides previously set config element on Secondary

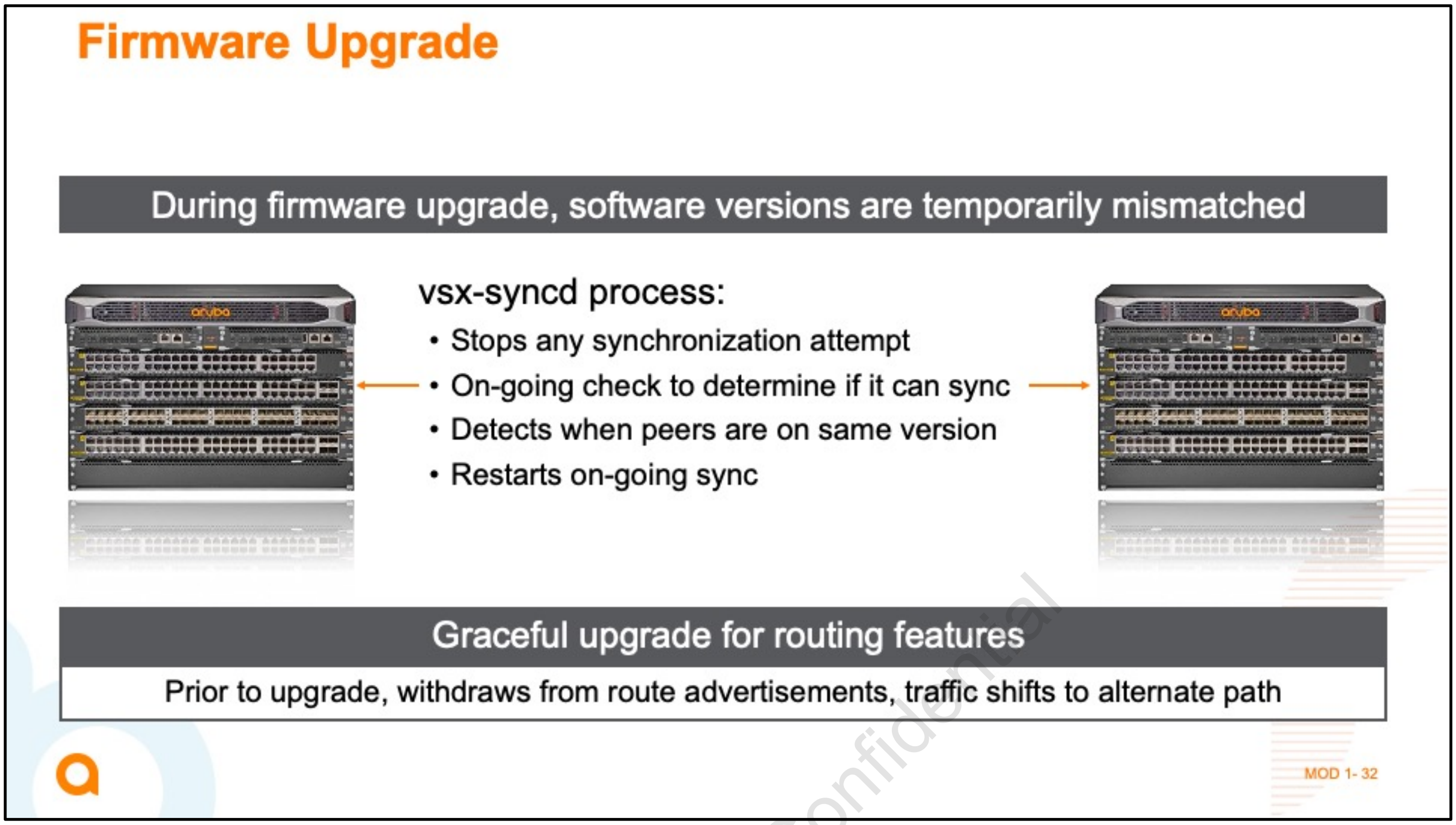
vsx-sync removed from config element on Primary

Config item stays on Secondary - without vsx-sync keyword

MOD 1- 31

Aruba Training-Confidential

It is perfectly valid to configure a setting manually on the secondary switch without that setting being synched from the primary switch. However, if vsx-synch is now enabled for this feature on the primary switch, the setting on the primary switch overrides the previous setting on the secondary switch. For example, you could add VLAN 10 on the secondary switch with description “employees.” You then add VLAN 10 with no description on the primary switch and enable vsx-sync. The secondary switch now has VLAN 10, but with no description.



Firmware Upgrade

During firmware upgrade, software versions are temporarily mismatched

vsx-syncd process:

- Stops any synchronization attempt
- On-going check to determine if it can sync
- Detects when peers are on same version
- Restarts on-going sync

Graceful upgrade for routing features

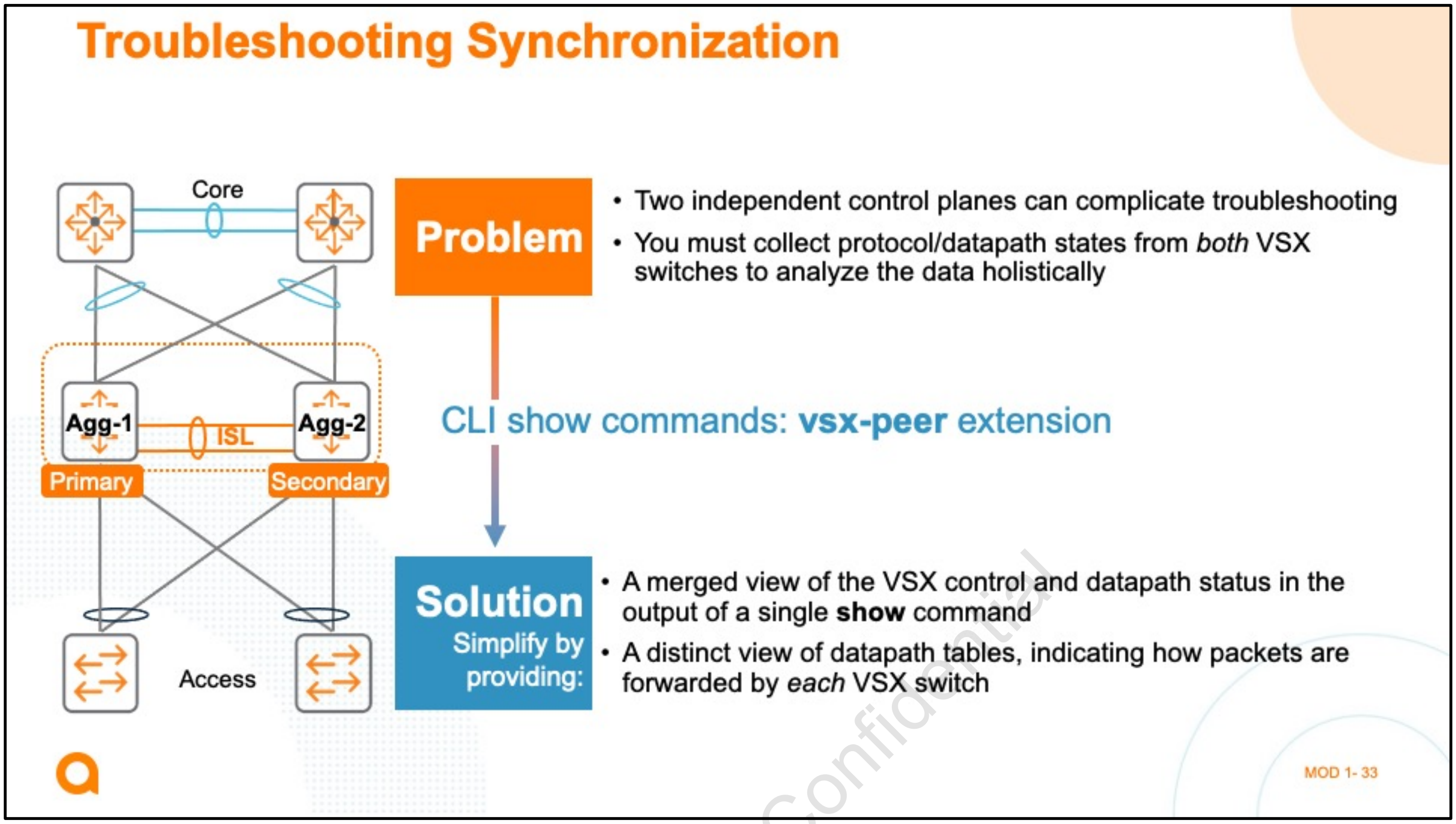
Prior to upgrade, withdraws from route advertisements, traffic shifts to alternate path

MOD 1- 32

Aruba Training-Confidential

During a firmware upgrade, software versions will temporarily mismatch on the VSX peers. Therefore, the vsx-syncd process, which is still running, stops any synchronization attempts. It then performs an ongoing check to determine if it can sync. Once both VSX peers are running the same firmware, this ongoing check sees that it can restore the synchronization. The vsx-syncd process then continues the on-going sync. The database connection between the peers works if the firmware versions are the same.

As of AOS-CX v10.2, VSX also supports graceful upgrade for routing features. When a member is about to be updated, it takes action such as withdrawing itself from route advertisements, so traffic is shifted away from it during the upgrade process.



Troubleshooting Synchronization

Problem

- Two independent control planes can complicate troubleshooting
- You must collect protocol/datapath states from *both* VSX switches to analyze the data holistically

CLI show commands: **vsx-peer** extension

Solution
Simplify by providing:

- A merged view of the VSX control and datapath status in the output of a single **show** command
- A distinct view of datapath tables, indicating how packets are forwarded by *each* VSX switch

MOD 1- 33

While independent control planes and management planes for the VSX switches help to provide greater availability, they can complicate troubleshooting. You must collect protocol and datapath states from both VSX switches to analyze the issue holistically.

To simplify troubleshooting, the AOS-CX CLI offers a **vsx-peer** option for many show commands. When you enter a show command with this option, the output merges data from both the local switch and its VSX peer in a single output, making it easier to collect information and troubleshoot. The output also shows the distinct datapath tables on each switch, indicating how each switch will forward packets, which again helps pinpoint issues.

The **vsx-peer** option for show commands is not available during a firmware upgrade process.

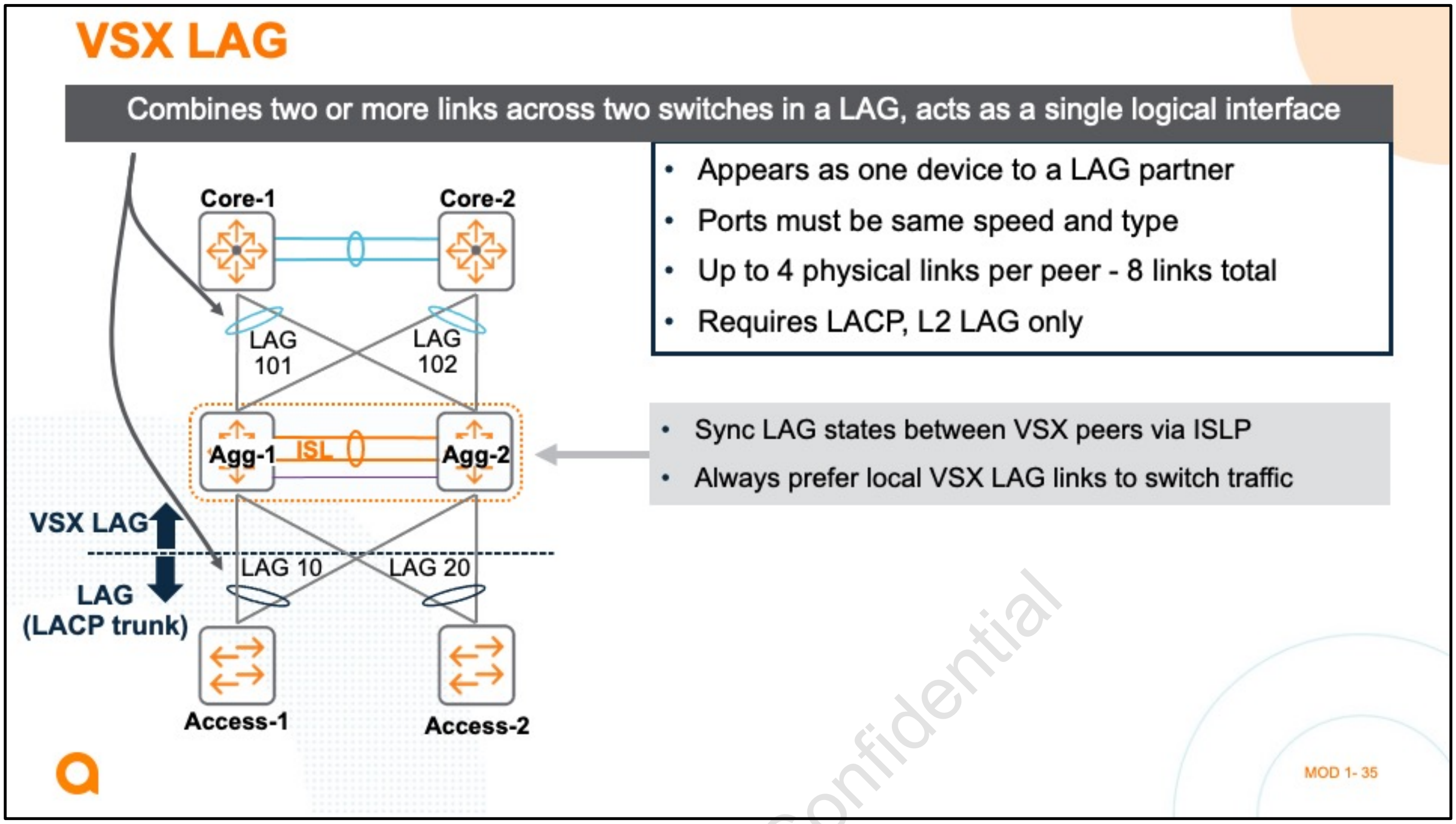
Aruba Training-Confidential

Traffic Forwarding

MOD 1- 34

This section introduces how a VSX cluster processes and forwards traffic.

Aruba Training-Confidential



VSX LAG

Combines two or more links across two switches in a LAG, acts as a single logical interface

- Appears as one device to a LAG partner
- Ports must be same speed and type
- Up to 4 physical links per peer - 8 links total
- Requires LACP, L2 LAG only

- Sync LAG states between VSX peers via ISLP
- Always prefer local VSX LAG links to switch traffic

VSX LAG allows you to connect a switch to both physical peers in a VSX fabric. Here switch Access-1 connects to both Agg-1 and Agg-2 of the VSX pair and perceives a standard LAG link – a single logical interface to a single logical switch. The same holds true for Access-2, and the core switches.

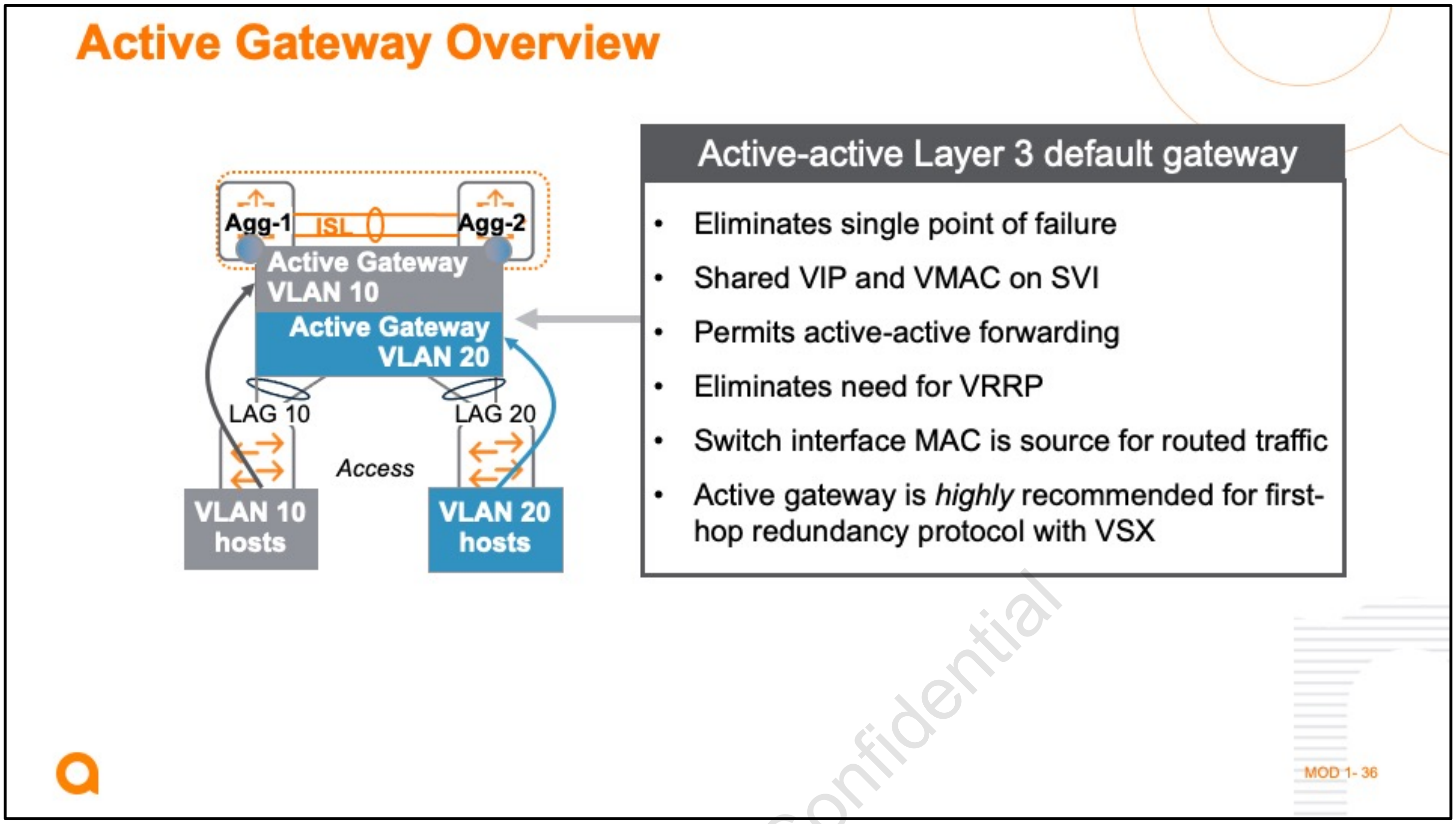
The two VSX switches appear as a single peer ID to the upstream or downstream switches. Ports must have the same speed. The LAG can include up to four physical links per peer for eight links max per VSX LAG. As of the publication of this course, VSX LAGs are layer 2 only, and VSX LAGs must be LACP-based. A non-LACP or static VSX LAG is not supported.

The VSX peer switches synchronize their LAG states using ISLP. You can configure a lag-hash scheme, which dictates how the switches load-share traffic across the links in the LAG. (Note: The 8400 provides L4-based hashing, but the 8320 does not.) VSX LAGS are locally optimized, meaning if a switch with local links is in a LAG then it constrains the LAG hash-scheme to look at only local links, and it will only forward traffic over the ISL for transmission on the peer links in the LAG if local links of a VSX LAG are down.

The number of VSX LAGs supported is dependent on the switch model: 255 on 8400, 53 on 8320, 55 on 8325.

Note: Originally, Aruba used the term multi-chassis LAG (MC-LAG) to describe a LAG between an external device and a VSX switch pair. Today, Aruba uses the term *VSX LAG*.

Aruba Training-Confidential



Active Gateway Overview

Active-active Layer 3 default gateway

- Eliminates single point of failure
- Shared VIP and VMAC on SVI
- Permits active-active forwarding
- Eliminates need for VRRP
- Switch interface MAC is source for routed traffic
- Active gateway is *highly* recommended for first-hop redundancy protocol with VSX

MOD 1- 36

Aruba Training-Confidential

Active gateway is a first-hop redundancy protocol that eliminates a single point of failure for default gateway services on an access network. The active gateway feature improves the reliability and performance of the host network by enabling a virtual router to act as the default gateway for that network.

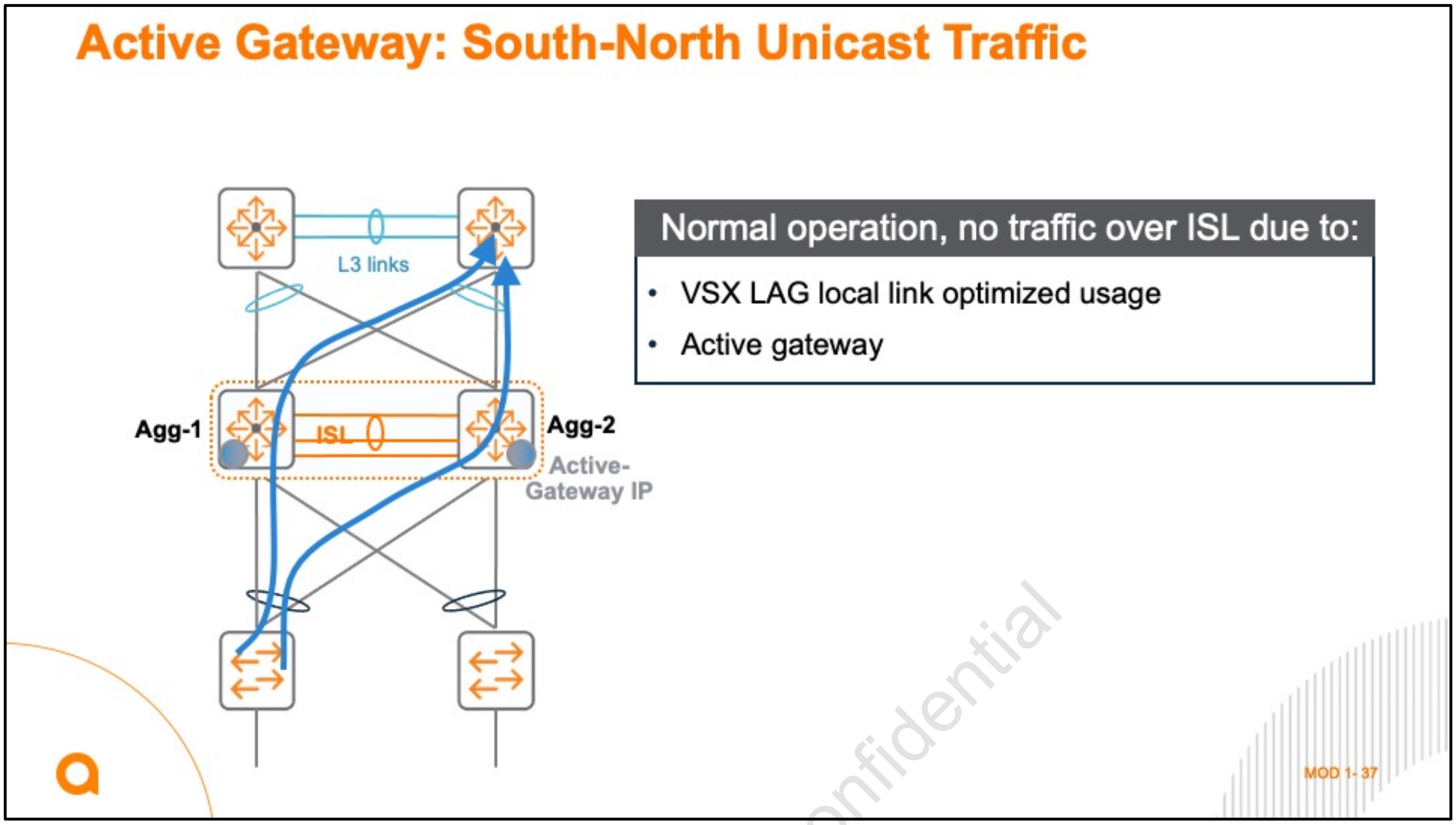
To configure this feature, you simply configure a shared virtual IP address (VIP) and a shared virtual MAC address (VMAC) on an SVI on both VSX switches. The VIP/VMAC serves as the default gateway for each access VLAN. Which peer in the VSX pair receives a particular piece of traffic from an access VLAN depends on which link in the LAG the downstream access switch's LAG has selects. Whichever peer receives the traffic routes it across to the layer 3 domain.

When you use active gateway on a VLAN, the VLAN has no need for VRRP; the features are mutually exclusive. As with VRRP, routed traffic from the VSX peer is sourced from the switch interface MAC, not the VMAC address. Every three minutes (non-configurable), each active gateway broadcasts a hello packet from the VMAC to avoid VMAC aging on the access switches. The hello from both peers transmits on the same VSX LAG, so no issues occur with the same MAC address appear on different ports. The hello is a non-IP Ethernet packet using HP EtherType, 080009 Hewlett-Packard, and a destination reserved multicast address.

The VSX pair supports up to 16 different VMACs, which are not shareable between IPv4 and IPv6. For example, the pair could use eight VMACs for IPv4 simultaneously with eight VMACs for IPv6.

You should understand some key differences between VRRP and the active gateway feature. VRRP requires multiple configuration options, including the VRID on the VLAN, the roles on each VRID, the virtual IP address on each VRID, the advertisement timer configurations, and so on. Active gateway only requires a single line of configuration per VLAN interface: the configuration of the virtual IP/MAC. VRRP and active gateway also differ in terms of how they forward traffic. VRRP uses an active-standby dataplane with only the VRRP master routing traffic. With active gateway, both devices forward traffic.

Aruba Training-Confidential



Active Gateway: South-North Unicast Traffic

Normal operation, no traffic over ISL due to:

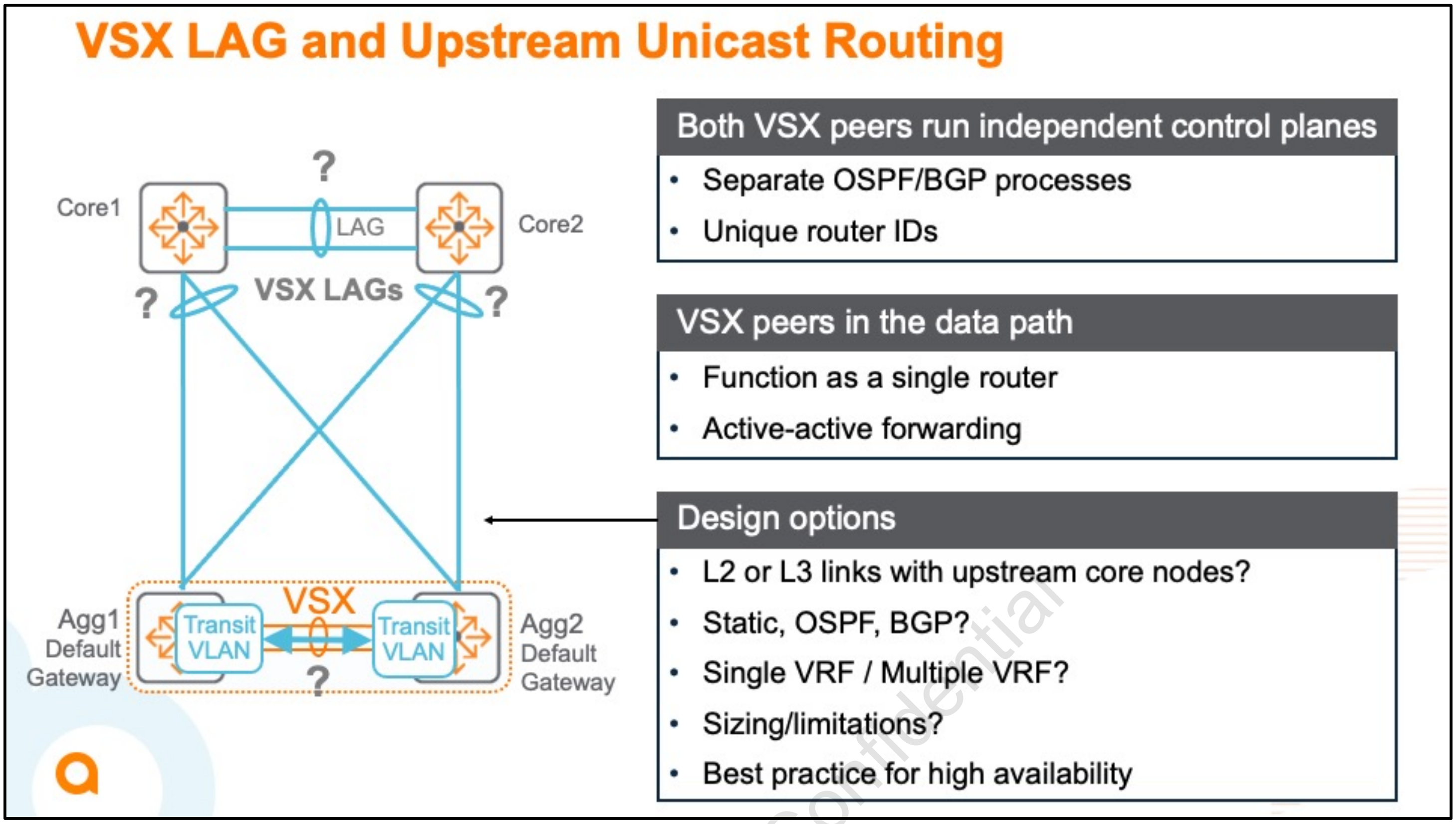
- VSX LAG local link optimized usage
- Active gateway

MOD 1- 37

Aruba Training-Confidential

If you configure VSX LAGs with links on each peer, the normal operation unicast traffic sent between the access layer and the core (south-north) uses all access-to-aggregation links and does not flow over the ISL. From a Layer 2 perspective, VSX LAGs span two switches and operate in active-active mode. Traffic between the access layer and aggregation layer switches can be forwarded to any of the active links. There are no loops and no need for spanning tree protocol or blocked ports.

From a datapath perspective, local link optimized usage ensures that when a VSX switch receives traffic, it always uses its local LAG links to forward the traffic to its destination. The VSX switch only uses the ISL link if the local LAG links are down. An active gateway ensures that the first VSX switch that receives traffic from the access layer can route the traffic across to the Layer 3 domain.



VSX LAG and Upstream Unicast Routing

Both VSX peers run independent control planes

- Separate OSPF/BGP processes
- Unique router IDs

VSX peers in the data path

- Function as a single router
- Active-active forwarding

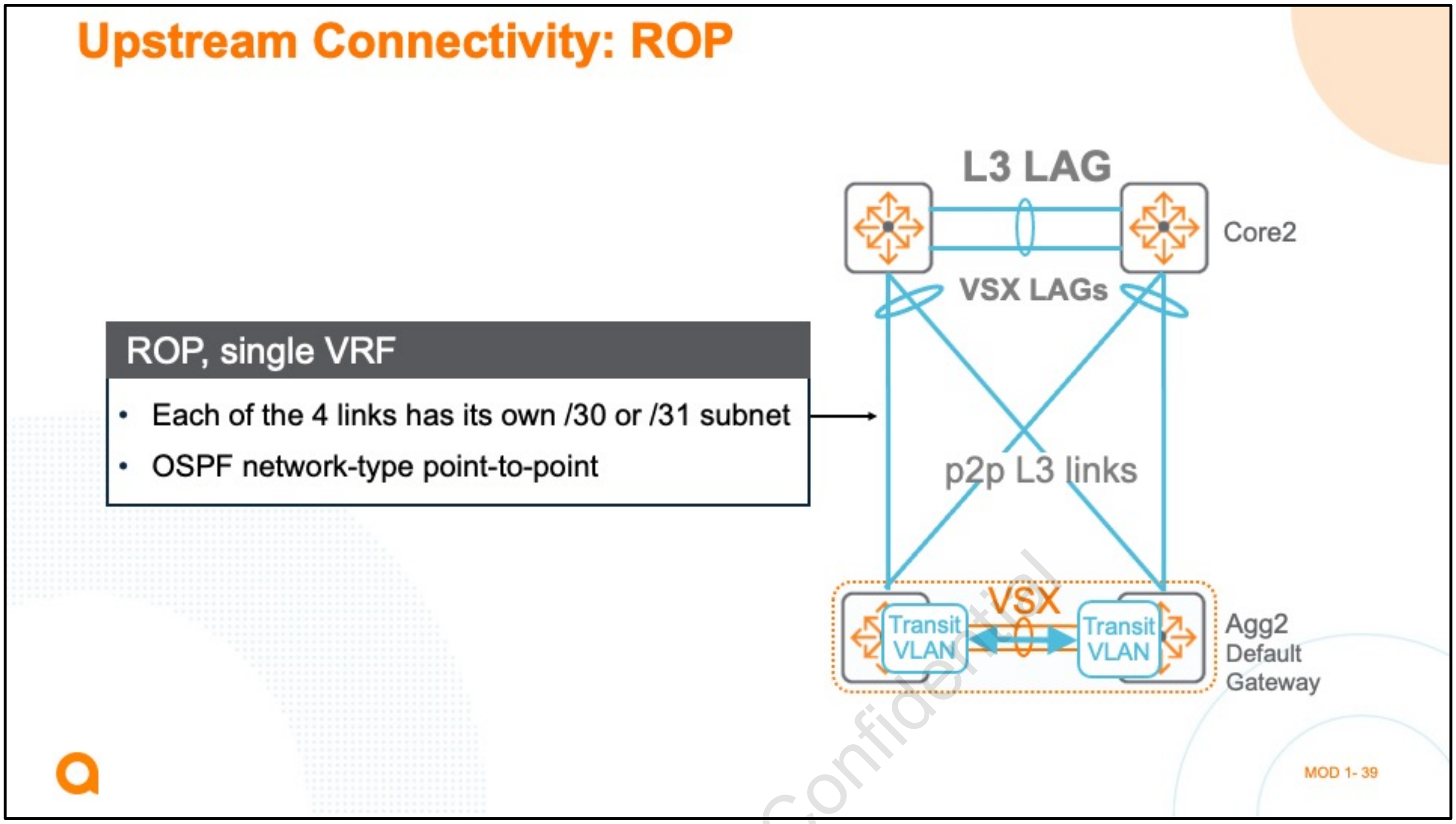
Design options

- L2 or L3 links with upstream core nodes?
- Static, OSPF, BGP?
- Single VRF / Multiple VRF?
- Sizing/limitations?
- Best practice for high availability

Aruba Training-Confidential

You will now examine how the VSX pair operates on the connections to upstream core switches, which are generally routed connections. Remember that each VSX peer runs an independent control plane with its own OSPF or BGP processes. Therefore, each VSX peer appears as a different router with its own router ID to the core switches. However, as far as the data path is concerned, the VSX peers function as a single router and support active-active forwarding. This is possible because a core switch uses Equal Cost Multipath (ECMP) routing to learn routes to the subnets behind the VSX pair through both VSX peers and because the VSX pair supports a feature called “active forwarding,” about which you will soon learn. Also remember that VSX LAGs can only operate at Layer 2.

Keeping these principles in mind, you can consider the options that you have for designing the upstream VSX LAGs and unicast routing. You could choose to connect with routed ports or Layer 2 ports associated with Layer 3 VLAN interfaces. Or you can use Layer 2 VSX LAGs, which are associated with Layer 3 VLAN interfaces. For routing, you could choose static routing, OSPF, or BGP. You could use a single VRF or multiple VRFs, depending on the requirements of the scenario. Your choices will be informed by sizing limitations and by best practices for high availability.



Upstream Connectivity: ROP

ROP, single VRF

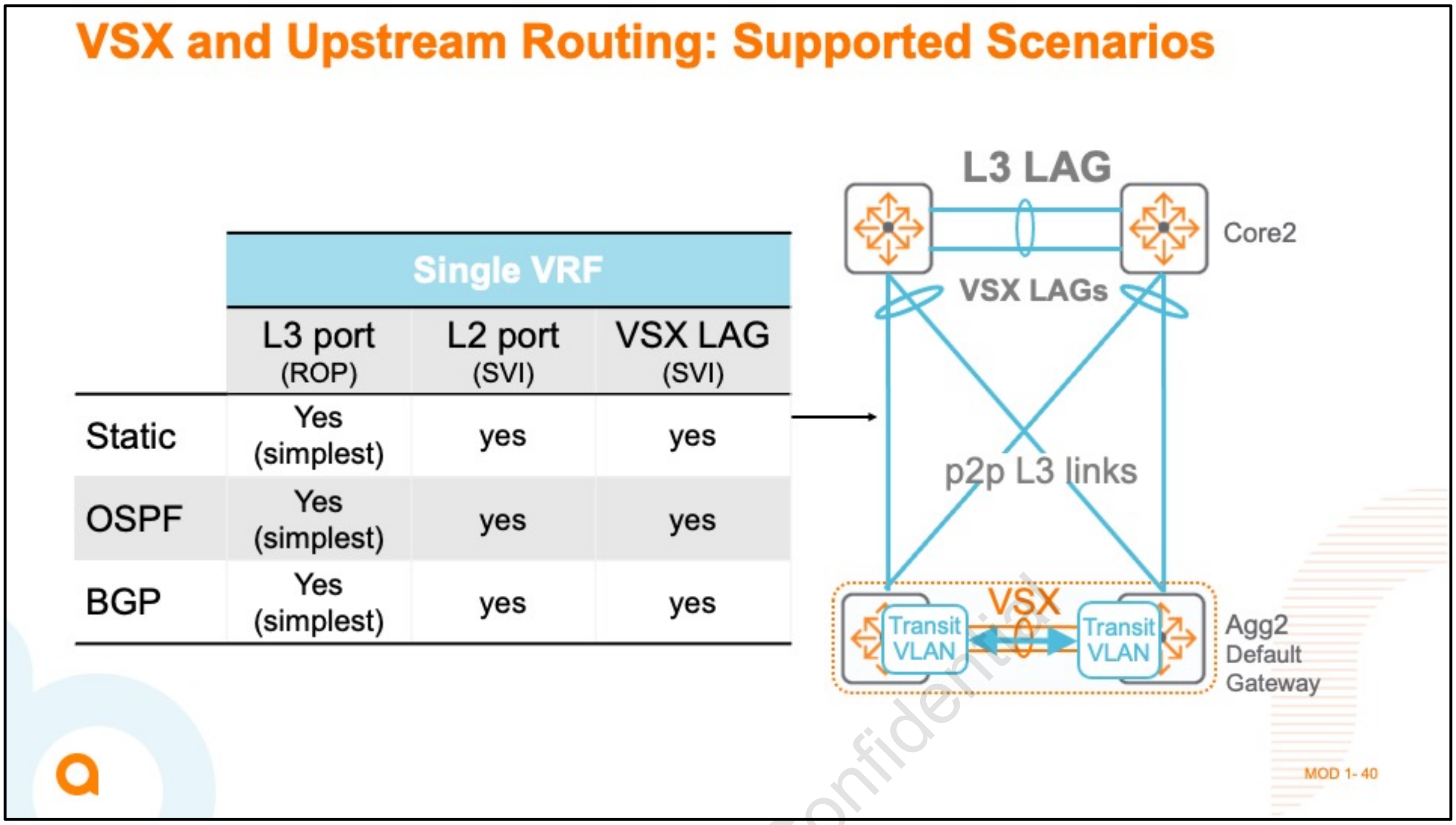
- Each of the 4 links has its own /30 or /31 subnet
- OSPF network-type point-to-point

There are several options for links between the aggregation VSX pair and the core: one involves a single VRF implementation while the other options involve multiple VRFs.

The first option uses Route-Only Ports (ROP). It could also use routed LAGs, but those routed LAGs would only have links on the local switch; they cannot be multi-chassis. Each VSX switch has its own routed link or LAG to each core switch, and each of these four links is its own /30 or /31 subnet. In OSPF, this would be a point-to-point network. This simple VLAN-free configuration is best suited for a network that runs on a single VRF domain. If the network had multiple VRFs, each VRF would need its own ROP, which is less feasible.

The other options are beyond the scope of this course.

Aruba Training-Confidential



VSX and Upstream Routing: Supported Scenarios

	Single VRF		
	L3 port (ROP)	L2 port (SVI)	VSX LAG (SVI)
Static	Yes (simplest)	yes	yes
OSPF	Yes (simplest)	yes	yes
BGP	Yes (simplest)	yes	yes

MOD 1- 40

The following chart summarizes the supported scenarios for the upstream connections between a VSX pair and the core. When you have a single VRF, ROPs provide the simplest option. However, Layer 2 ports associated with dedicated SVIs, as well as Layer 2 VSX LAGs associated with dedicated SVIs, work as well.

All of these options work with static routing, OSPF, or BGP.

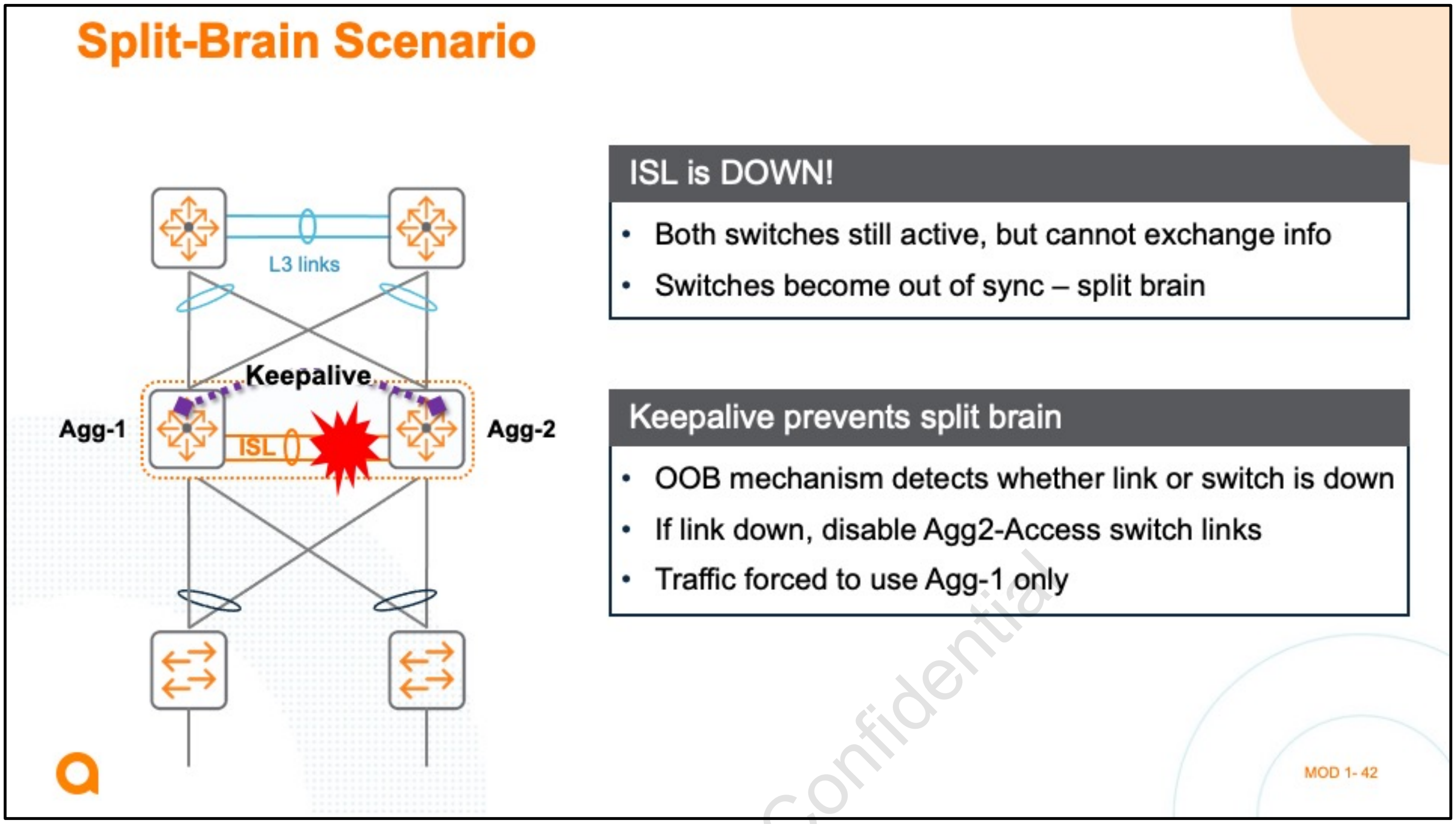
Aruba Training-Confidential

Split-Brain Condition

MOD 1- 41

This section introduces how a pair of VSX switches determine how an ISL link failure is detected and how the switches deal with the issue, given that they have the same virtual IP and MAC addresses for data forwarding processes.

Aruba Training-Confidential



Split-Brain Scenario

ISL is DOWN!

- Both switches still active, but cannot exchange info
- Switches become out of sync – split brain

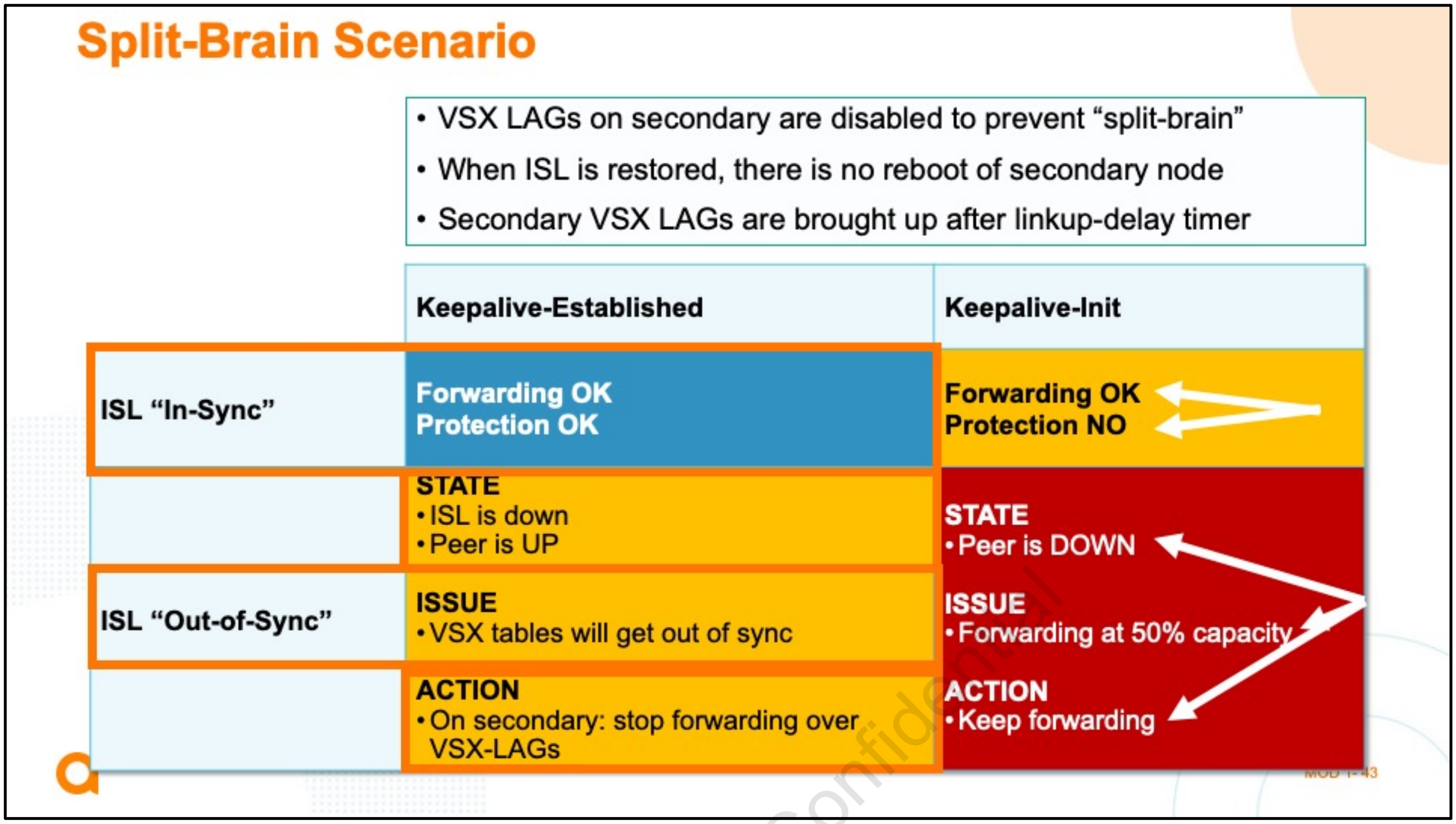
Keepalive prevents split brain

- OOB mechanism detects whether link or switch is down
- If link down, disable Agg2-Access switch links
- Traffic forced to use Agg-1 only

MOD 1- 42

The diagram shows a situation in which the ISL between the two VSX switches is down. Both switches are still active, but the switches cannot exchange information, which causes the switches to become out of sync. This situation is called a split brain, and it occurs if the keepalive function is not enabled. The keepalive function prevents split brain issues by providing an out-of-band mechanism to detect if the link between the VSX switch is down or if the VSX switches are not working. The keepalive functionality brings down the link between Agg2 and Access Switches. The traffic is forced to go from access switches to Agg1 only.

Aruba Training-Confidential



Split-Brain Scenario

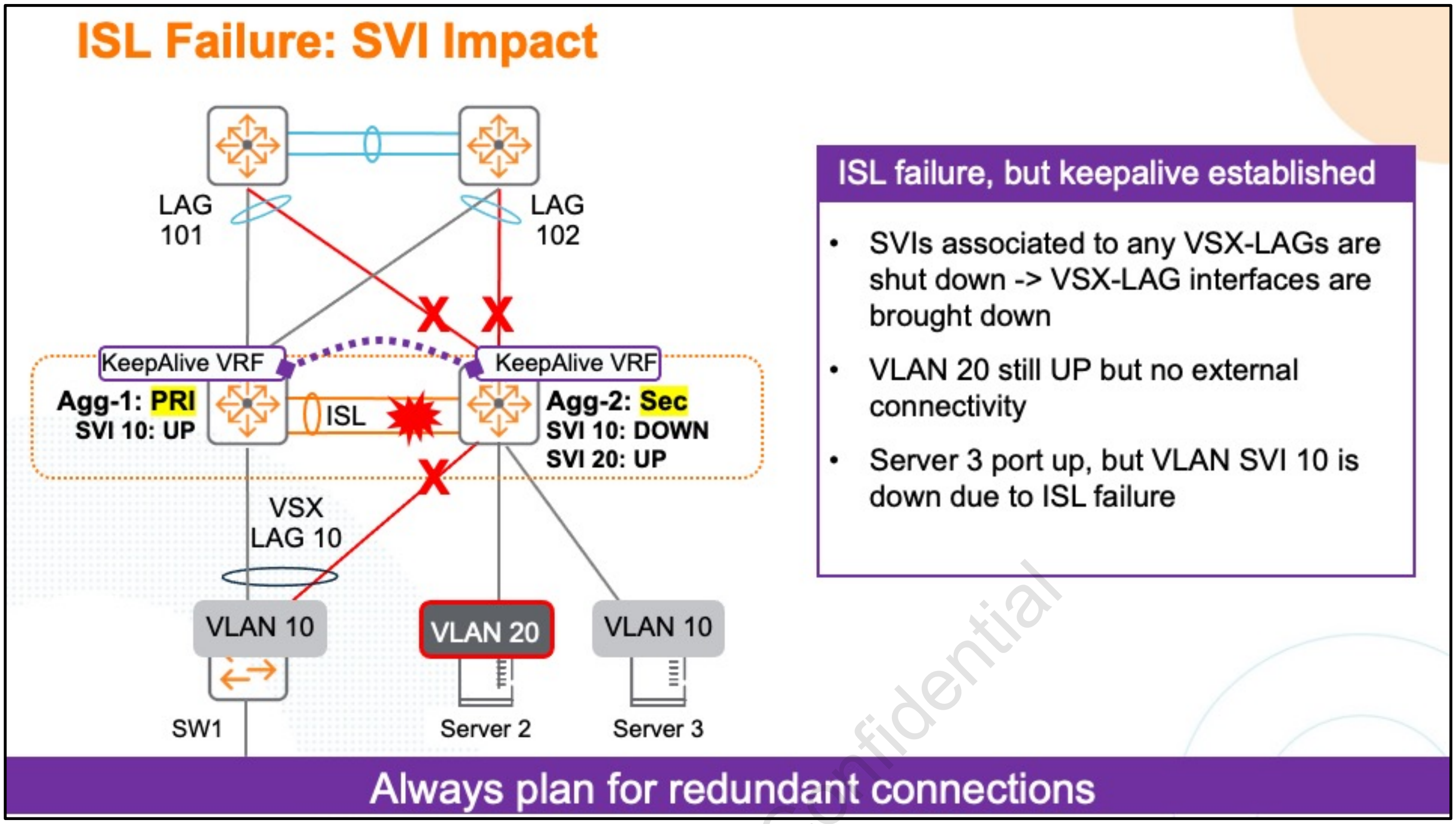
- VSX LAGs on secondary are disabled to prevent "split-brain"
- When ISL is restored, there is no reboot of secondary node
- Secondary VSX LAGs are brought up after linkup-delay timer

	Keepalive-Established	Keepalive-Init
ISL "In-Sync"	Forwarding OK Protection OK	Forwarding OK Protection NO
	STATE • ISL is down • Peer is UP	STATE • Peer is DOWN
ISL "Out-of-Sync"	ISSUE • VSX tables will get out of sync	ISSUE • Forwarding at 50% capacity
	ACTION • On secondary: stop forwarding over VSX-LAGs	ACTION • Keep forwarding

Aruba Training-Confidential

The table shows what happens with various ISL and keepalive statuses on the VSX pair, as described below.

- If ISL is up and the peers are in sync: VSX pair forwards traffic normally
- If keepalive connection is up, system is protected from split-brain condition
- If keepalive connection is down, but ISL is fine: VSX pair can forwards traffic normally, but system is at risk. (**Protection No** in the table)
- If ISL goes down and peers are out of sync, but keepalive is up: Each switch knows that their peer is still up. They know that they are at risk of becoming out of sync, so the secondary switch shuts down its VSX LAGs.
- If ISL goes down and keepalive is down: The switch assumes that its peer is down and that the system is functional but forwarding at 50 percent capacity. It keeps forwarding even if it is the secondary peer. It is important that you form the keepalive connection correctly so that this assumption is correct.



ISL failure, but keepalive established

- SVIs associated to any VSX-LAGs are shut down -> VSX-LAG interfaces are brought down
- VLAN 20 still UP but no external connectivity
- Server 3 port up, but VLAN SVI 10 is down due to ISL failure

Always plan for redundant connections

An ISL cut has no effect on the up/down state of VLANs and associated SVIs that are not part of any VSX LAG but part of at least one orphan port (beside the ISL LAG, which is not a VSX LAG).

Aruba Training-Confidential

<u>During ISL cut (before initial sync)</u>

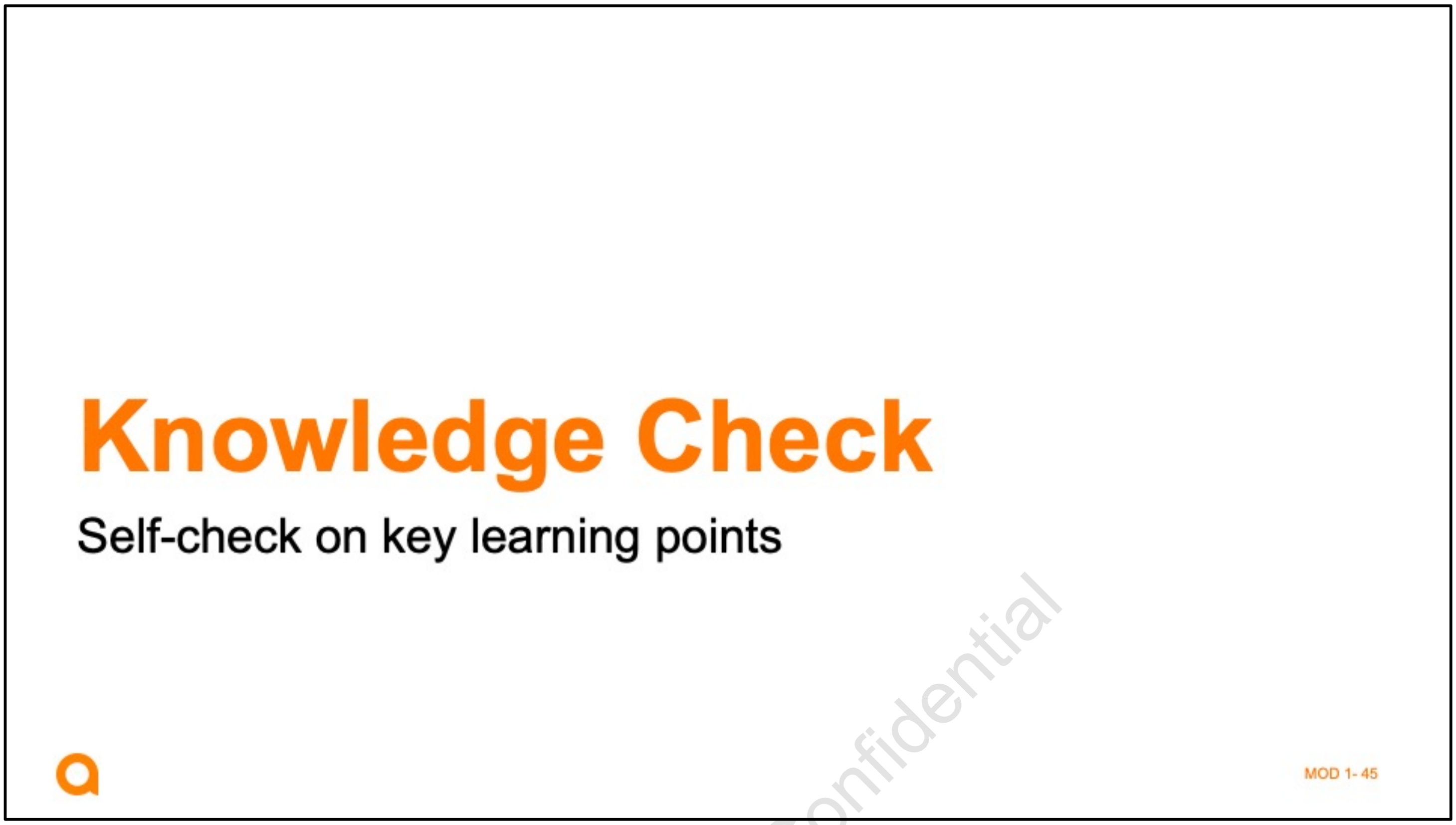
If the secondary VSX node has a port that is a member of a VSX LAG, the associated SVI of the VLAN – transported by said VSX LAG – has the VSX secondary node shut down, whether or not there are orphan ports carrying that given VLAN.

<u>During link-up delay timer (after initial sync)</u>

As long as the secondary VSX node has a port that is a member of a VSX LAG, the associated SVI of the VLAN – transported by said VSX LAG – is kept shut down on the VSX secondary node, whether there is an orphan port(s) carrying that given VLAN or not. The associated SVI of the VLAN transported by a VSX LAG restores ON/UP on the VSX secondary node, only if the following conditions are met:

- You have used the command **link-up delay timer exclude lag** to exclude the said VSX LAG, and
- No VSX LAG not part of the exclusion set allows the given VLAN.

In the topology above, after an ISL failure, ports and SVI associated to a VSX LAG are automatically shut down, ports and SVIs not related to and VSX LAG still operating. In this case, server 2 and VLAN 20 are still operating, but since there are no external uplinks and routes, server 2 is isolated. The server's 3 port remains UP, but VLAN 10 SVI shutdown due to ISL failure, as a best practice, always plan for redundant connections at the data center.

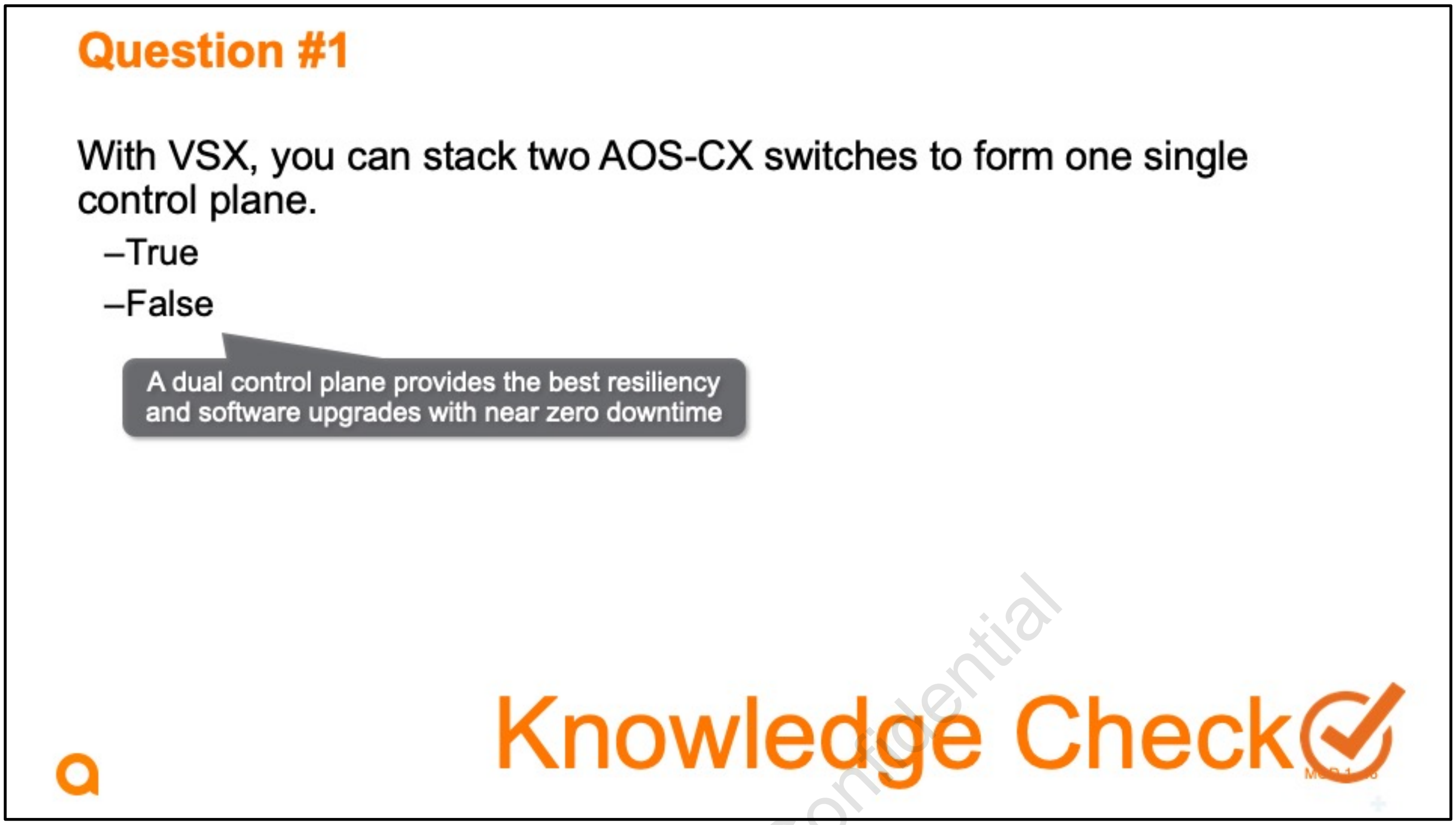


Knowledge Check

Self-check on key learning points

MOD 1- 45

Aruba Training-Confidential



Question #1

With VSX, you can stack two AOS-CX switches to form one single control plane.

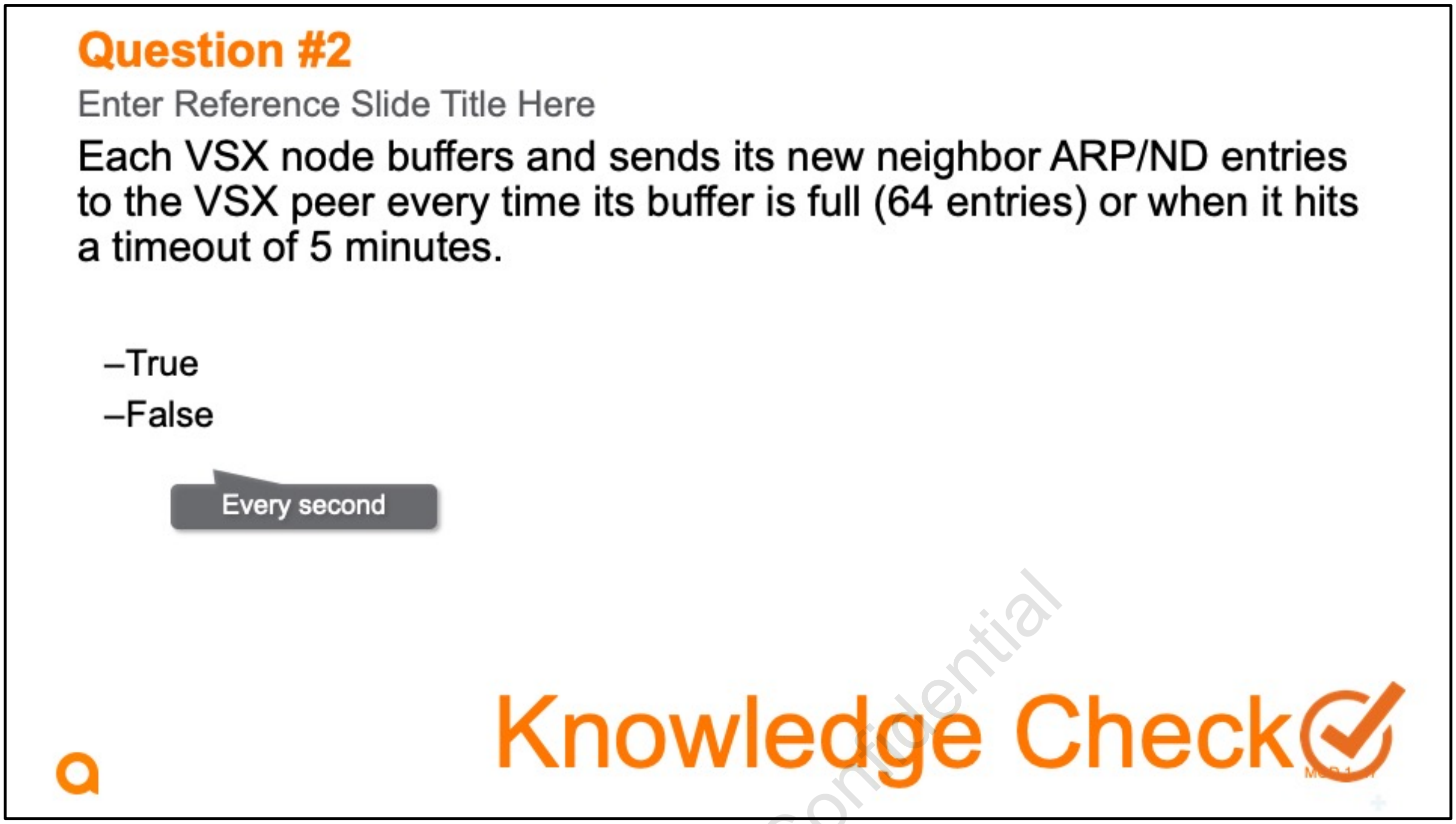
–True

–False

A dual control plane provides the best resiliency and software upgrades with near zero downtime

Knowledge Check

Aruba Training-Confidential



Question #2

Enter Reference Slide Title Here

Each VSX node buffers and sends its new neighbor ARP/ND entries to the VSX peer every time its buffer is full (64 entries) or when it hits a timeout of 5 minutes.

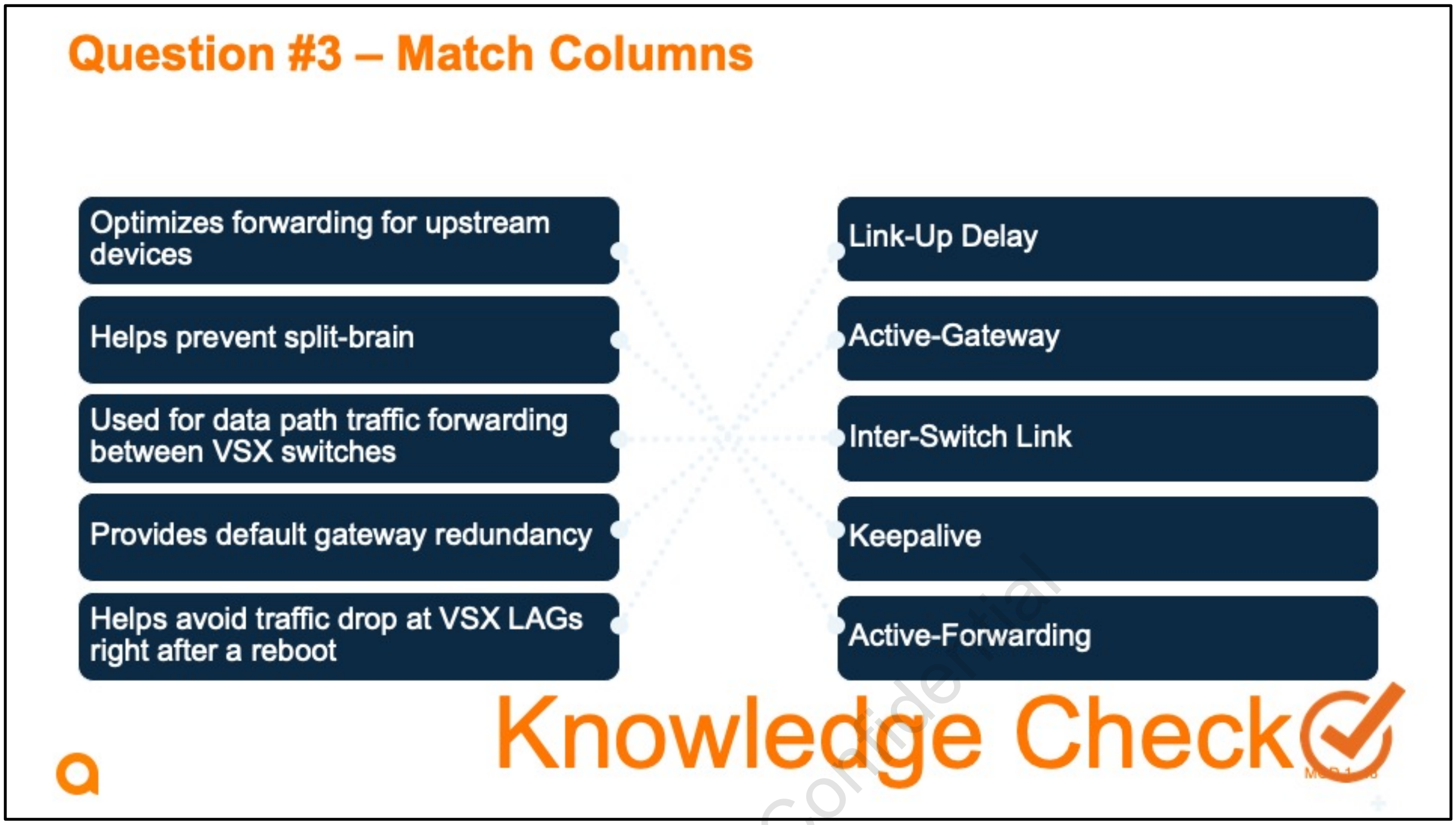
–True

–False

Every second

Knowledge Check

Aruba Training-Confidential



Question #3 – Match Columns

Optimizes forwarding for upstream devices	Link-Up Delay
Helps prevent split-brain	Active-Gateway
Used for data path traffic forwarding between VSX switches	Inter-Switch Link
Provides default gateway redundancy	Keepalive
Helps avoid traffic drop at VSX LAGs right after a reboot	Active-Forwarding

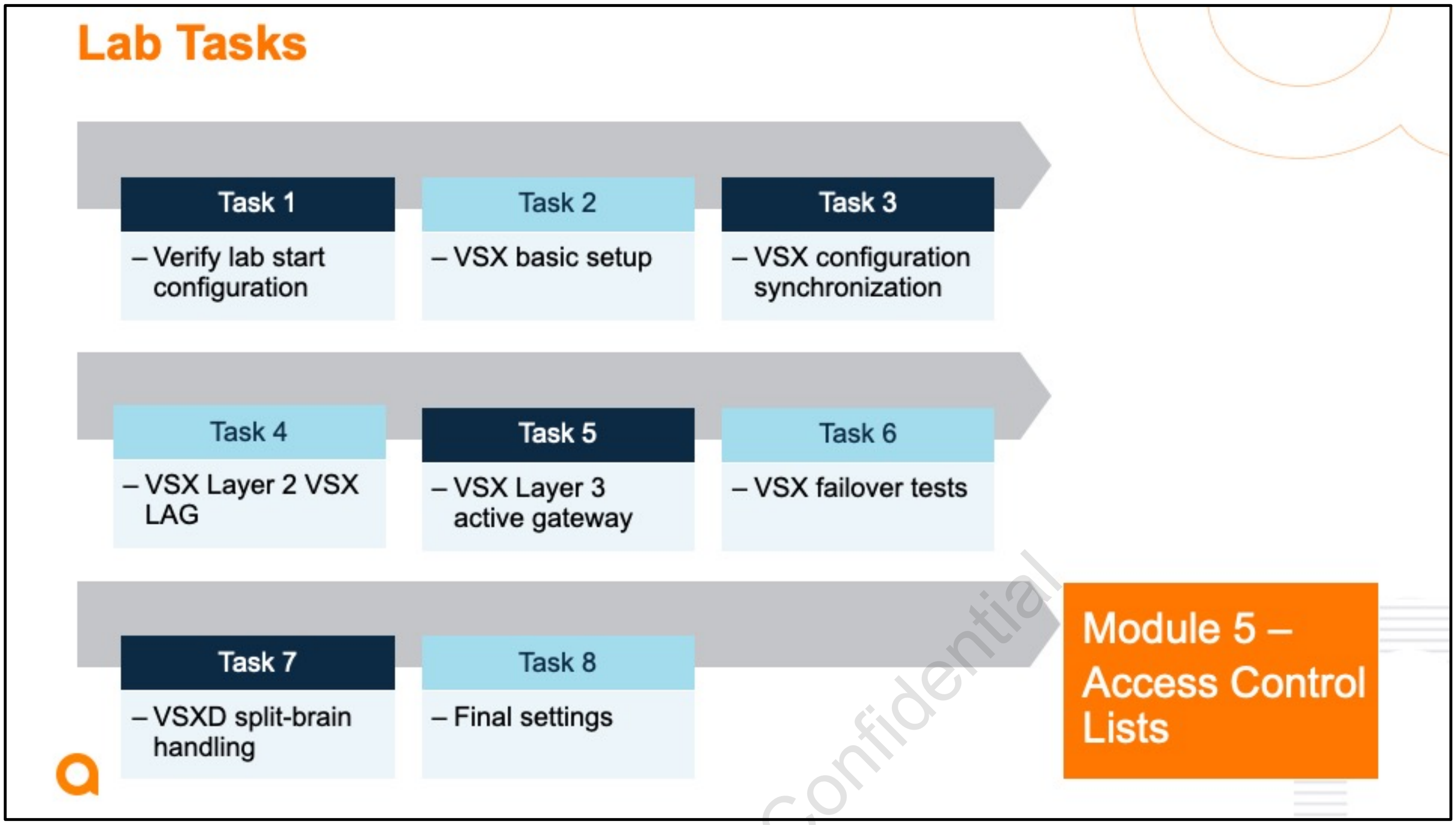
Knowledge Check

Aruba Training-Confidential

Lab Activity

You should do the lab activity.

Aruba Training-O



Lab Tasks

Task 1
– Verify lab start configuration

Task 2
– VSX basic setup

Task 3
– VSX configuration synchronization

Task 4
– VSX Layer 2 VSX LAG

Task 5
– VSX Layer 3 active gateway

Task 6
– VSX failover tests

Task 7
– VSXD split-brain handling

Task 8
– Final settings

Module 5 – Access Control Lists

The figure provides a brief review of lab tasks. Please see your lab guide for details.

Aruba Training-Confidential

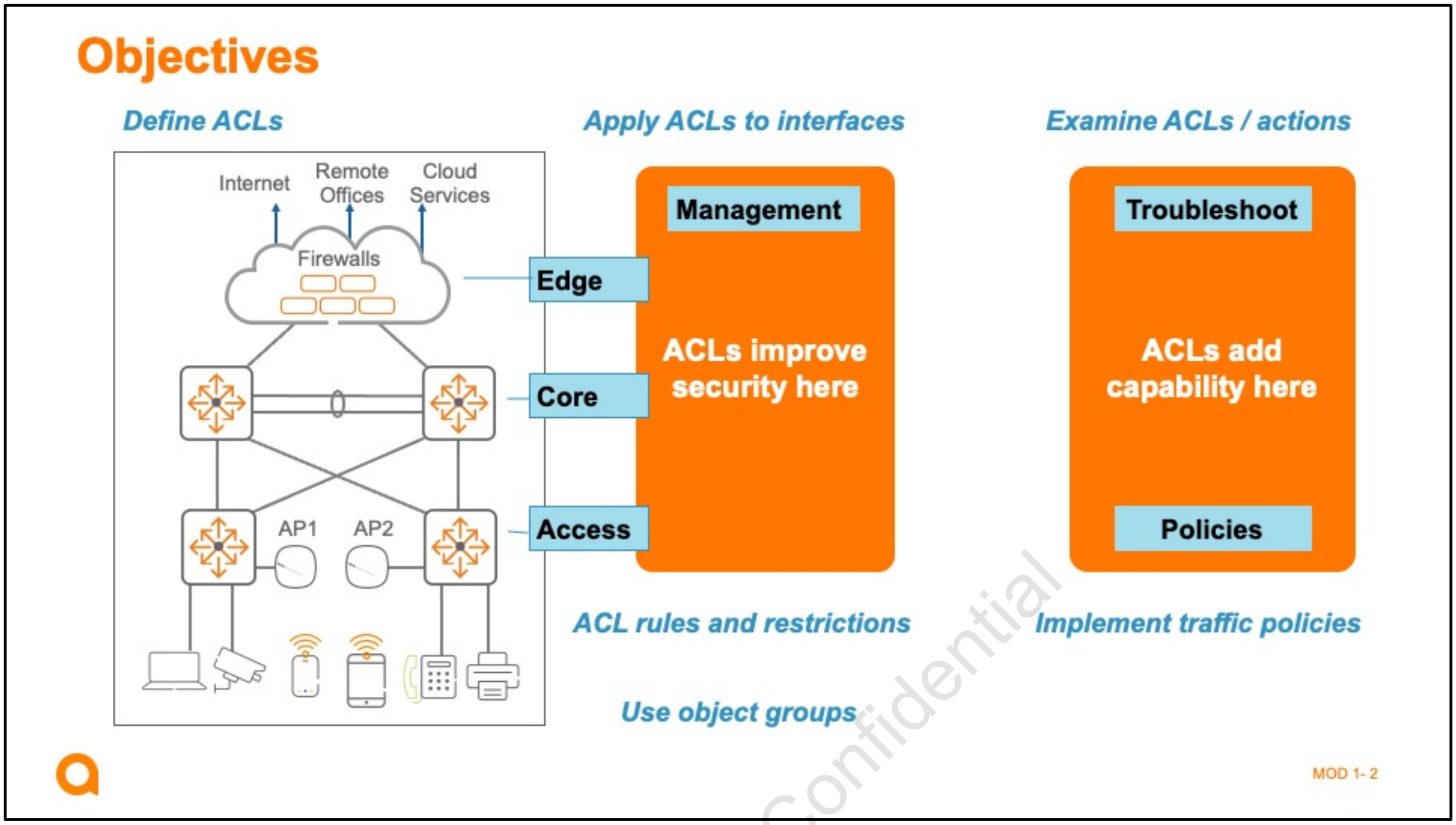
Rev # 20.21

ACLs

Implementing ArubaOS-CX Switching

Hi team, and welcome back! This is Module 5 – Access Lists.

Aruba Training-Confidential



Objectives

Define ACLs

Apply ACLs to interfaces

Management

ACLs improve security here

Examine ACLs / actions

Troubleshoot

ACLs add capability here

Policies

ACL rules and restrictions

Implement traffic policies

Use object groups

MOD 1- 2

After completing this module, you should be able to:

Define, apply, and examine ACLs and actions

Understand ACL rules and restrictions

Leverage object groups for efficiency

Implement traffic policies

Aruba Training-Confidential

Overview

- ACL Overview
- Creating Rules
- Applying ACLs
- Object Groups
- Restrictions/Resources
- Classifier Policies
- Lab Activity

Aruba Training-Confidential

With internal threats so common, you must secure your network's internal infrastructure and IT resources. So here you examine one of the basic mechanisms for controlling traffic: access control lists (ACLs).

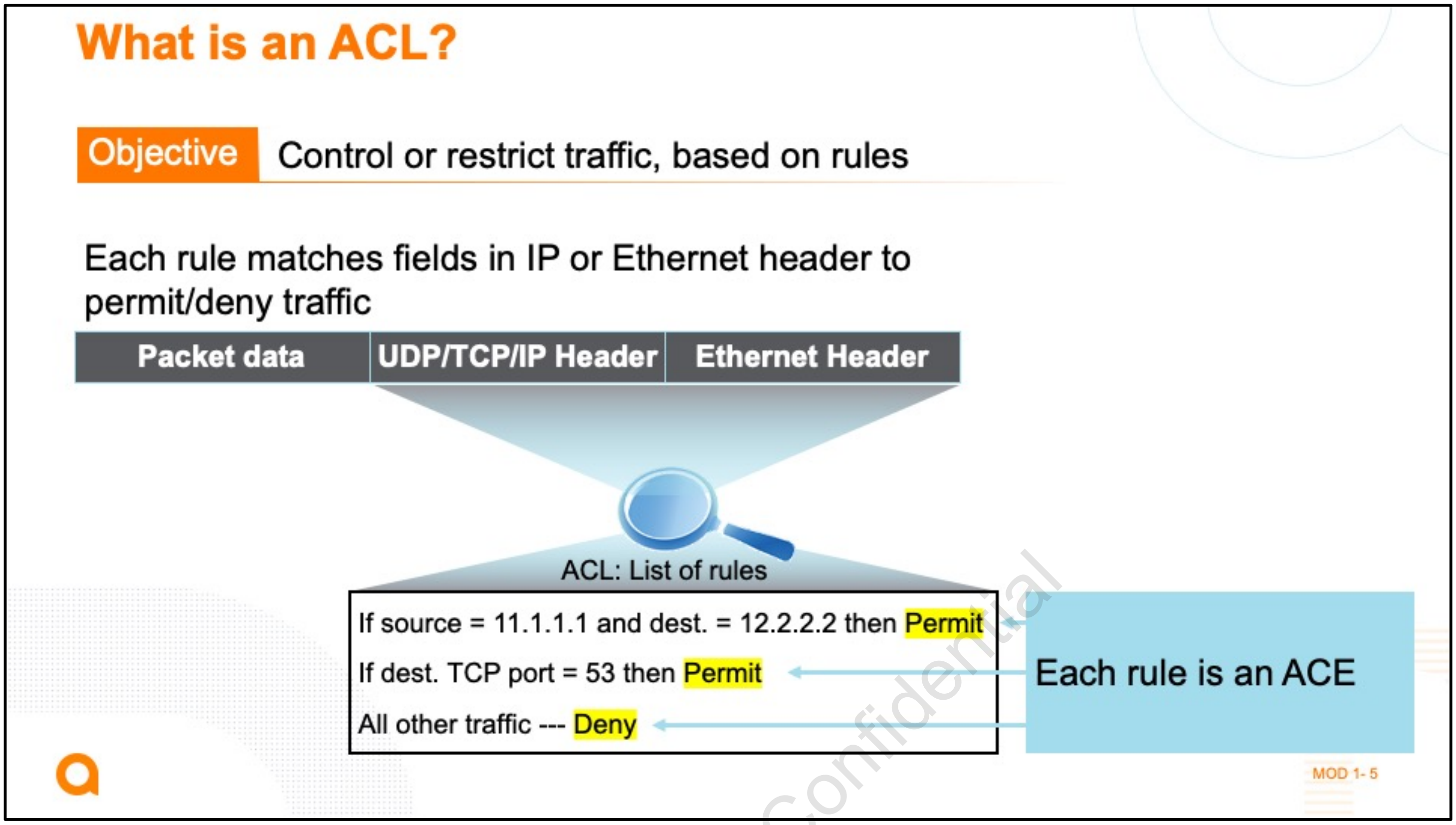
You begin with an overview of services that ACLs perform, how they work, and how to configure and apply them to AOS-CX switch interfaces

ACL Overview

aruba
a Hewlett Packard Enterprise company

Lets explore usage, components, and application of ACLs.

Aruba Training-Confidential



What is an ACL?

Objective Control or restrict traffic, based on rules

Each rule matches fields in IP or Ethernet header to permit/deny traffic

If source = 11.1.1.1 and dest. = 12.2.2.2 then Permit

If dest. TCP port = 53 then Permit

All other traffic --- Deny

Each rule is an ACE

MOD 1- 5

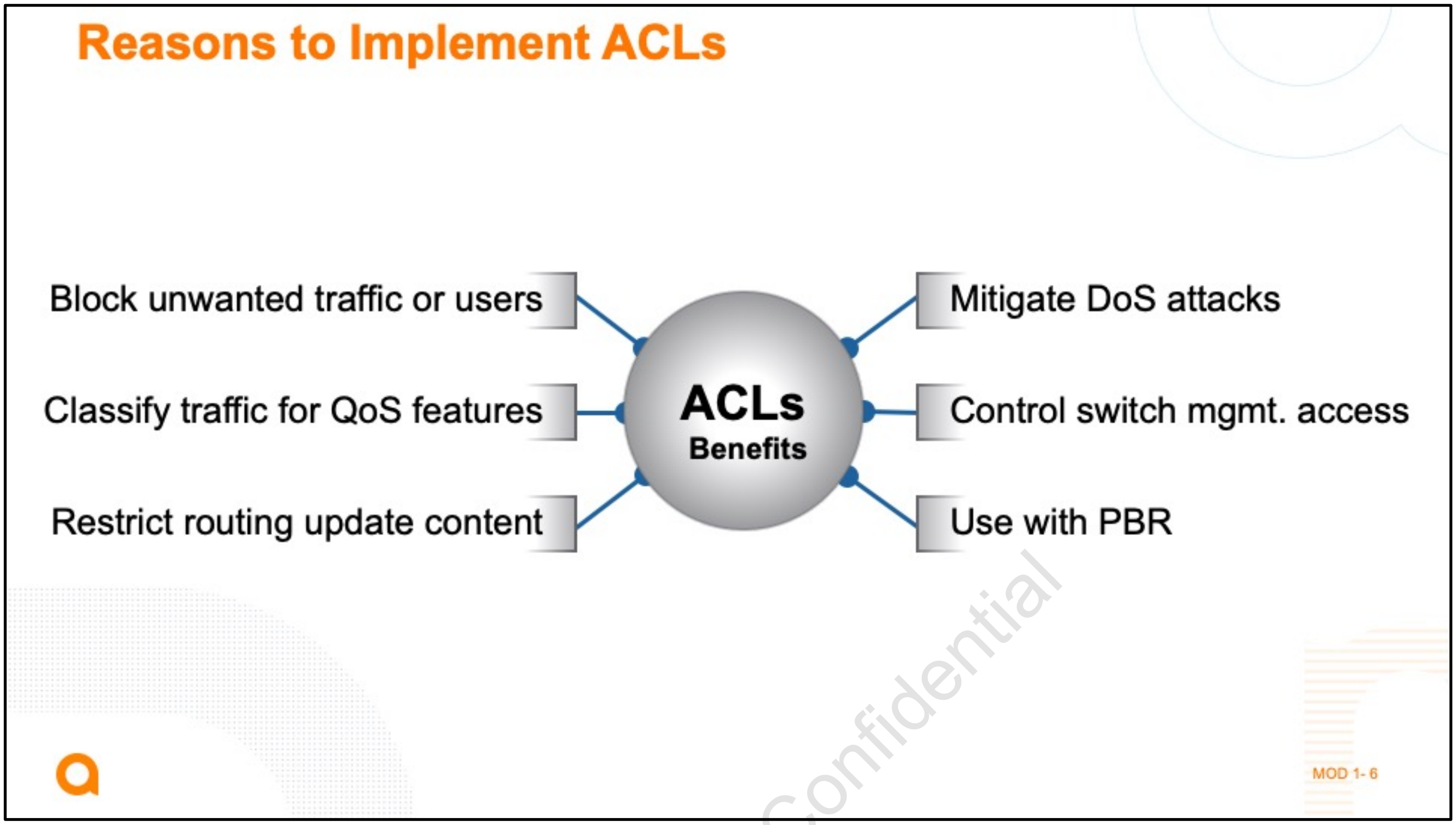
Aruba Training-Confidential

You use ACLs to control traffic, based on a list of rules. Each rule matches fields in an IP or Ethernet header to permit or deny traffic. Each rule consists of specific criteria - source and/or destination IP address, TCP or UDP port numbers, and more. A firewall or routing switch compares ingress or egress packets against ACL rules. If it finds a match some action is taken – such as permit the traffic to enter or exit a router interface, or deny it.

ACLs have several use cases, but this module focuses on packet filtering ACLs—ACLs that are applied to routing switch interfaces to filter incoming and outgoing traffic.

Note: You might also see rules referred to as access control entries (ACEs).

The example ACL shown is not proper syntax – it represents a simplified ACL logic, just so you get the general idea.



Reasons to Implement ACLs

Block unwanted traffic or users
Classify traffic for QoS features
Restrict routing update content
ACLs
Benefits
Mitigate DoS attacks
Control switch mgmt. access
Use with PBR

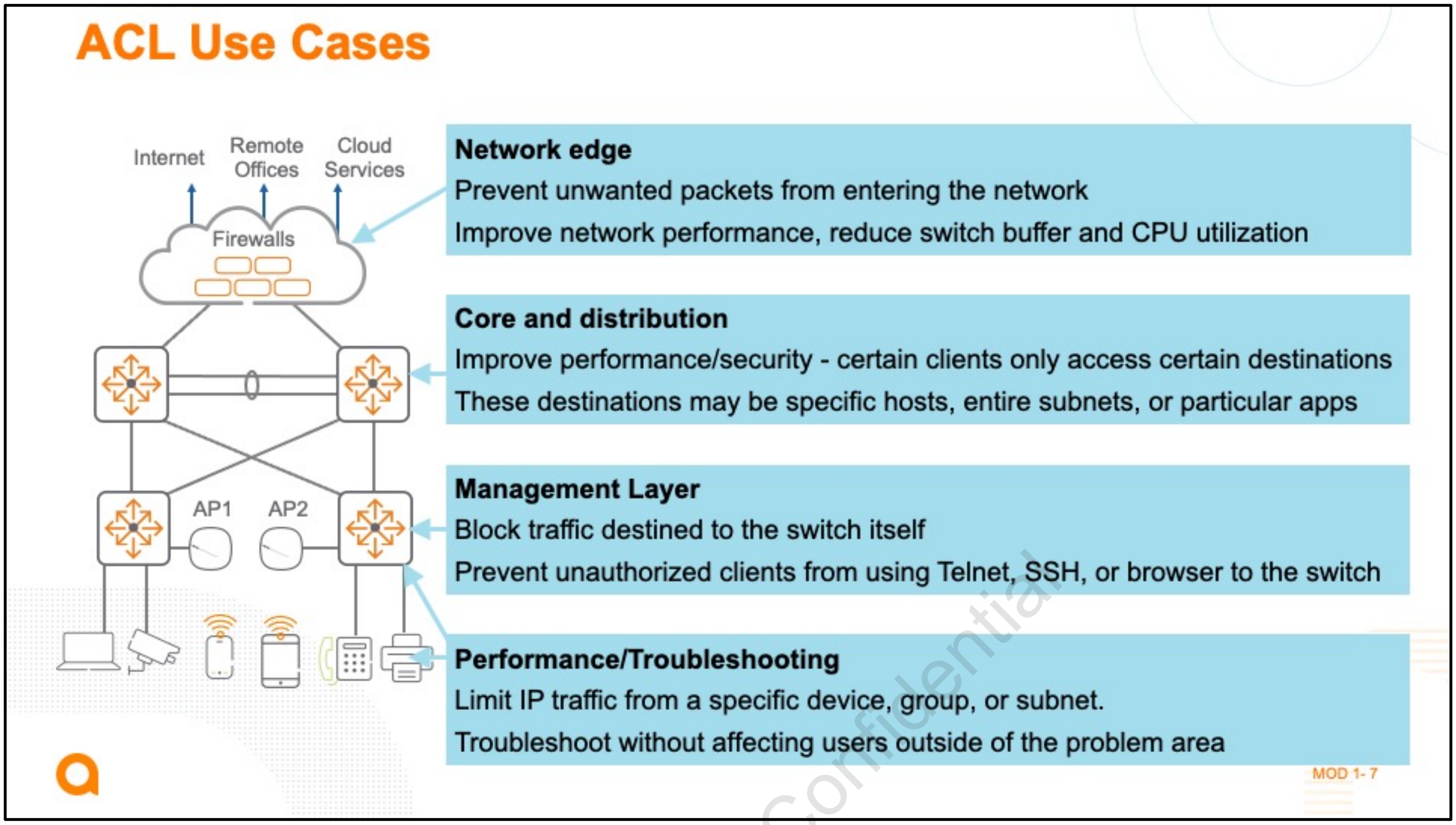
MOD 1- 6

ACLs have several use cases, but this module focuses on packet filtering ACLs— applied to routing switch interfaces to filter in- and outgoing traffic.

Without such packet filtering ACLs, each routing switch interface accepts all packets and forwards that traffic based on its forwarding tables. The ACLs enable you to implement more control, improving network performance and enhancing security. As you will see, you can apply packet filtering ACLs to filter traffic to or from an endpoint, a group of endpoints, or entire subnets.

The figure summarizes the many benefits of using ACLs.

Aruba Training-Confidential



ACL Use Cases

Network edge

Prevent unwanted packets from entering the network

Improve network performance, reduce switch buffer and CPU utilization

Core and distribution

Improve performance/security - certain clients only access certain destinations

These destinations may be specific hosts, entire subnets, or particular apps

Management Layer

Block traffic destined to the switch itself

Prevent unauthorized clients from using Telnet, SSH, or browser to the switch

Performance/Troubleshooting

Limit IP traffic from a specific device, group, or subnet.

Troubleshoot without affecting users outside of the problem area

MOD 1- 7

You might implement packet filtering ACLs for reasons such as those shown in the figure. Click each box in the figure to explore.

At the network edge, ACLs can prevent unwanted IP packets from entering the network infrastructure. You can take restrictive approaches in which you permit the required traffic and deny all other traffic. Or you can take a permissive approach in which you deny specific traffic and permit all other traffic. By combining permit and deny rules, you can create highly flexible policies.

Implementing ACLs at the network edge can also improve network performance and reduce switch buffer and CPU utilization by reducing the volume of packets that upstream switches and routers handle.

At the network core and distribution layer, ACLs can similarly improve performance and security. For example, you can use ACLs to ensure various collections of clients only have access to selected destinations. These destinations may be specific hosts, entire subnets, or even particular applications. For example, you may want to ensure that all hosts and servers in a given VLAN are only allowed to communicate within that VLAN or with a limited number of other specific VLANs.

Management Layer: You can also use ACLs to block IP traffic that has the switch as the destination address. For example, ACLs can deny unauthorized clients attempting to access the switch using functions such as Telnet, Secure Shell (SSH), or a web browser.

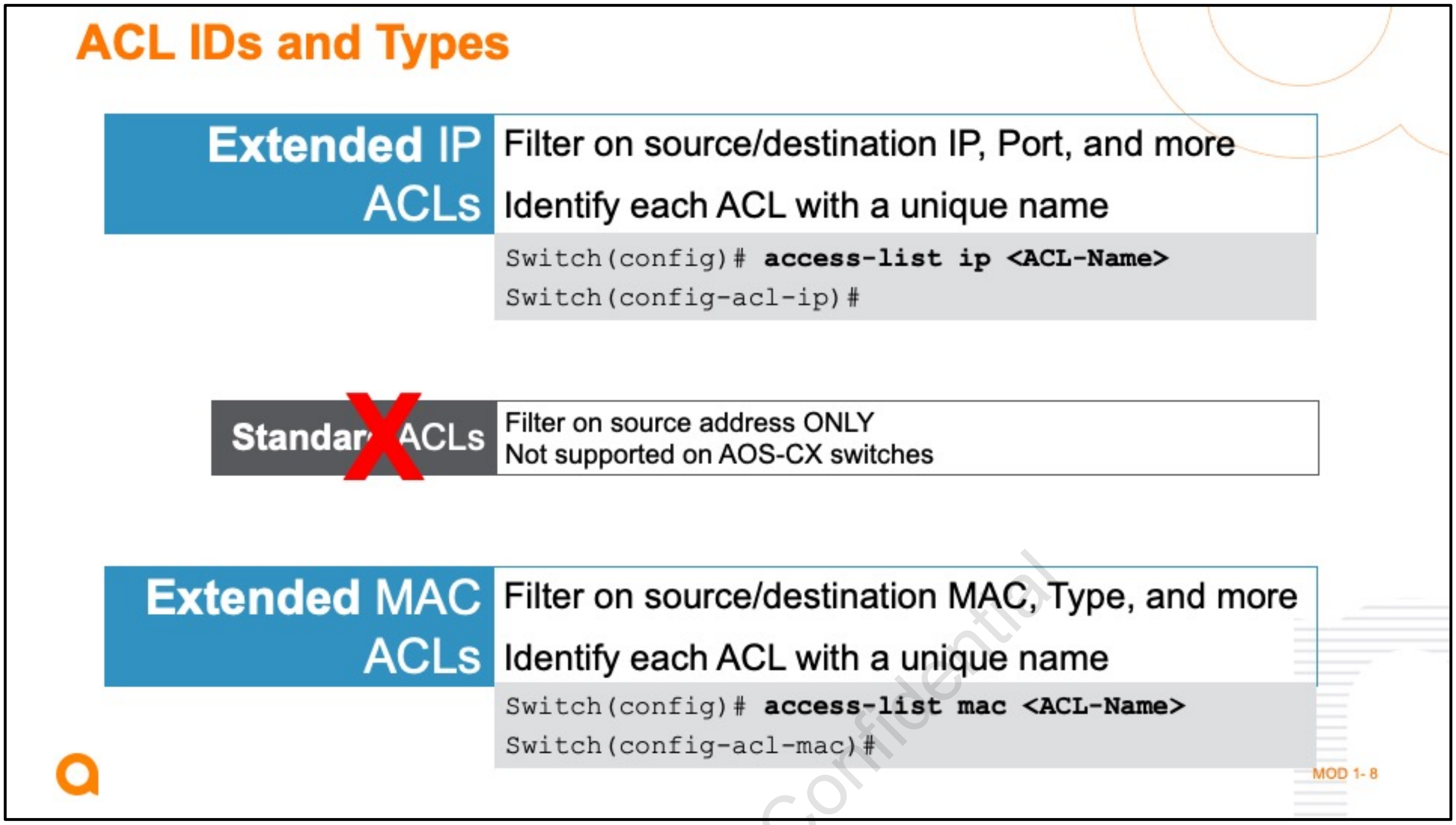
Performance/Troubleshooting: In addition to securing the network infrastructure, you can also use ACLs to control a performance problem quickly by limiting IP traffic generated by a specific

Aruba Training-Confidential

device, group of devices, or a subnet. This gives you an opportunity to troubleshoot without sacrificing performance for users outside of the problem area.

Note: AOS-CX switches ACL feature implements basic packet filtering. The switches do not support stateful (allowing returning traffic for a permitted connection) or application (payload) inspection functions. If these features are required, then a different product should be used, like Aruba's Mobility Controllers (MCs) or an HPE partner firewall solution.

Aruba Training-Confidential



ACL IDs and Types

Extended IP ACLs

Filter on source/destination IP, Port, and more

Identify each ACL with a unique name

```
Switch(config)# access-list ip <ACL-Name>
Switch(config-acl-ip)#
```

Standard ACLs

Filter on source address ONLY
Not supported on AOS-CX switches

Extended MAC ACLs

Filter on source/destination MAC, Type, and more

Identify each ACL with a unique name

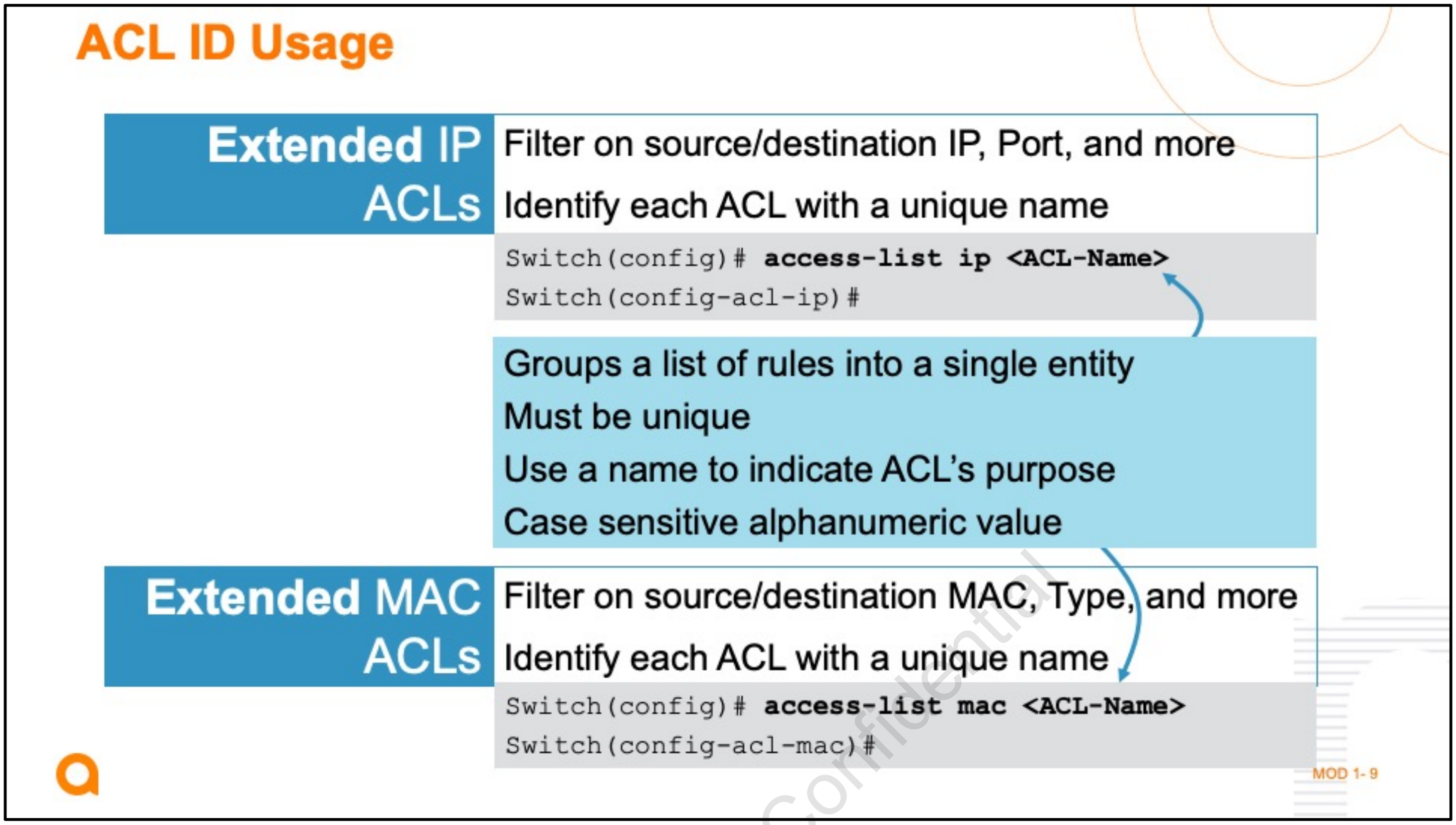
```
Switch(config)# access-list mac <ACL-Name>
Switch(config-acl-mac)#
```

MOD 1- 8

Aruba Training-Confidential

When you create an ACL, you give it a valid ACL identifier (ID) and determine its type. This course covers IPv4 extended ACLs which of course select traffic based on the IP header. This course also covers Extended MAC ACLs, which select traffic based on the MAC header. You can also create IPv6 ACLs, which is not covered in this course.

Standard ACLs allow you to create rules based solely on the source address, and are not currently supported by AOS-CX switches.



ACL ID Usage

Extended IP ACLs

Filter on source/destination IP, Port, and more

Identify each ACL with a unique name

```
Switch(config)# access-list ip <ACL-Name>
Switch(config-acl-ip)#
```

Groups a list of rules into a single entity

Must be unique

Use a name to indicate ACL's purpose

Case sensitive alphanumeric value

Extended MAC ACLs

Filter on source/destination MAC, Type, and more

Identify each ACL with a unique name

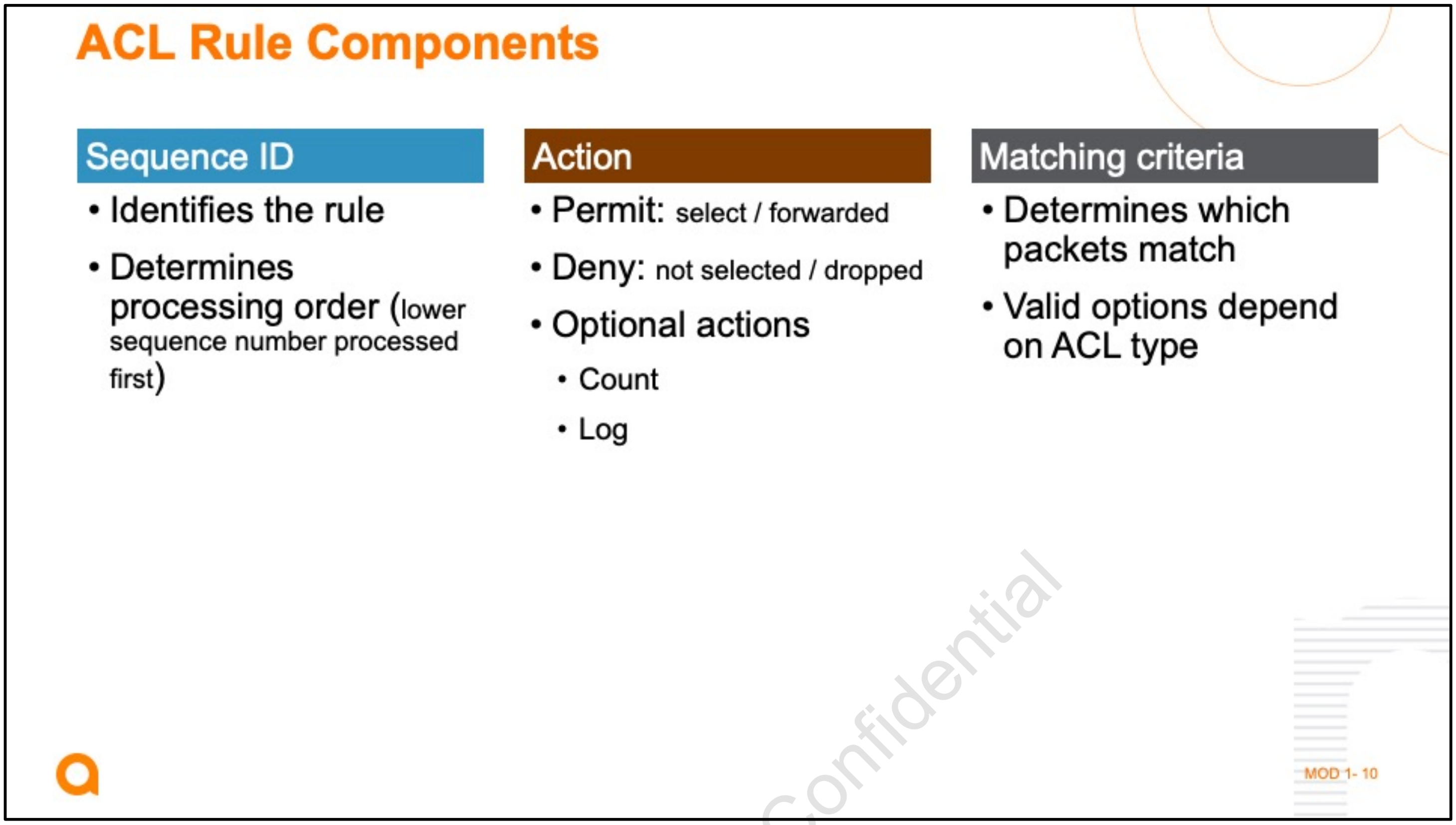
```
Switch(config)# access-list mac <ACL-Name>
Switch(config-acl-mac)#
```

MOD 1- 9

The ID groups the rules in the ACL, so you can view and modify it, and each ACL ID on a switch must be unique. With AOS-CX, you use a name to for ACLs, to indicate its intended purpose. The name is a case-sensitive alphanumeric value.

Note: You can re-use IDs between IPv4 and IPv6 lists. In other words, an ACL's ID must be unique in IPv4, and its ID must be unique in IPv6, but an IPv4 and IPv6 ACL can use the same ID.

Aruba Training-Confidential



ACL Rule Components

Sequence ID

- Identifies the rule
- Determines processing order (lower sequence number processed first)

Action

- Permit: select / forwarded
- Deny: not selected / dropped
- Optional actions
  - Count
  - Log

Matching criteria

- Determines which packets match
- Valid options depend on ACL type

MOD 1- 10

Aruba Training-Confidential

An ACL rule consists of three basic components, as shown in the figure.

Sequence ID

The sequence ID identifies the rule within the list, enabling you to identify and, if necessary, remove it. The ID also tells the switch when to process this rule in relation to other rules. The ID with the lowest number is processed first. You'll look at the implications of processing order in more detail later. Specifying a sequence ID is optional. If you do not specify the sequence ID, the switch assigns an ID automatically, placing the rule at the end of the list. You will learn more about the automatic sequence IDs a bit later in this module.

Action

ACLs have the following actions that are performed on rule match conditions:

Mandatory actions:

- permit – packets matching statements with a permit action will be allowed to progress through the switch.
- deny - packets matching statements with a deny action will be dropped.

Optional actions:

- count - packets matching statements with a count action will increment hit counts for the specific entry.
- log (available on permit and deny statements depending on platforms) – packets matching

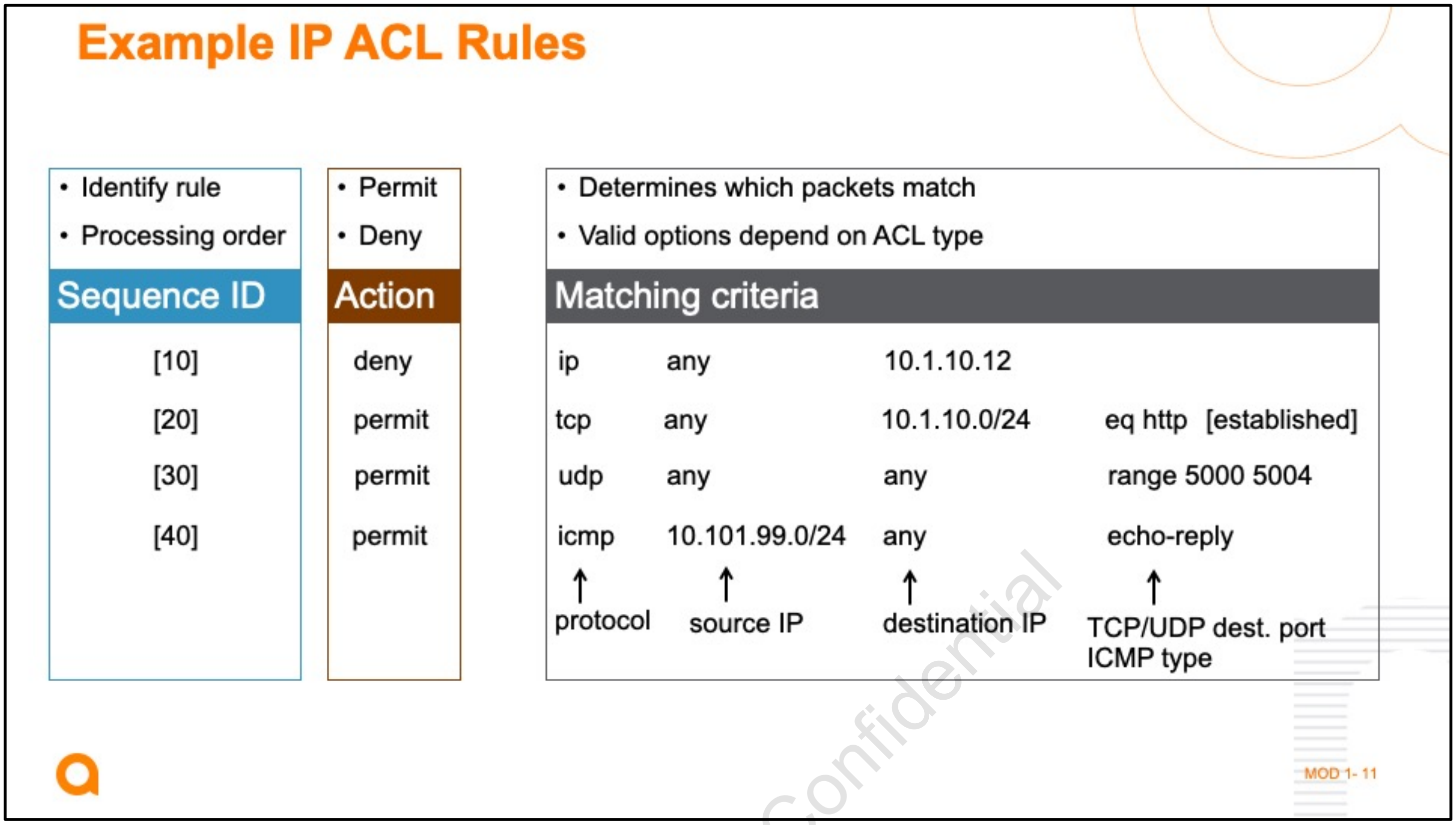
statements with a log action will increment hit counts as well for the specified entries (i.e. log performs an implicit count). The first packet that hits any log entry will be copied to the CPU and logged to the operator's specified destination (either console or syslog server). This first packet will start the 5-minute ACL logging timer (the ACL log-timer is configurable from 30s to default: 300s). When the timer expires, the summary of all the hit counts per ACL entry will be sent to the specified logging destination.

Matching criteria

The switch matches traffic against the rule's matching criteria to determine whether the traffic matches that rule.

Standard (IPv4) ACLs and standard MAC ACLs simply select traffic basic on its source IP or MAC address, so the rules are quite simple. Extended IP and MAC extended ACLs include more matching criteria, giving you more flexibility. In either case, AOS-CX switches currently only support extended ACL.

Aruba Training-Confidential



Example IP ACL Rules

- Identify rule
- Processing order

- Permit
- Deny

- Determines which packets match
- Valid options depend on ACL type

Sequence ID	Action	Matching criteria			
[10]	deny	ip	any	10.1.10.12	
[20]	permit	tcp	any	10.1.10.0/24	eq http [established]
[30]	permit	udp	any	any	range 5000 5004
[40]	permit	icmp	10.101.99.0/24	any	echo-reply
		↑ protocol	↑ source IP	↑ destination IP	↑ TCP/UDP dest. port ICMP type

MOD 1- 11

Aruba Training-Confidential

IPv4 extended ACLs can select packets based on several IP header fields. All rules include a source and destination IP address, specified like standard ACLs - any, host <IP address>, or as a subnet/prefix length. If you only care about the source IP, specify the destination as “any”. Or specify any as the source, and a specific destination.

You can further categorize IPv4 extended ACL rules by the IP protocol type, which is the first option that you specify after permit or deny:

IP - filter traffic based only on source and/or destination IP address. The IP protocol and other information in the IP header do not matter. See rule 10 in the figure.

tcp or udp – filter based on source and/or destination IP address, plus source and/or destination TCP or UDP port numbers. When you specify the port, you must also include a keyword that indicates whether to select that specific port, a range of ports, any port except that port. Rather than use the port number, you can use the keyword for many well-known ports. For example, if you specify http, the switch automatically interprets the port number as 80.

For TCP, you can filter traffic based on the TCP flag. For example, you might want to drop TCP RST packets from untrusted IP addresses. You can also include the established keyword to indicate that the rule applies only to sessions that are already established. Assume that an ACL is filtering traffic from an untrusted zone to a trusted zone. A rule with the established keyword can prevent untrusted endpoints from establishing certain types of TCP sessions, but allow sessions that were established by trusted endpoints to flow. See ACE’s 20 and 30 for examples.

Icmp - Filter ICMP traffic based on source and/or destination IP address, and optionally based on ICMP type. For example, you could permit ICMP echo replies but prohibit ICMP unreachable packets. See ACE 40 in the figure.

igmp—These rules filter IGMP traffic, optionally based on type and by source and destination address.

Other protocols—These rules let you permit or deny specific IP protocols such as OSPF. The CLI provides keywords for specifying several protocols, as shown in the job aid. If you do not see a keyword for the protocol that you want to control, use the IP protocol number. Unlike with TCP, UDP, ICMP, and IGMP, you can only specify the IP protocol and the source and destination IP addresses, not other options.

The command syntax depends on the type of protocol in the rule.

For any of these types of rules, you can specify a QoS mark so that the rule selects only traffic that meets the other criteria and is also marked for a specific priority. (You will learn more about QoS priorities later in this course.) You can specify the traffic's IP precedence value, its Type of Service (ToS) value, or both. You can specify the more modern Differentiated Services Code Points (DSCP) values by specifying both an IP precedence and ToS value.

Aruba Training Confidential

Here is the ACE syntax for IP and less common protocols:

[<SEQUENCE-NUMBER>]

{permit|deny}

{any|ip|ah|gre|esp|igmp|ospf|pim|<IP-PROTOCOL-NUM>}

{any|<SRC-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}

{any|<DST-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}

[dscp <DSCP-SPECIFIER>] [ecn <ECN-VALUE>] [ip-precedence <IP-PRECEDENCE-VALUE>]

[tos <TOS-VALUE>] [fragment] [vlan <VLAN-ID>] [ttl <TTL-VALUE>] [count] [log]

Here is the ACE syntax for SCTP, TCP, and UDP protocols:

```
[<SEQUENCE-NUMBER>]
{permit|deny}
{sctp|tcp|udp}
{any|<SRC-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}
[{eq|gt|lt} <PORT>|range <MIN-PORT> <MAX-PORT>|group <PORT-GROUP>]
{any|<DST-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}
[{eq|gt|lt} <PORT>|range <MIN-PORT> <MAX-PORT>|group <PORT-GROUP>]
[urg] [ack] [psh] [rst] [syn] [fin] [established]
[dscp <DSCP-SPECIFIER>] [ecn <ECN-VALUE>] [ip-precedence <IP-PRECEDENCE-VALUE>]
[tos <TOS-VALUE>] [fragment] [vlan <VLAN-ID>] [ttl <TTL-VALUE>] [count] [log]
```

Aruba Training-Confidential

Table 9-1: Syntax for UDP or TCP port

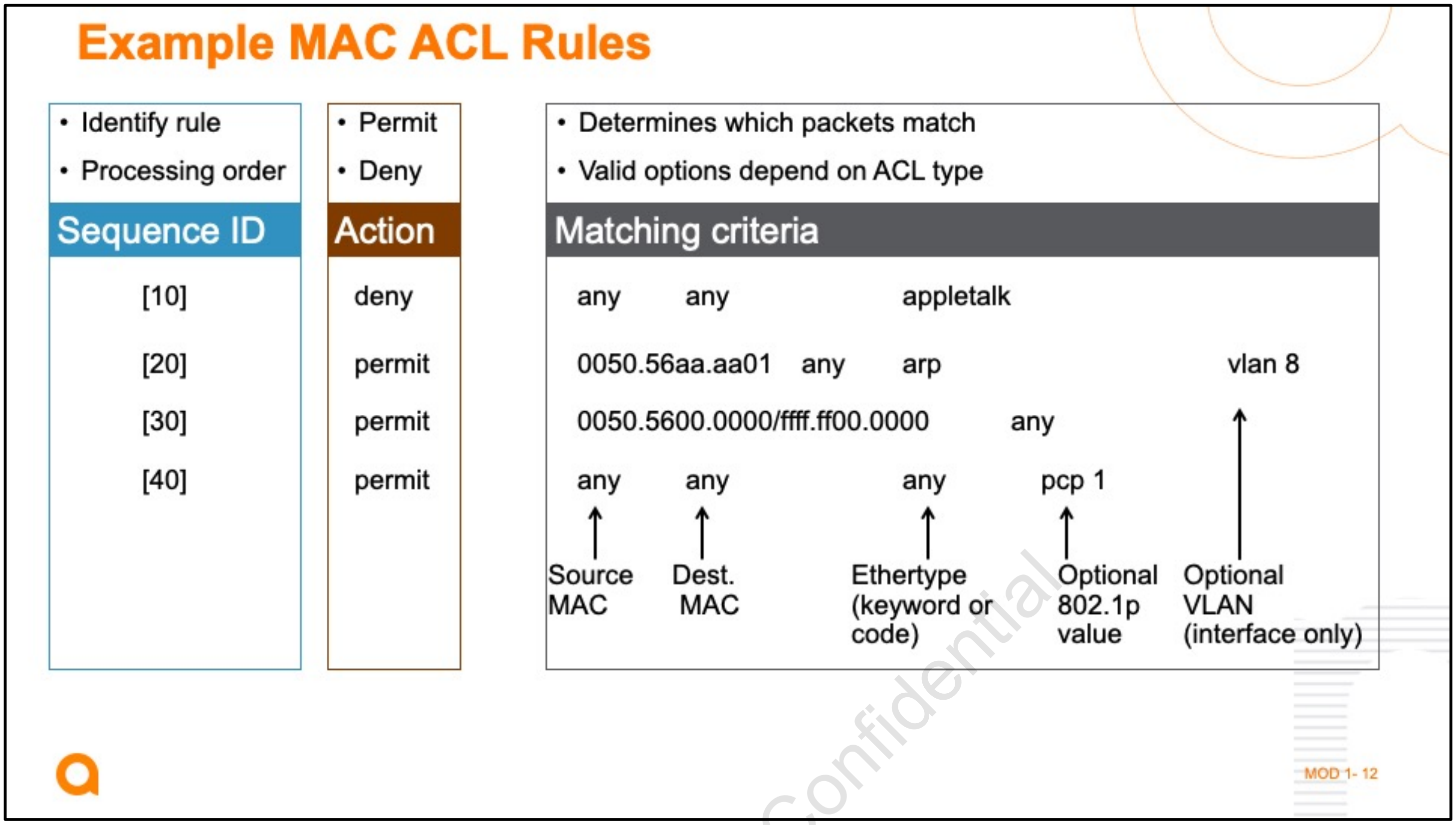
Need	Keyword	Example
Specify one port	eq	eq 80 Any traffic with port 80 is selected.
Specify a range of ports	range	range 50000 50100 Any traffic with port 50000-50100 is selected (including 50000 and 50100).
Specify any port with a smaller value than that listed	lt	lt 20 Any traffic with port 1-19 (including 1 and 19) is selected.
Specify the listed port and any port with a greater value	gt	gt 49999 Any traffic with port 50000-65535 (including 50000 and 65535) is selected.

Here is the ACE syntax for the IMCP protocol:

```
[<SEQUENCE-NUMBER>]
{permit|deny}
{icmp}
{any|<SRC-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}
{any|<DST-IP-ADDRESS>[/{<PREFIX-LENGTH>|<SUBNET-MASK>}]|<ADDRESS-GROUP>}
[icmp-type {echo|echo-reply|<ICMP-TYPE-VALUE>}] [icmp-code <ICMP-CODE-VALUE>]
[dscp <DSCP-SPECIFIER>] [ecn <ECN-VALUE>] [ip-precedence <IP-PRECEDENCE-VALUE>]
[tos <TOS-VALUE>] [fragment] [vlan <VLAN-ID>] [ttl <TTL-VALUE>] [count] [log]
```

Note: Please see the AOS-CX ACLs and Classifier Policies Guide for a more in-depth coverage of the particular ACE parameters.

Aruba Training-Confidential



Example MAC ACL Rules

- Identify rule
- Processing order

- Permit
- Deny

- Determines which packets match
- Valid options depend on ACL type

Sequence ID	Action	Matching criteria
[10]	deny	any any appletalk
[20]	permit	0050.56aa.aa01 any arp vlan 8
[30]	permit	0050.5600.0000/ffff.ff00.0000 any
[40]	permit	any any any pcp 1

Aruba Training-Confidential

Use MAC ACLS to filter traffic based on MAC address, or to filter various types of non-IP traffic. Match criteria for extended ACLs include:

Source and/or destination MAC addresses (any, a single address, or a range indicated with masking bits)

Ethertype—You can specify one of several well-known Ethernet protocols such as IP, IPv6, ARP, FCoE, and so on. You can see a list of these protocols by pressing ? in the CLI when you reach the portion of the command for specifying the Ethertype. If the protocol that you want to select does not appear in the list, enter the protocol number. You must specify the Ethertype in all rules. If you want the rule to match all Ethernet traffic, specify any.

Optional 802.1Q QoS Priority Code Point (PCP) value

- Note: You will learn more about QoS later. Often you do not want to deny traffic with a specific PCP (802.1p) value entirely, but instead remove or change the value.

Optional VLAN number

- Note: The VLAN value is only applicable when the MAC ACL is applied to a port or link aggregation (trunk) interface.

Here is the subcontext syntax for creating a MAC ACE entry:

```
[<SEQUENCE-NUMBER>] {permit|deny}
{any|<SRC-MAC-ADDRESS>[/<ETHERNET-MASK>}]}
{any|<DST-MAC-ADDRESS>[/<ETHERNET-MASK>}]}
```

```
{any|aarp|appletalk|arp|fcoe|fcoe-init|ip|ipv6|
ipx-arpa|ipx-non-arpa|is-is|lldp|mpls-multicast|mpls-unicast|q-in-q|
rbridge|trill|wake-on-lan|<NUMERIC-ETHERTYPE>}
[pcp <PCP-VALUE>] [vlan <VLAN-ID>] [count] [log]
```

This example rule denies all Appletalk traffic:

```
Switch(config-ext-macl)# 10 deny any any appletalk
```

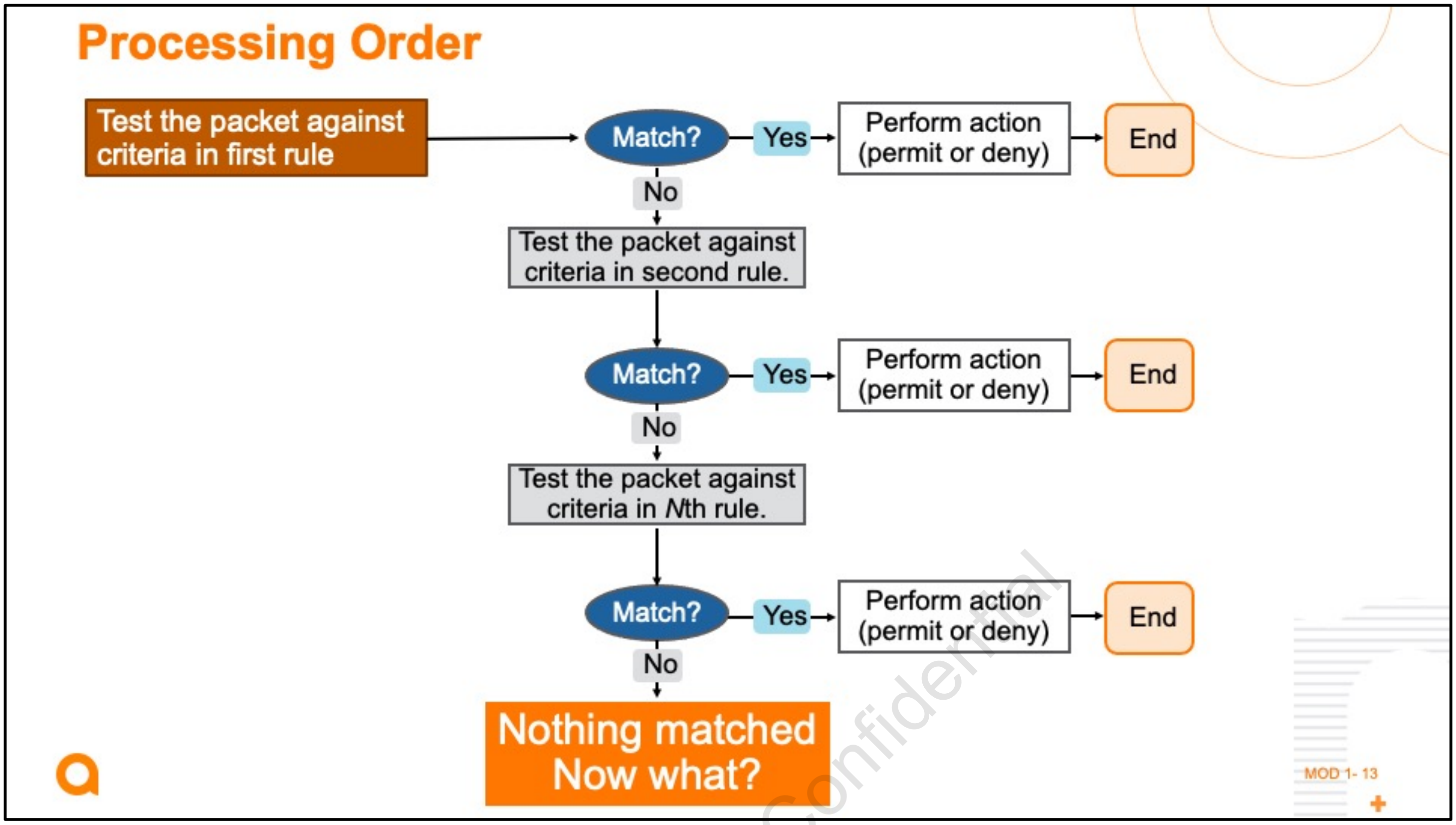
This example rule permits ARP traffic from a specific MAC address in a specific VLAN:

```
Switch(config-ext-macl)# 20 permit host AAAA.AAAA.AAAA any arp vlan 8
```

This example rule denies all traffic with 802.1p value 1:

```
Switch(config-ext-macl)# 30 deny any any any pcp 1
```

Aruba Training-Confidential



Processing Order

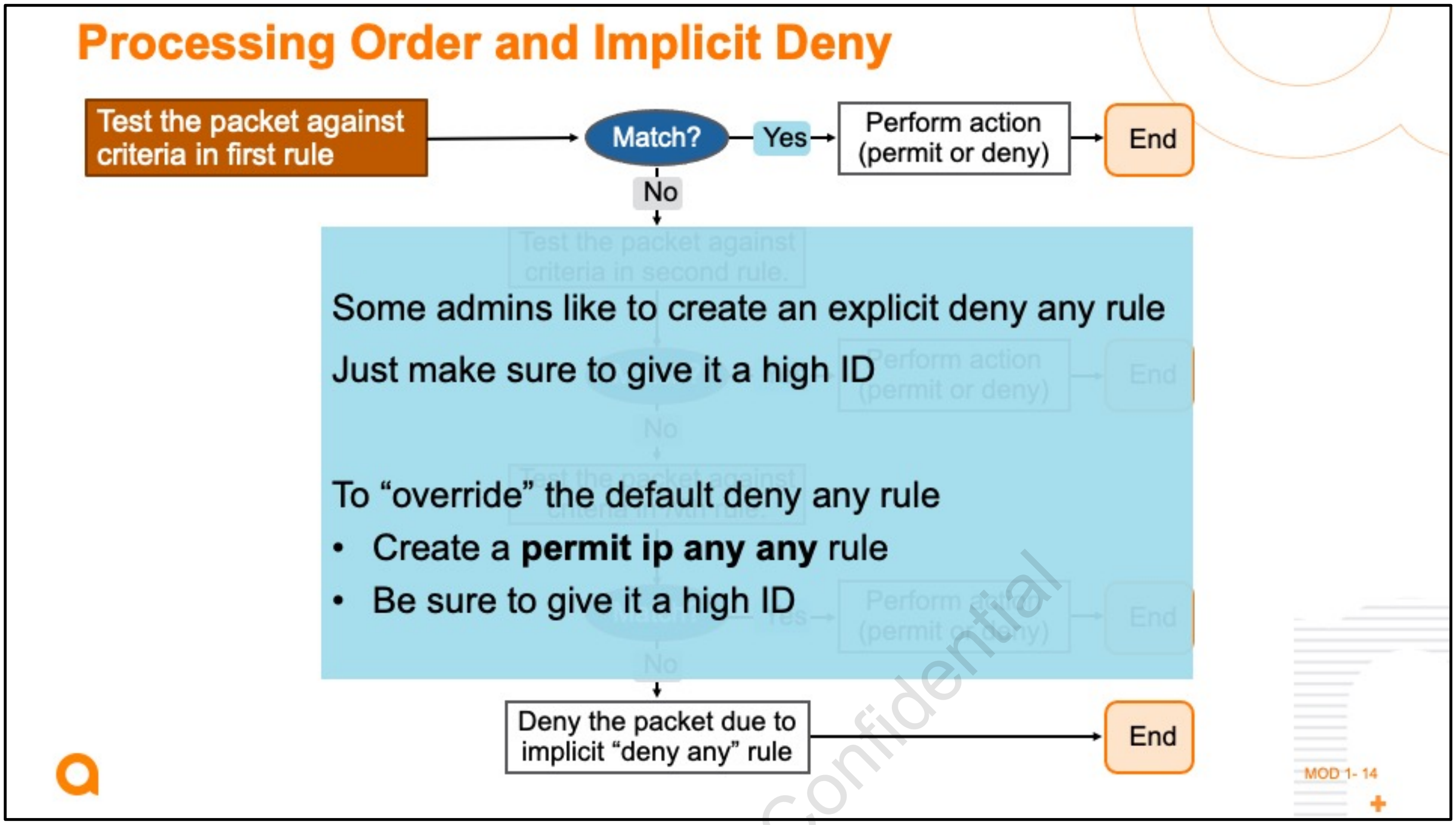
Test the packet against criteria in first rule
Match?
Yes
Perform action (permit or deny)
End
No
Test the packet against criteria in second rule.
Match?
Yes
Perform action (permit or deny)
End
No
Test the packet against criteria in *N*th rule.
Match?
Yes
Perform action (permit or deny)
End
No
Nothing matched Now what?
MOD 1- 13

Aruba Training-Confidential

The switch processes rules in a top-down manner, from lowest to highest rule ID number. It compares a packet against the first rule's matching criteria. If it matches, it performs the specified permit or deny action, and stops processing the ACL.

| If no match, then it compares the packet against the next rule in sequence until it finds a match, and so on.

| But if it processes all rules and still doesn't find a match? What happens to the packet?



Processing Order and Implicit Deny

Test the packet against criteria in first rule → Match? — Yes → Perform action (permit or deny) → End

No ↓

Some admins like to create an explicit deny any rule

Just make sure to give it a high ID

To “override” the default deny any rule

- Create a **permit ip any any** rule
- Be sure to give it a high ID

Deny the packet due to implicit “deny any” rule → End

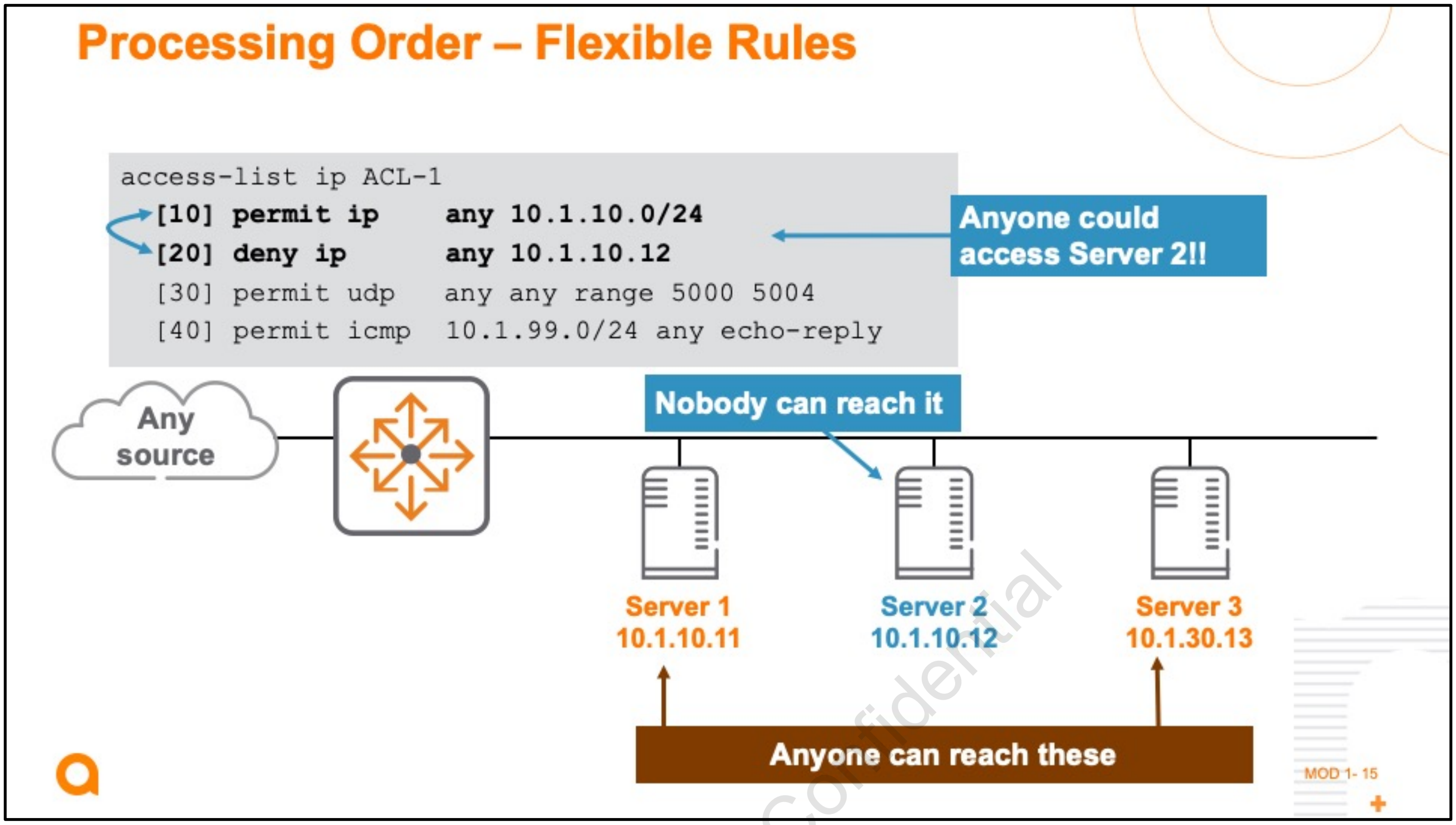
MOD 1- 14

Aruba Training-Confidential

There is an implicit deny any at the end of all ACLs – a deny any any statement. If you apply an ACL to an interface and the switch finds no matches between an inbound packet and the ACL rules, the packet is dropped. Its an “implicit” rule because you can’t see it when viewing an ACL.

| Some folks like to explicitly define the implicit rule at the end of an ACL to clarify the behavior. If you do so, make sure to assign the rule a very high ID so that you can add as many rules as you want.

If you need to override the implicit deny so that a packet that does not have a match is permitted, then you should create a rule that permits all traffic as the last rule in the ACL. Again, assign this rule a high ID.



Processing Order – Flexible Rules

```
access-list ip ACL-1
  [10] permit ip    any 10.1.10.0/24
  [20] deny ip      any 10.1.10.12
  [30] permit udp   any any range 5000 5004
  [40] permit icmp  10.1.99.0/24 any echo-reply
```

Anyone could access Server 2!!

Nobody can reach it

Anyone can reach these

Aruba Training-Confidential

You can combine permit and deny rules to create highly flexible policies.

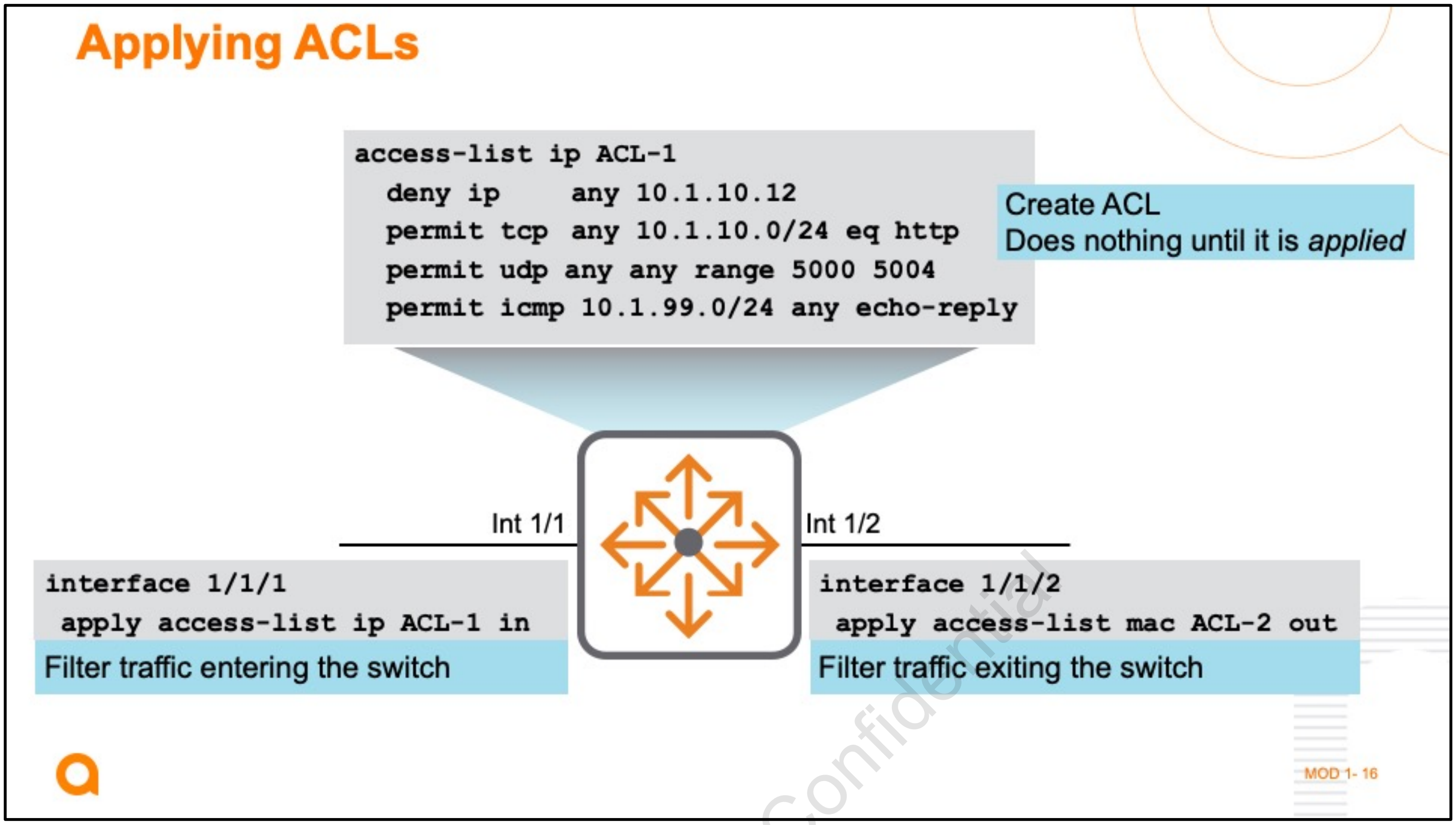
| For example, you can deny everyone from accessing one device within a subnet (such as 10.1.10.12).

| Then permit them to access everyone else on the complete subnet (10.1.10.0/24).

This scenario highlights the importance of rule processing order.

| For example, suppose you swapped order of rules 10 and 20.

Destination 10.1.10.12 would match 10.1.10.0/24, and so rule 20 would never be processed. Anyone would be able to access server 2!



Applying ACLs

```
access-list ip ACL-1
  deny ip     any 10.1.10.12
  permit tcp  any 10.1.10.0/24 eq http
  permit udp any any range 5000 5004
  permit icmp 10.1.99.0/24 any echo-reply
```

Create ACL
Does nothing until it is *applied*

```
interface 1/1/1
 apply access-list ip ACL-1 in
```

Filter traffic entering the switch

```
interface 1/1/2
 apply access-list mac ACL-2 out
```

Filter traffic exiting the switch

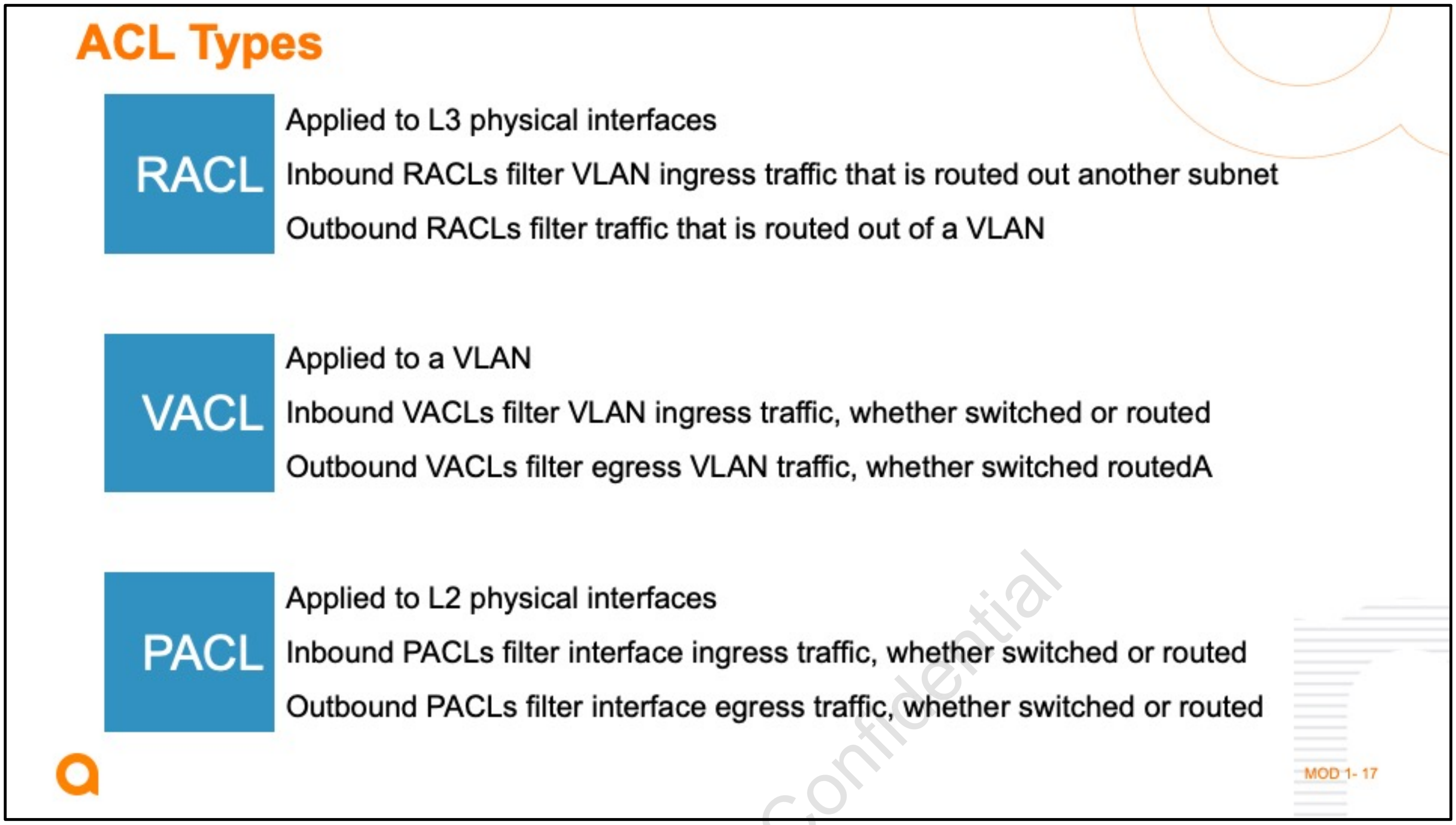
Aruba Training-Confidential

When you create an ACL, it does nothing until you apply it to one or more interfaces or VLANs, as shown here. ACLs configured on an interface using this syntax are called static ACLs – they remain on the interface permanently until you modify or remove them.

AOS-CX switches can also apply dynamic Port-based ACLs (PACLs) as part of an Authentication, Authorization, and Accounting (AAA) solution. This module focuses on static ACLs.

You have three choices as you apply a static ACL. First, you choose the interface (Layer-2 or Layer-3) on which to apply the ACL. Then choose the direction, which determines whether the ACL filters traffic that arrives on that interface or that is destined out that interface. As shown in the figure, you can apply both IP and MAC ACLs to VLAN or physical interfaces.

Note: Currently, AOS-CX does not support applying ACLs to logical Layer-3 interfaces, like VLAN SVI or GRE tunnel interfaces.



ACL Types

RACL

Applied to L3 physical interfaces

Inbound RACLs filter VLAN ingress traffic that is routed out another subnet

Outbound RACLs filter traffic that is routed out of a VLAN

VACL

Applied to a VLAN

Inbound VACLs filter VLAN ingress traffic, whether switched or routed

Outbound VACLs filter egress VLAN traffic, whether switched routedA

PACL

Applied to L2 physical interfaces

Inbound PACLs filter interface ingress traffic, whether switched or routed

Outbound PACLs filter interface egress traffic, whether switched or routed

MOD 1- 17

Aruba Training-Confidential

How you apply the ACL determines its type: a Layer-3 interface or Routed ACL (RACL), a VLAN ACL (VACL), or a Port-based ACL (PACL) (applied on a layer-2 interface or a link aggregation or LAG). Whether the switch applies the ACL as a RACL, VACL, or PACL in turn determines which packets and frames the switch processes against the ACL.

RACLs are IP ACLs applied to routed Layer 3 physical interfaces in either the outbound or inbound direction. Inbound RACLs filter traffic that arrives on a VLAN and is routed out another subnet while outbound RACLs filter traffic that is routed out of a VLAN.

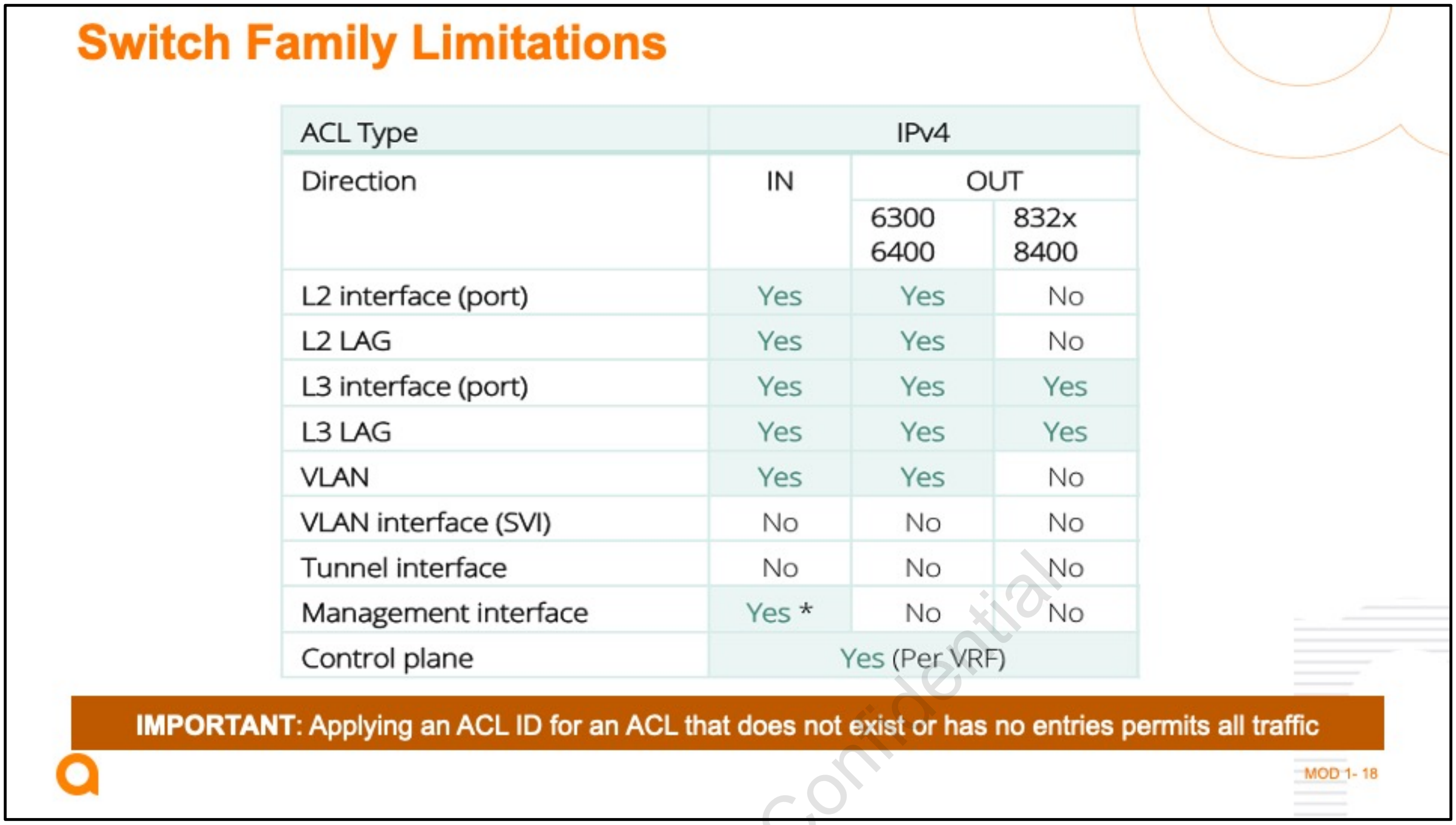
VACLs are applied to a VLAN within the VLAN context. Inbound VACLs filter all traffic that arrives on a VLAN, whether switched or routed. Outbound VACLs filter all traffic that is forwarded out a VLAN, whether the source is within the same subnet (switched) or another subnet (routed). Because both inbound and outbound VACLs filter traffic that arrives on a VLAN and is switched out that VLAN, some of the traffic that they filter overlaps.

Port-based ACLs (PACLs) are any ACLs applied to Layer-2 physical interfaces. Inbound PACLs filter all traffic that arrives on the interface, whether switched or routed. Outbound PACLs filter all traffic forwarded on the interface, whether switched or routed.

You will explore the different ways these ACLs filter traffic in more detail throughout this module.

You should also keep in mind that you can apply an ACL to multiple different interfaces, but each interface supports a maximum of one ACL in each direction. On VLAN interfaces, you are allowed to use the same ACL for both the inbound and outbound directions, but often you should not do so because, of course, the source and destination addresses for your rules might not be valid in both directions.

Aruba Training-Confidential



Switch Family Limitations

ACL Type	IPv4		
Direction	IN	OUT	
		6300 6400	832x 8400
L2 interface (port)	Yes	Yes	No
L2 LAG	Yes	Yes	No
L3 interface (port)	Yes	Yes	Yes
L3 LAG	Yes	Yes	Yes
VLAN	Yes	Yes	No
VLAN interface (SVI)	No	No	No
Tunnel interface	No	No	No
Management interface	Yes *	No	No
Control plane	Yes (Per VRF)		

IMPORTANT: Applying an ACL ID for an ACL that does not exist or has no entries permits all traffic

MOD 1- 18

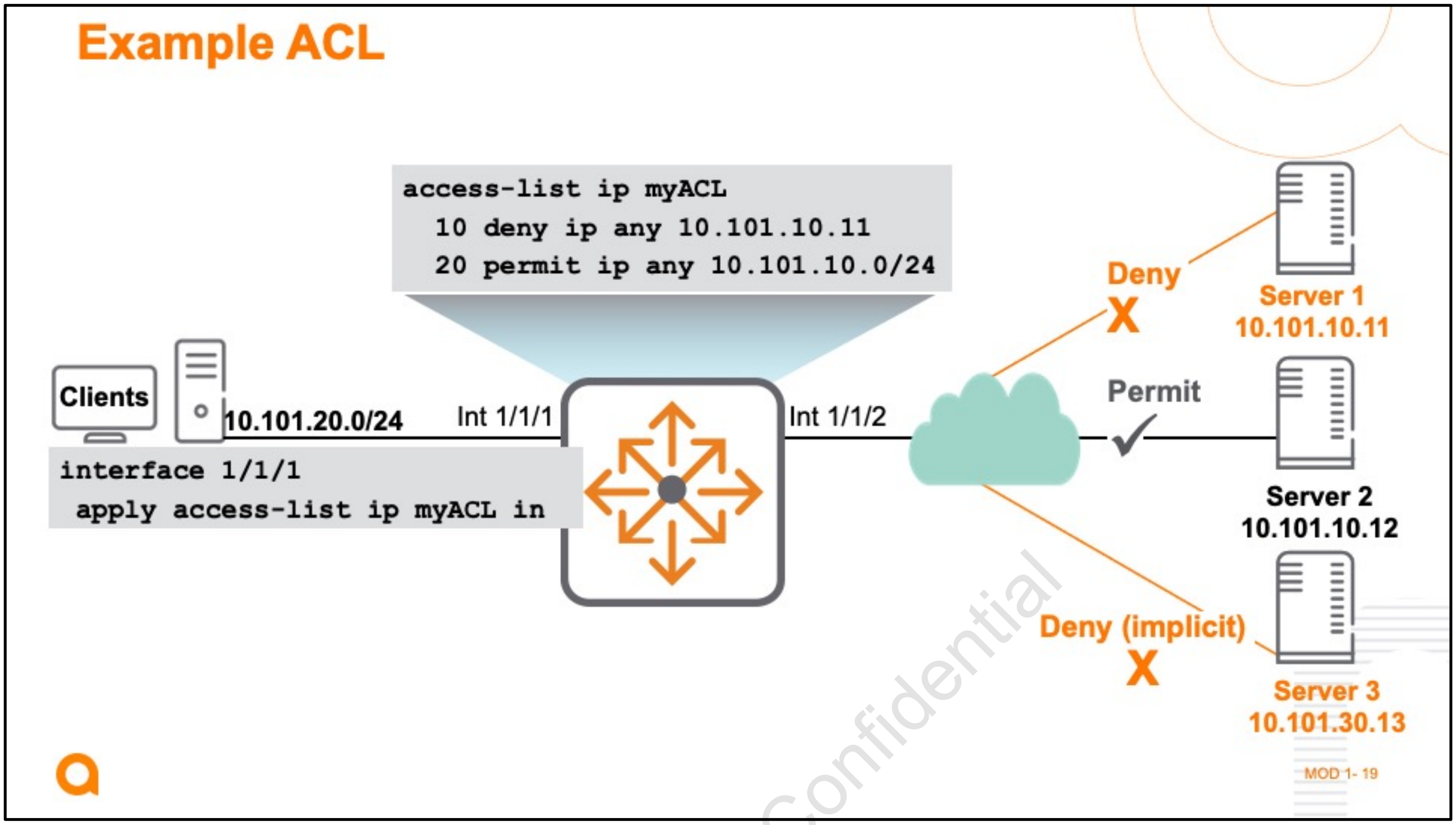
Aruba Training-Confidential

The extent of ACL support varies among the AOS-CX switch families. Based on the hardware ASIC used by the switch, there might be limitations in how an ACL can be applied, as shown in the figure, as of AOS-CX 10.4.

Tips: Keep the following points in mind when applying ACLs:

AOS-CX switches allow you to assign a nonexistent ACL name to an interface. When you apply a nonexistent or empty ACL to an interface, the interface permits all traffic until you create an ACL that uses that name and add a rule to the ACL. The implicit “deny any” does not take effect until at least one rule is defined. In short, you should be careful to specify the correct ACL name.

On AOS-CX switches, if you delete an assigned ACL from the switch without subsequently removing the ACL from an interface to which it is assigned, the ACL assignment remains on the interface. The empty ACL does not affect traffic (all traffic is permitted). However, the switch will automatically begin applying any new ACL you create with the same identifier (name). You should clean up your configurations to ensure that ACLs are applied when and where you expect.



Example ACL

```
access-list ip myACL
  10 deny ip any 10.101.10.11
  20 permit ip any 10.101.10.0/24
```

```
interface 1/1/1
 apply access-list ip myACL in
```

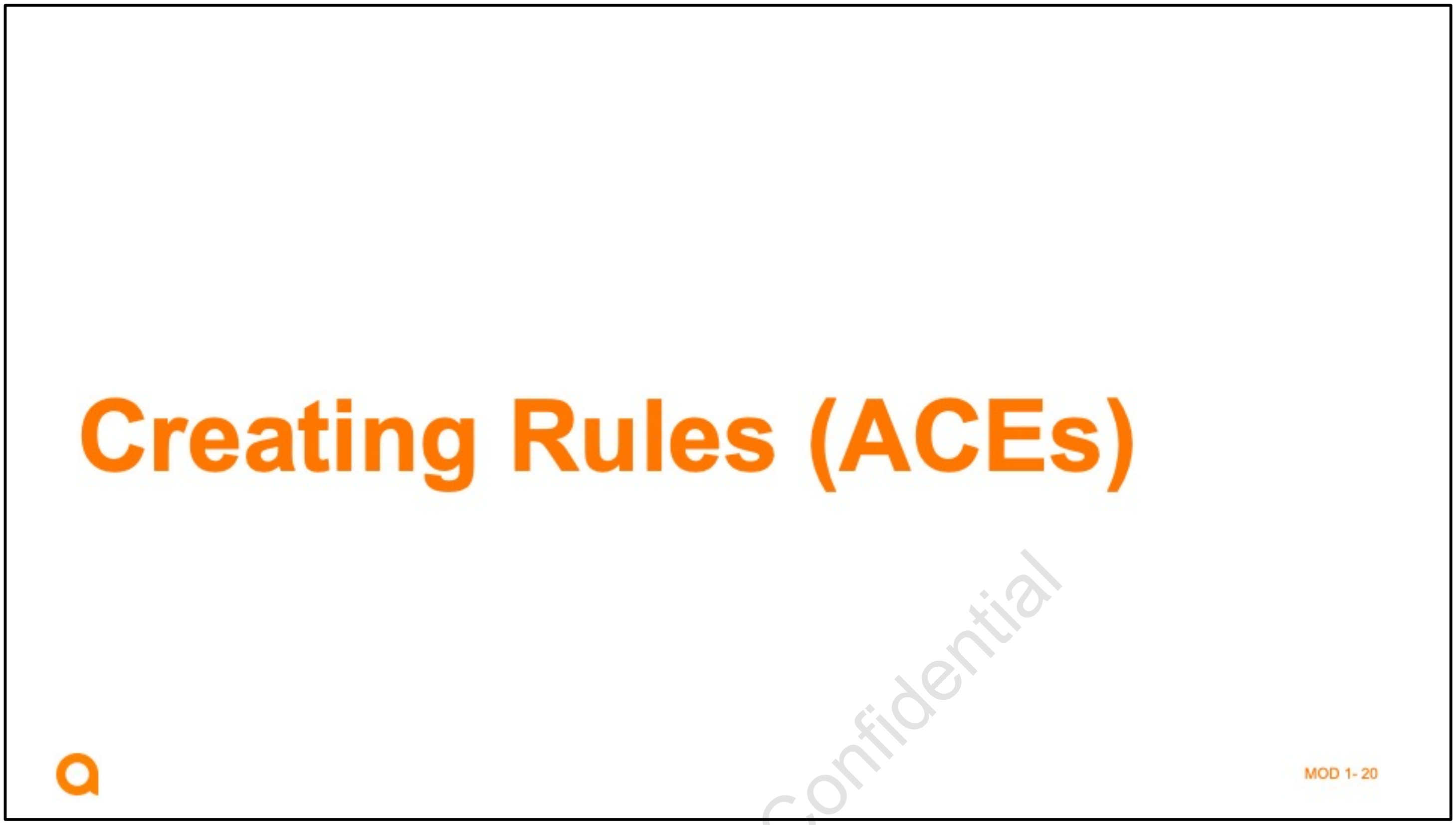
Aruba Training-Confidential

The figure shows how you can combine ACL rules and apply them to an interface. In this example, you want to strictly control traffic from devices connected to interface 1/1/1. You configure an IPv4 extended ACL with the two rules shown and apply it inbound on the interface.

When a device tries to send IP traffic to Server 1, the switch processes the traffic against the ACL. This traffic matches the first rule, which denies all IP traffic destined to Server 1 at 10.101.10.11. The traffic is dropped.

Next the device tries to send IP traffic to Server 2 at 10.101.10.12. Again, the switch processes the traffic against the first rule, but the traffic does not match this rule. The switch then processes the traffic against the second rule, which it does match because 10.101.10.12/30 is within 10.101.10.0/24. The second rule is a permit rule, so the switch forwards the traffic.

Finally, the device tries to send IP traffic to Server 3 at 10.101.30.13. This traffic matches neither rule. Therefore, it is processed by the implicit deny any, and the switch drops it.

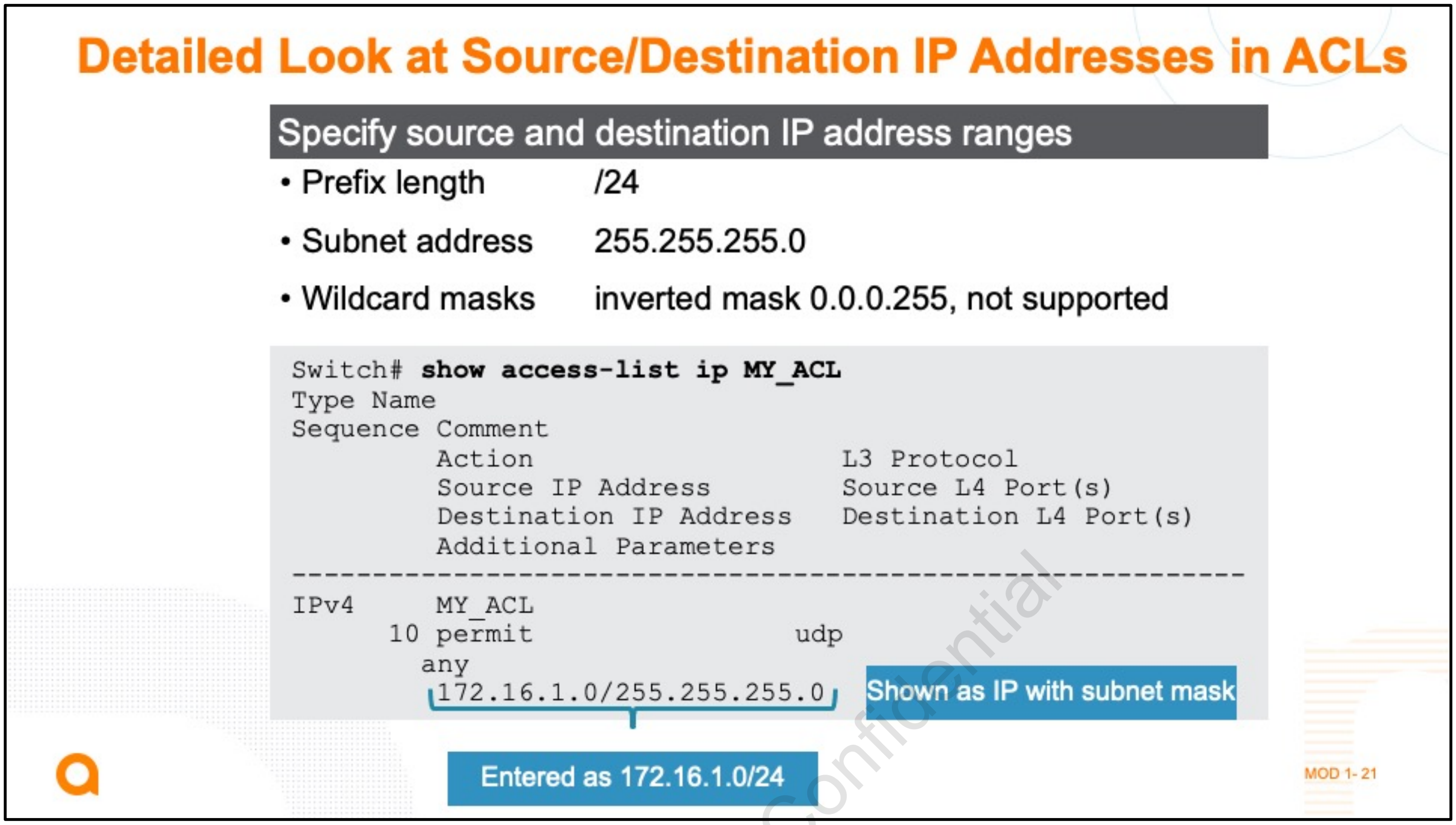


Creating Rules (ACEs)

MOD 1- 20

Aruba Training-Confidential

Now that you have had some practice creating an ACL, you will dive down deeper into the specifics of creating flexible rules, commonly referred to as access control entries (ACEs).



Detailed Look at Source/Destination IP Addresses in ACLs

Specify source and destination IP address ranges

• Prefix length	/24
• Subnet address	255.255.255.0
• Wildcard masks	inverted mask 0.0.0.255, not supported

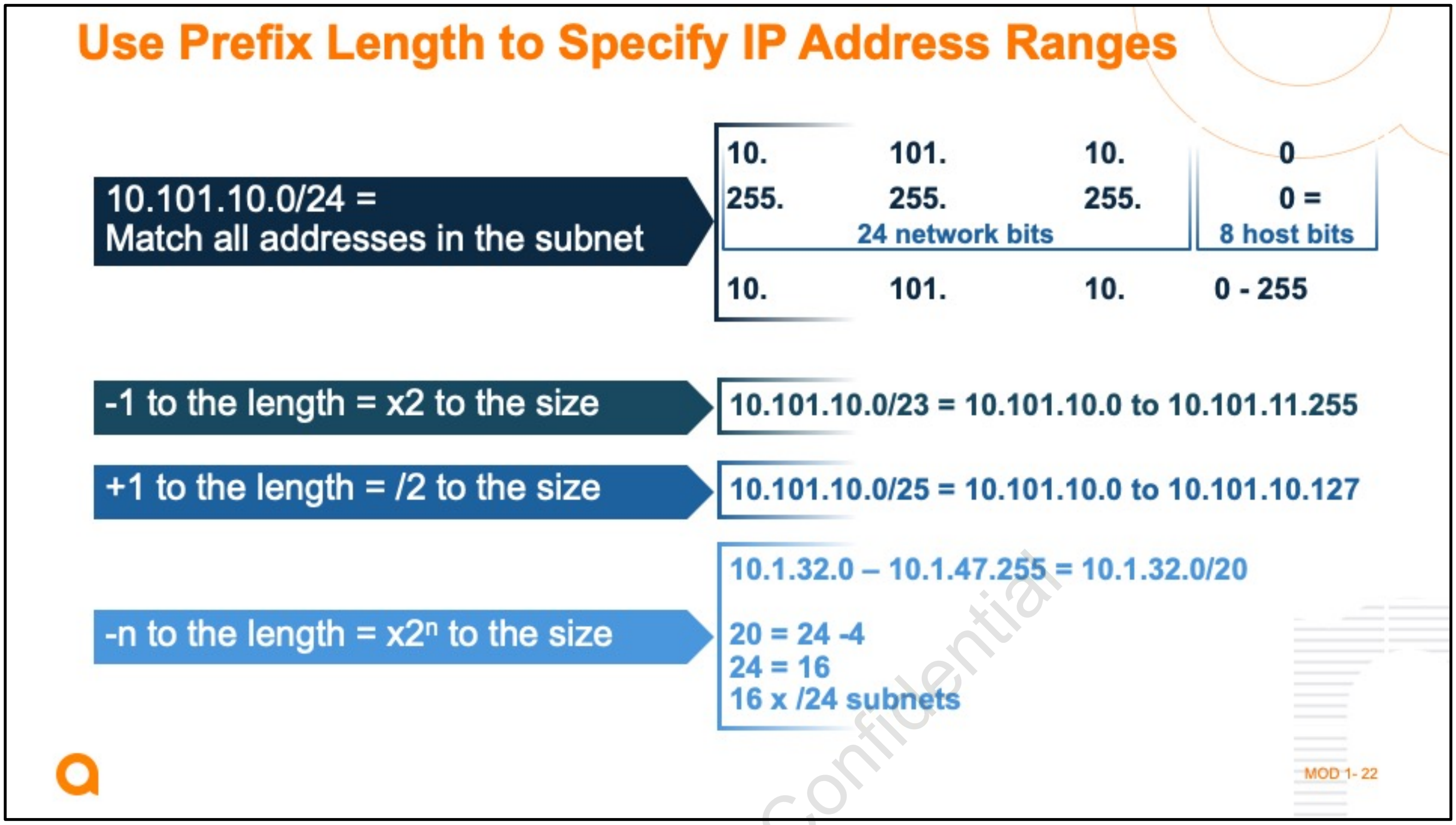
```
Switch# show access-list ip MY_ACL
Type Name
Sequence Comment
         Action                    L3 Protocol
         Source IP Address         Source L4 Port(s)
         Destination IP Address    Destination L4 Port(s)
         Additional Parameters
-----------------------------------------------------------
IPv4     MY_ACL
      10 permit                  udp
        any
         172.16.1.0/255.255.255.0
```

Aruba Training-Confidential

You can use a network address and prefix length to define a range of IP addresses in an ACL rule. This option works well to permit or deny an entire subnet. When you looked at the ACL, though, you saw that the range was denoted as shown in the figure. Instead of 10.101.10.0/24, you saw IP: 10.101.10.0 and Mask: 255.255.255.0. This mask is a subnet mask and you can actually specify the IP address and subnet mask, instead, when you create the rule.

To master both creating and interpreting rules, you need to explore prefix lengths and the relationships between them in more depth.

AOS-CX switches do not support wildcard masks—only prefixes or subnet masks—when creating ACEs.



Use Prefix Length to Specify IP Address Ranges

10.101.10.0/24 =
Match all addresses in the subnet

10.	101.	10.	0
255.	255.	255.	0 =
24 network bits			8 host bits
10.	101.	10.	0 - 255

-1 to the length = x2 to the size

10.101.10.0/23 = 10.101.10.0 to 10.101.11.255

+1 to the length = /2 to the size

10.101.10.0/25 = 10.101.10.0 to 10.101.10.127

-n to the length = $x2^n$ to the size

10.1.32.0 – 10.1.47.255 = 10.1.32.0/20

20 = 24 -4
24 = 16
16 x /24 subnets

MOD 1- 22

Aruba Training-Confidential

When you specify a subnet such as 10.101.10.0/24 in an ACL rule, the rule matches all IP addresses in the subnet from 10.101.10.0 to 10.101.10.255. (Of course, 10.101.10.0 is the subnet address and 10.101.10.255 is the broadcast address, so typical traffic will not use the addresses at either end of the range for the source.)

Subnets with the /24 prefix length are familiar and easy to use. However, you might want to conserve the number of rules within an ACL by specifying larger ranges of IP addresses. Or you might want to specify a subset of a /24 network in one rule rather than several identical rules with different host addresses.

To understand how to specify the desired range of IP addresses, briefly review prefix lengths and their relationship with subnet masks.

Subnet masks

In common IP addressing, a network (or subnet) mask defines which part of the IP address to use for the network (or subnet) address and which part to use for the hosts on the network:

The bits set to 1 define the part of the IP address to use for the network address.

The bits set to 0 define the part of the IP address to use for the host address.

In the example shown, the subnet mask is 255.255.255.0, or all 1s within the first three octets. Therefore, 10.101.10.0 is the network address. Valid host numbers in the fourth octet are between 1 and 254. Therefore, valid IP addresses that could be assigned to devices are from 10.101.10.1 through 10.101.10.254. The IP addresses 10.101.10.0 and 10.101.10.255 are reserved for identifying the subnet and broadcast addresses, respectively.

Note: Sometimes the bits are referred to as wildcard bits: 1 means to match and 0 can be ignored. For example, a mask of 255.255.0.0 would mean that the first two octets of the packet must match what's in the ACE rule, and the last two octets don't matter.

Aruba Training-Confidential

Prefix lengths

Classless Inter-Domain Routing (CIDR) notation represents a subnet mask by the prefix length. In other words, the prefix length specifies the number of bits set to 1 in the subnet mask. For example, the subnet mask in the slide has 24 1s and defines a /24 subnet.

Rules of thumb for specifying ranges with prefix lengths

You are probably familiar with /8, /16, and /24 subnets. These subnets have their boundaries at the octet border in decimal format, so it is easy to understand the range of addresses included in these subnets.

Sometimes, though, you want to specify ranges of IP addresses of different lengths than these.

As a rule of thumb, you double the size of a subnet when you subtract one from the prefix length. You halve the size when you add one.

For example, 10.101.10.0/23 includes IP addresses in two /24 subnets—IP addresses from 10.101.10.0 to 10.101.11.255. 10.101.10.0/25 includes IP addresses from 10.101.10.0 to 10.101.10.127.

If you want to multiply or divide by more than 2, choose a factor that is a power of 2 such as 4, 8, 16, and so on, and use the related exponent as the number that you add or subtract from the length. For example, you want a rule to select traffic with IP addresses between 10.1.32.0 and 10.1.47.255. This range has 16 /24 subnets, so you subtract 4 from 24 (24 = 16). You would specify 10.1.32.0/20.

As another example, you want a permit rule to select traffic with source IP addresses between 192.168.4.32 and 192.168.4.39. This range is eight times the size of 192.168.4.32/32, so you would specify 192.168.4.32/29 (23 = 8).

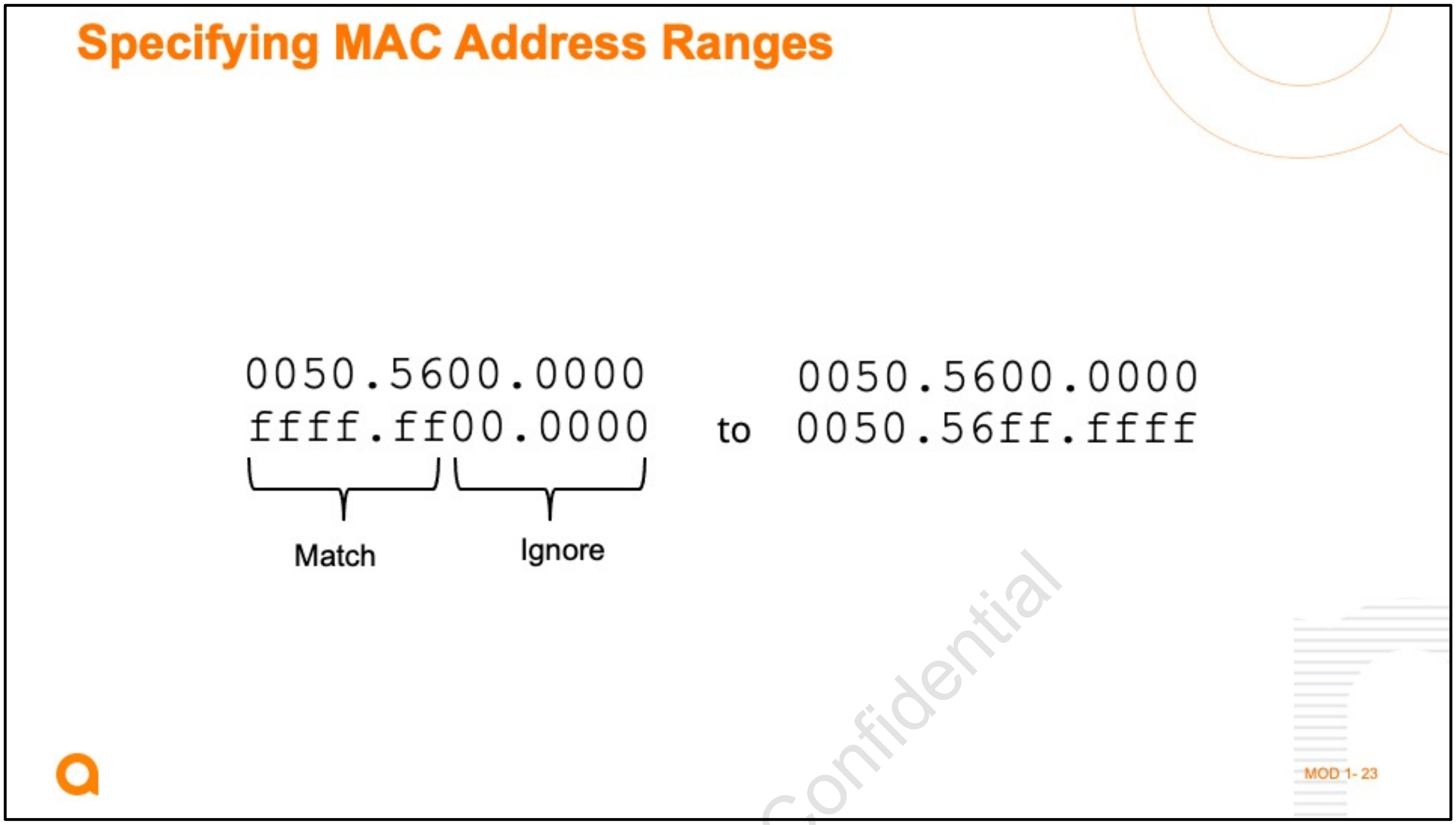
When you are specifying ranges in this way, you must remember that you are specifying the network address for a subnet. For example, you could not specify 192.168.4.30/29 to select the eight addresses from 192.168.4.30-192.168.4.37. This range includes addresses in two different /29 subnets, and 192.168.4.30 is not a valid network address for a /29 subnet. An easy way to check if an IP address is a valid network address for a subnet with a length between /24 and /32 is to determine the subnet's size and make sure that the final octet is divisible by that size. For example, a /29 subnet has 8 IP addresses. 32 is divisible by 8, but 30 is not. You also need to make sure that your range includes a number of addresses that is a power of two (such as 8, 16, 32, and so on).

If you want a range of a different size, you must enter multiple rules. Sometimes it is easier to specify a broader range than you need in the rule and create rules with lower sequence numbers that define exceptions.

You can apply similar logic to larger subnets, paying attention to the most significant octet for that subnet. For example, a /20 subnet includes 16 /24 subnets, so the third octet of the network address must be divisible by 16, and the last octet should be zero.

You can find prefix length calculators on the Internet to help you check your addresses.

Aruba Training-Confidential



Specifying MAC Address Ranges

```
0050.5600.0000          0050.5600.0000
ffff.ff00.0000    to    0050.56ff.ffff
```

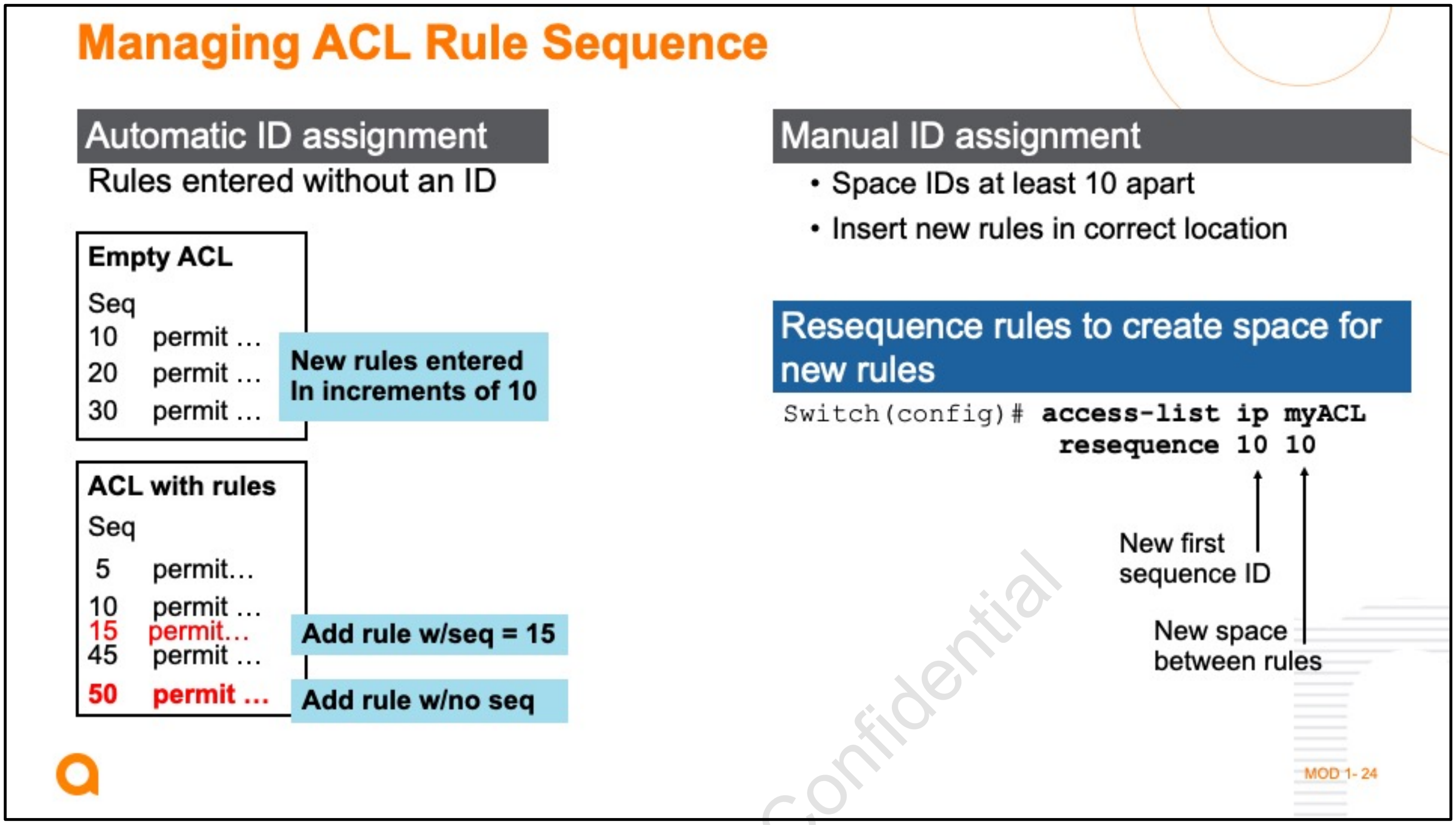
Aruba Training-Confidential

You must use wildcard bits to specify a range of MAC addresses in a MAC ACL. However, these bits use hexadecimal format. Each f indicates that a frame’s source or destination MAC address must match the corresponding hexadecimal character in the address in the rule. A zero tells the switch to ignore the corresponding character in frame addresses.

For example, to match 0050.5600.0000 ffff.ff00.0000, a frame can have any source MAC address that begins with 0050.56.

You can create more complicated rules using different hexadecimal characters in the mask. Determine the exact bits with source or destination MAC addresses that you want the switch to check or ignore. Set checked bits to 0s and ignored bits to 1; then convert to hexadecimal. For example, you enter 0050.5600.0000 ffff.ffff.fffc. This mask tells the switch to check all bits for all hexadecimal characters until the last character. At that point, the switch checks the first two bits but ignores the last two bits (1100 = hexadecimal c). This rule selects packets with destination MAC addresses 0050.5600.0000 through 0050-5600-0003. As you see, you can use the guideline of subtracting the first address from the last address to obtain the wildcard bits.

However, complicated rules such as these are usually unnecessary.



Managing ACL Rule Sequence

Automatic ID assignment

Rules entered without an ID

Empty ACL

Seq	
10	permit …
20	permit …
30	permit …

New rules entered In increments of 10

ACL with rules

Seq	
5	permit…
10	permit …
15	permit…
45	permit …
50	permit …

Add rule w/seq = 15

Add rule w/no seq

Manual ID assignment

- Space IDs at least 10 apart
- Insert new rules in correct location

Resequence rules to create space for new rules

```
Switch(config)# access-list ip myACL
                  resequence 10 10
```

Aruba Training-Confidential

If you do not manually assign ACE sequence numbers, the AOS-CX switch automatically assigns them. This means that rules are processed in the order that you enter them. Manual or automatic, you must plan the ACL carefully and configure rules for processing in the correct order, generally entering more specific rules with a lower ID than the ID for less specific rules.

Automatic numbering

If you do not enter a manual ID, the switch automatically adds rule IDs in order, leaving spaces. New rules are added to the end of the list, even if spaces exist earlier in the list.

For example, if the first rule that you enter does not have a manual ID, an AOS-CX switch numbers the rule as 10. Subsequent rules without rule numbers are numbered in increments of 10: 10, 20, 30, and so on. The switch always adds a rule without a manual ID after the final explicit rule. If this rule has an ID that is not divisible by 10, such as 45, the switch adds the new rule with the next ID that is divisible by 10, such as 50.

Manual numbering

When you enter rules manually, it is best practice for you to space rule IDs, preferably by 10s, so that you can insert additional rules at a later point, if necessary. Then, when you do not want to add a rule to the end of the list, you can choose a rule ID between existing rules to insert the rule in the correct location.

Sequence renumbering

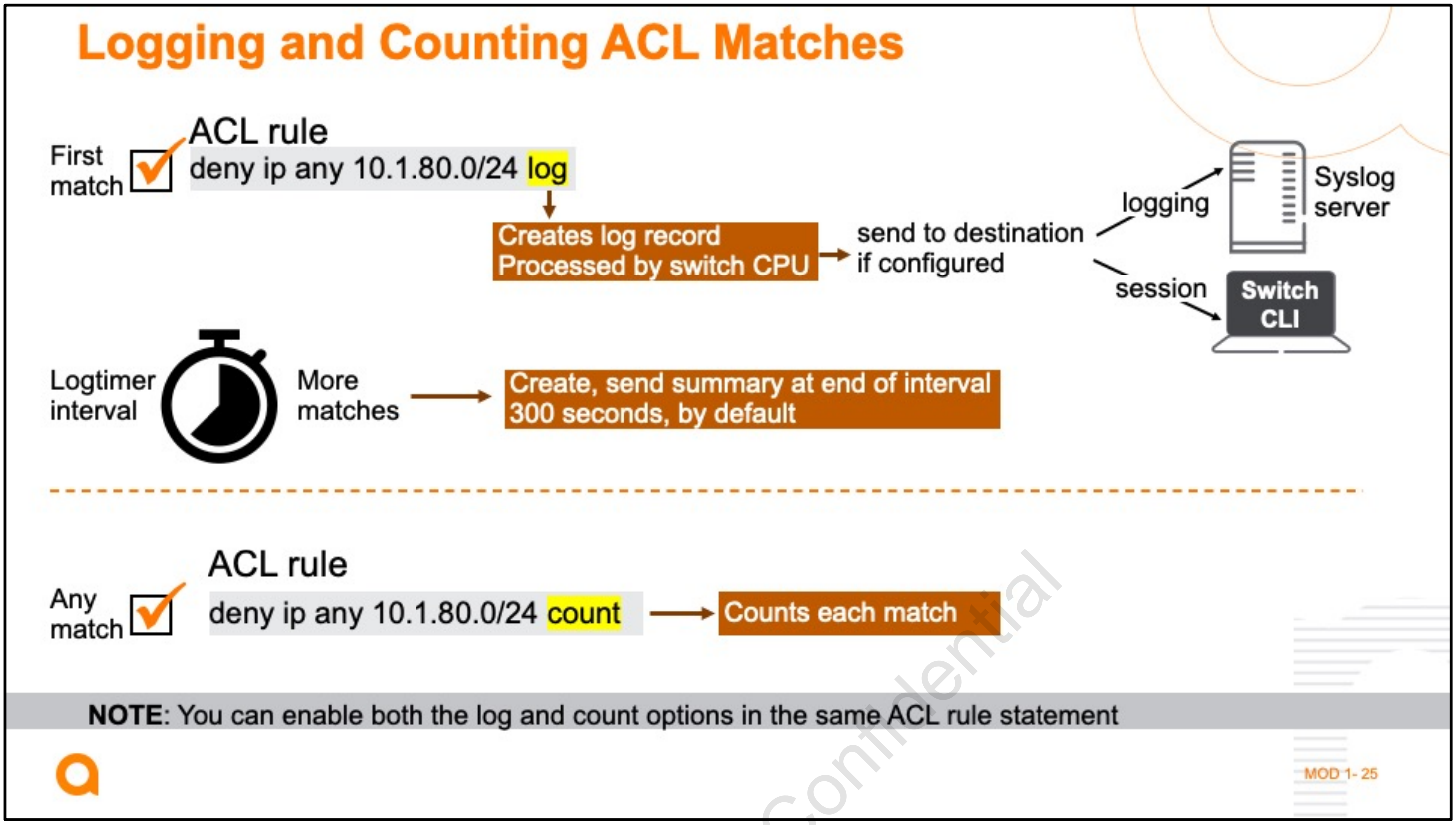
Sometimes, despite your best planning, you need to change rule IDs to place rules in the correct order, as shown in the figure and described below.

Switch(config)# access-list {ip | mac} <ACL-name> resequence <start-sequence> <increment>

<start-sequence> is the sequence number you want to assign to the first rule, and <increment> is the incrementing value by which subsequent rules will be numbered. This command leaves the rules in the same relative order. However, with more spaces between the rules, you can remove rules and add them in the appropriate location as necessary.

Remove a rule using the no <sequence-number> command.

Aruba Training-Confidential



Logging and Counting ACL Matches

NOTE: You can enable both the log and count options in the same ACL rule statement

Aruba Training-Confidential

You now have a good sense of how to create rules. In addition to the other options, an ACL rule on an AOS-CX switch can include the log or count options at the end.

Logging

The log option tells the switch to generate an Event Log message when a packet matches a rule. The log is created if the action is “deny” or “permit,” if there is a match, and if you have enabled ACL logging.

Depending on platforms, the log keyword also increments hit counts for the specified entries (i.e. log performs an implicit count). The first packet that hits any log entry is copied to the CPU and logged to the operator’s specified destination - either console or syslog server. This first packet starts the 5-minute ACL logging timer, configurable from 30s to the default of 300s. When the timer expires, the summary of all hit counts per ACL entry is sent to the specified logging destination. This capability allows throttling of logging ACL hits. To change the default logging period, use the following command:

```
Switch(config)# access-list log-timer {default|<VALUE>}
```

When using multiple ACL types (IPv4, IPv6, or MAC) with logging on the same interface, the first packet that matches an ACE with the log option is logged. Any packets, matching other

ACL types do not create a log until the log-timer wait-period is over. At the end of the wait-period, a summary log is made of all the ACLs that were matched, regardless of type.

Note: You may see a minor discrepancy between the ACL logging statistics and the hit counts statistics due to the time required to record the log message.

To enable ACL logging to a syslog server, you must define a syslog server and a logging level:

```
Switch(config)# logging {<IPV4-ADDR> | <IPV6-ADDR> | <HOSTNAME>}
[udp [<PORT-NUM>] |
            tcp [<PORT-NUM>]| tls [<PORT-NUM>]] [include-auditable-events]
            [severity <LEVEL>] [vrf <VRF-NAME>]
```

The default port for UDP, the most common protocol used for syslog, is 514.

The supported logging levels include the following:

alert: Forwards syslog messages with the severity of alert (6) and emergency (7).

crit: Forwards syslog messages with the severity of critical (5) and above.

debug: Forwards syslog messages with the severity of debug (0) and above.

emerg: Forwards syslog messages with the severity of emergency (7) only.

err: Forwards syslog messages with the severity of err (4) and above

info: Forwards syslog messages with the severity of info (1) and above. Default.

notice: Forwards syslog messages with the severity of notice (2) and above.

warning: Forwards syslog messages with the severity of warning (3) and above.

Certificate and TLS logging are supported. See the AOS-CX Command-Line Interface Guide for more information.

Hit counts

The count key word will keep track of each match on an ACE entry. To reset

Aruba Training Confidential

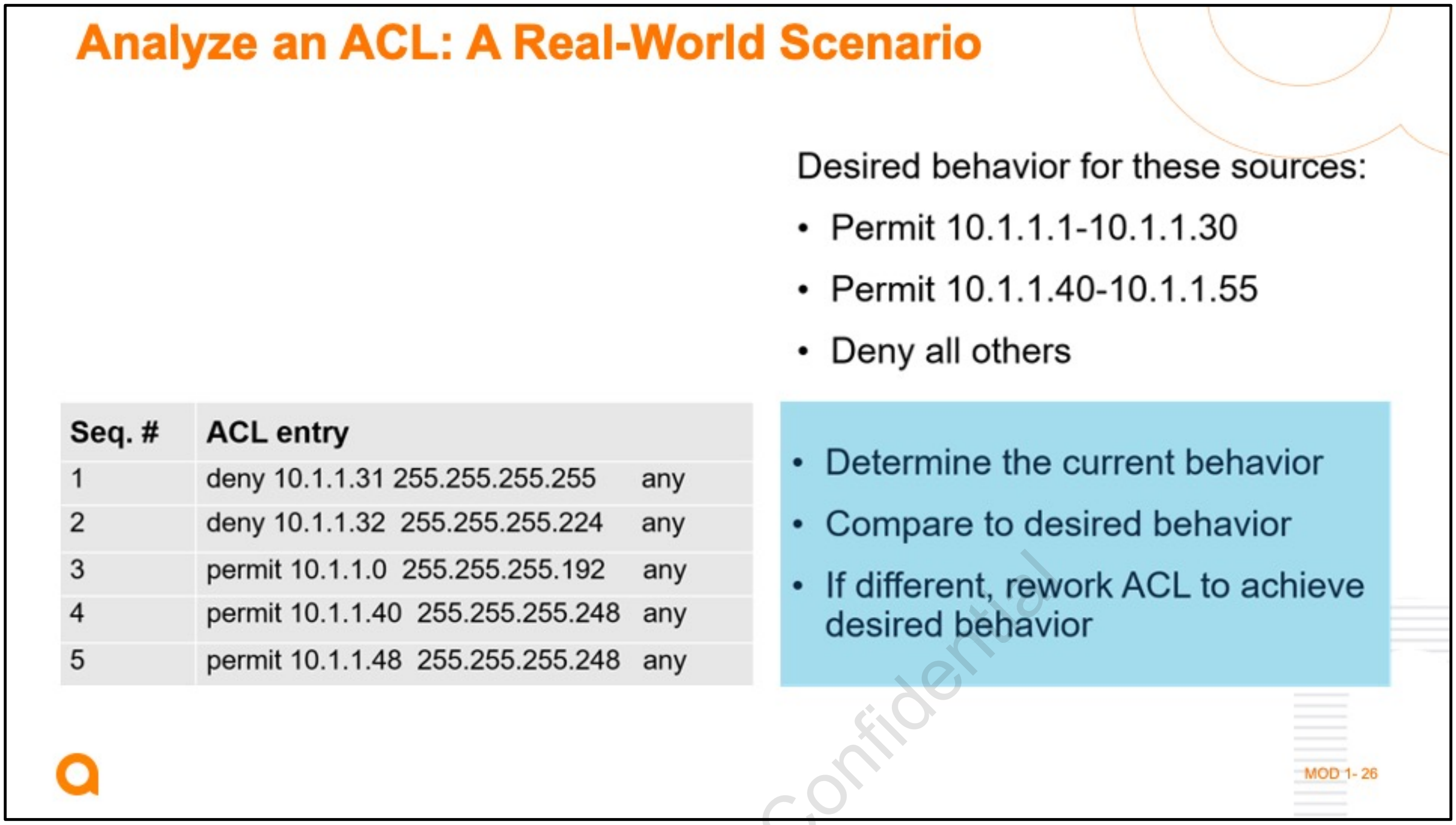
the hit counts for all ACLs, a specific ACL, or a specific interface, use the following command:

```
Switch# clear access-list hitcounts {all | [{ip|ipv6|mac} <ACL-NAME>]
[interface <IFNAME>|
      vlan <VLAN-ID>] [in|out]}
```

Note: You can add the log and count functions in a single ACL rule statement.

Important: Logging information is processed by the CPU of the switch.

Aruba Training-Confidential



Analyze an ACL: A Real-World Scenario

Desired behavior for these sources:

- Permit 10.1.1.1-10.1.1.30
- Permit 10.1.1.40-10.1.1.55
- Deny all others

Seq. #	ACL entry	
1	deny 10.1.1.31 255.255.255.255	any
2	deny 10.1.1.32 255.255.255.224	any
3	permit 10.1.1.0 255.255.255.192	any
4	permit 10.1.1.40 255.255.255.248	any
5	permit 10.1.1.48 255.255.255.248	any

- Determine the current behavior
- Compare to desired behavior
- If different, rework ACL to achieve desired behavior

MOD 1- 26

Aruba Training-Confidential

Scenario

Assume that the ACL shown in the figure is applied outbound to a VLAN interface, and a switch is now comparing packets being routed out that VLAN against the ACL.

You want the switch to permit only packets with these source IP addresses:

10.1.1.1-10.1.1.30. Understand that 10.1.1.0 is the network address in this case, so it is okay if the rule permit packets from the address. The switch will not route packets from it in any case.

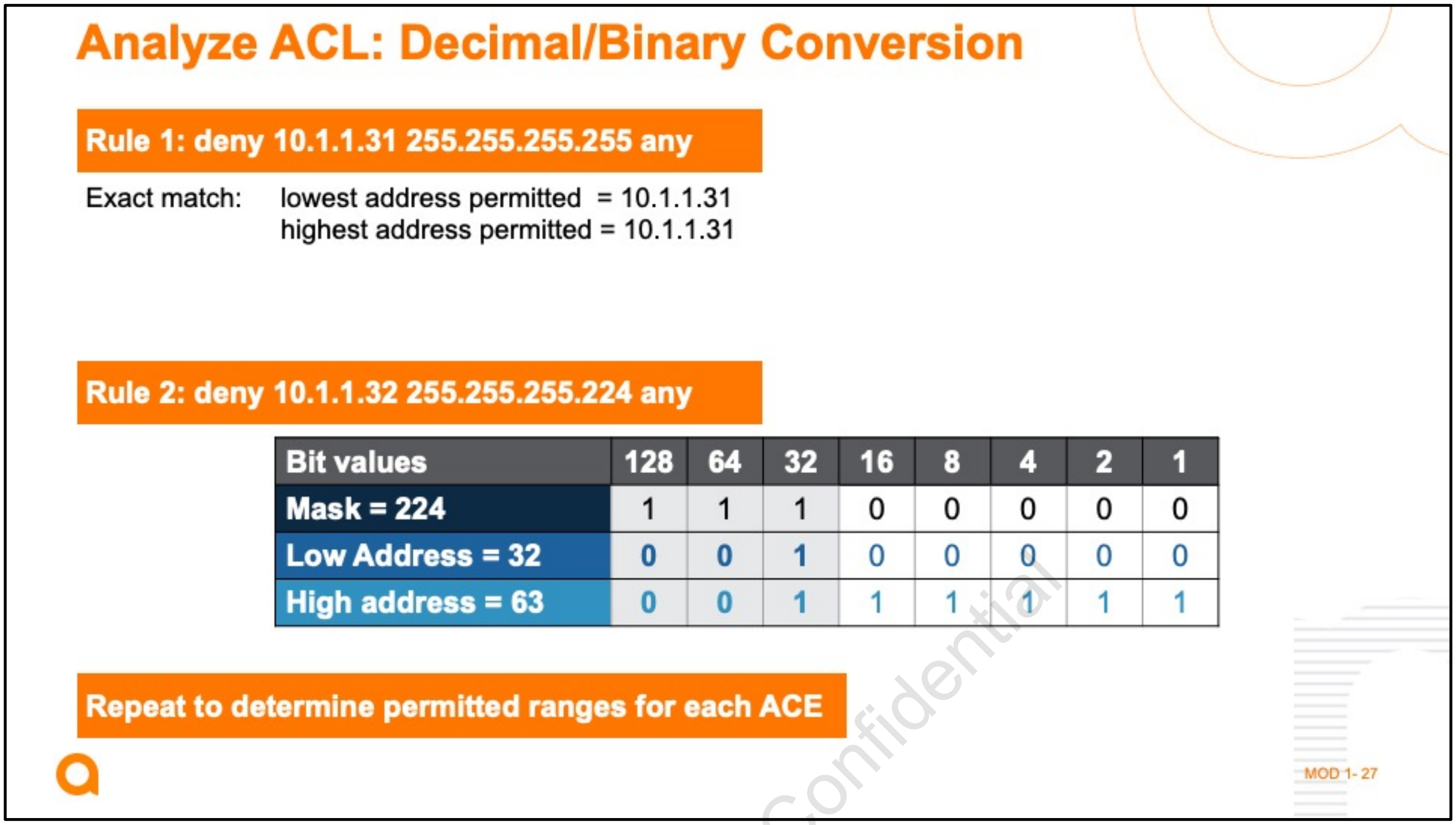
10.1.1.40-10.1.1.55

Objectives

You have two objectives for this exercise:

Describe the actual, current behavior of this scenario.

Write a new ACL to get the desired behavior.



Analyze ACL: Decimal/Binary Conversion

Rule 1: deny 10.1.1.31 255.255.255.255 any

Exact match: lowest address permitted = 10.1.1.31
highest address permitted = 10.1.1.31

Rule 2: deny 10.1.1.32 255.255.255.224 any

Bit values	128	64	32	16	8	4	2	1
Mask = 224	1	1	1	0	0	0	0	0
Low Address = 32	0	0	1	0	0	0	0	0
High address = 63	0	0	1	1	1	1	1	1

Repeat to determine permitted ranges for each ACE

MOD 1- 27

Aruba Training-Confidential

Rule 1 is deny 10.1.1.31 255.255.255.255 any

No real need to convert this line to binary, because the mask is 255.255.255.255 - this means that all 32 bits of the specified address must match. Thus, the lowest address permitted by this line is 10.1.1.31, and the highest address permitted is also 10.1.1.31.

Rule 2 is deny 10.1.1.32 255.255.255.224 any

For this and all remaining rules, the first three octets of the mask = 255. This means that the first three octets must exactly match the specified pattern of 10.1.1. It is that fourth octet that must be converted to binary, along with its associated mask. To do this, setup a grid on your scratch paper with 8 headings - the weighted value of each column in an 8-bit binary number:
128 64 32 16 8 4 2 1

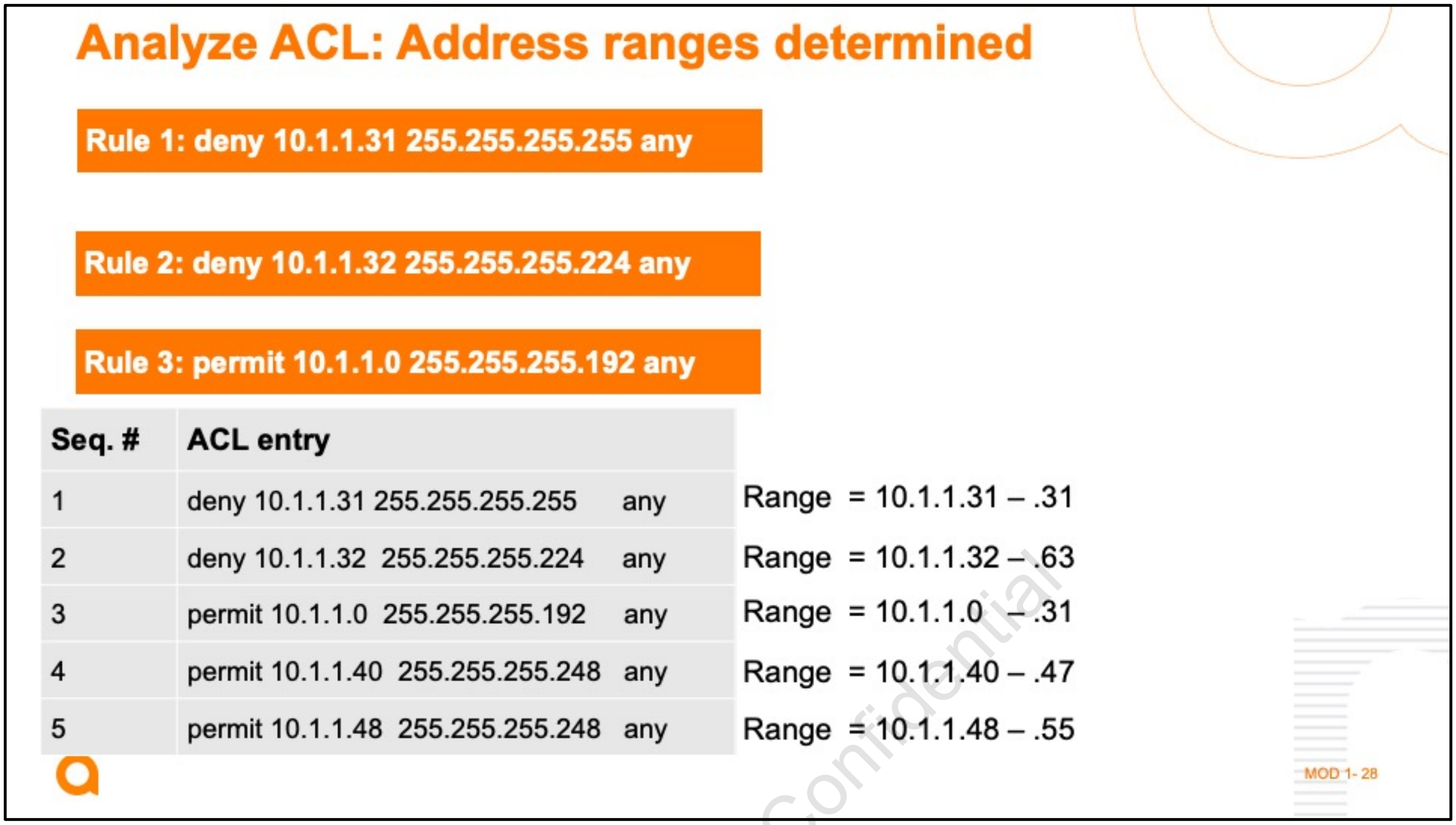
Now use these headings to convert the mask - decimal 224 into binary. Ask yourself, "Is 224 greater than 128?" If so, place a binary 1 in that column. If not, place a 0. Of course the 128's column = 1. Now add the next column - 128 = 64 = 192. Is 224 greater than 192? Yes, so the 64's column gets a 1.Add the next column 192+32 = 224, and that's our target. So, decimal 224 = binary 11100000.

Now, convert the 4th octet of the specified address (32) into binary. 32 is not greater than 128, nor 64, and so those columns get a 0. 32 is equal to the next column, so you get a 1 in the 32's column, and 0's on the remaining columns. Decimal 32 = binary 00100000. The 224 mask means that the first 3 bits are 1s. This tells us that the first 3 bits must exactly match the specified pattern of 001 (the high order bits of decimal 32). The rest of the bits are zero, and you can't be lower than that, so the lowest address is 32.

Now determine the highest address denied by this address/mask combination. You know the first three bits must be 001 (as dictated by the mask). So, simply place 001 in the 3 high-order bit positions, and place 1s in the remaining bit positions - you can't be higher than all 1s. Now convert this number into decimal. There are 0s in the 128 and 64's column, so those are not added. There are 1s in the remaining columns, so add the bit values - 32+16+8+4+2+1 = 63.

So, the address/mask 10.1.1.32 255.255.255.224, the range of addresses to acted upon (denied in this case) is 10.1.1.32 - 10.1.1.63.Use this technique to determine the remaining values.

Aruba Training-Confidential



Analyze ACL: Address ranges determined

Rule 1: deny 10.1.1.31 255.255.255.255 any

Rule 2: deny 10.1.1.32 255.255.255.224 any

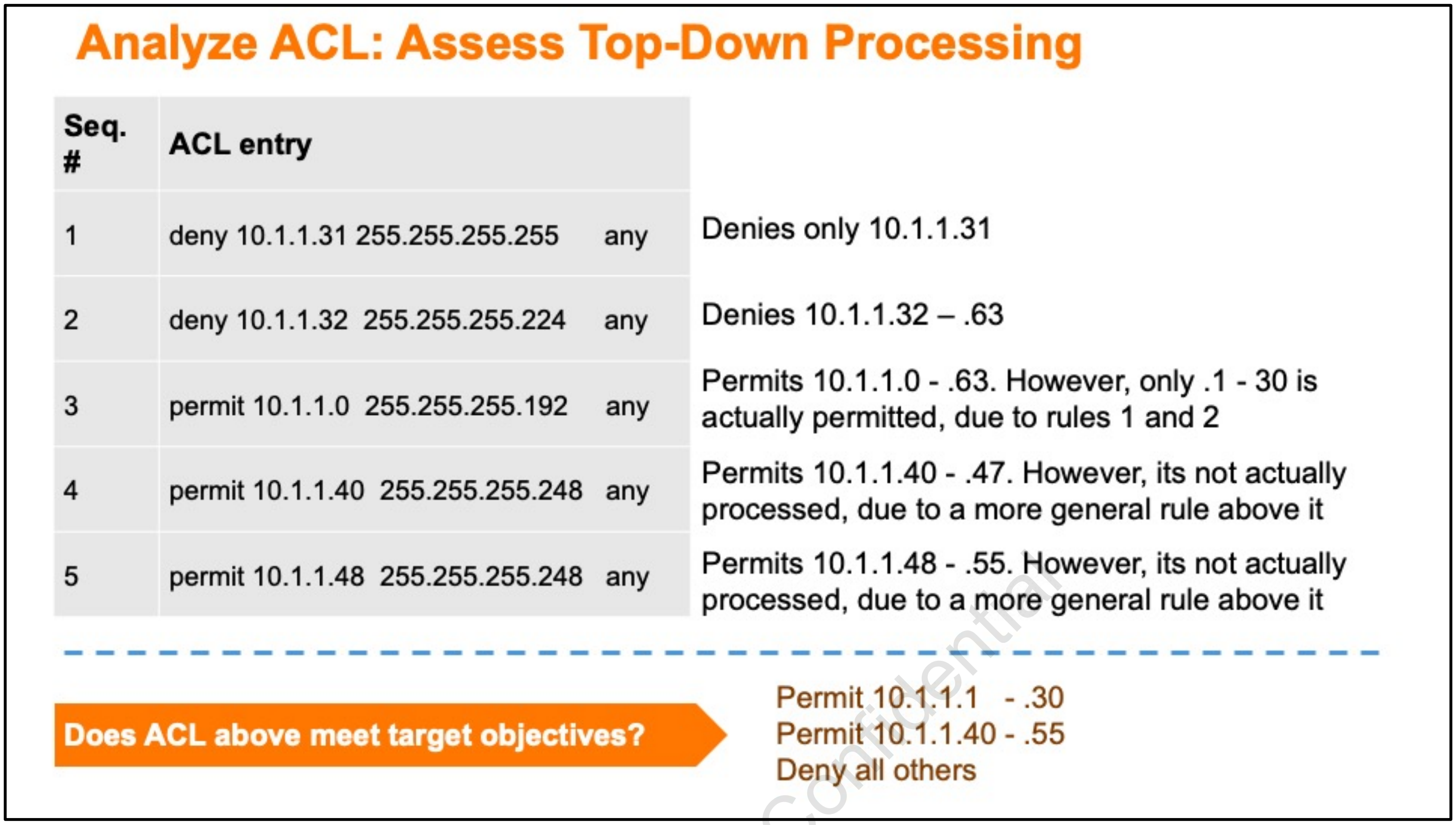
Rule 3: permit 10.1.1.0 255.255.255.192 any

Seq. #	ACL entry		
1	deny 10.1.1.31 255.255.255.255	any	Range = 10.1.1.31 – .31
2	deny 10.1.1.32 255.255.255.224	any	Range = 10.1.1.32 – .63
3	permit 10.1.1.0 255.255.255.192	any	Range = 10.1.1.0 – .31
4	permit 10.1.1.40 255.255.255.248	any	Range = 10.1.1.40 – .47
5	permit 10.1.1.48 255.255.255.248	any	Range = 10.1.1.48 – .55

MOD 1- 28

Aruba Training-Confidential

This is the result of using the previously described technique to determine the address ranges permitted or denied by each rule in the current-state ACL.



Analyze ACL: Assess Top-Down Processing

Seq. #	ACL entry		
1	deny 10.1.1.31 255.255.255.255	any	Denies only 10.1.1.31
2	deny 10.1.1.32 255.255.255.224	any	Denies 10.1.1.32 – .63
3	permit 10.1.1.0 255.255.255.192	any	Permits 10.1.1.0 - .63. However, only .1 - 30 is actually permitted, due to rules 1 and 2
4	permit 10.1.1.40 255.255.255.248	any	Permits 10.1.1.40 - .47. However, its not actually processed, due to a more general rule above it
5	permit 10.1.1.48 255.255.255.248	any	Permits 10.1.1.48 - .55. However, its not actually processed, due to a more general rule above it

Does ACL above meet target objectives?

Permit 10.1.1.1 - .30
Permit 10.1.1.40 - .55
Deny all others

Aruba Training-Confidential

This is the result of using the previously described technique to determine the address ranges permitted or denied by each rule in the current-state ACL.

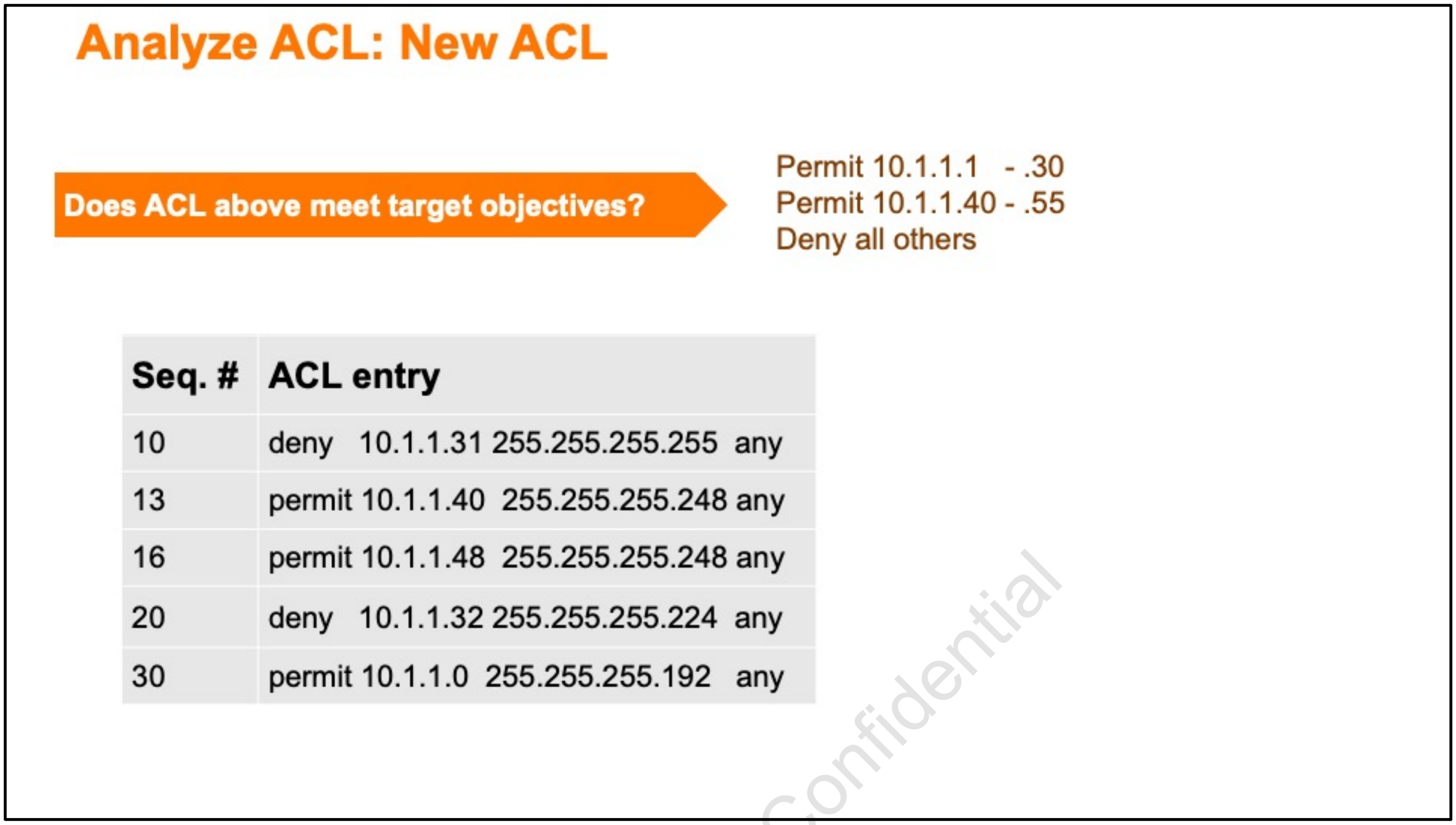
Given what you have just determined, does the current behavior achieve the desired objective?

Desired objective:

Permit 10.1.1.1-10.1.1.30

Permit 10.1.1.40-10.1.1.55

Deny all others



Analyze ACL: New ACL

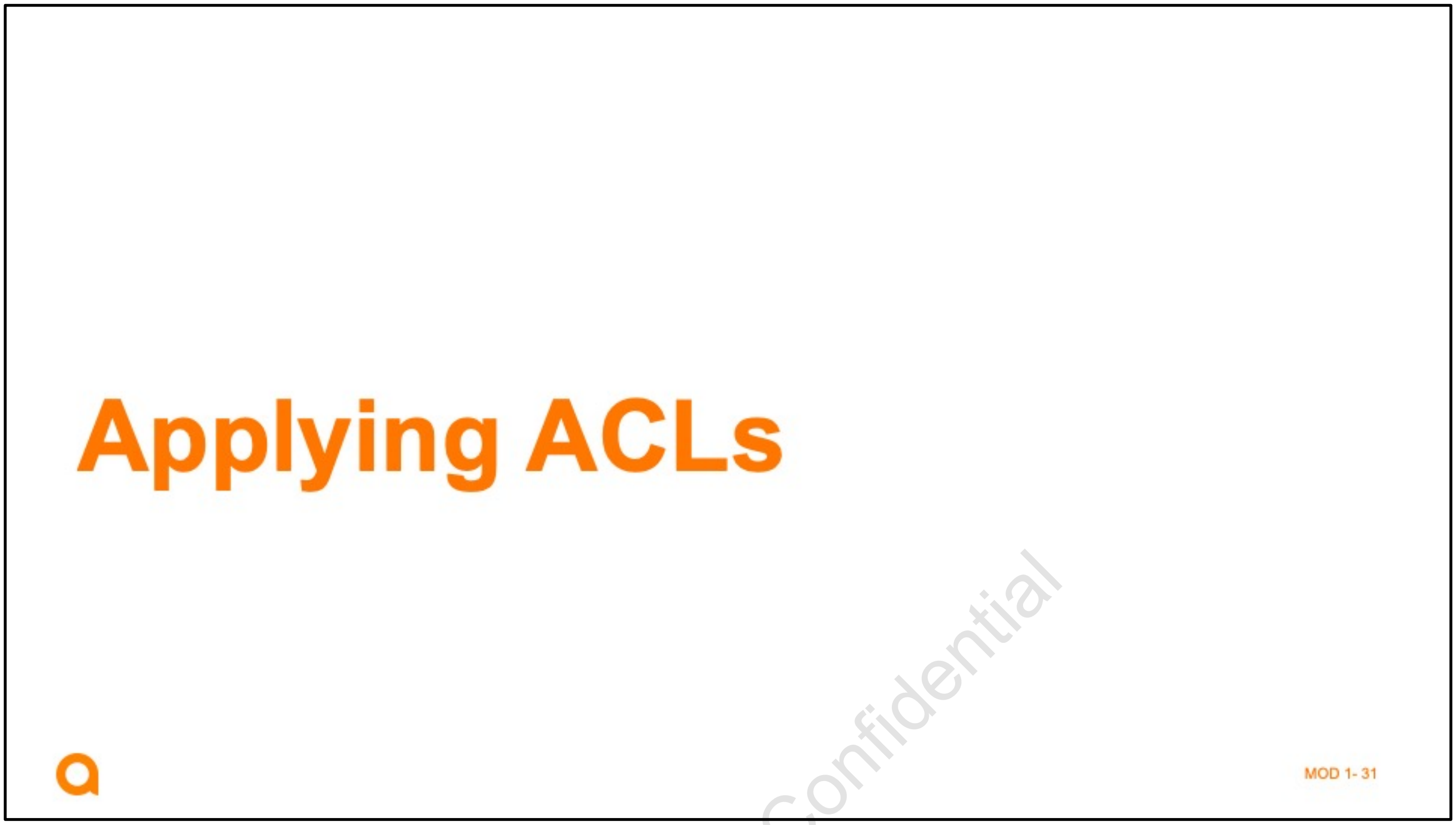
Does ACL above meet target objectives?

Permit 10.1.1.1 - .30
Permit 10.1.1.40 - .55
Deny all others

Seq. #	ACL entry
10	deny 10.1.1.31 255.255.255.255 any
13	permit 10.1.1.40 255.255.255.248 any
16	permit 10.1.1.48 255.255.255.248 any
20	deny 10.1.1.32 255.255.255.224 any
30	permit 10.1.1.0 255.255.255.192 any

The ACL shown achieves the desired objective.

Aruba Training-Confidential

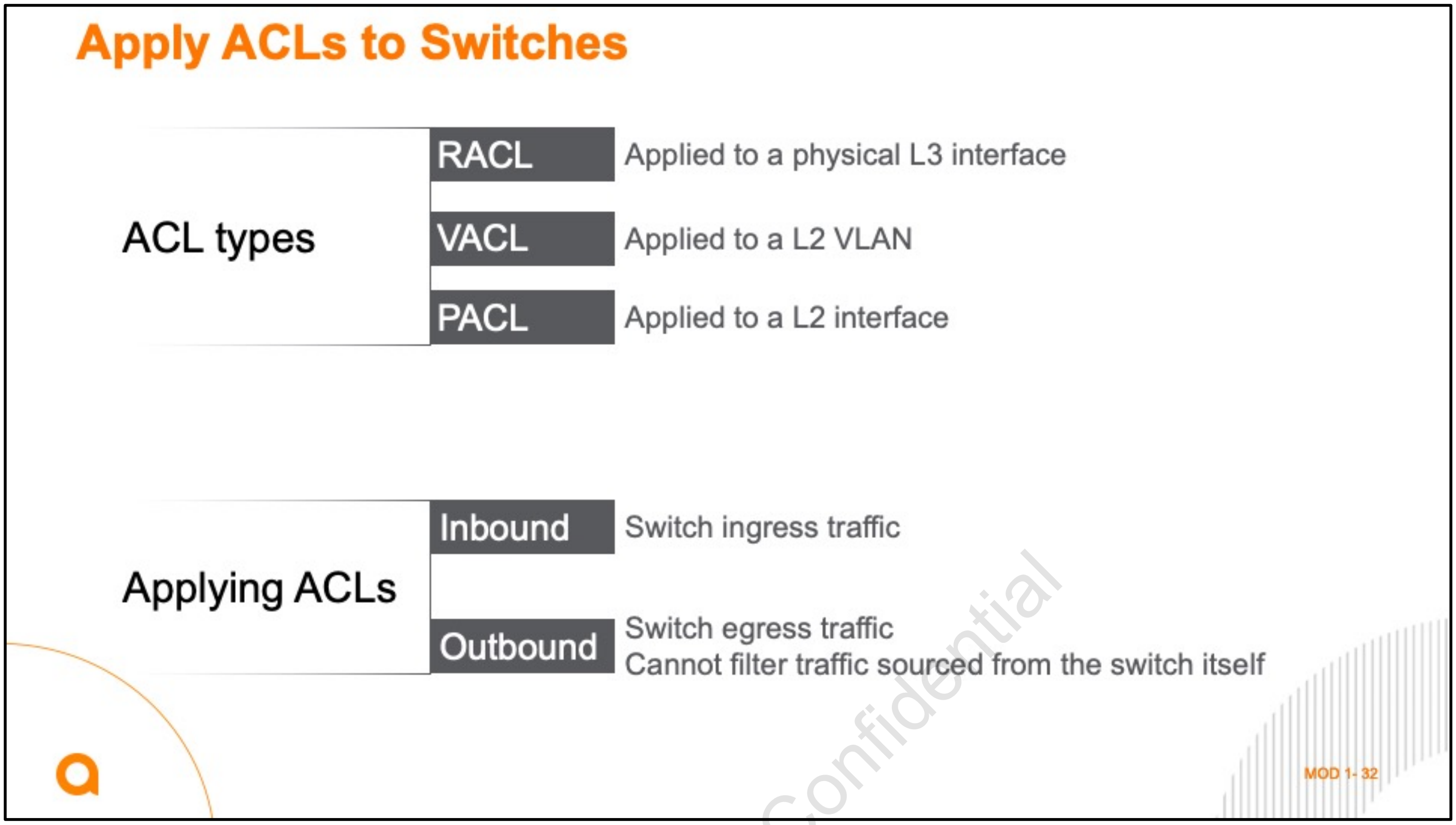


Applying ACLs

MOD 1- 31

You explored differences between RACLs and VACLs. But some differences only become clear in a topology of Layer 2 access switches and routes at the distribution layer or core. Let’s explore details about the traffic filtered by RACLs, VACLs, and PACLs.

Aruba Training-Confidential



Apply ACLs to Switches

ACL types

- RACL: Applied to a physical L3 interface
- VACL: Applied to a L2 VLAN
- PACL: Applied to a L2 interface

Applying ACLs

- Inbound: Switch ingress traffic
- Outbound: Switch egress traffic
Cannot filter traffic sourced from the switch itself

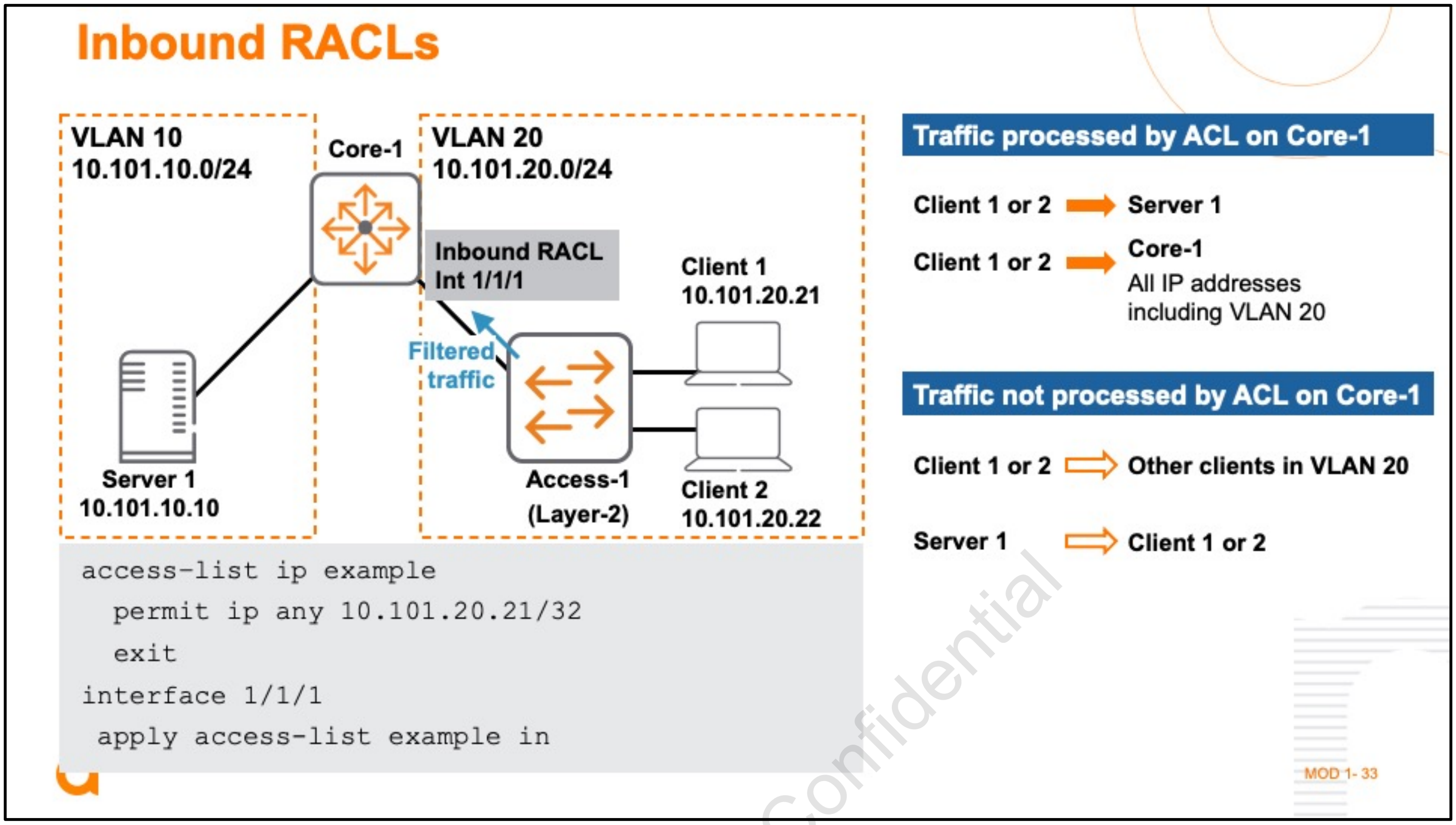
MOD 1- 32

Recall that you can apply ACLs to routed ports (RACLs), VLANs (VACLs), and Layer-2 ports (PACLs). Apply ACLs inbound to filter traffic entering the switch, and/or outbound to filter traffic leaving the switch. An outbound ACL cannot filter traffic sourced from the switch itself.

Note: As of AOS-CX 10.4, you cannot apply RACLs to Layer-3 VLAN interfaces. If you need to filter traffic for a Layer-3 VLAN interface, you must implement a classifier policy, discussed at the end of this module.

Let’s explore some scenarios that focus on each of these ACL types – both inbound and outbound.

Aruba Training-Confidential



Inbound RACLs

```
access-list ip example
  permit ip any 10.101.20.21/32
  exit
interface 1/1/1
 apply access-list example in
```

Traffic processed by ACL on Core-1

Client 1 or 2 → Server 1

Client 1 or 2 → Core-1
All IP addresses
including VLAN 20

Traffic not processed by ACL on Core-1

Client 1 or 2 ⇒ Other clients in VLAN 20

Server 1 ⇒ Client 1 or 2

MOD 1- 33

Aruba Training-Confidential

You should configure a RACL to control traffic at VLAN boundaries and filter routed traffic between subnets. RACLs apply to Layer-3 VLAN interfaces, and you should typically apply them on a routing switch that handles routing for that VLAN.

For example, Core-1 in the figure is the default gateway for VLAN 20, which is also configured on several access layer switches. To control traffic routed to or from VLAN 20, apply the RACL to VLAN 20 on Core-1.

Apply RACLs inbound or outbound - one RACL per-direction per-physical interface (Layer-2 or Layer-3). You apply the ACL to the interface in the interface context by specifying the direction: apply access-list <name> <in | out>.

An inbound RACL processes the following traffic:

Traffic received on this physical interface and routed out another subnet— In the figure, you applied the RACL inbound to routed interface 1/1/1, which is associated to VLAN 20 - the 10.101.20.0/24 subnet. The ACL filters packets that ingress interface 1/1/1 and are routed to other subnets. This includes traffic from Client 1 or 2 to Server 1. Understand this. Packets that “ingress interface 1/1/1” are packets that arrive untagged for VLAN 20 or with a VLAN tag of 20 on a port that permits such a tag. Because the ACL applies to traffic that this network infrastructure device routes out of this interface, the device must be routing the traffic.

Traffic received on this physical interface and destined to this switch—A RACL does not filter switched IP traffic unless the switch's own IP address on the interface is the destination. For example, as shown above, the ACL processes traffic from Client 1 or 2 to all IP addresses on Core-1, including the VLAN 20 IP address. This helps you to protect access to the switch itself.

Note: The RACL does not filter traffic generated on the interface by the switch itself.

Inbound RACLs do not process the following traffic:

Traffic received on the VLAN and switched on the same subnet (except to the switch itself). So, the inbound ACL does not process traffic between Client 1 and Client 2.

Traffic received on other VLANs and routed out this VLAN. The inbound ACL also does not process traffic from Server 1 to either of the clients. That would require the RACL to be applied outbound.

Important: A RACL filters traffic only in one direction; returning traffic may or may not be affected by ACLs in the return direction.

In the example above, if traffic is allows from VLAN 20 to VLAN 10 by the inbound RACL on Core-1, and since there is no ACL in the return direction, the return traffic from VLAN 10 to VLAN 20 would be allowed.

There is no ACL that affects traffic from VLAN 10 to reach VLAN 20; however, the returning traffic would be processed by the inbound RACL on interface 1/1/1. The figure shows an example ACL configuration for this scenario.

In the example this traffic would be processed by the ACL and allowed:

Client1 > Server1

Client1 > Core1

And this traffic would also be processed by the ACL, but denied:

Client2 > Server1

Client2 > Core1

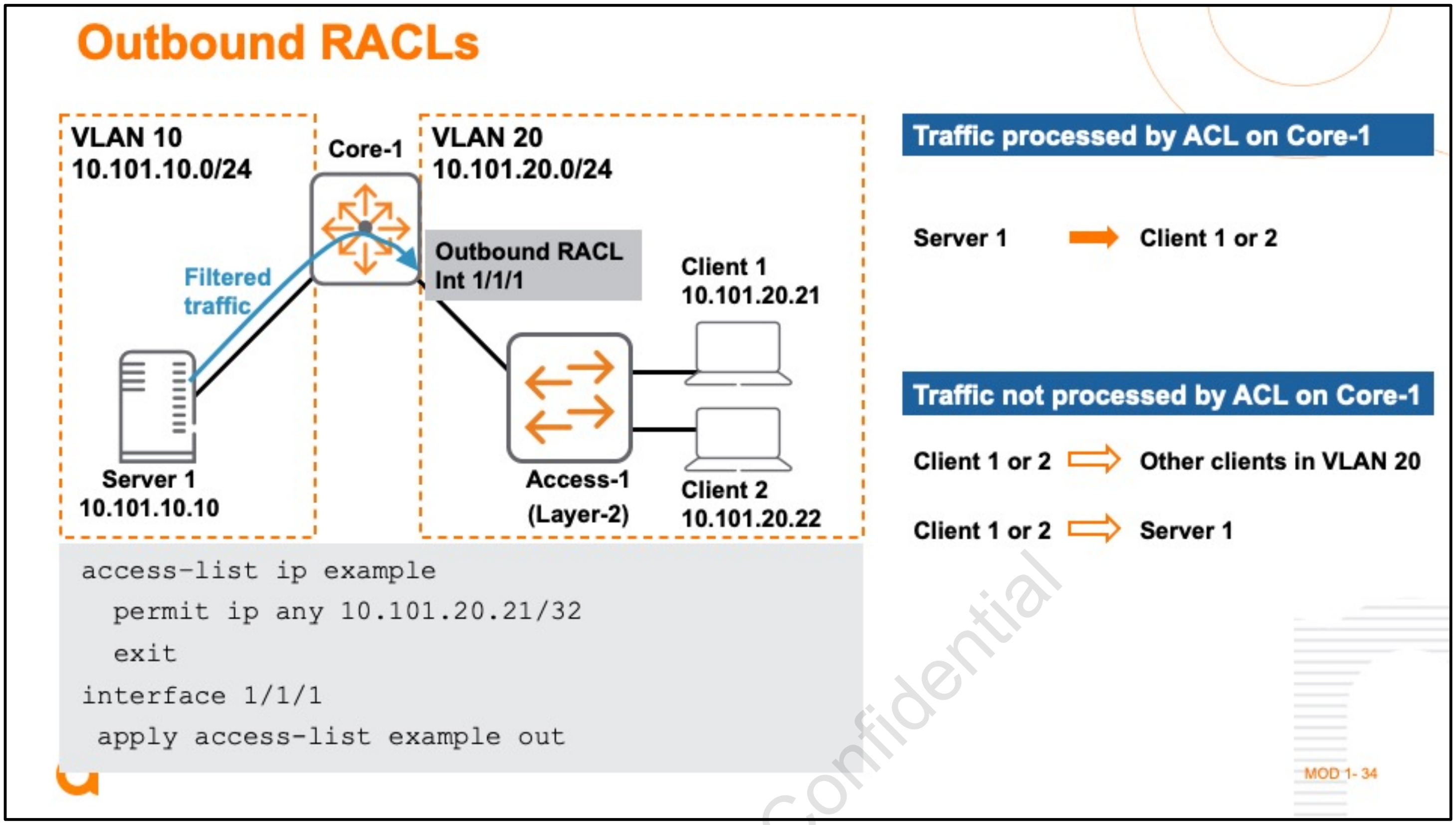
Aruba Training Confidential

And this traffic is not even processed by the ACL:

Client1 or Client2 > any other client in VLAN 20

Server1 > Client1 or Client 2 (but the returning traffic from these clients will be processed by the ACL)

Aruba Training-Confidential



Outbound RACLs

```
access-list ip example
  permit ip any 10.101.20.21/32
  exit
interface 1/1/1
 apply access-list example out
```

Traffic processed by ACL on Core-1

Server 1 → Client 1 or 2

Traffic not processed by ACL on Core-1

Client 1 or 2 ⇨ Other clients in VLAN 20

Client 1 or 2 ⇨ Server 1

MOD 1- 34

Aruba Training-Confidential

Outbound RACLs filter traffic received from any subnet and routed to a subnet on this VLAN interface. Here you have applied a RACL outbound to Core-1's 1/1/1 routed interface - subnet 10.101.20.0/24. The switch receives a packet from Server 1 on VLAN interface 10, and the packet is destined to Client 1 at 10.101.20.21. This ACL processes that packet because the switch routes the packet to VLAN 20.

An outbound RACL does not process the following traffic:

Traffic generated by the switch and routed out this VLAN interface

Traffic received on this VLAN and routed out another subnet. That would require an inbound RACL.

For example, the ACL shown in the figure does not filter traffic from the clients to the server. In other words, the ACL does not process any traffic from the clients.

You can apply only one RACL per-interface to inbound traffic and one RACL per-interface to outbound traffic. Plan the ACL rules accordingly. The figure shows an example ACL configuration for this scenario. Compare this to the previous example. What is the only difference?

Based on this example, this traffic would be processed by the ACL and allowed:

Server1 > Client1

And this traffic would also be processed by the ACL, but denied:

Server1 > Client2

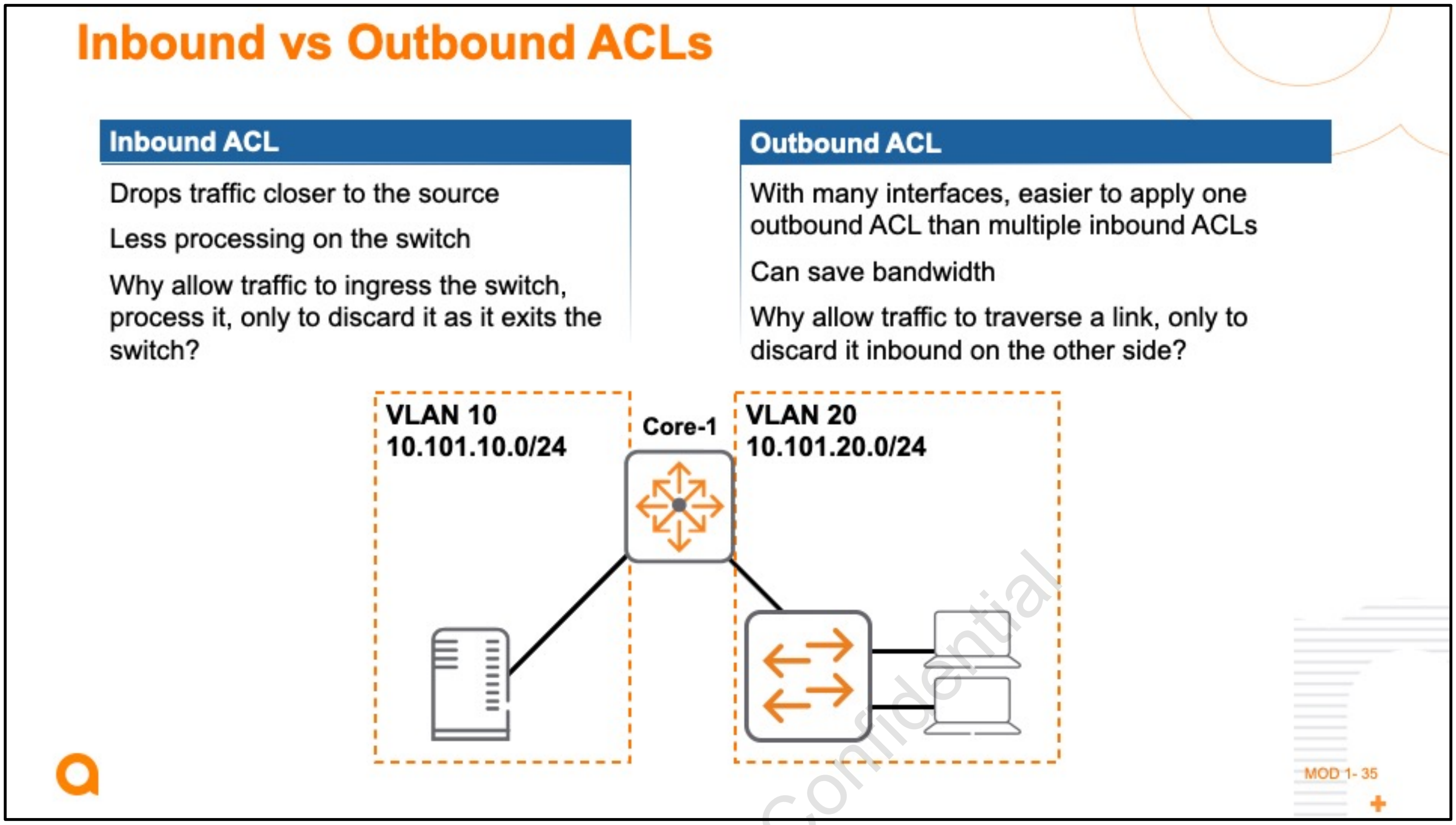
And this traffic is not even processed by the ACL:

Client1 or Client2 > any other client in VLAN 20

Client 1 or Client2 > Server1 (however, the returning traffic from Server1 back to the clients would be processed by the ACL)

Note: the ACL example is looking at traffic from the source to the destination. Remember that returning traffic might be affected the ACL, based the rules and how it is applied.

Aruba Training-Confidential



Inbound vs Outbound ACLs

Inbound ACL

Drops traffic closer to the source

Less processing on the switch

Why allow traffic to ingress the switch, process it, only to discard it as it exits the switch?

Outbound ACL

With many interfaces, easier to apply one outbound ACL than multiple inbound ACLs

Can save bandwidth

Why allow traffic to traverse a link, only to discard it inbound on the other side?

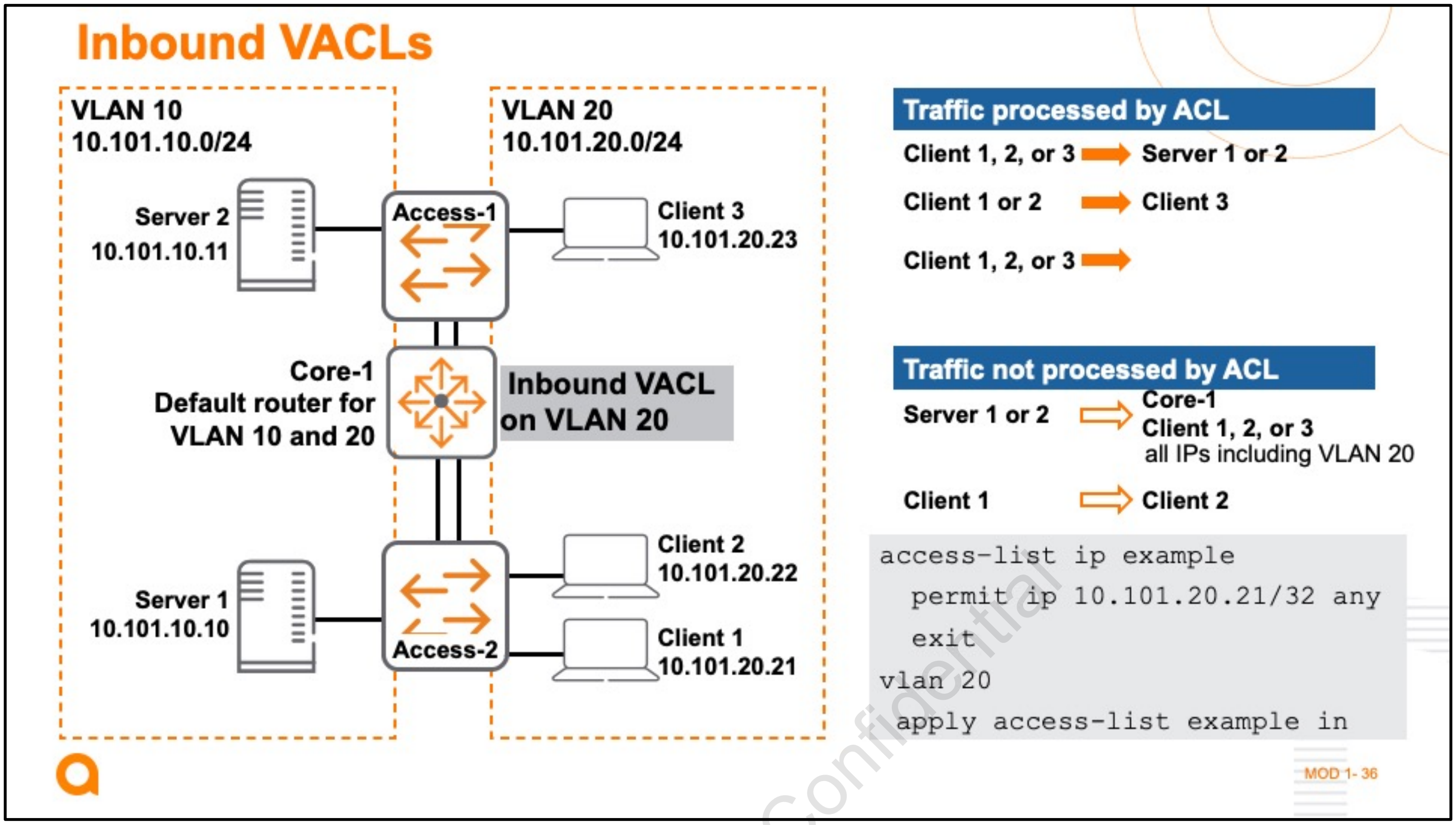
MOD 1- 35

So when should an ACL be applied inbound versus outbound, given that in many instances you could implement your policy with either?

The advantage of applying an ACL inbound is that you are dropping the traffic closer to the source, and thus less processing has to be done by the switch.

| However, if you have many, many interfaces, but only one exit interface, like an uplink, it might be easier to manage the application of the ACL by applying it outbound on the uplink interface. Also, ACLs do consume resources, as discussed towards at the end of the module, so for an uplink, applying it outbound on an uplink will consume less resources than applying it inbound on a multitude of ports. These are things you should consider when determining when creating an ACL policy as to whether it is better to use an inbound versus and outbound ACL solution.

Aruba Training-Confidential



Inbound VACLs

Traffic processed by ACL

Client 1, 2, or 3	→	Server 1 or 2
Client 1 or 2	→	Client 3
Client 1, 2, or 3	→	

Traffic not processed by ACL

Server 1 or 2	⇨	Core-1 Client 1, 2, or 3 all IPs including VLAN 20
Client 1	⇨	Client 2

```
access-list ip example
  permit ip 10.101.20.21/32 any
  exit
vlan 20
 apply access-list example in
```

MOD 1- 36

Aruba Training-Confidential

As you learned, RACLs applied to VLAN interfaces do not filter traffic that flows within the same subnet. When you want to control traffic switched within a subnet, you must apply a different type of ACL.

A PACL (PACL) filters both routed and switched traffic and meets this need. However, when you want to control all devices in a VLAN in the same way, a VACL is the most efficient mechanism. One VACL applies to all ports that are members of the VLAN. The VACL filters all IP traffic entering a switch on the specific VLAN interface. On a multinetted VLAN (a VLAN with multiple subnets), this includes inbound IP traffic from any subnet.

An inbound MAC ACL is always an inbound VACL. You can apply inbound or outbound IP VACL on 6300 and 6400 switches, but only inbound on 8400 and 832x switches.

Inbound VACLs process routed traffic like an inbound RACL but also filters switched traffic. The VACL processes this traffic:

Traffic received on this VLAN interface and routed out another subnet. In the figure this is traffic from clients to servers.

Traffic received on this VLAN interface and forwarded in the same subnet. This is traffic between the clients. However, the VACL placement affects which precise traffic is filtered, as discussed below.

Traffic received on this VLAN interface and destined to this network infrastructure device (on any interface)

The VACL does not process:

Traffic received on another interface and routed to this VLAN interface. However, the return traffic from a device in this VLAN would be filtered. For example, traffic from the servers to the clients is not processed by the ACL.

Traffic generated by the switch itself using its IP address on this VLAN and routed out another VLAN

The switch does not require an ACL rule to permit the DHCP traffic that it itself relayed. However, if the ACL were applied to a switch operating at Layer 2 and not implementing DHCP, you must permit DHCP (ports 67 and 68) to and from any address.

You can assign the same ACL to multiple VLANs. AOS-CX switches, however, allow you to configure just one inbound VACL per VLAN. This is in addition to any other ACLs of different types you assign to the VLAN or to ports in the VLAN. You also need to consider the correct locations to apply VACLs.

RACLs affect routed traffic, so you naturally apply them at the point at which the traffic that you want to control is routed. But VACLs, both inbound and outbound, control switched traffic, which does not necessary flow through a single point. For example, suppose that you convert the inbound RACL on VLAN 20 to a VACL. The VACL filters all the routed traffic that an inbound RACL would, and it also filters some switched traffic, such as that between Client 1 and Client 3. But traffic between Client 1 and Client 2 does not reach Core-1, so it is not filtered.

The figure shows an example ACL configuration for this scenario, applied inbound for VLAN 20 on Core-1.

Based on this example, this traffic would be processed by the ACL and allowed:

Client1 > Server1 or Server2

Client1 > Client3

Aruba Training-Confidential

Client1 > Core1

And this traffic would also be processed by the ACL, but denied:

Client2 and Client3 > Server1 or Server2

Client3 > Client1

Client2 > Client3

Client2 or Client3 > Core1

And this traffic is not even processed by the ACL:

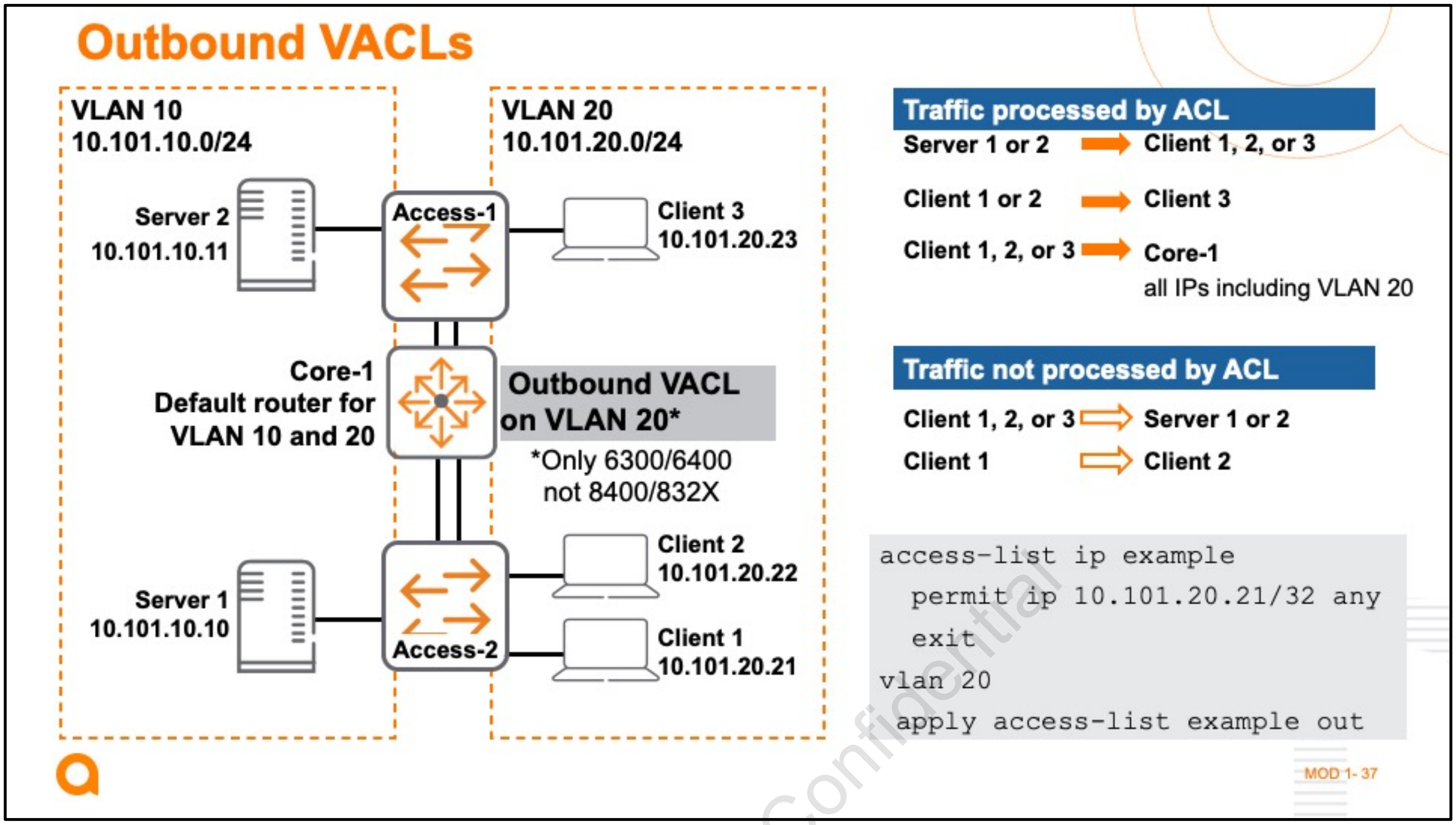
Client1 > Client2

Client2 > Client1

Server1 or Server2 > Client1, Client2, or Client 3 (however, the returning traffic would be processed by the ACL)

Note: the ACL example looks at traffic from source to destination. Remember that returning traffic might be affected by the ACL, based on the rules and how they are applied.

Aruba Training-Confidential



Outbound VACLs

Traffic processed by ACL		
Server 1 or 2	➡	Client 1, 2, or 3
Client 1 or 2	➡	Client 3
Client 1, 2, or 3	➡	Core-1 all IPs including VLAN 20

Traffic not processed by ACL		
Client 1, 2, or 3	⇨	Server 1 or 2
Client 1	⇨	Client 2

```
access-list ip example
  permit ip 10.101.20.21/32 any
  exit
vlan 20
 apply access-list example out
```

Aruba Training-Confidential

In a moment, you will see how VACLs work when applied closer to the edge. First lets examine an outbound VACL. IP ACLs must be applied to VLAN context with the out option. Remember, outbound VACLs are currently not supported on the 8400 and 832x switches.

Outbound VACLs filter traffic much like an outbound RACL, but it also filters traffic switched out the VLAN. Because, by definition, switched traffic arrives on and is forwarded in the same VLAN, this means that an outbound VACL also filters some traffic that arrives on the VLAN.

The outbound VACL processes this traffic:

Traffic received on another VLAN interface and routed out this VLAN. In the example, this server-to-client traffci.

Traffic received on this VLAN interface and switched out this VLAN - traffic between the clients. But remember, traffic between clients connected to the same access switch does not reach the ACL. The ACL also processes traffic from the clients to Core-1’s own IP address on VLAN 20.

If multinetting applies, traffic received on one subnet in this VLAN and routed out another subnet on this VLAN

The VACL does not process traffic received on this VLAN interface and routed out another VLAN. Here, this includes traffic from the clients to the servers.

The same considerations for applying inbound VACLs apply to outbound VACLs; each VLAN supports just one outbound VACL. And remember, Core-1's outbound VACL on VLAN 20 does not filter traffic between Client 1 and Client 2 - their traffic does not pass through Core-1. You must apply the VACL to Access-2 for this purpose.

The figure shows an example ACL configuration for this scenario, applied outbound for VLAN 20 on Core-1.

In the example, this traffic is processed by the ACL and allowed:

Client1 > Server1 or Server2

Client1 > Client3

Client1 > Core1

Server1 or Server2 > Client1 (the returning traffic might be filtered, however)

This traffic is processed by the ACL, but denied:

Client 2 > Client 3

Client2 and Client3 > Core-1 (all IP addresses including VLAN 20)

Client2 > Client3

Client2 or Client3 > Core1

And this traffic is not processed by the ACL:

Client1, Client2, or Client3 > Server1 or Server2

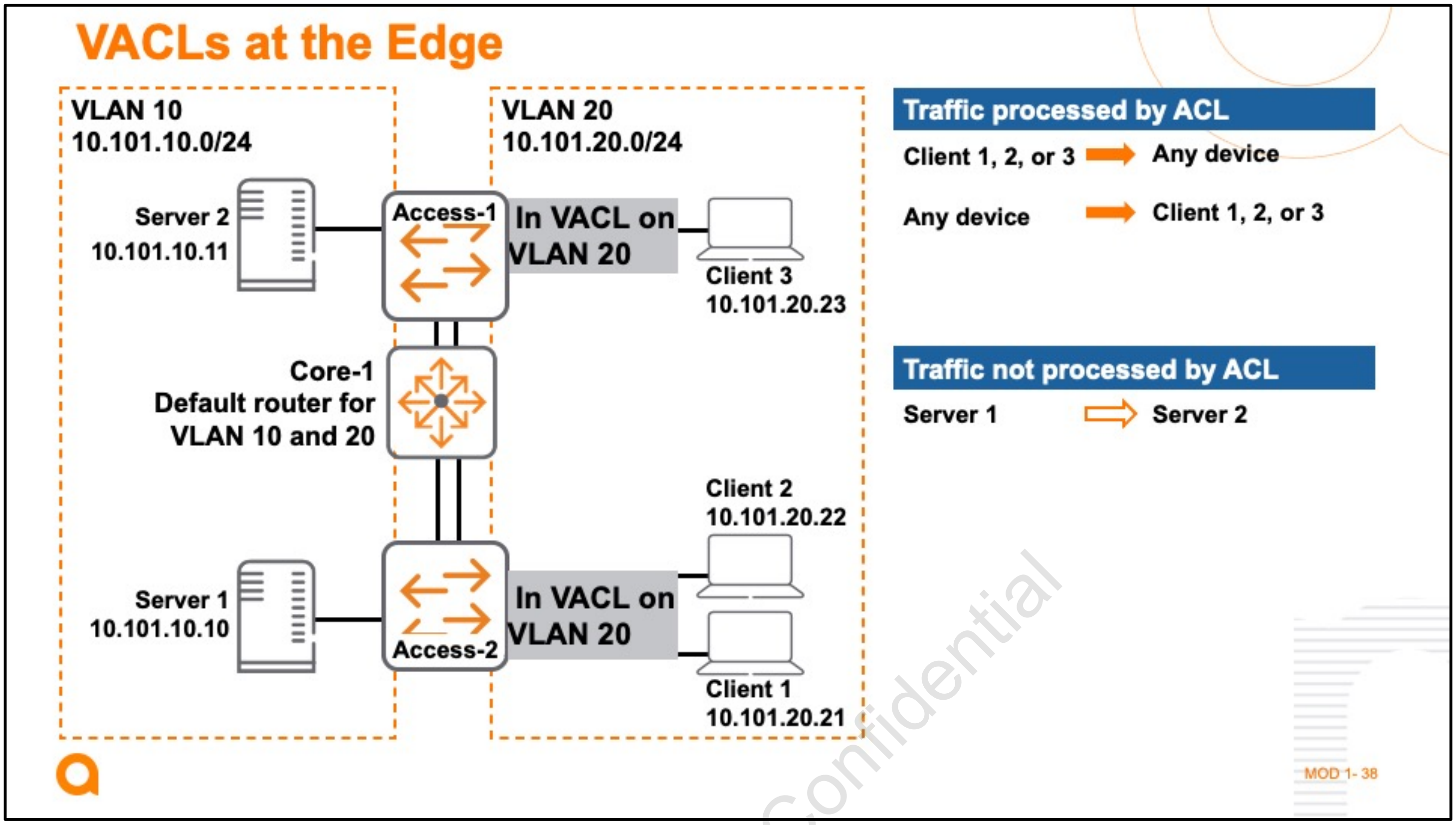
Client1 > Client2

Client2 > Client1

Server1 or Server2 > Client1, Client2, or Client 3 (however, the returning traffic would be processed by the ACL)

Note: the ACL example is looking at traffic from the source to the destination. Remember that returning traffic might be affected the ACL, based the rules and how it is applied.

Aruba Training-Confidential



VACLs at the Edge

VLAN 10
10.101.10.0/24

Server 2
10.101.10.11

Core-1
Default router for
VLAN 10 and 20

Server 1
10.101.10.10

Access-1

Access-2

VLAN 20
10.101.20.0/24

In VACL on VLAN 20

Client 3
10.101.20.23

Client 2
10.101.20.22

In VACL on VLAN 20

Client 1
10.101.20.21

Traffic processed by ACL

Client 1, 2, or 3 ➡ Any device

Any device ➡ Client 1, 2, or 3

Traffic not processed by ACL

Server 1 ⇨ Server 2

MOD 1- 38

You should typically apply VACLs at the access layer where devices connect to the VLAN in question, as shown here. You can then be certain to filter all switched traffic, as well as the routed traffic. An Aruba switch filters routed traffic for the VLAN to which the VACL is applied, regardless of whether it is the default gateway. It simply processes all traffic that ingresses this VLAN.

Aruba Training-Confidential

Be careful! Consider what it means for traffic to arrive in the VLAN. The previous example shows that the VACL did not process traffic from the servers to the clients because a VACL applies to traffic that arrives on the VLAN. The traffic from servers arrives on Core-1 in VLAN 10.

However, in this scenario, Core-1 routes the traffic to VLAN 20, and the traffic from servers to clients arrives on Access-1 and Access-2 in VLAN 20. That means that the VACLs applied on the access layer switches now process traffic routed by Core-1 into VLAN 20, as well as traffic that arrives on VLAN 20 on the access layer switches. In the example above, the VACL processes traffic from Server 1 to Client 1 and even traffic from Server 2 to Client 1. Even though both Server 2 and Client 1 are connected to Access-2, the traffic from Server 2 passes to Core-1 for routing and then returns to Access-2 in VLAN 20.

If you intend to apply a VACL to a VLAN on a switch that does not route traffic for that VLAN, be very careful to plan the rules to permit return traffic to devices in that VLAN as well as the traffic from the devices.

For example, the FromGuests ACL for guest access might have these rules:

```
50 permit udp any 10.255.0.5/32 eq 67
150 permit udp any 10.255.0.5/32 eq 53
250 permit icmp any 10.255.0.5/32
350 deny ip any 10.0.0.0/8
450 permit ip any any
```

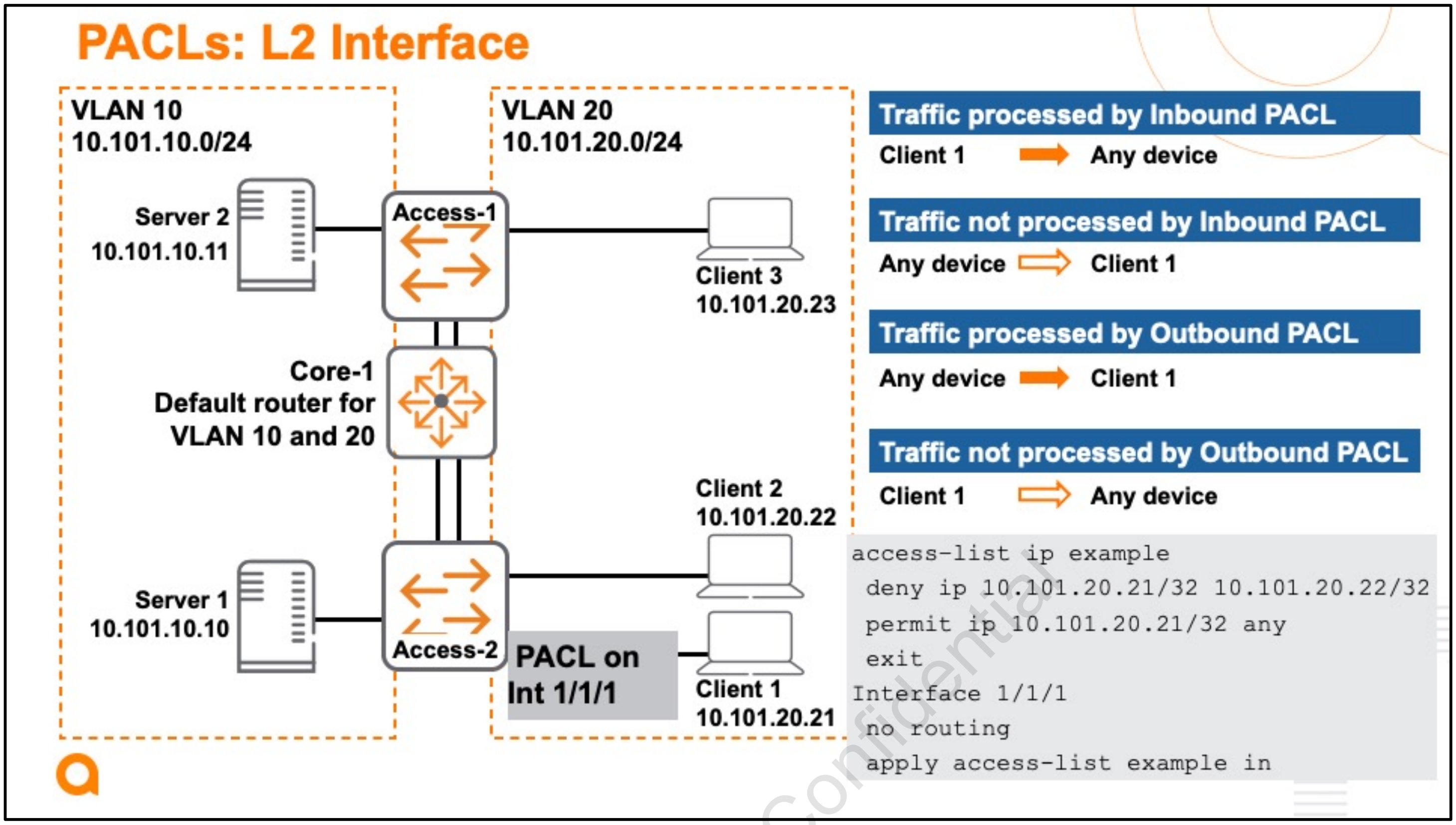
Access-1 was routing the traffic, so it only filtered traffic that arrived on VLAN 99. Because the ACL could only filter traffic assigned to VLAN 99, which terminated on this switch, specifying 10.101.99.0/24 or any as the source had much the same effect.

However, the ACL behavior changes if you are using a topology more like the one shown in the figure above, in which Core-1 routes the traffic, but Access-1 still implements the ACL as a VACL. You would then need to adjust your approach to ensure that the ACL properly filters return traffic. For example, rule 350 above would deny return traffic from the Internet to guests in 10.101.99.0/24. You might want to change this rule to 350 deny ip 10.101.99.0/8 10.0.0.0/8. Then return traffic from the Internet to guests would match rule 450 instead and be permitted. Another way that you should to adjust the ACL would be to permit DHCP traffic (UDP ports 67 and 68) to and from all IP addresses.

Because, in this example, the access layer switches are not routing VLAN 20 traffic and do not have IP addresses on VLAN 20, applying the ACL as outbound VACL has the same effect as applying it as an inbound VACL. All traffic that arrives in the VLAN is also forwarded in that VLAN.

Note: the ACL example is looking at traffic from the source to the destination. Remember that returning traffic might be affected the ACL, based the rules and how it is applied.

Aruba Training Confidential



PACLs: L2 Interface

Traffic processed by Inbound PACL

Client 1 → Any device

Traffic not processed by Inbound PACL

Any device ⇨ Client 1

Traffic processed by Outbound PACL

Any device → Client 1

Traffic not processed by Outbound PACL

Client 1 ⇨ Any device

```
access-list ip example
 deny ip 10.101.20.21/32 10.101.20.22/32
 permit ip 10.101.20.21/32 any
 exit
Interface 1/1/1
 no routing
 apply access-list example in
```

Aruba Training-Confidential

On an AOS-CX switch, you can apply an IPv4 or MAC ACL to a port or link aggregation to filter traffic inbound or outbound on that physical interface. (Some switches support only inbound PACLs).

An inbound PACL filters all inbound traffic regardless of whether it is routed or switched. The PACL also applies to traffic inbound on the interface with a destination on the switch itself. An outbound PACL filters all outbound traffic on the port regardless of whether it is routed or switched.

You apply the PACL with the same command as a RACL.

The figure shows an example PACL configuration for this scenario, applied inbound to port 1/1/1 on Access-2. It filters all traffic to or from the connected client.

Based on this example, this traffic would be processed by the ACL and allowed:

Client1 > Client3

Client1 > Server1 and Server2

Client1 > Core1

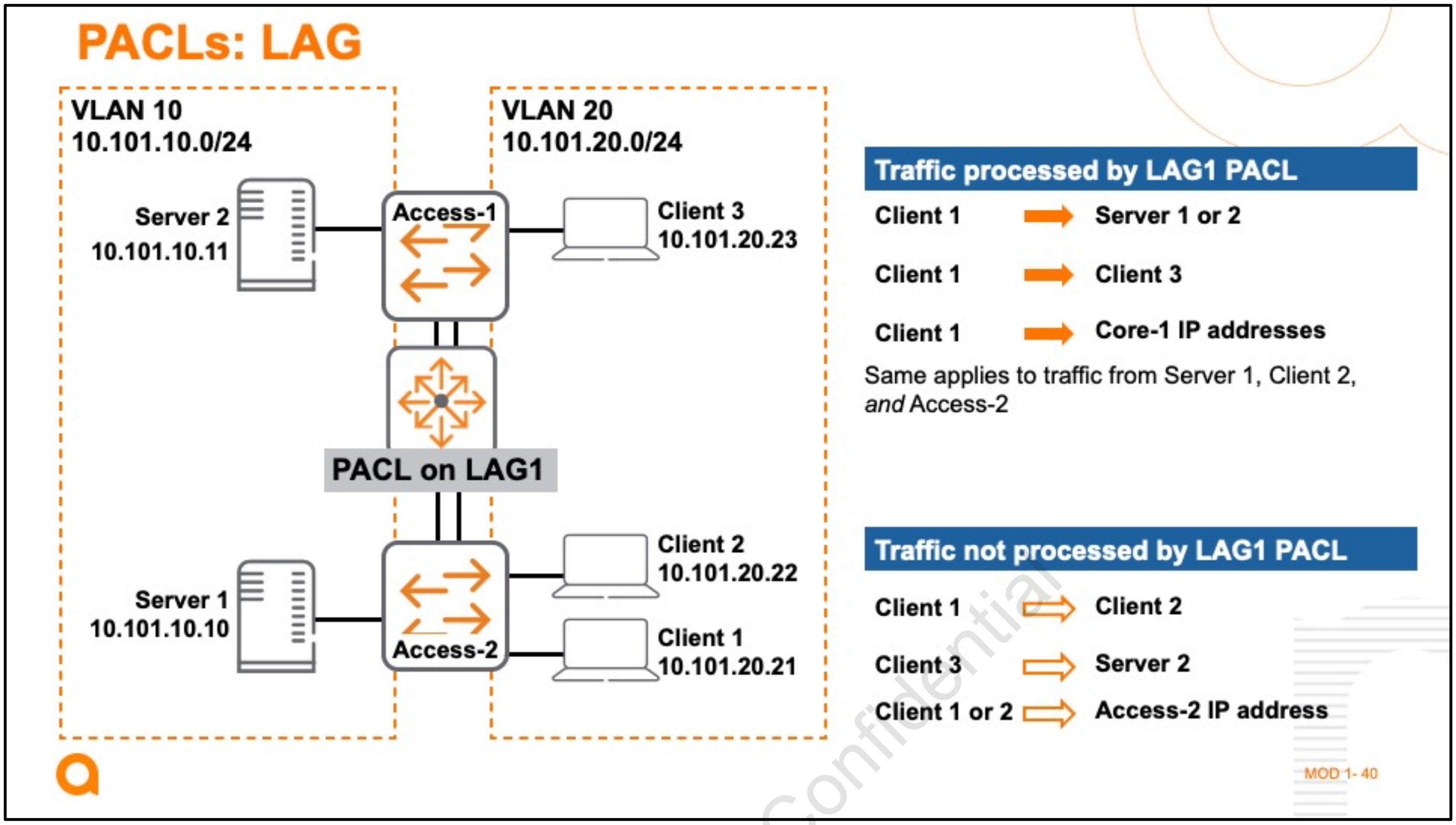
And this traffic would also be processed by the ACL, but denied:

Client1 > Client2

And this traffic is not even processed by the ACL:
Any device > Client1

Note: the ACL example is looking at traffic from the source to the destination. Remember that returning traffic might be affected the ACL, based the rules and how it is applied.

Aruba Training-Confidential



PACLs: LAG

Traffic processed by LAG1 PACL

Client 1 → Server 1 or 2

Client 1 → Client 3

Client 1 → Core-1 IP addresses

Same applies to traffic from Server 1, Client 2, *and* Access-2

Traffic not processed by LAG1 PACL

Client 1 ⇨ Client 2

Client 3 ⇨ Server 2

Client 1 or 2 ⇨ Access-2 IP address

Aruba Training-Confidential

As always, you must carefully consider where you apply the ACL. If you apply an ACL to a switch-to-switch link, the ACL affects all incoming traffic from that link. Here you see how an ACL applied inbound on the link aggregation between Access-2 and Core-1 filters all traffic from Client 1, Client 2, and Server 1 that passes through Core-1. The filtering even extends to any traffic that Access-2 sends itself—an important implication for you to keep in mind. However, the ACL, of course, cannot filter any communications that do not pass through it. As long as Core 1 is routing the traffic, traffic between Client 1 and Server 1 is filtered, but traffic between Client 1 and Client 2 is not.

In many cases when you want to control both routed and switched traffic, you should apply the ACL as close to the edge as possible, as in the previous example. You can then ensure that all appropriate traffic is filtered. If you only want to control routed traffic, you should apply the ACL as a RACL on the VLAN interface on the VLAN's default router.

This example shows that you can apply PACLs to link aggregations. Keep a few guidelines in mind. You must apply the PACLs to the link aggregation (trunk) rather than to individual ports within the link aggregation. It then automatically applies to the member ports. Similarly, removing the PACL from the link aggregation, removes it from member ports. Removing a port from the link aggregation removes the PACL from that port.

Best practice is that you should also remove any ACLs from a port before adding it to a link aggregation.

Note: the ACL example is looking at traffic from the source to the destination. Remember that returning traffic might be affected the ACL, based the rules and how it is applied.

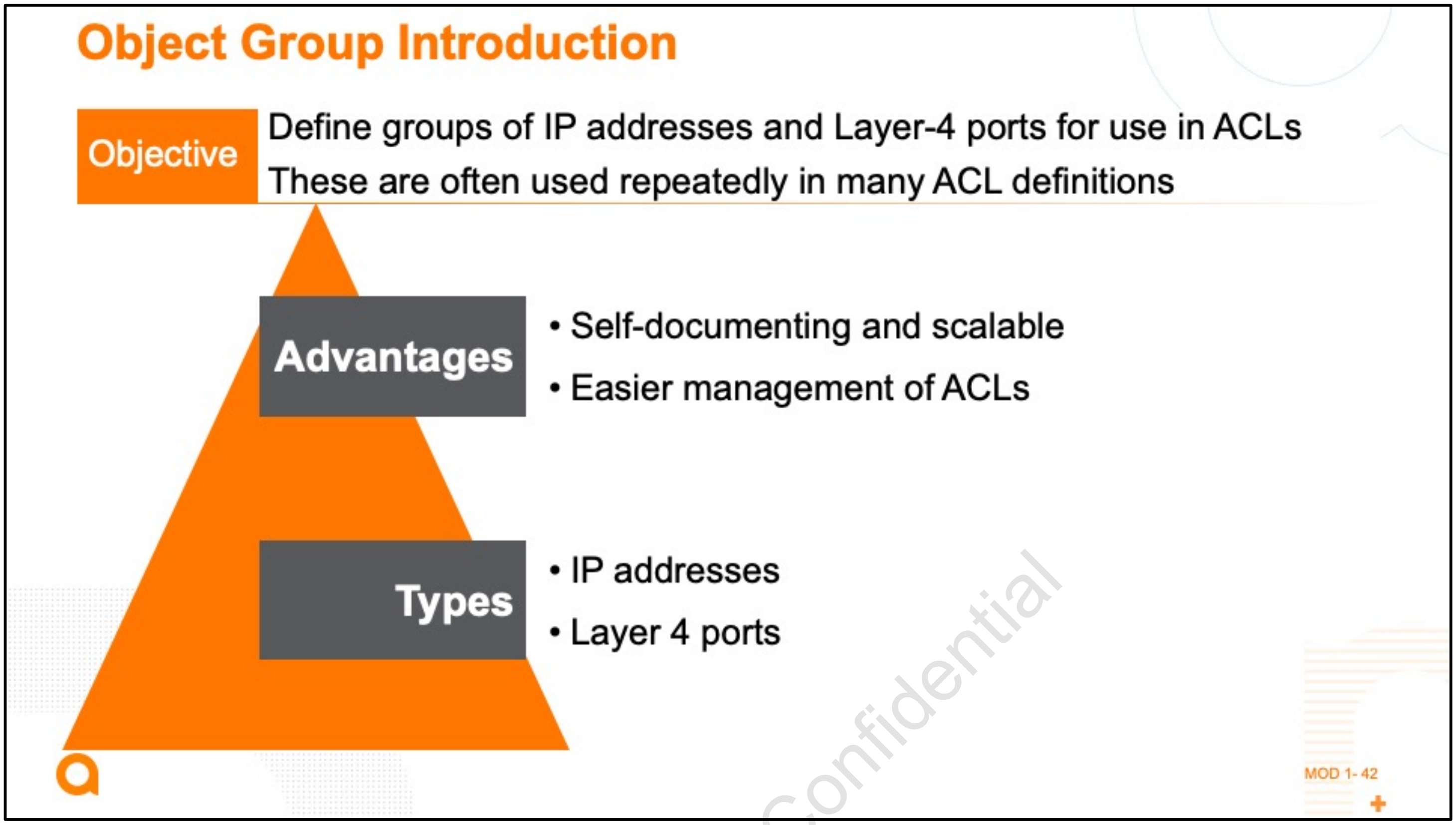
Aruba Training-Confidential

Object Groups

MOD 1- 41

Aruba Training-Confidential

Let’s explore object groups.



Object Group Introduction

Objective

Define groups of IP addresses and Layer-4 ports for use in ACLs
These are often used repeatedly in many ACL definitions

Advantages

- Self-documenting and scalable
- Easier management of ACLs

Types

- IP addresses
- Layer 4 ports

MOD 1- 42

Aruba Training-Confidential

Object groups are useful for defining groups of IP addresses and Layer-4 ports for use exclusively in the ACE rules. Often, common groups of addresses and ports or port ranges are use repeatedly in many ACL definitions.

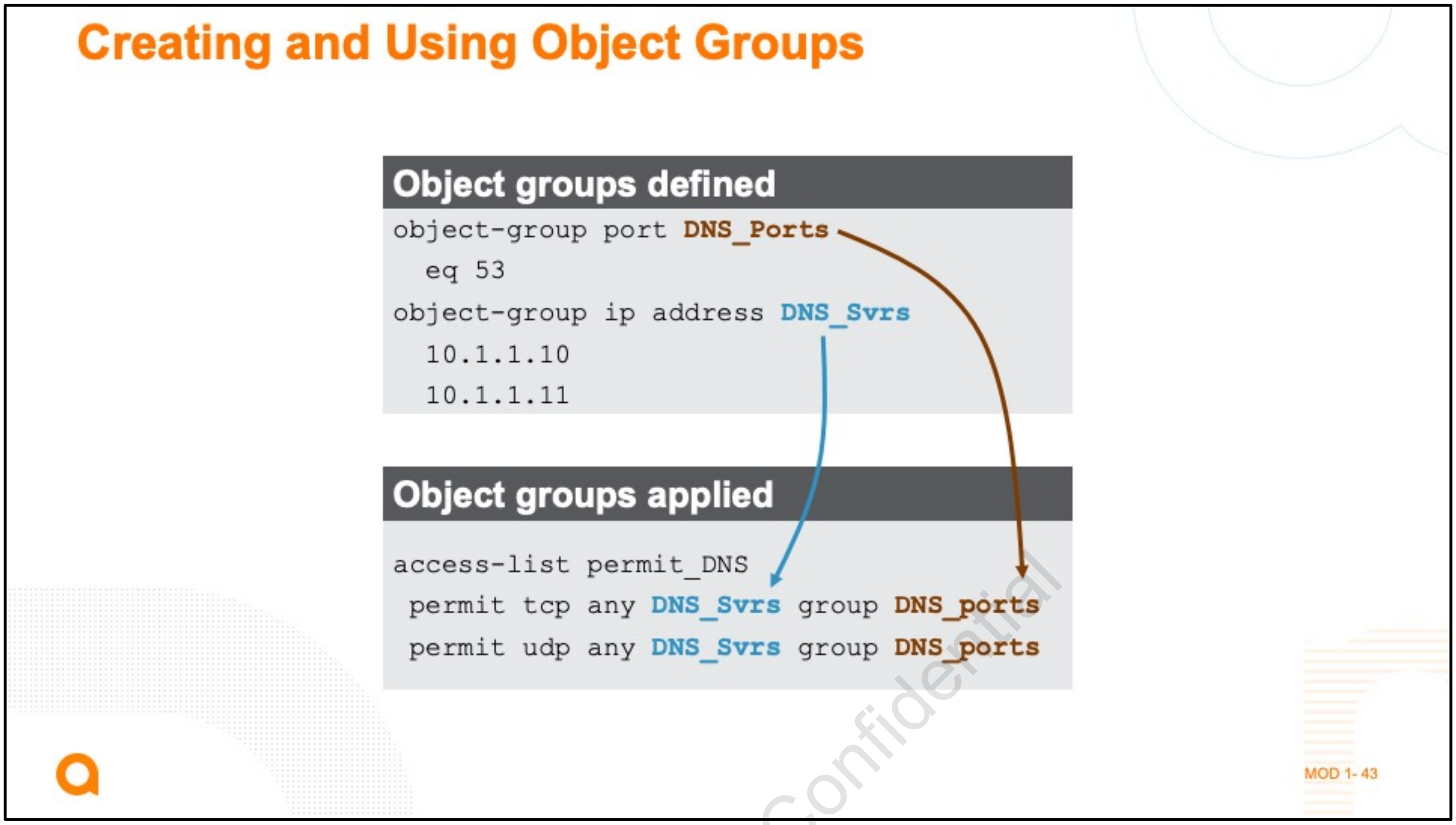
| The nice thing is they are self-documenting, scalable, and makes for easier ACL management.

| There are two object group types:

IP addresses

Layer-4 ports

Without address and port object groups, the same addresses and ports must be repeated in each ACL definition that uses them.



Creating and Using Object Groups

Object groups defined

```
object-group port DNS_Ports
  eq 53
object-group ip address DNS_Svrs
  10.1.1.10
  10.1.1.11
```

Object groups applied

```
access-list permit_DNS
 permit tcp any DNS_Svrs group DNS_ports
 permit udp any DNS_Svrs group DNS_ports
```

MOD 1- 43

Creating an object group is done with the object-group command. You can include many objects in a group, but they must be the same type: either addresses and/or networks for an IP address object or a port or ports for a port object.

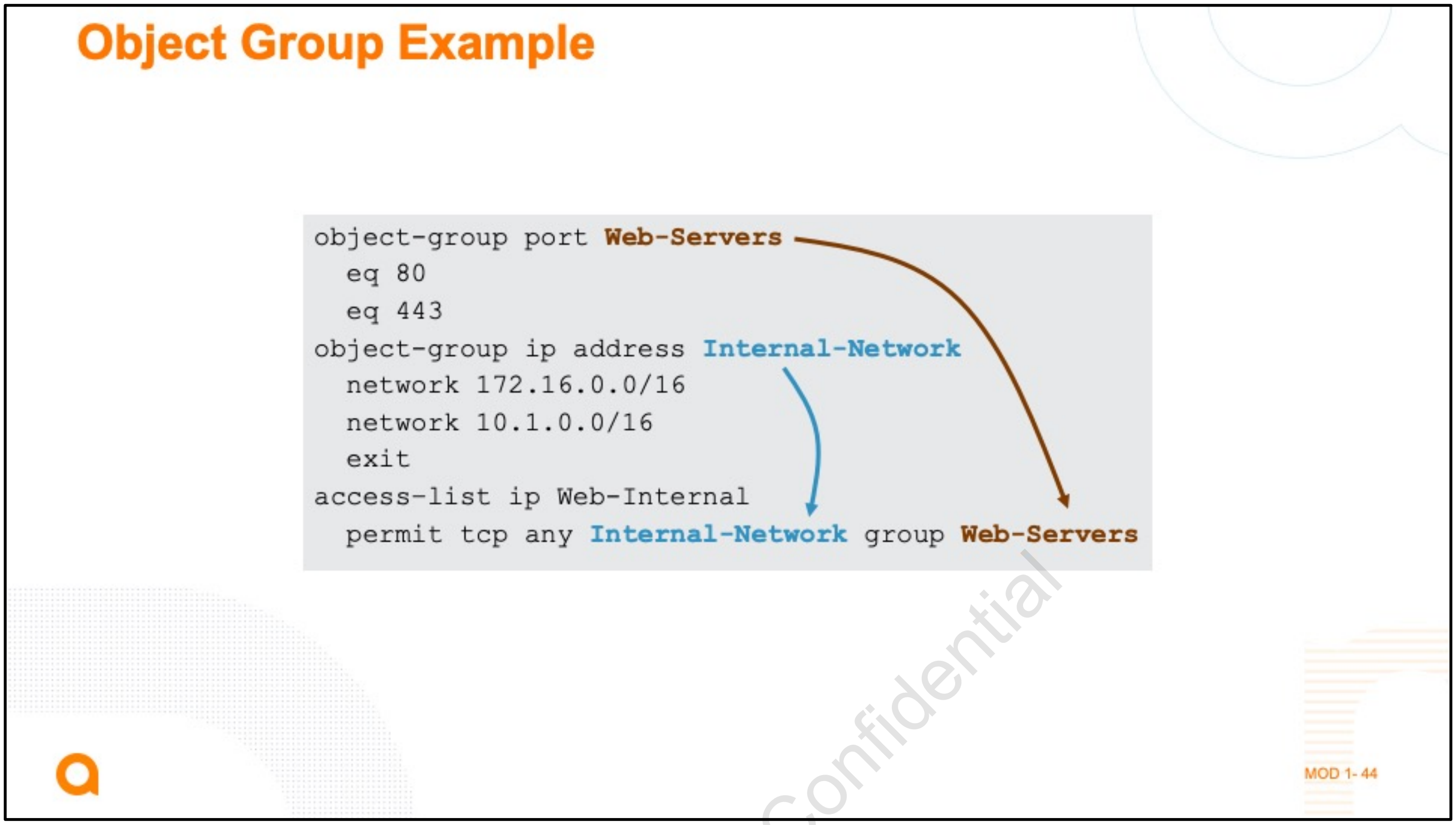
Once they are created, you can reference them in an ACE entry at the appropriate place:

Source and/or destination IP address

Source and/or destination Layer-4 port

The figure shows an example of configuring both IP and port object groups, which are then applied to an ACL. Note that you precede a port group in the ACL with the keyword group.

Aruba Training-Confidential



Object Group Example

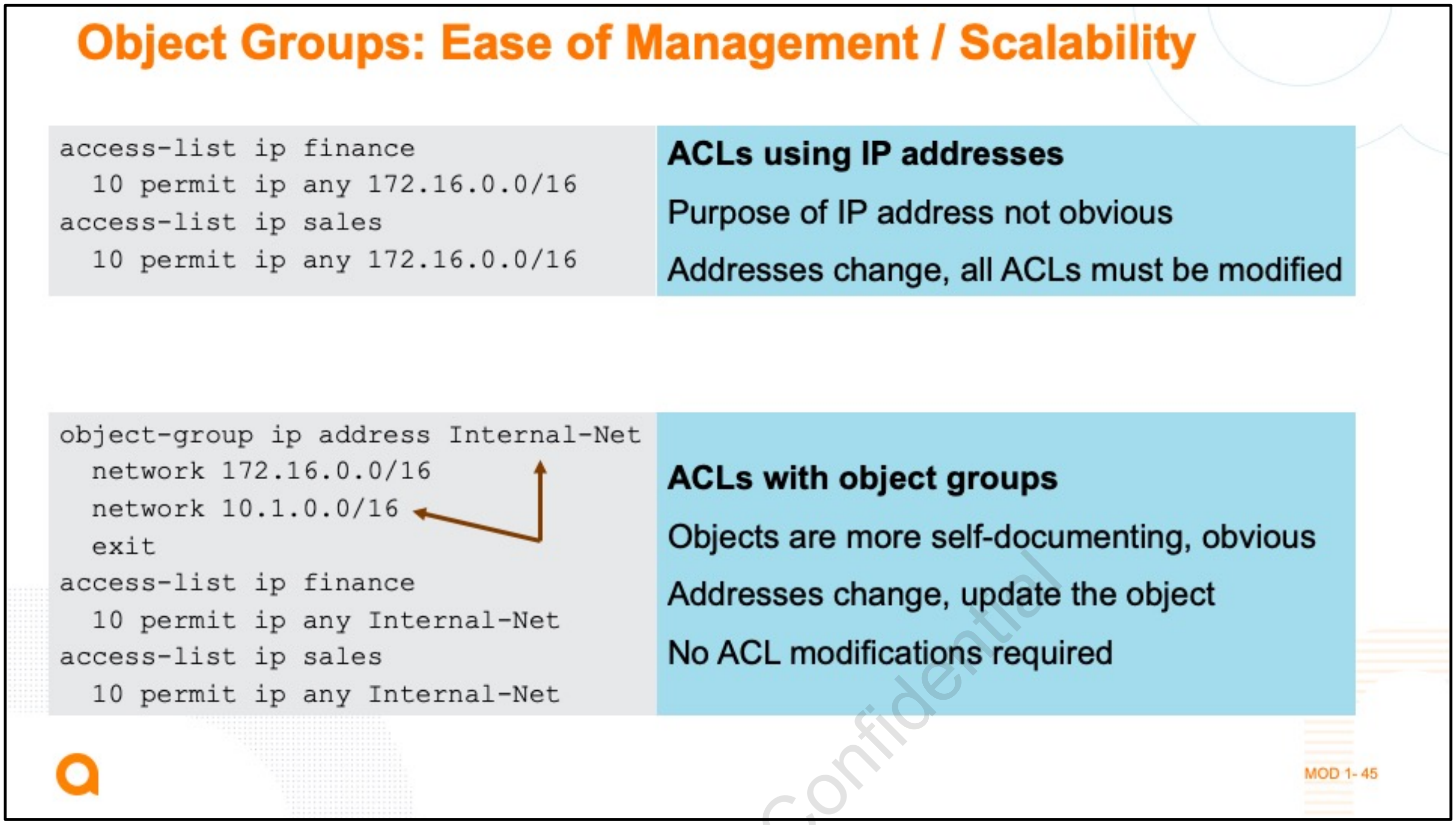
```
object-group port Web-Servers
  eq 80
  eq 443
object-group ip address Internal-Network
  network 172.16.0.0/16
  network 10.1.0.0/16
  exit
access-list ip Web-Internal
  permit tcp any Internal-Network group Web-Servers
```

MOD 1- 44

The figure shows another example that combines both address and ports:

In this example, anyone could open a web connection (port 80 or 443) to any server with an address in 172.16.0.0/16 or 10.1.0.0/16.

Aruba Training-Confidential



Object Groups: Ease of Management / Scalability

```
access-list ip finance
  10 permit ip any 172.16.0.0/16
access-list ip sales
  10 permit ip any 172.16.0.0/16
```

ACLs using IP addresses

Purpose of IP address not obvious

Addresses change, all ACLs must be modified

```
object-group ip address Internal-Net
  network 172.16.0.0/16
  network 10.1.0.0/16
  exit
access-list ip finance
  10 permit ip any Internal-Net
access-list ip sales
  10 permit ip any Internal-Net
```

ACLs with object groups

Objects are more self-documenting, obvious

Addresses change, update the object

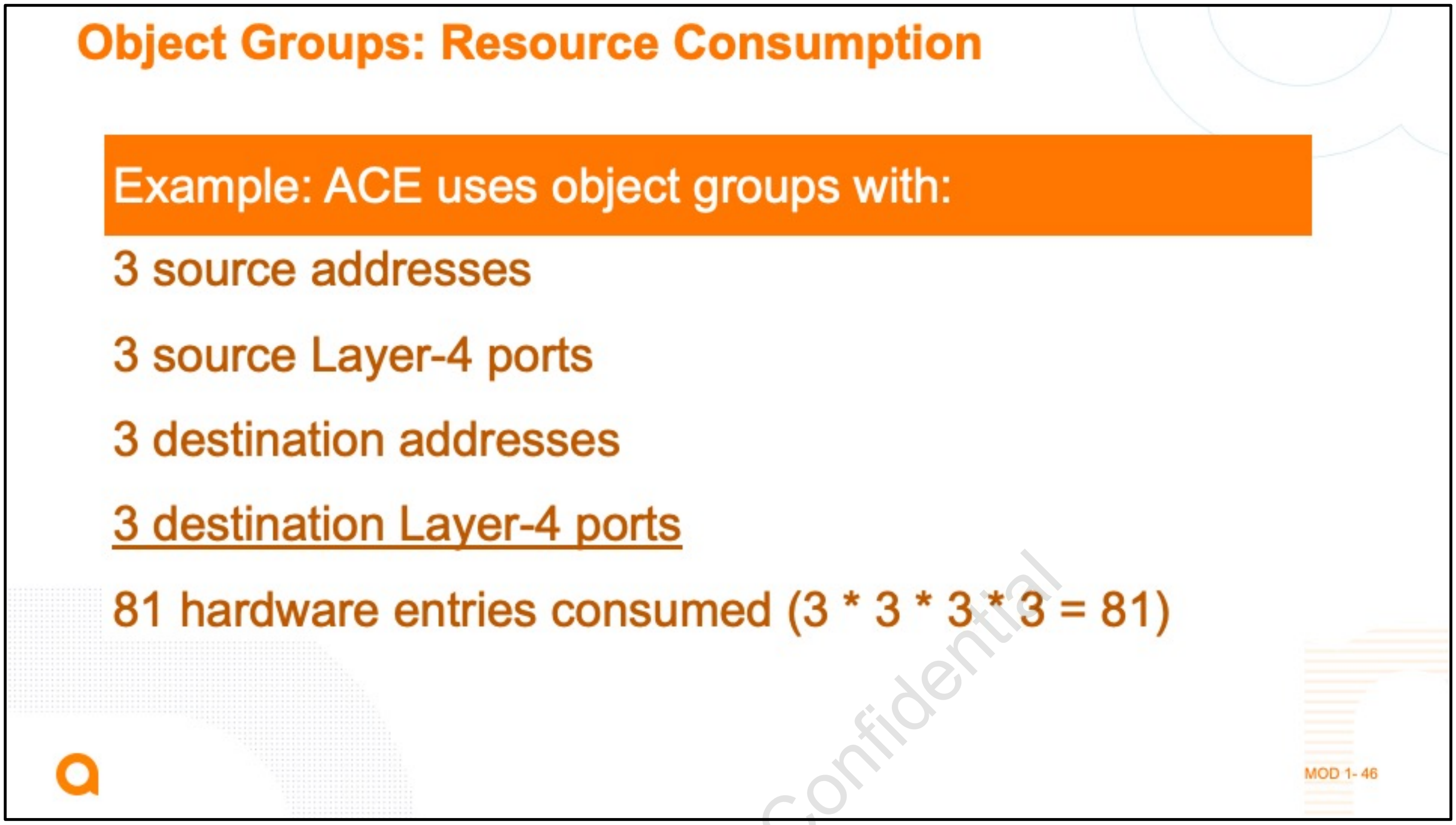
No ACL modifications required

Aruba Training-Confidential

The figure compares ACLs that use IP addresses directly to ACLs with object groups.

The employees in the various VLANs have an ACE entry that allows them access to the internal network that is the 172.16.0.0/16 network. Without prior knowledge, the purpose and use of subnet 172.16.0.0/16 is not obvious. You must reference other documents to know that it is the corporation’s internal network. Now suppose the network grows, and you must add an additional 10.1.0.0/16 subnet. With this traditional method you must modify every ACL in the corporation. This could be dozens, even hundreds of ACLs spread across various Layer 3 switches.

Object groups are self-documenting. You originally created an object group named Internal-Net, with network 172.16.0.0/16 assigned. This descriptive name makes the use more obvious. Later, when the network grows, you simply add network 10.1.0.0/16 to the object group, and it instantly becomes active in any ACL where that object group is used. No ACL modifications are required.



Object Groups: Resource Consumption

Example: ACE uses object groups with:

3 source addresses

3 source Layer-4 ports

3 destination addresses

3 destination Layer-4 ports

81 hardware entries consumed (3 * 3 * 3 * 3 = 81)

Aruba Training-Confidential

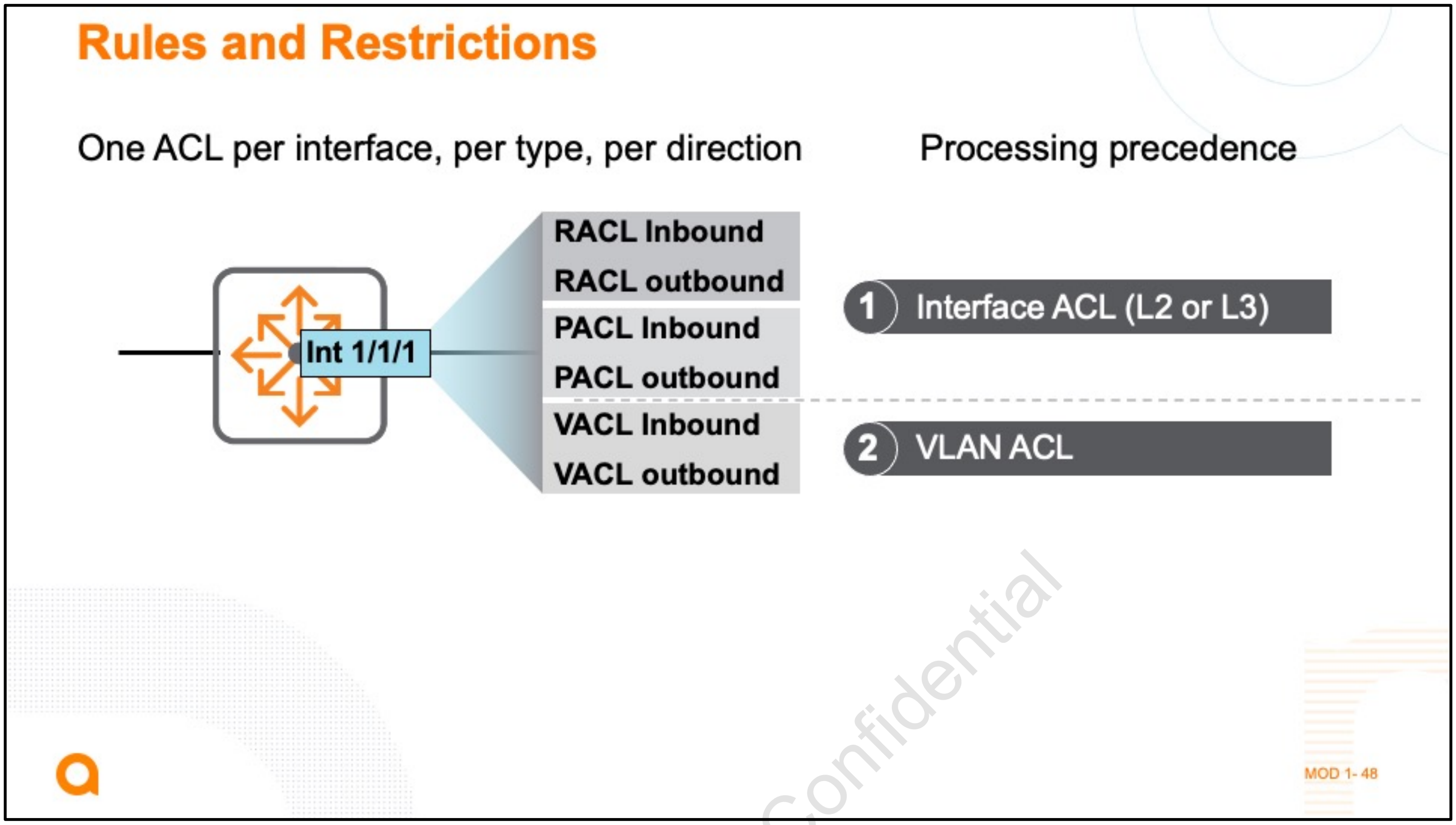
It is possible to consume many hardware resource entries when using the object group commands. For example, in a typical situation, an ACE that uses object groups with 3 source addresses, 3 source Layer-4 ports, 3 destination addresses, and 3 destination Layer-4 ports, a total of 81 hardware entries are consumed (3 * 3 * 3 * 3 = 81). You learn about this in the next section.

Restrictions and Resources

MOD 1- 47

This section introduces the rules and restrictions regarding ACLs and the software and hardware resources that they consume.

Aruba Training-Confidential



Rules and Restrictions

One ACL per interface, per type, per direction

Processing precedence

1 Interface ACL (L2 or L3)

2 VLAN ACL

MOD 1- 48

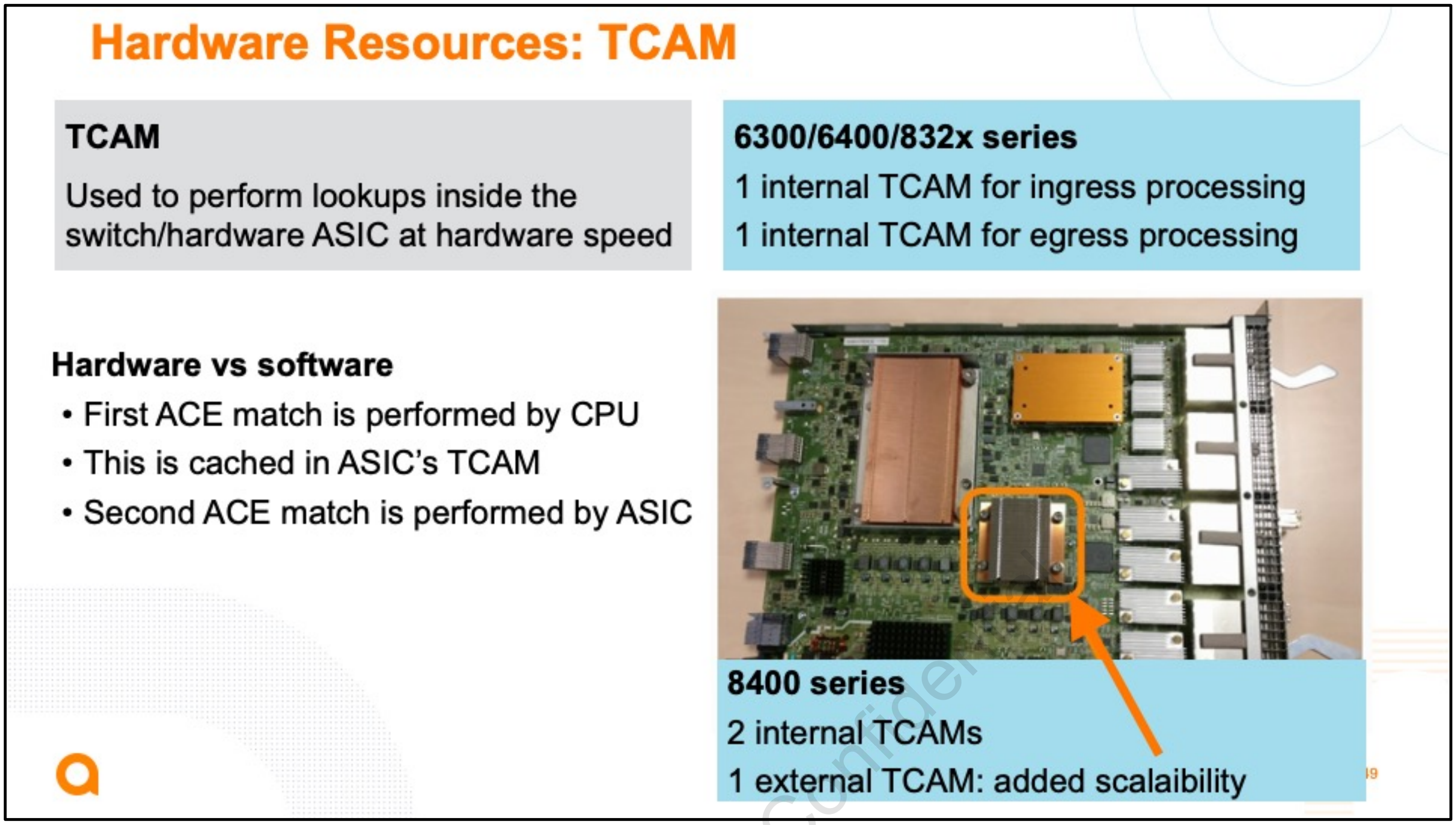
Aruba Training-Confidential

Only one access-list of each type (IPv4, IPv6, MAC) can be applied on the same interface/VLAN in the same direction (IN or OUT). As shown in the figure, you can have one and only one RACL inbound, and one outbound. You can also have one inbound and one outbound PACL, and an inbound and outbound VACL. If you attempt to apply more than one ACL (or policy) on an interface, the new list replaces the old one.

For the precedence of ACLs, if traffic is to be processed by both an interface (Layer-2 or Layer-3) ACL and a VLAN ACL (based on the VLAN the interface is associated with, the ACLs would be processed in this order:

Interface ACL

VLAN ACL



Hardware Resources: TCAM

TCAM

Used to perform lookups inside the switch/hardware ASIC at hardware speed

6300/6400/832x series

1 internal TCAM for ingress processing

1 internal TCAM for egress processing

Hardware vs software

- First ACE match is performed by CPU
- This is cached in ASIC's TCAM
- Second ACE match is performed by ASIC

8400 series

2 internal TCAMs

1 external TCAM: added scalaibility

Aruba Training-Confidential

Switching/routing ASICs have different types of tables where they keep MAC addresses, ARP/ND entries, IPv4/IPv6 routes, ACLs and policies. The switch uses Ternary Content Addressable Memory (TCAM) to perform lookups at hardware speed and to provide an index corresponding to a given search-key.

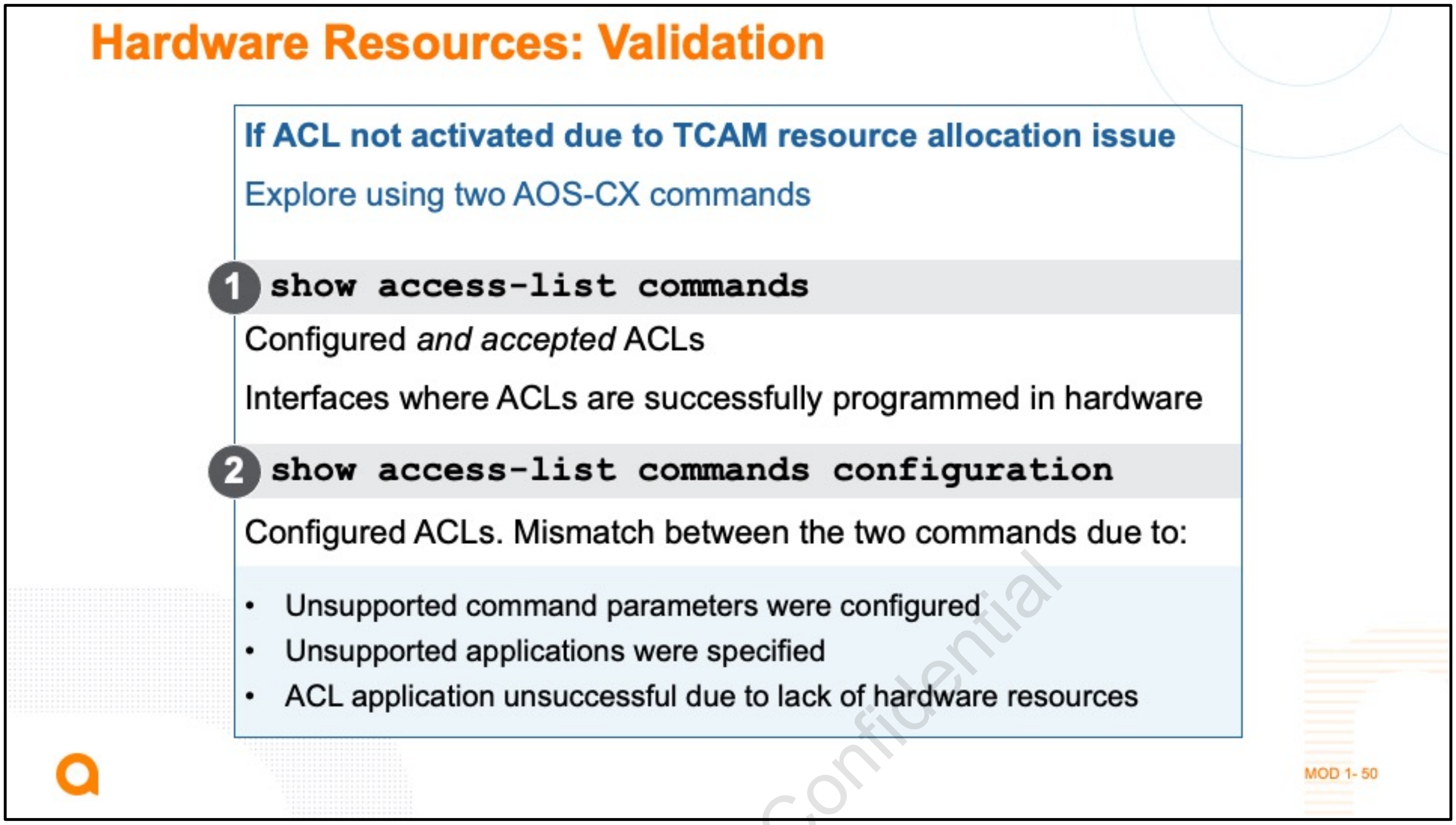
All AOS-CX switches have "internal TCAMs" - responsible for look-ups inside the switch or line-card ASIC. The Aruba 6300/6400/832X series switches have 1 internal TCAM for ingress processing, and 1 for egress processing. Aruba 8400s also have an "external TCAM" – external to the packet-processor ASIC, that gives higher scale capacity.

When processing access lists, the first ACE match is performed by the CPU using a software process. However, this is cached in the ASIC's TCAM, so that all subsequent matches can be performed quicker, and with less CPU utilization in the hardware ASICs.

Therefore you must have ample switch resources to support the number of ACLs you need. Understand that when applying ACLs or policies, hardware resources are consumed to enforce these rules on the data plane:

Management Plane Resources: Used when CREATED

ASIC / TCAM (data-plane) Resources: Used when APPLIED



Hardware Resources: Validation

If ACL not activated due to TCAM resource allocation issue

Explore using two AOS-CX commands

1 `show access-list commands`

Configured *and accepted* ACLs

Interfaces where ACLs are successfully programmed in hardware

2 `show access-list commands configuration`

Configured ACLs. Mismatch between the two commands due to:

- Unsupported command parameters were configured
- Unsupported applications were specified
- ACL application unsuccessful due to lack of hardware resources

MOD 1- 50

Aruba Training-Confidential

In some cases, the configured ACL might not be activated due to TCAM resource allocation issue. Here's what you can do.

First use show access-list commands to see the active switch configuration - what is currently being used in RAM and TCAM. The active configuration is about:

The ACLs that have been configured and accepted by the system

The interfaces on which ACLs have successfully been programmed in hardware.

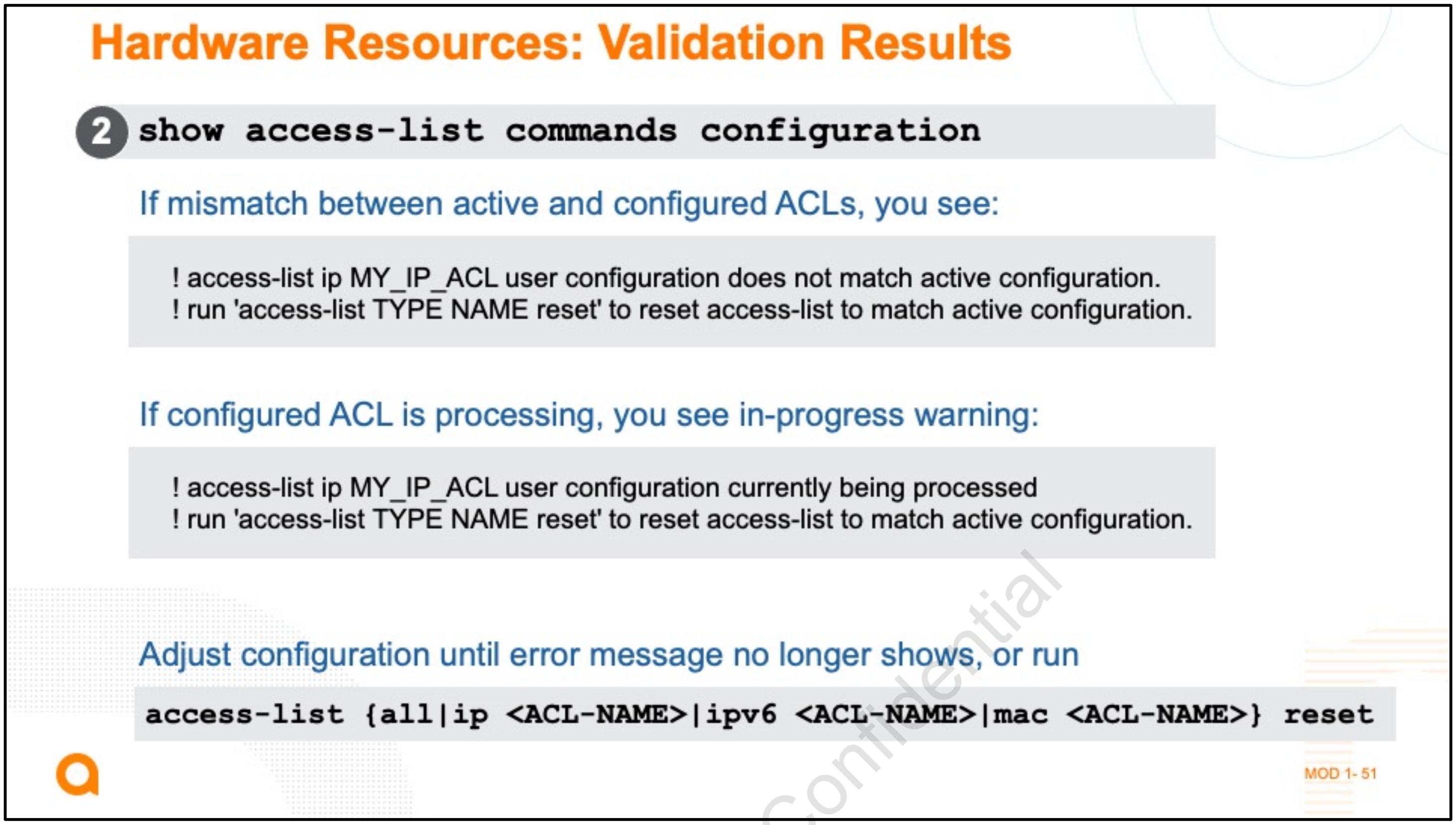
Next, use show access-list commands configuration to see configured ACLs - but not necessarily accepted.

The second command's output may not match what was programmed in hardware or what is active on the switch. This may be because:

Unsupported command parameters might have been configured

Unsupported applications might have been specified

Applying an ACL might have been unsuccessful due to lack of hardware resources



Hardware Resources: Validation Results

2 `show access-list commands configuration`

If mismatch between active and configured ACLs, you see:

```
! access-list ip MY_IP_ACL user configuration does not match active configuration.
! run 'access-list TYPE NAME reset' to reset access-list to match active configuration.
```

If configured ACL is processing, you see in-progress warning:

```
! access-list ip MY_IP_ACL user configuration currently being processed
! run 'access-list TYPE NAME reset' to reset access-list to match active configuration.
```

Adjust configuration until error message no longer shows, or run

```
access-list {all|ip <ACL-NAME>|ipv6 <ACL-NAME>|mac <ACL-NAME>} reset
```

MOD 1- 51

Aruba Training-Confidential

So, to determine if a discrepancy exists between what was configured and what is active, run the show access-list command configuration.

If the active ACLs and configured ACLs are not the same, the switch shows a warning message in the output of the show command:

! access-list ip MY_IP_ACL user configuration does not match active configuration.

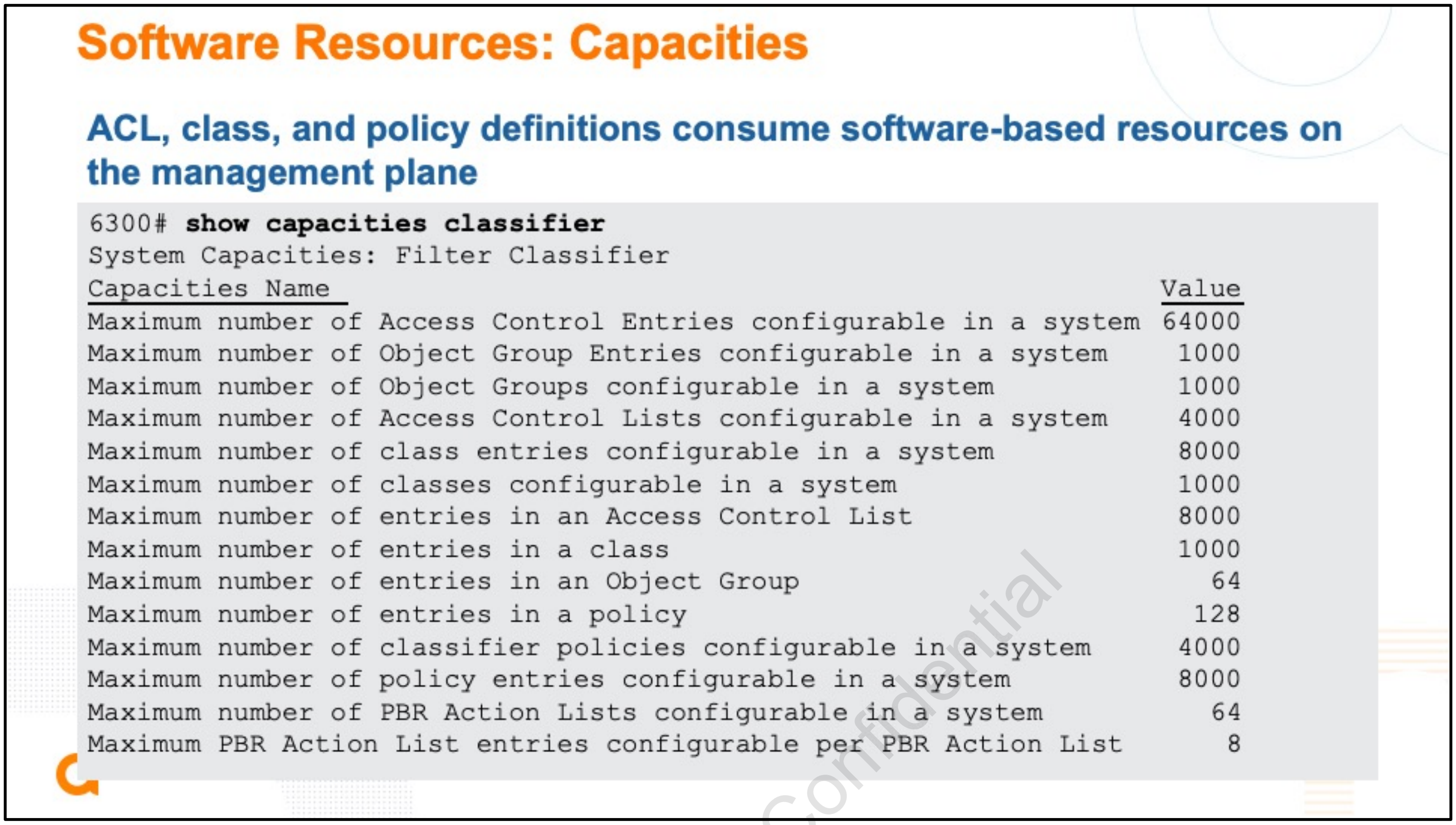
! run 'access-list TYPE NAME reset' to reset access-list to match active configuration.

If the configured ACL is processing, the switch shows an in-progress warning.

! access-list ip MY_IP_ACL user configuration currently being processed

! run 'access-list TYPE NAME reset' to reset access-list to match active configuration.

If the switch shows a warning message or in-progress message, additional changes can be made until the error message is no longer shown in the show command, or you can run the access-list {all|ip <ACL-NAME>|ipv6 <ACL-NAME>|mac <ACL-NAME>} reset command. The access-list reset command changes the user-specified configuration to match the active configuration. For details, see “Configuring ACLs and Classifier Policies Guide for AOS-CX”.



Software Resources: Capacities

ACL, class, and policy definitions consume software-based resources on the management plane

```
6300# show capacities classifier
System Capacities: Filter Classifier
Capacities Name                                                    Value
Maximum number of Access Control Entries configurable in a system 64000
Maximum number of Object Group Entries configurable in a system     1000
Maximum number of Object Groups configurable in a system            1000
Maximum number of Access Control Lists configurable in a system     4000
Maximum number of class entries configurable in a system            8000
Maximum number of classes configurable in a system                  1000
Maximum number of entries in an Access Control List                 8000
Maximum number of entries in a class                                1000
Maximum number of entries in an Object Group                          64
Maximum number of entries in a policy                                128
Maximum number of classifier policies configurable in a system      4000
Maximum number of policy entries configurable in a system           8000
Maximum number of PBR Action Lists configurable in a system           64
Maximum PBR Action List entries configurable per PBR Action List       8
```

Aruba Training-Confidential

AOS-CX allows you to configure more ACLs and policies than will be applied, for flexibility - like migrating security from one ACL to another one. ACL, class, and policy definitions consume software-based resources on the management plane. These resources are generally increased along with major AOS-CX releases.

The switch capabilities can be seen with the show capacities and show capacities-status commands. The former command shows the capabilities. The figure shows an example of show capacities classifier for a 6300 switch

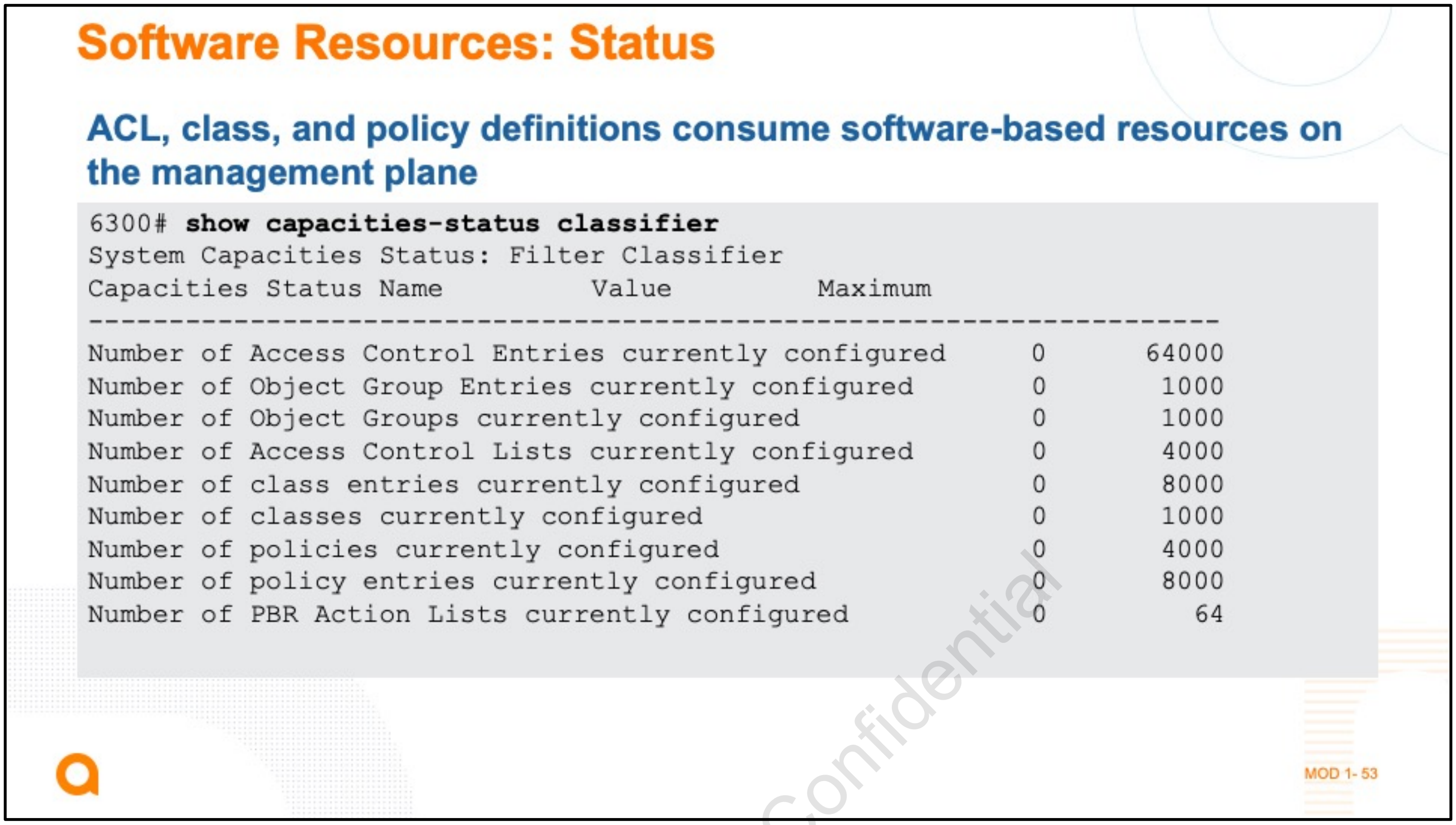
Here is an example of the latter command for a 6300 switch:

```
6300# show capacities-status classifier
System Capacities Status: Filter Classifier
Capacities Status Name                                   Value      Maximum
--------------------------------------------------------------------------------
Number of Access Control Entries currently configured              0            64000
Number of Object Group Entries currently configured     0            1000
Number of Object Groups currently configured            0            1000
Number of Access Control Lists currently configured     0            4000
Number of class entries currently configured            0            8000
Number of classes currently configured                  0            1000
Number of policies currently configured              0          4000
Number of policy entries currently configured        0          8000
Number of PBR Action Lists currently configured      0            64
```

It displays the capabilities and the resources being used on the switch.

Note: Remember that the amount of resources available is dependent on the switch model and will typically vary. Also, available resource limits might increase based on the version of AOS-CX code running on the switch.

Aruba Training-Confidential



Software Resources: Status

ACL, class, and policy definitions consume software-based resources on the management plane

```
6300# show capacities-status classifier
System Capacities Status: Filter Classifier
Capacities Status Name          Value         Maximum
----------------------------------------------------------------------
Number of Access Control Entries currently configured     0      64000
Number of Object Group Entries currently configured       0       1000
Number of Object Groups currently configured              0       1000
Number of Access Control Lists currently configured       0       4000
Number of class entries currently configured              0       8000
Number of classes currently configured                    0       1000
Number of policies currently configured                   0       4000
Number of policy entries currently configured             0       8000
Number of PBR Action Lists currently configured           0         64
```

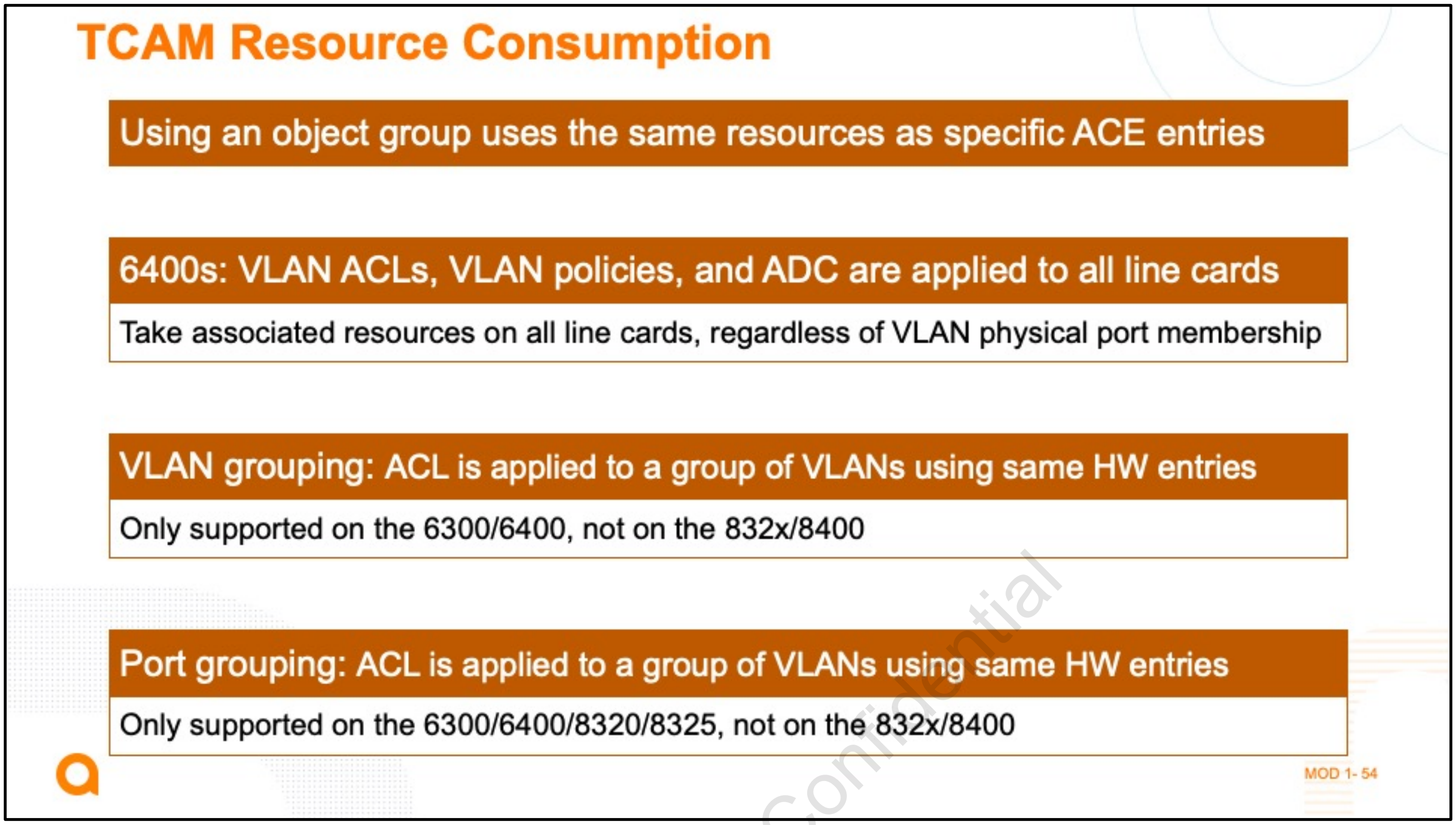
MOD 1- 53

Here is an example of show capacities-status classifier for a 6300 switch

It displays the capabilities and the resources being used on the switch.

Note: Remember that the amount of resources available is dependent on the switch model and will typically vary. Also, available resource limits might increase based on the version of AOS-CX code running on the switch.

Aruba Training-Confidential



TCAM Resource Consumption

Using an object group uses the same resources as specific ACE entries

6400s: VLAN ACLs, VLAN policies, and ADC are applied to all line cards

Take associated resources on all line cards, regardless of VLAN physical port membership

VLAN grouping: ACL is applied to a group of VLANs using same HW entries

Only supported on the 6300/6400, not on the 832x/8400

Port grouping: ACL is applied to a group of VLANs using same HW entries

Only supported on the 6300/6400/8320/8325, not on the 832x/8400

MOD 1- 54

Aruba Training-Confidential

The figure summarizes key aspects of TCAM resource consumption. Click each heading to explore.

Object groups ease management tasks – they do not reduce resource consumption. The switch expands the ACE entries with the actual values, so the resource consumption is the same as if you didn't use object groups. For example, an ACE with groups containing 3 source addresses, 3 destination addresses, 3 source TCP ports and 3 destination TCP ports will expand to 3x3x3x3 = 81 hardware entries.

6400s

For 6400s: VLAN ACLs, VLAN policies, and ADC are applied to all line cards, and take associated resources on all line cards, regardless of physical port membership.

When applying Layer-4 port range ACEs that use 'lt', 'gt', 'range', or port groups, they may use more than one hardware entry. When applying an ACL to an interface, each interface to be matched will allocate one extra hardware entry to store the invisible implicit deny rule (ACE) configured to the end of ACL list by default.

Note: There may be additional restrictions or resource consumption based on the switch model.

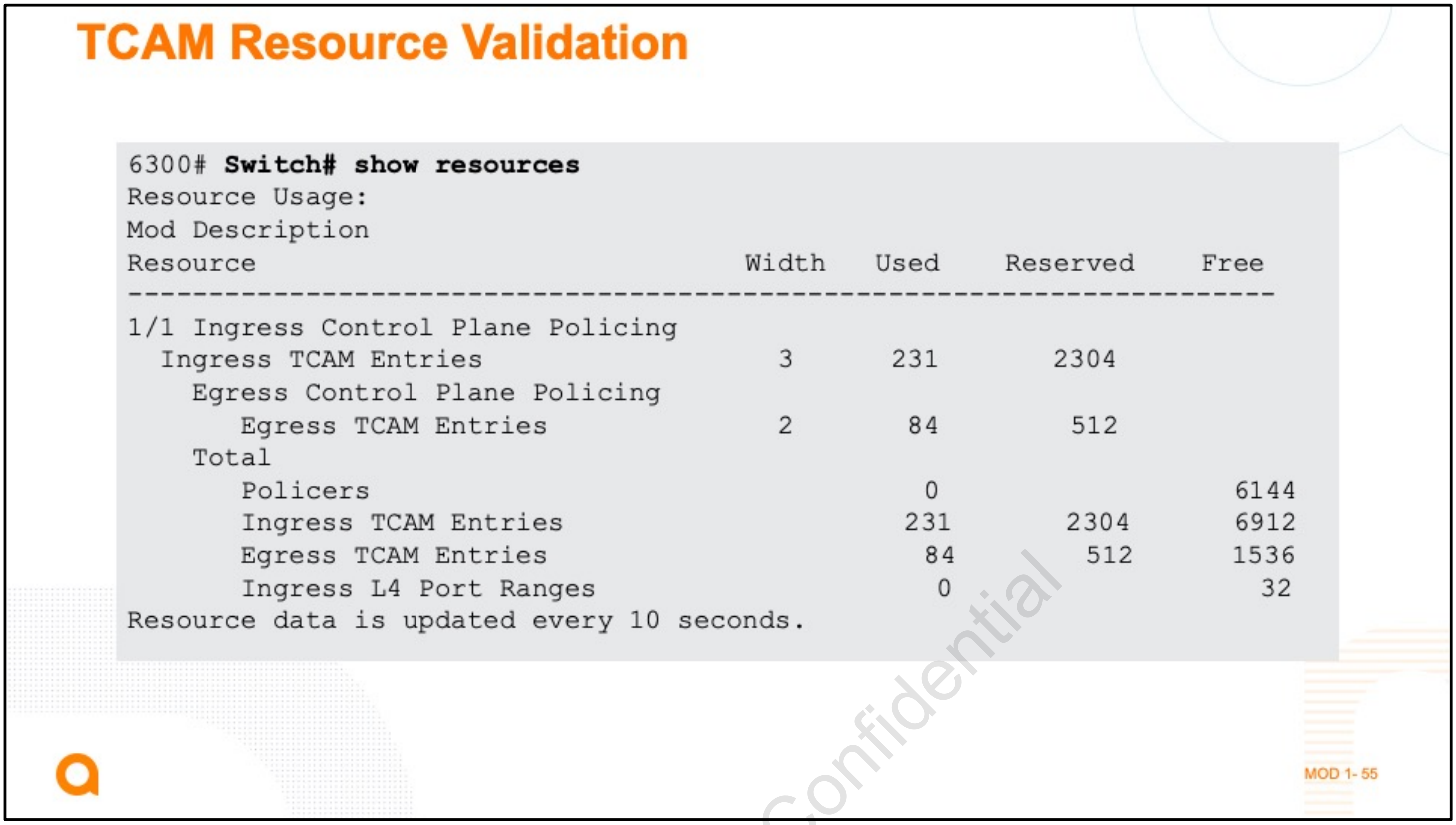
VLAN grouping

The VLAN grouping functionality allows an ACL or Policy to be applied to a group of VLANs and to leverage the same HW entries for multiple VLANs (applying the same ACL to multiple VLANs is just a VLAN bitmapping). Such optimization is available only on 6300/6400, not on 832x, 8400.

Port grouping

The port grouping functionality allows an ACL or Policy to be applied to a group of Ports and to leverage the same HW entries for multiple ports (applying the same ACL to multiple ports is just a port bitmapping). Such optimization is possible on 6300/6400/8320/8325, not on 8400.

Aruba Training-Confidential



TCAM Resource Validation

```
6300# Switch# show resources
Resource Usage:
Mod Description
Resource                                  Width   Used    Reserved    Free
----------------------------------------------------------------------
1/1 Ingress Control Plane Policing
  Ingress TCAM Entries                      3      231       2304
    Egress Control Plane Policing
       Egress TCAM Entries                  2       84        512
    Total
       Policers                                      0                 6144
       Ingress TCAM Entries                        231        2304     6912
       Egress TCAM Entries                          84         512     1536
       Ingress L4 Port Ranges                        0                   32
Resource data is updated every 10 seconds.
```

MOD 1- 55

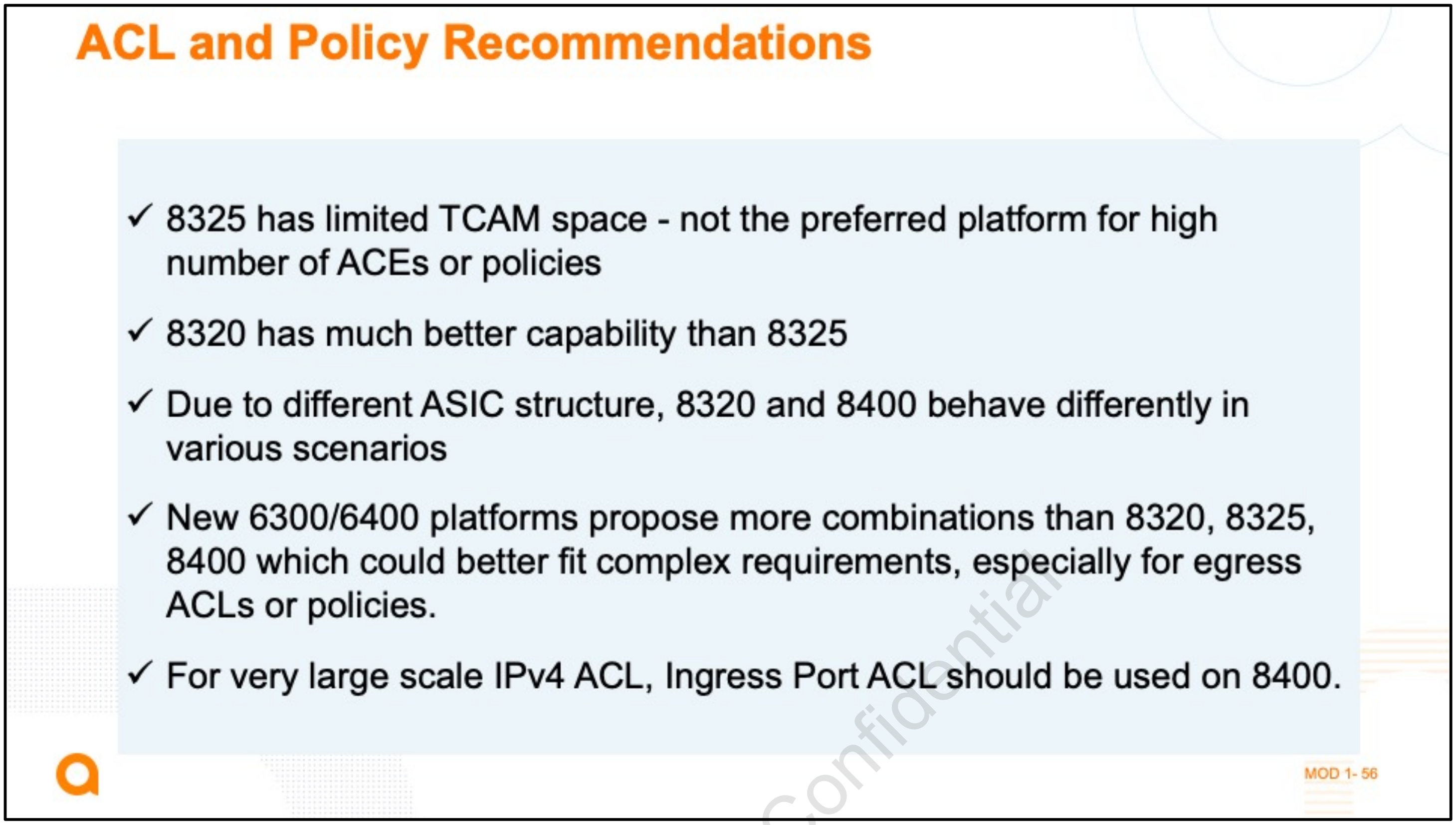
Aruba Training-Confidential

To verify allocated TCAM resources, two important commands are used:

From 10.0 to 10.3 release: diag-dump acl basic
To run this command you must enable the diagnostics function (diagnostics).

From 10.4 release: show resources - a more user-friendly version - no need to enter diagnostics.

The figure shows an example of the show resources command, used to troubleshoot the installation of ACLs and/or policies.



ACL and Policy Recommendations

- ✓ 8325 has limited TCAM space - not the preferred platform for high number of ACEs or policies
- ✓ 8320 has much better capability than 8325
- ✓ Due to different ASIC structure, 8320 and 8400 behave differently in various scenarios
- ✓ New 6300/6400 platforms propose more combinations than 8320, 8325, 8400 which could better fit complex requirements, especially for egress ACLs or policies.
- ✓ For very large scale IPv4 ACL, Ingress Port ACL should be used on 8400.

MOD 1- 56

The various AOS-CX switch models have different restrictions and capabilities. The figure shows a list of Aruba recommendations related to ACLs and policy implementation and usage

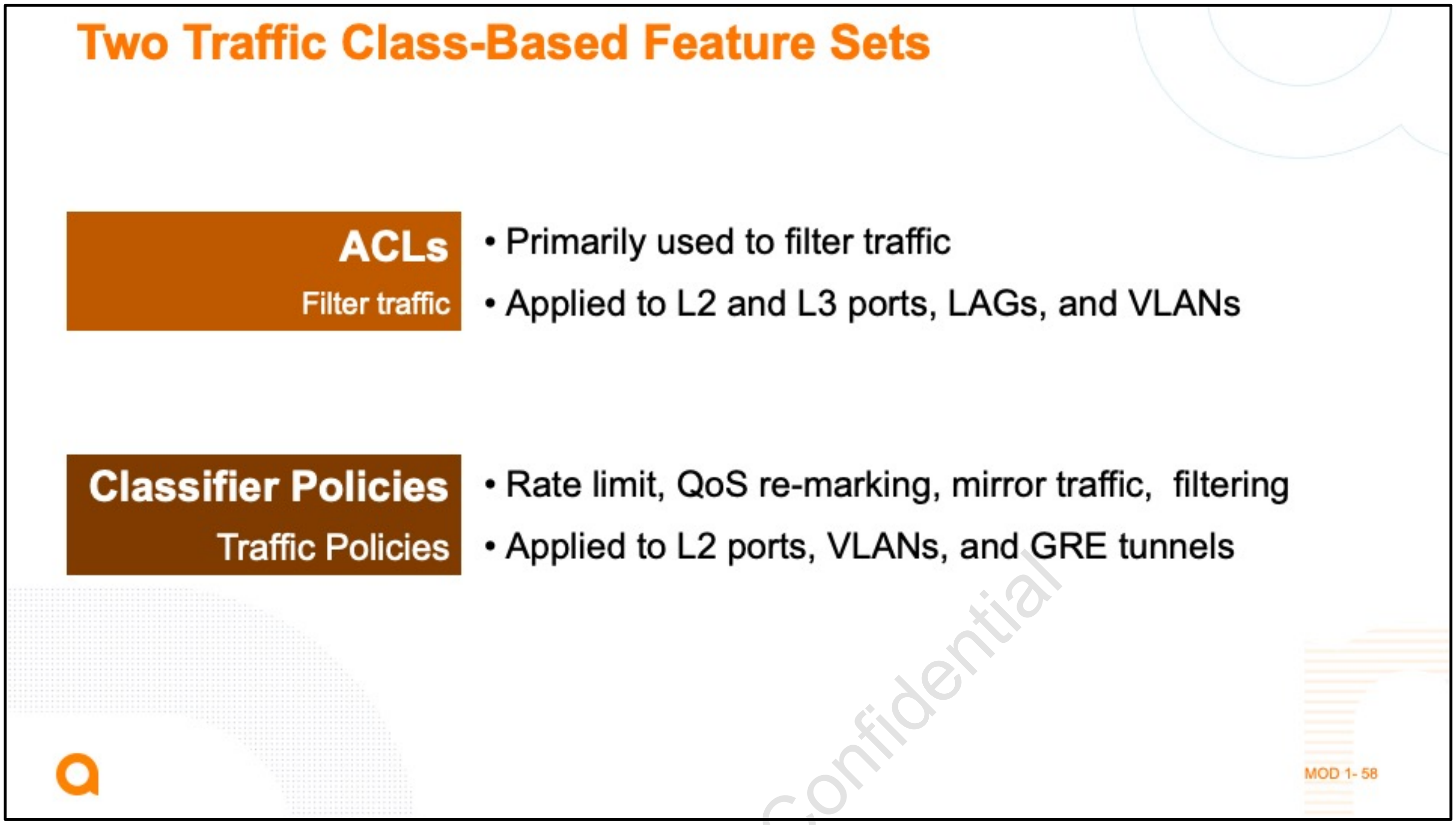
Aruba Training-Confidential

Classifier Policies

MOD 1- 57

This section introduces how to use classes and policies to define very specific QoS and other traffic polices for protocols and connections at the global or interface level.

Aruba Training-Confidential



Two Traffic Class-Based Feature Sets

ACLs
Filter traffic

- Primarily used to filter traffic
- Applied to L2 and L3 ports, LAGs, and VLANs

Classifier Policies
Traffic Policies

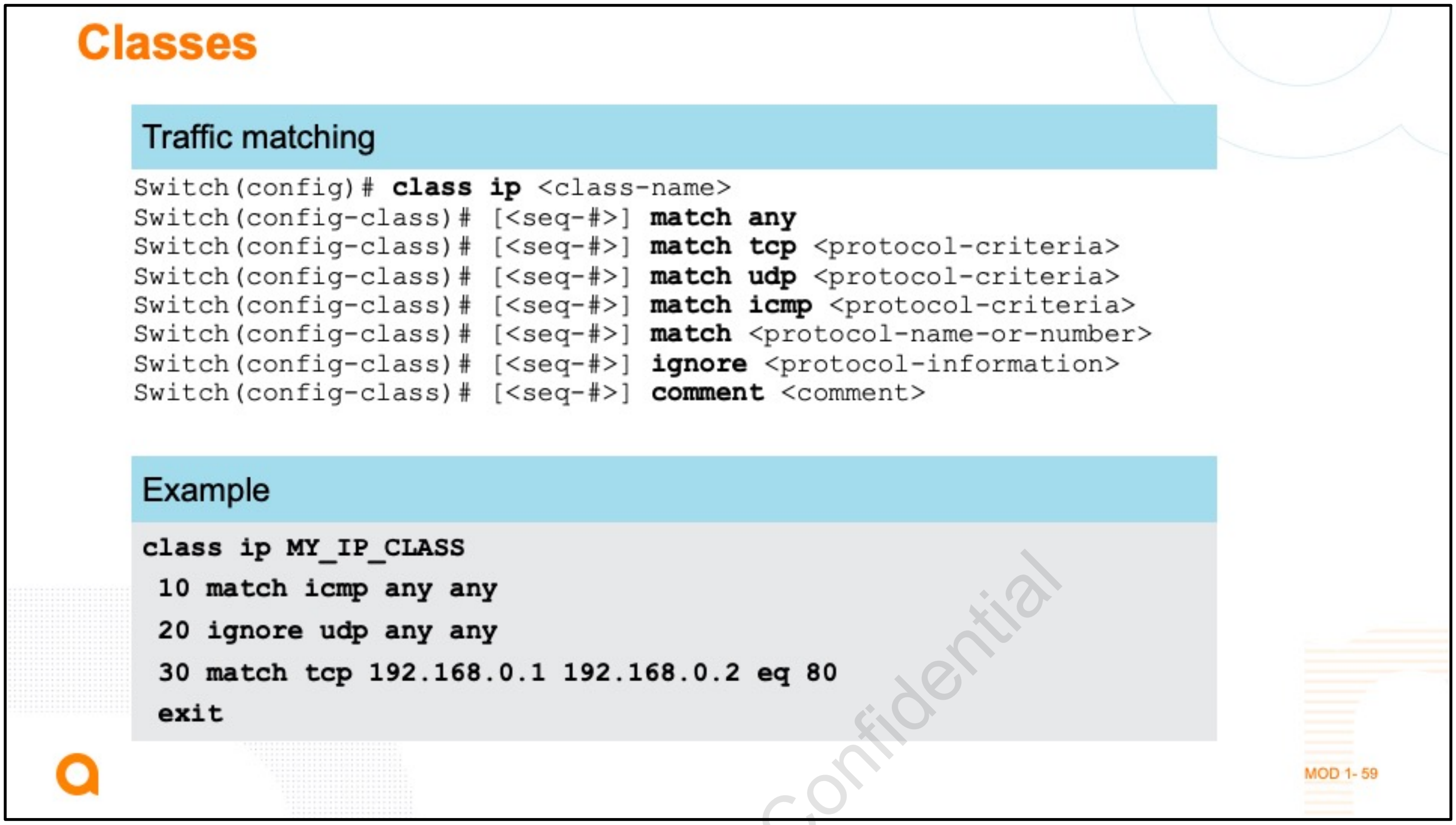
- Rate limit, QoS re-marking, mirror traffic, filtering
- Applied to L2 ports, VLANs, and GRE tunnels

MOD 1- 58

AOS-CX switches support two feature sets related to traffic policy implementation. You use ACLs primarily to filter traffic. You use classifier policies to implement a multitude of policies - traffic filtering, rate limiting traffic, changing (re-marking) QoS markings for DSCP or 802.1P, and traffic mirroring.

The focus here is on Classifier Policies.

Aruba Training-Confidential



Classes

Traffic matching

```
Switch(config)# class ip <class-name>
Switch(config-class)# [<seq-#>] match any
Switch(config-class)# [<seq-#>] match tcp <protocol-criteria>
Switch(config-class)# [<seq-#>] match udp <protocol-criteria>
Switch(config-class)# [<seq-#>] match icmp <protocol-criteria>
Switch(config-class)# [<seq-#>] match <protocol-name-or-number>
Switch(config-class)# [<seq-#>] ignore <protocol-information>
Switch(config-class)# [<seq-#>] comment <comment>
```

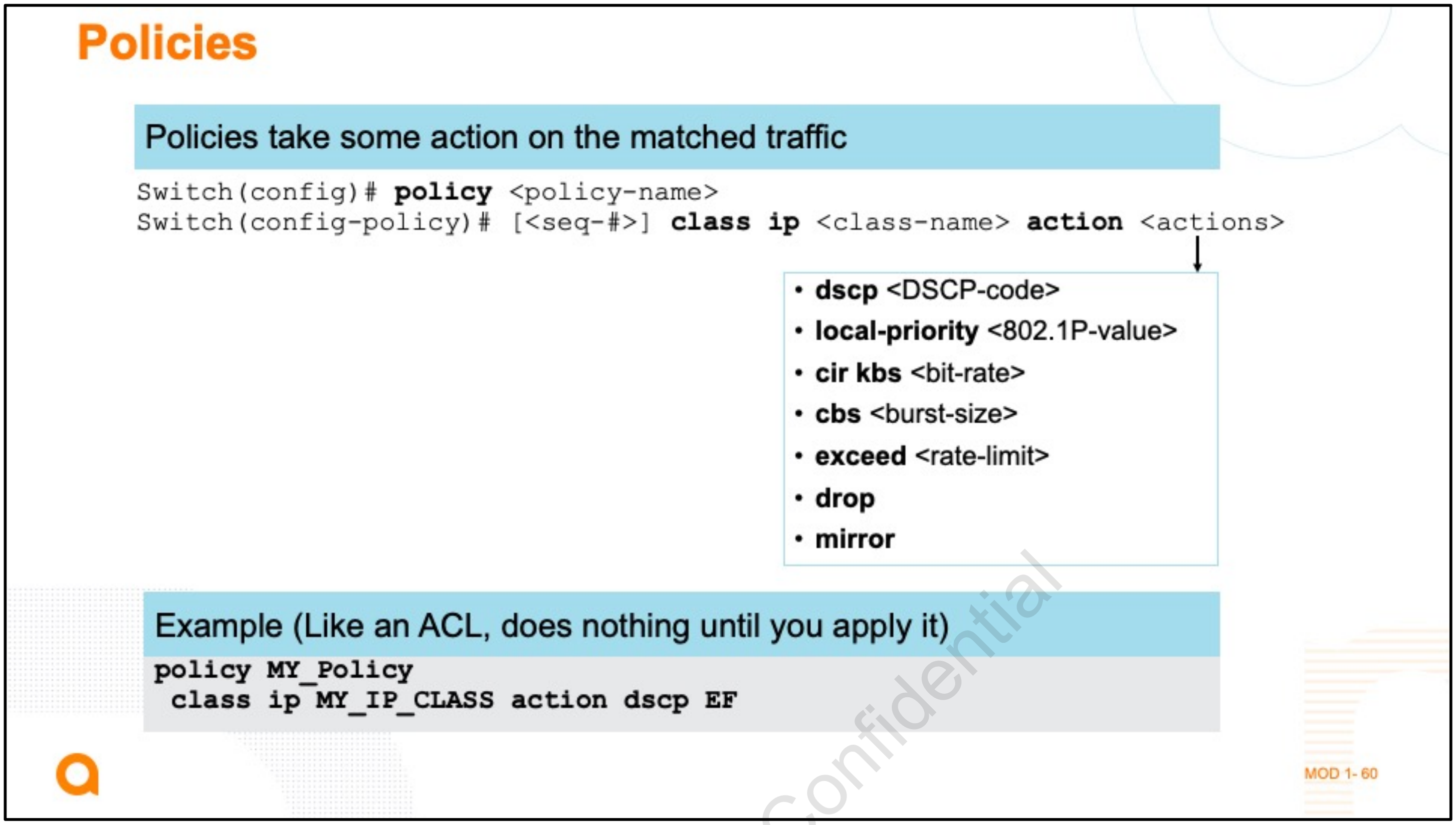
Example

```
class ip MY_IP_CLASS
 10 match icmp any any
 20 ignore udp any any
 30 match tcp 192.168.0.1 192.168.0.2 eq 80
 exit
```

MOD 1- 59

Aruba Training-Confidential

The figure shows how to create classes (called “class maps”) to identify traffic to be applied to a policy. Traffic to include is indicated by a match command, ignore commands indicate which traffic to ignore. The entries are processed in numerical order, so the sequencing of the entries is important. Use the show class command to verify your configuration. See the Command-Line Interface Guide for a more detailed description of the configuration of classes.



Policies

Policies take some action on the matched traffic

```
Switch(config)# policy <policy-name>
Switch(config-policy)# [<seq-#>] class ip <class-name> action <actions>
```

- **dscp** <DSCP-code>
- **local-priority** <802.1P-value>
- **cir kbs** <bit-rate>
- **cbs** <burst-size>
- **exceed** <rate-limit>
- **drop**
- **mirror**

Example (Like an ACL, does nothing until you apply it)

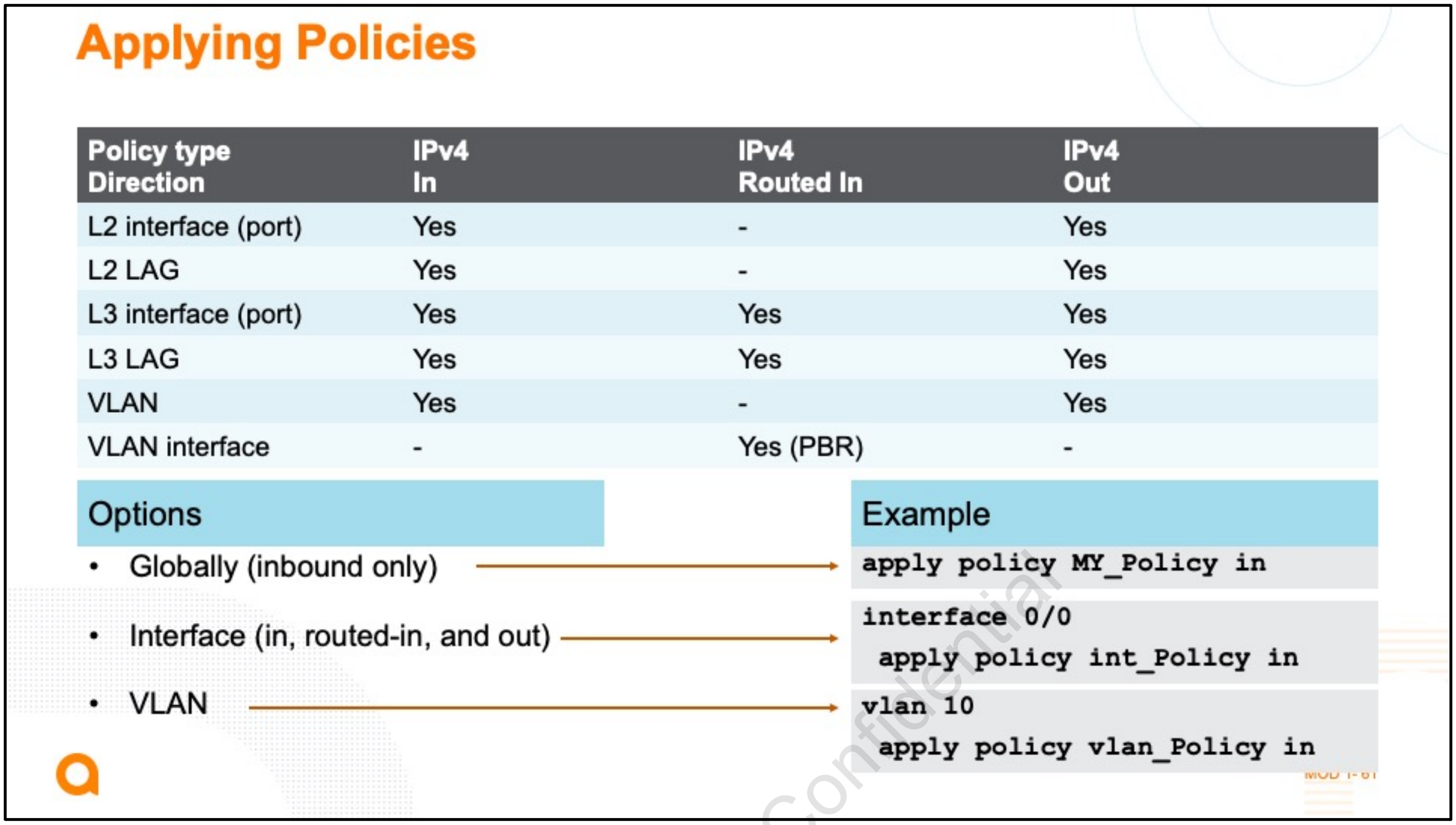
```
policy MY_Policy
 class ip MY_IP_CLASS action dscp EF
```

MOD 1- 60

Aruba Training-Confidential

The above slide shows how to reference classes in a policy and assign a policy action (or multiple actions) to a class. You typically use the dscp or local-priority parameters to define the policy. You can specify multiple classes within a policy, and that they are processed numerically based on each entry’s sequence number. Use the show policy command to verify your configuration. See the Command-Line Interface Guide for a more detailed description of the configuration of policies.

Note: If you change the DCSCP value, make sure that the DSCP value is correctly mapped to an appropriate LP value.



Applying Policies

Policy type Direction	IPv4 In	IPv4 Routed In	IPv4 Out
L2 interface (port)	Yes	-	Yes
L2 LAG	Yes	-	Yes
L3 interface (port)	Yes	Yes	Yes
L3 LAG	Yes	Yes	Yes
VLAN	Yes	-	Yes
VLAN interface	-	Yes (PBR)	-

Options

- Globally (inbound only)
- Interface (in, routed-in, and out)
- VLAN

Example

```
apply policy MY_Policy in
```

```
interface 0/0
 apply policy int_Policy in
```

```
vlan 10
 apply policy vlan_Policy in
```

MOD 1-61

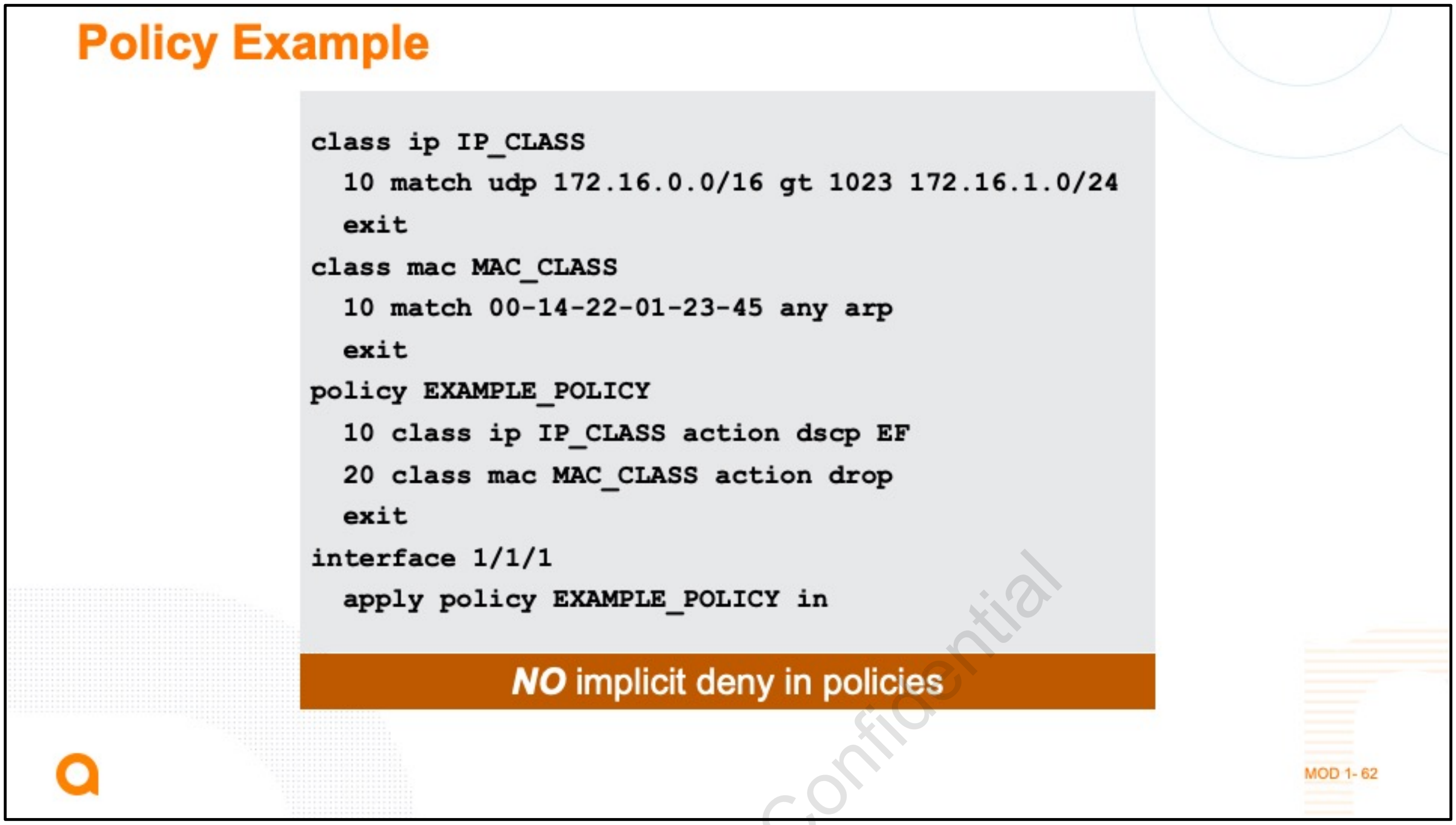
Aruba Training-Confidential

The table lists where policies can be applied. To apply a policy globally, configure this command:

Switch(config)# apply policy <POLICY-NAME> in

Globally policies can only be applied inbound. Only one policy can be globally applied at a time. Applying a policy globally again, replaces the previous globally applied policy.

Policies can also be applied to a VLAN or an interface. The apply command is used, but in the interface or VLAN context.



Policy Example

```
class ip IP_CLASS
  10 match udp 172.16.0.0/16 gt 1023 172.16.1.0/24
  exit
class mac MAC_CLASS
  10 match 00-14-22-01-23-45 any arp
  exit
policy EXAMPLE_POLICY
  10 class ip IP_CLASS action dscp EF
  20 class mac MAC_CLASS action drop
  exit
interface 1/1/1
  apply policy EXAMPLE_POLICY in
```

NO **implicit deny in policies**

MOD 1- 62

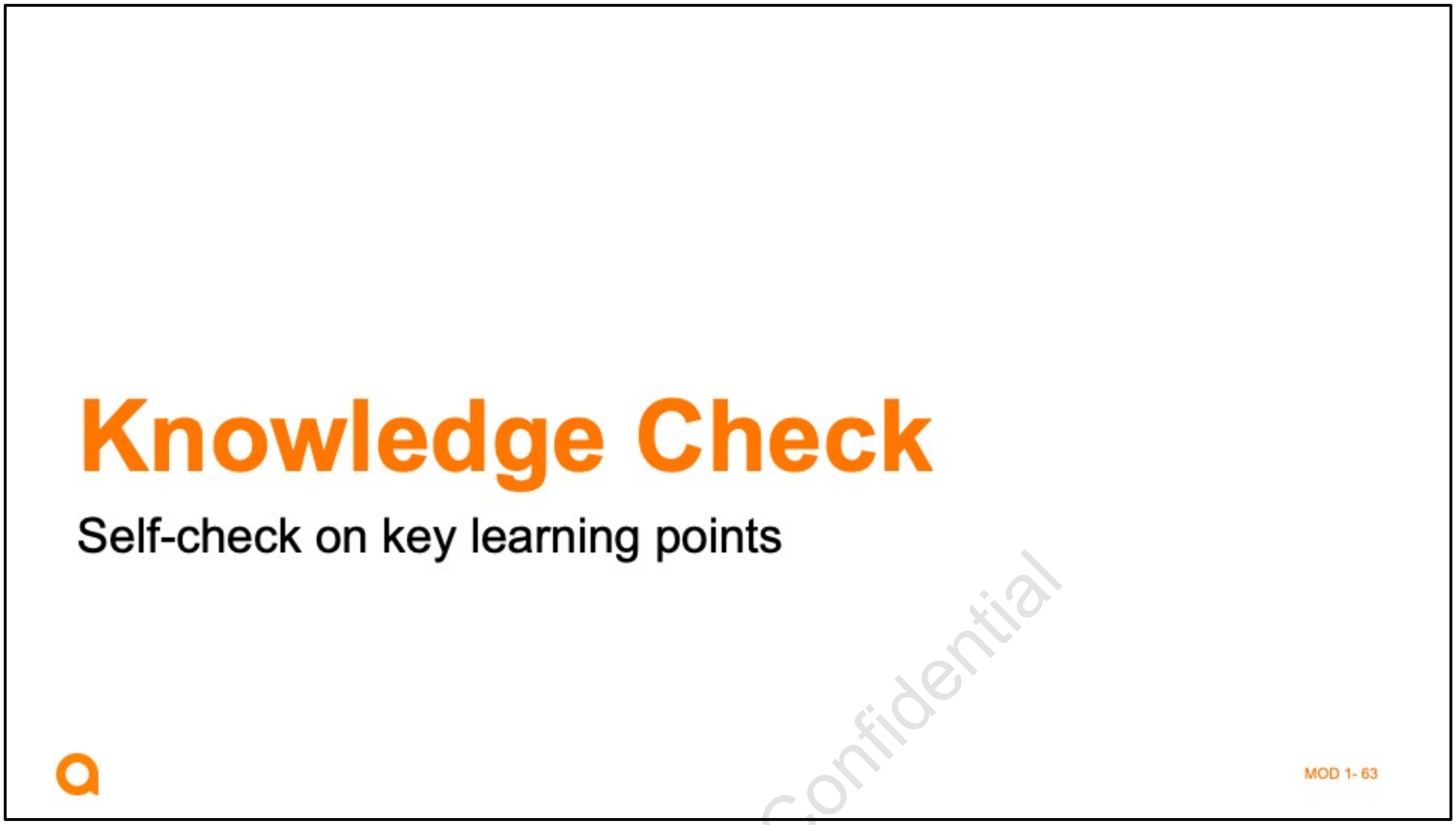
The figure shows an example policy configuration

In this example, traffic is possibly affected when it enters inbound on interface 1/1/1. For UDP traffic from 172.16.0.0/16 with a source port greater than 1023 that is destined to 172.16.1.0/24, it will be re-marked (re-classified with a DSCP value of EF. All other traffic will keep its assigned DSCP value, if any. Also, any ARP from 00-14-22-01-23-45 is dropped; but all other ARPs are allowed.

Aruba Training-Confidential

Note: There are no implicit deny in policies. If you want to apply a policy on traffic, it must match a permit. Matching on a deny statement exempts the traffic from the policy.

.

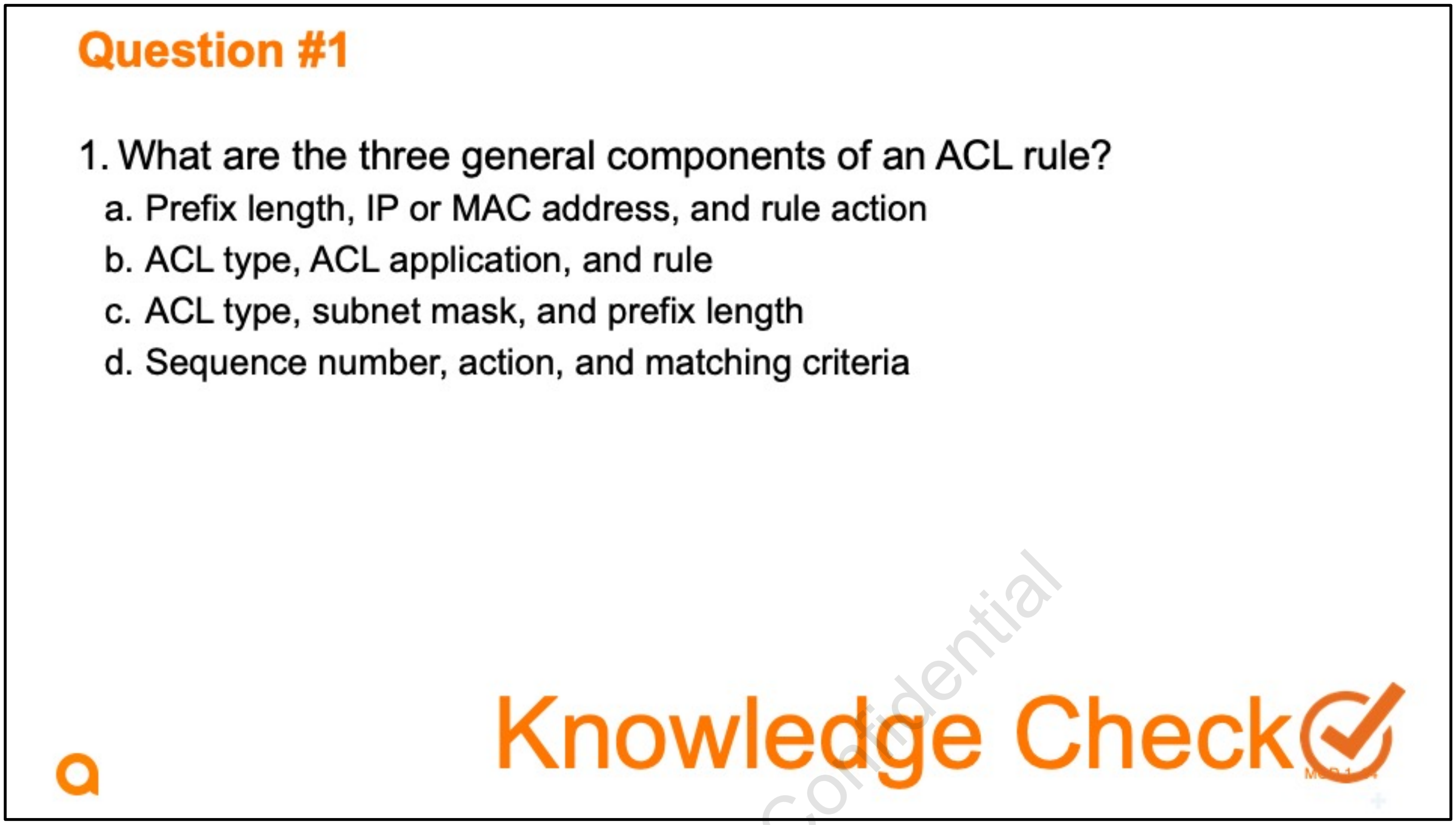


Knowledge Check

Self-check on key learning points

MOD 1- 63

Aruba Training-Confidential

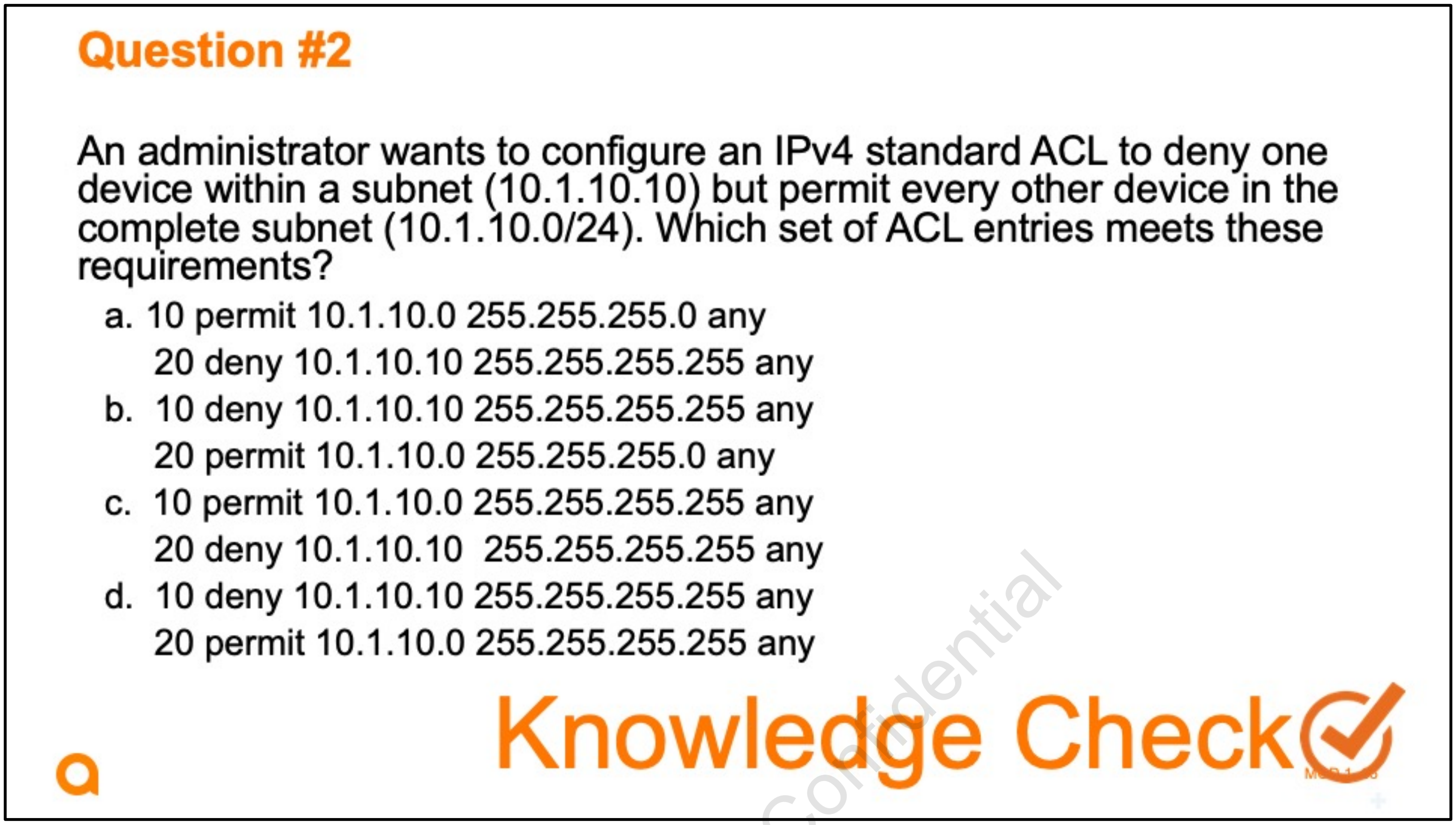


Question #1

1. What are the three general components of an ACL rule?
   a. Prefix length, IP or MAC address, and rule action
   b. ACL type, ACL application, and rule
   c. ACL type, subnet mask, and prefix length
   d. Sequence number, action, and matching criteria

Knowledge Check

Aruba Training-Confidential



Question #2

An administrator wants to configure an IPv4 standard ACL to deny one device within a subnet (10.1.10.10) but permit every other device in the complete subnet (10.1.10.0/24). Which set of ACL entries meets these requirements?

a. 10 permit 10.1.10.0 255.255.255.0 any
   20 deny 10.1.10.10 255.255.255.255 any

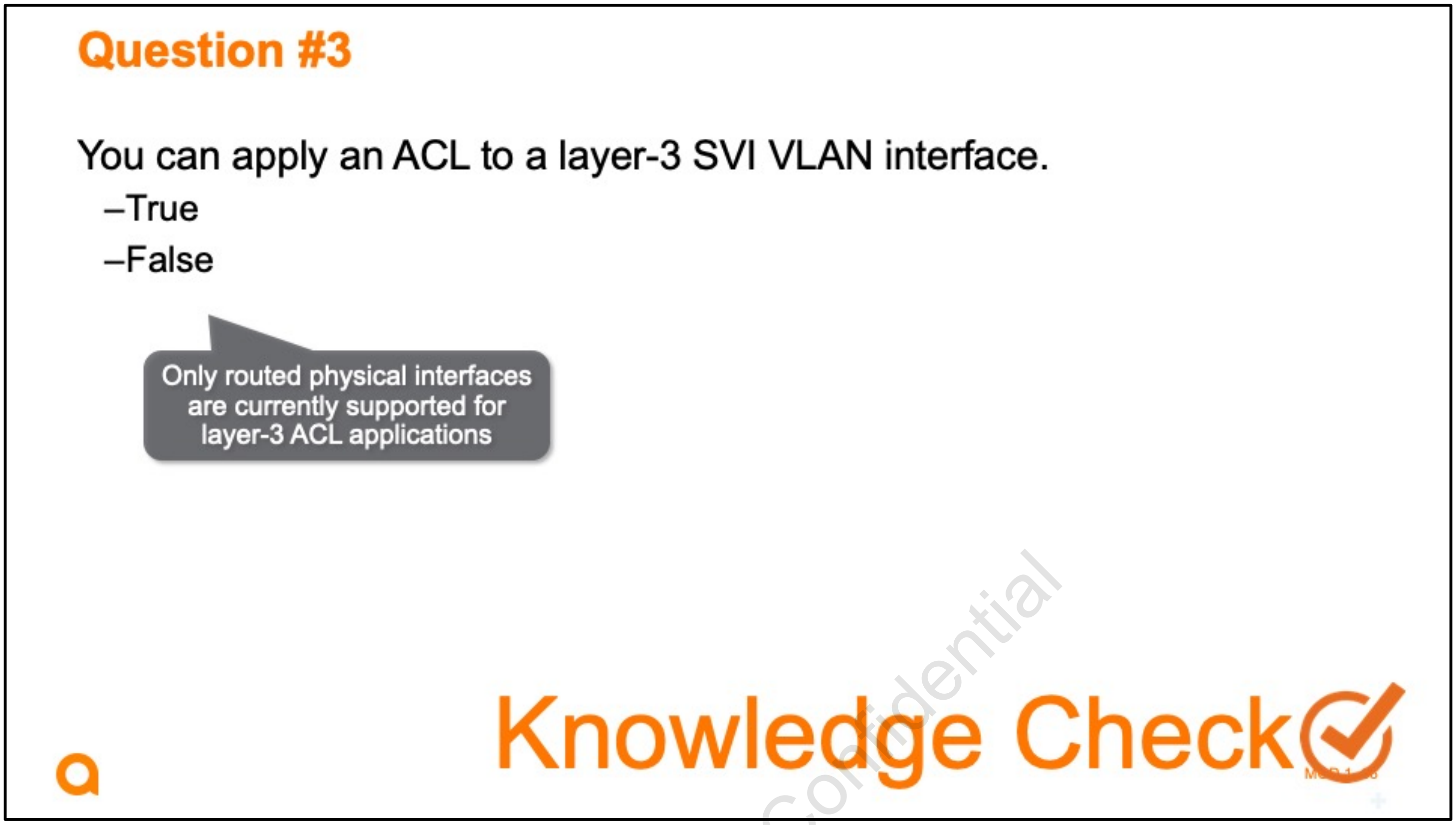
b. 10 deny 10.1.10.10 255.255.255.255 any
   20 permit 10.1.10.0 255.255.255.0 any

c. 10 permit 10.1.10.0 255.255.255.255 any
   20 deny 10.1.10.10 255.255.255.255 any

d. 10 deny 10.1.10.10 255.255.255.255 any
   20 permit 10.1.10.0 255.255.255.255 any

Knowledge Check ✓

Aruba Training-Confidential



Question #3

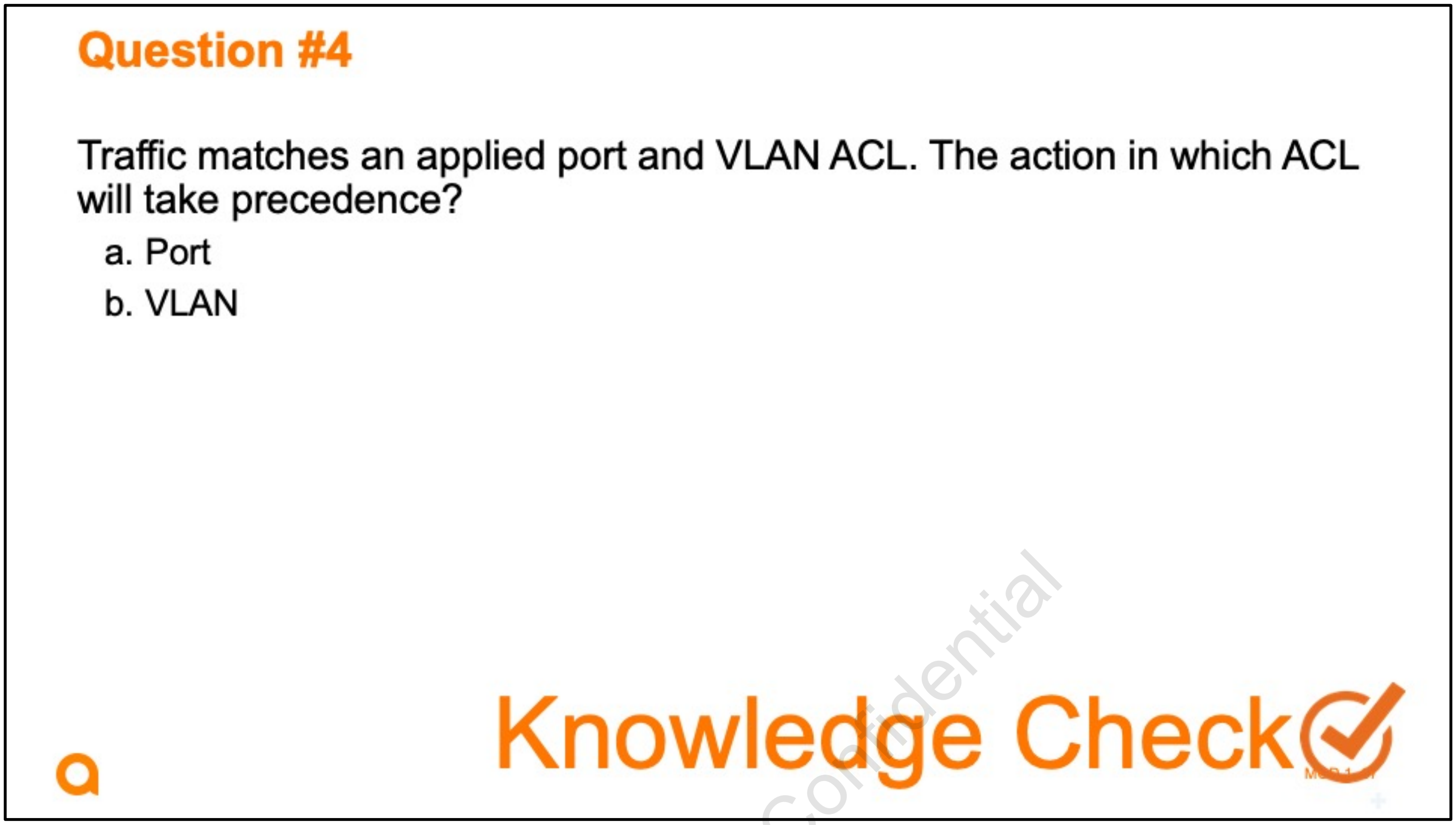
You can apply an ACL to a layer-3 SVI VLAN interface.

- –True
- –False

Only routed physical interfaces are currently supported for layer-3 ACL applications

Knowledge Check

Aruba Training-Confidential



Question #4

Traffic matches an applied port and VLAN ACL. The action in which ACL will take precedence?

a. Port

b. VLAN

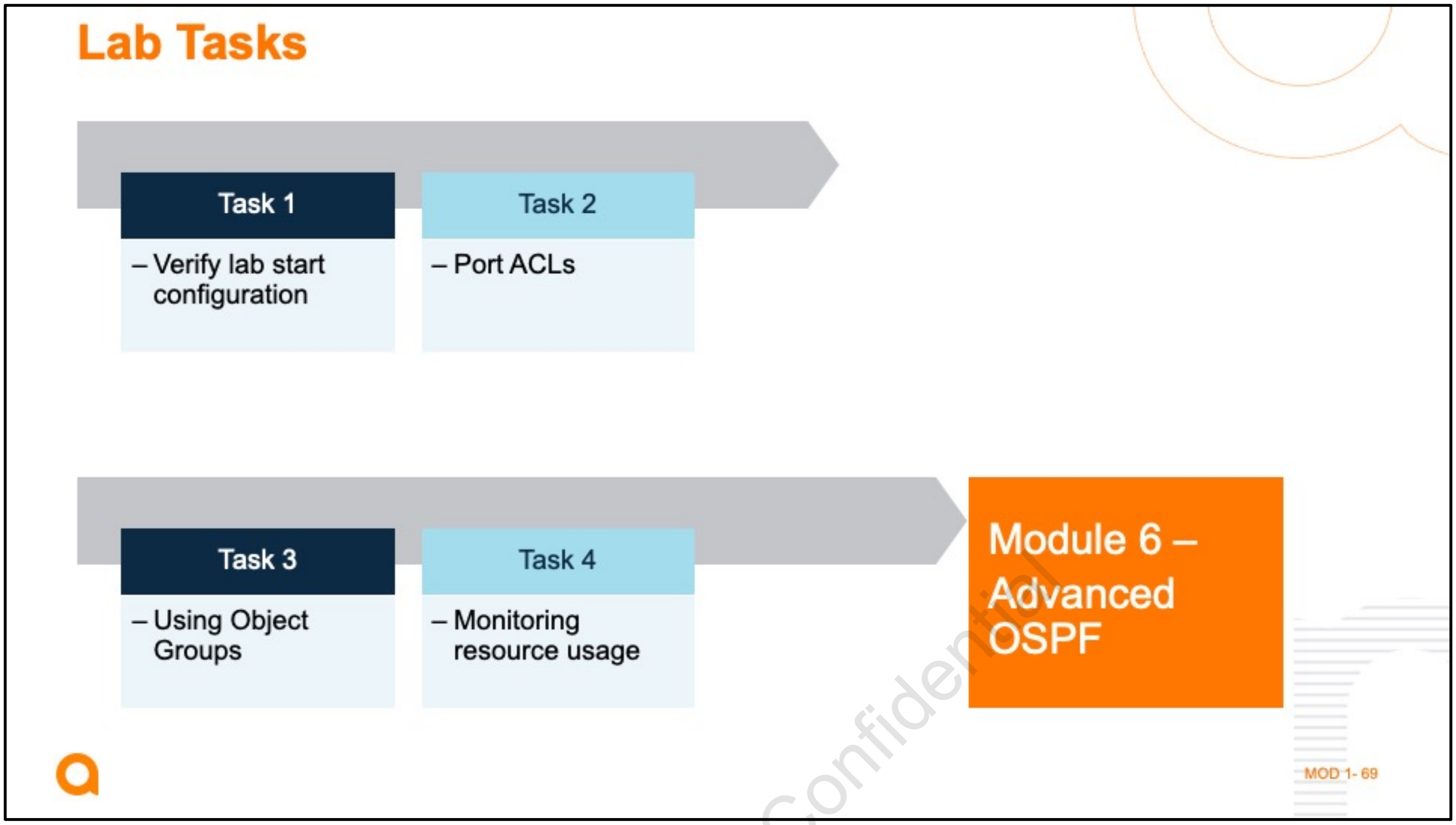
Knowledge Check

Aruba Training-Confidential

Lab Activity

How about a lab activity?

Aruba Training-O



Lab Tasks

Task 1

– Verify lab start configuration

Task 2

– Port ACLs

Task 3

– Using Object Groups

Task 4

– Monitoring resource usage

Module 6 – Advanced OSPF

MOD 1- 69

The figure provides a brief review of lab tasks. Please see your lab guide for details.

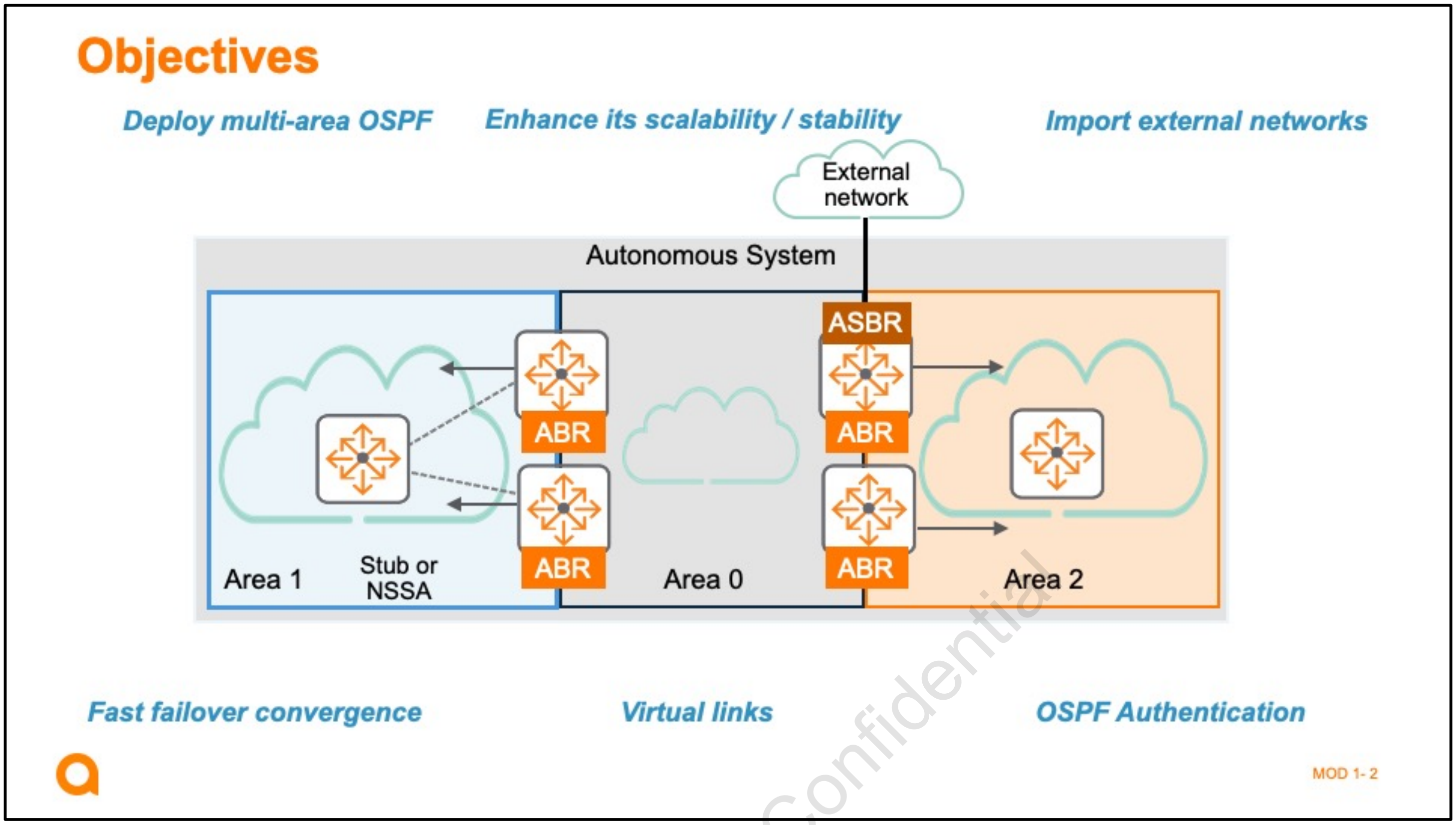
Aruba Training-Confidential

Rev # 20.21

Advanced OSPF

Implementing ArubaOS-CX Switching

Aruba Training-Confidential



Objectives

Deploy multi-area OSPF

Enhance its scalability / stability

Import external networks

Fast failover convergence

Virtual links

OSPF Authentication

MOD 1- 2

Aruba Training-Confidential

After completing this module, you should be able to:

Deploy multi-area OSPF within the enterprises networks

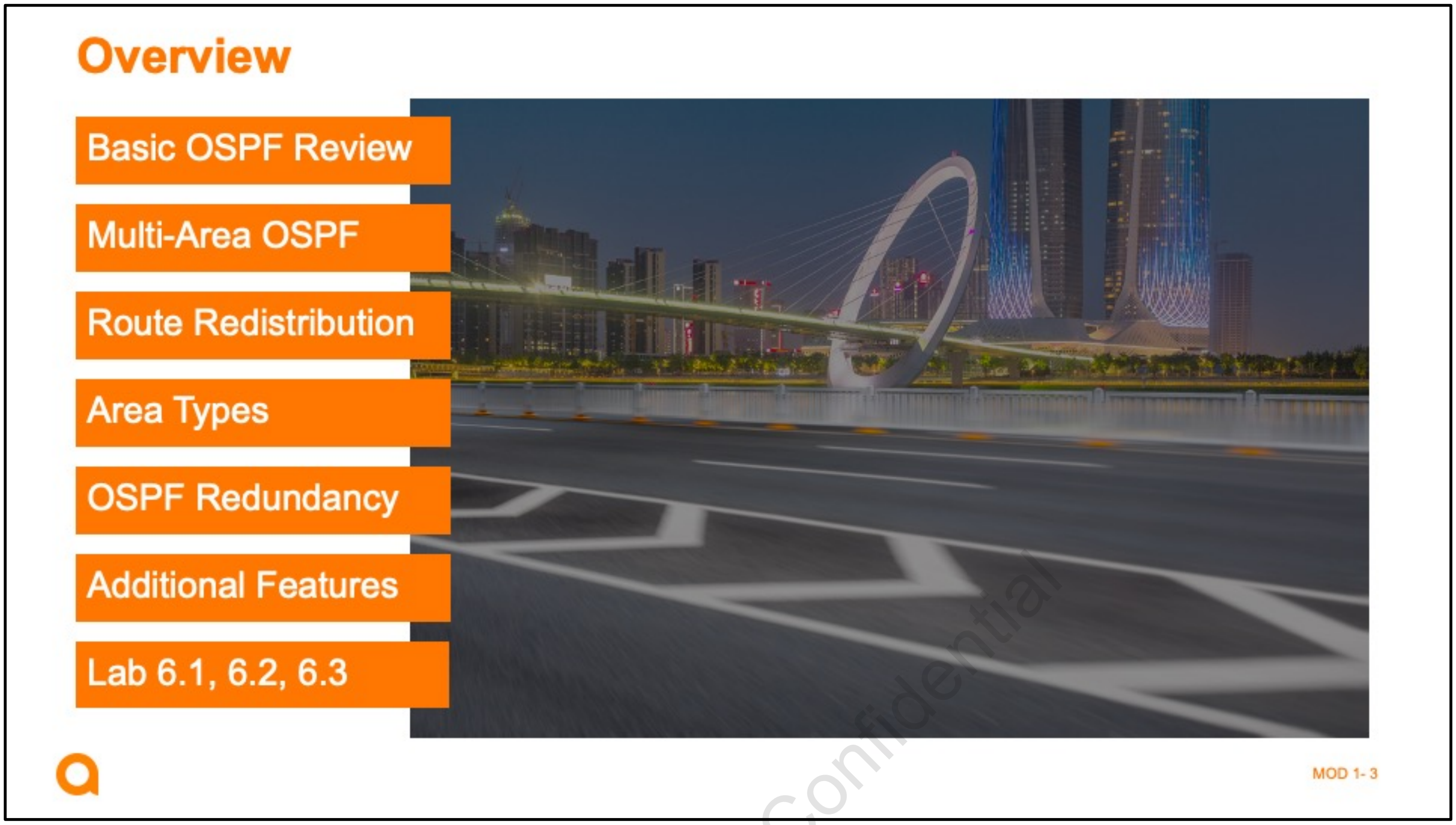
Efficient use of multi-area OSPF for scalability and stability

Import external networks in a variety of OSPF environments

How to achieve fast convergence during failover situations

Use of virtual links requirement and implementation

Implement OSFP authentication



Overview

- Basic OSPF Review
- Multi-Area OSPF
- Route Redistribution
- Area Types
- OSPF Redundancy
- Additional Features
- Lab 6.1, 6.2, 6.3

MOD 1- 3

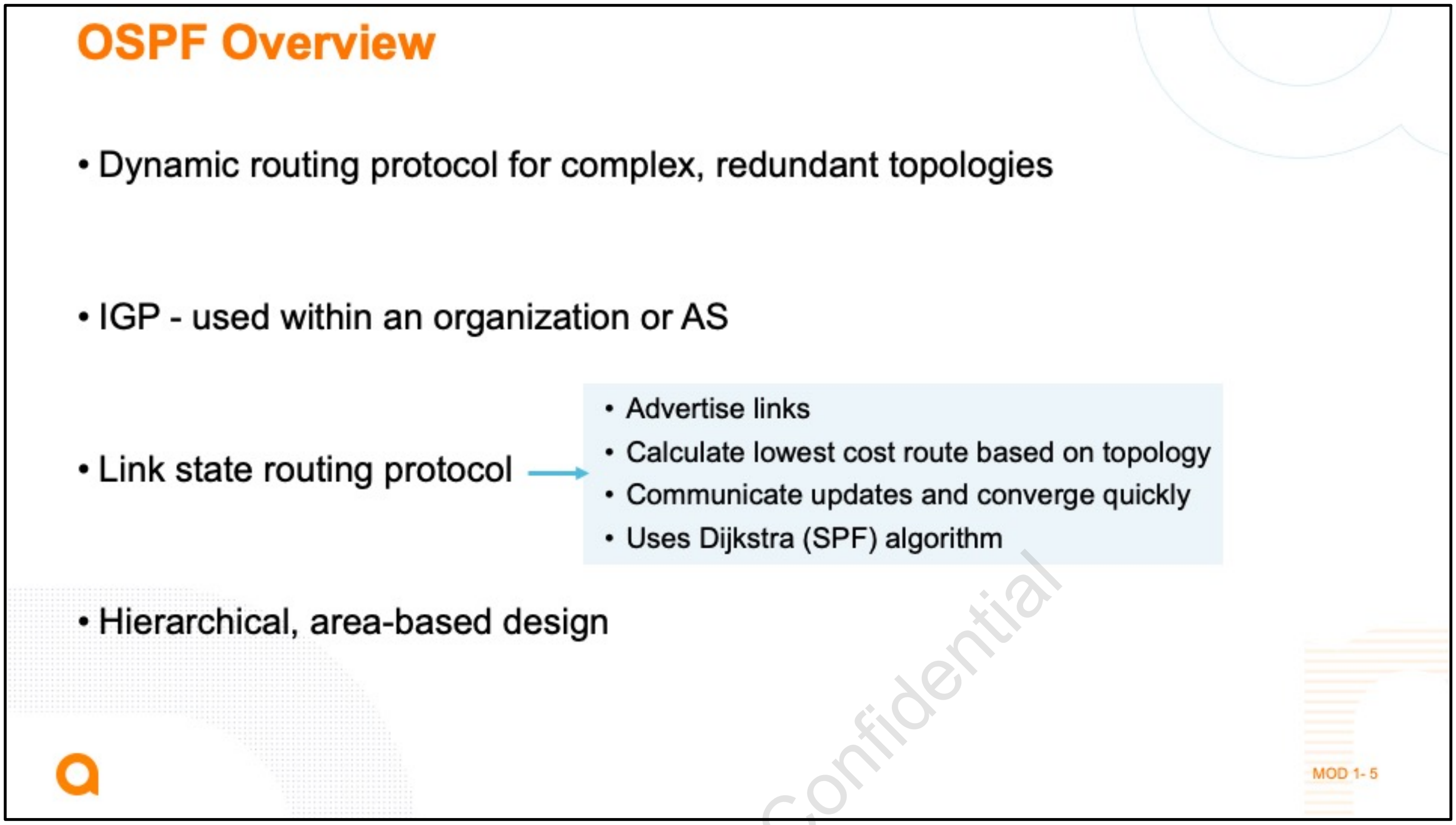
In this module, you will learn how to implement Open Shortest Path First (OSPF) routing in multiple area designs. You will examine several advanced scenarios and learn ways to configure OSPF to meet their needs. In the Fundamentals course, you learned how to implement OSPF in a small, single-area design. This module will quickly review what you learned, but then expand upon this for a large campus design. Design features that you'll explore including implementing multiple areas, area types, and redundancy. You'll also learn how to import (redistribute) routes from an external network (i.e., protocol), into OSPF as well as some additional OSPF features.

Aruba Training-Confidential

Basic OSPF Review

Let’s start with a review of OSPF operation in simple networks.

Aruba Training-Confidential



OSPF Overview

- Dynamic routing protocol for complex, redundant topologies
- IGP - used within an organization or AS
- Link state routing protocol →
  - Advertise links
  - Calculate lowest cost route based on topology
  - Communicate updates and converge quickly
  - Uses Dijkstra (SPF) algorithm
- Hierarchical, area-based design

MOD 1- 5

Aruba Training-Confidential

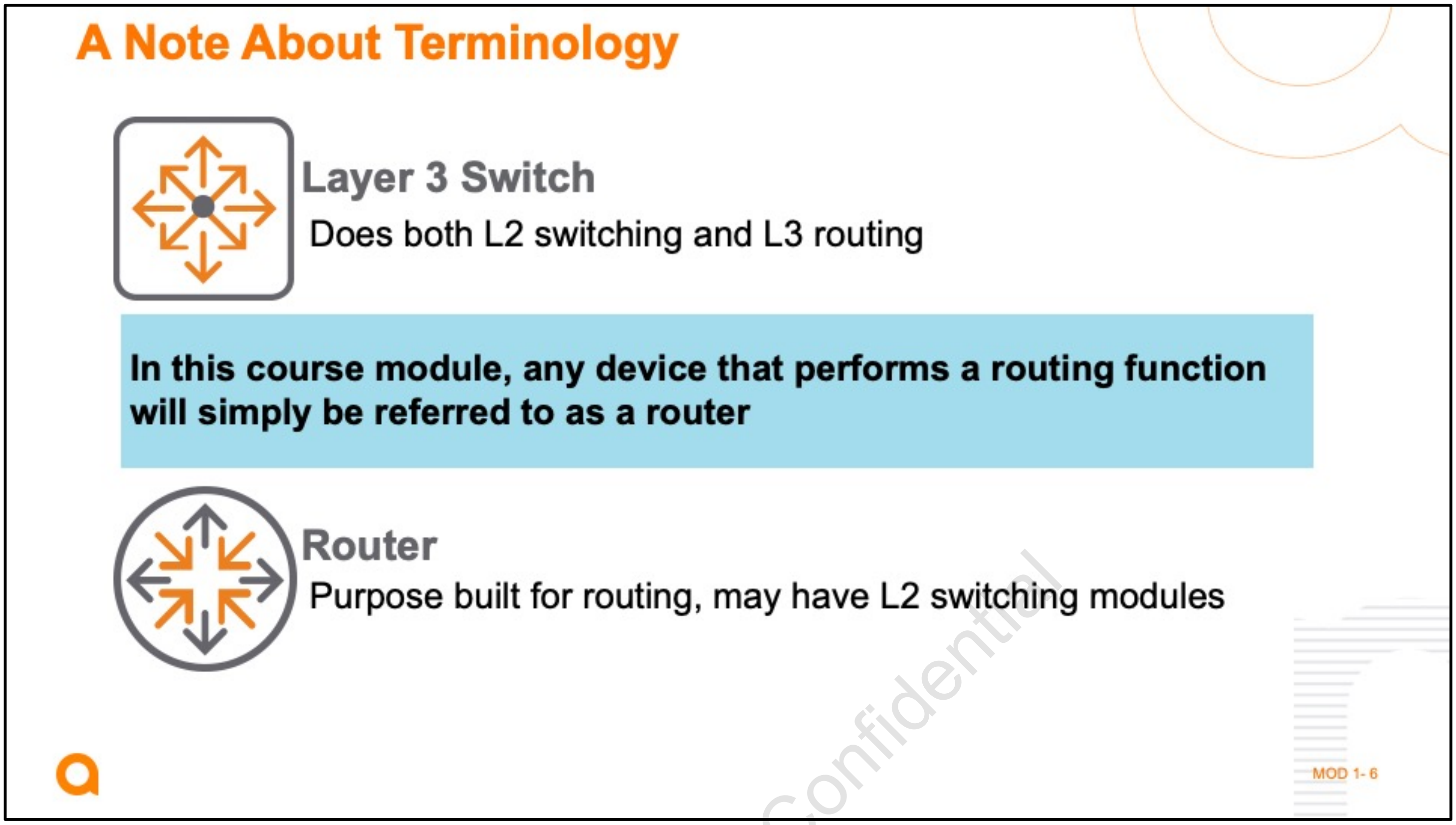
Many companies have multiple routers and routing switches that send traffic across the campus. Dynamic routing protocols enable routers to discover and update routes automatically.

As an Interior Gateway Protocol (IGP), OSPF enables route info exchange among routers in the same autonomous system (AS) – a system under a single organization's control – like your company.

OSPF is a link state protocol, where routers advertise information about their connected Layer 3 interfaces, or links, along with the cost (or metric) associated with that link. These link state advertisements (LSAs) propagate across participating routers, which use the advertisements to build a network topology – a map of the network - all known links, the routers connected to them, and associated costs. Using the topology, routers calculate their own lowest cost route to each destination subnet.

Link state protocols like OSPF tend to choose better routing paths and to converge faster than distance vector protocols like Routing Information Protocol (RIP). OSPF uses Dijkstras algorithm, also called the shortest-path first (SPF) algorithm to find the least cost, loop-free paths.

For scalability, OSPF supports a two layer hierarchical design.



A Note About Terminology

Layer 3 Switch

Does both L2 switching and L3 routing

In this course module, any device that performs a routing function will simply be referred to as a router

Router

Purpose built for routing, may have L2 switching modules

MOD 1- 6

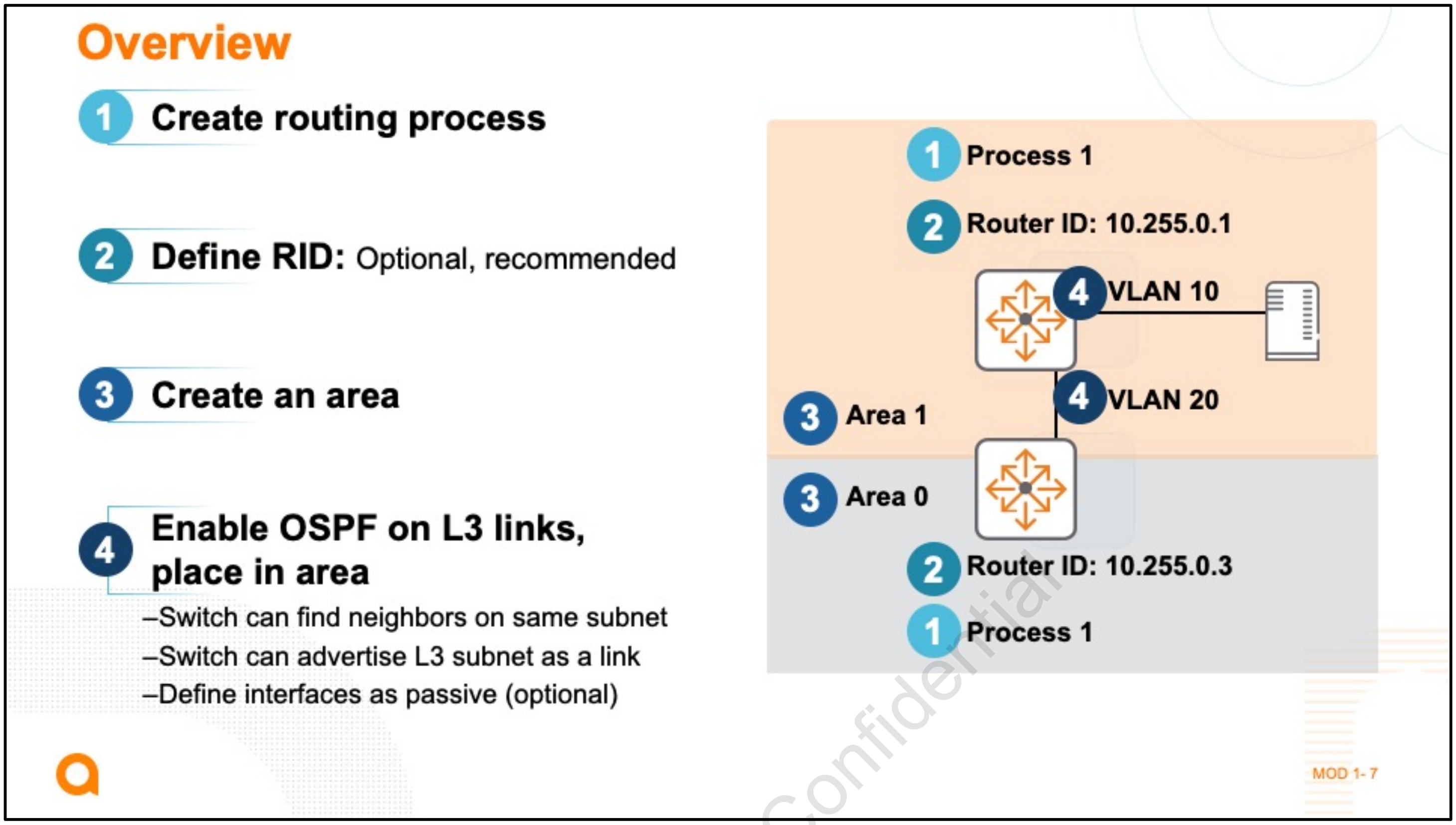
You may know that a Layer 3 switch is a device that can perform both Layer 2 switching functions and Layer 3 routing functions, depending on how you configure the device.

A router is a device that is purpose-built for routing, although some routers are modular, and may accommodate Layer 2 switching modules.

The point is, both devices perform a routing function. While performing these L3 functions, there is no significant functional difference between the two devices.

So, when discussing routing technology in this module, the term “router” is typically used to refer to any such device, unless there is some specific need to differentiate between the two.

Aruba Training-Confidential



Overview

1 **Create routing process**

2 **Define RID:** Optional, recommended

3 **Create an area**

4 **Enable OSPF on L3 links, place in area**

- –Switch can find neighbors on same subnet
- –Switch can advertise L3 subnet as a link
- –Define interfaces as passive (optional)

1 Process 1

2 Router ID: 10.255.0.1

4 VLAN 10

4 VLAN 20

3 Area 1

3 Area 0

2 Router ID: 10.255.0.3

1 Process 1

MOD 1- 7

Aruba Training-Confidential

Let's review the basic steps for setting up OSPF. Of course, configuring OSPF to meet your organization's needs has many more nuances—many of which are covered in this module. However, these basic steps can get your AOS-CX switches routing traffic across the network:

Create an OSPF routing process – an arbitrary number that has local significance only. Although they need not match between routers, its best to do so for consistency. Most folks simply set it to 1.

Optionally define a router ID (highly recommended). Unlike the routing process, this must be unique for each router in your OSPF system. For stability, use a loopback interface. If you use a physical interface as the router ID, and that link fails, the router ID changes – which is sub-optimal.

Create an area.

- OSPF uses areas for scalability - to segment the network into smaller pieces and to enable route aggregation. Your initial configuration will use one area, area 1. You will then learn about multiple area solutions.

Enable OSPF on layer 3 interfaces. For AOS-CX switches, these are either routed or VLAN interfaces. Enabling OSPF on an interface has two effects:

- The routing switch sends OSPF messages on the interface and can form neighbor relationships with other routers on the same interface. Neighbors that fully exchange their topology information are called fully adjacent.
- The router advertises this interface in OSPF. For example, if the interface is associated with

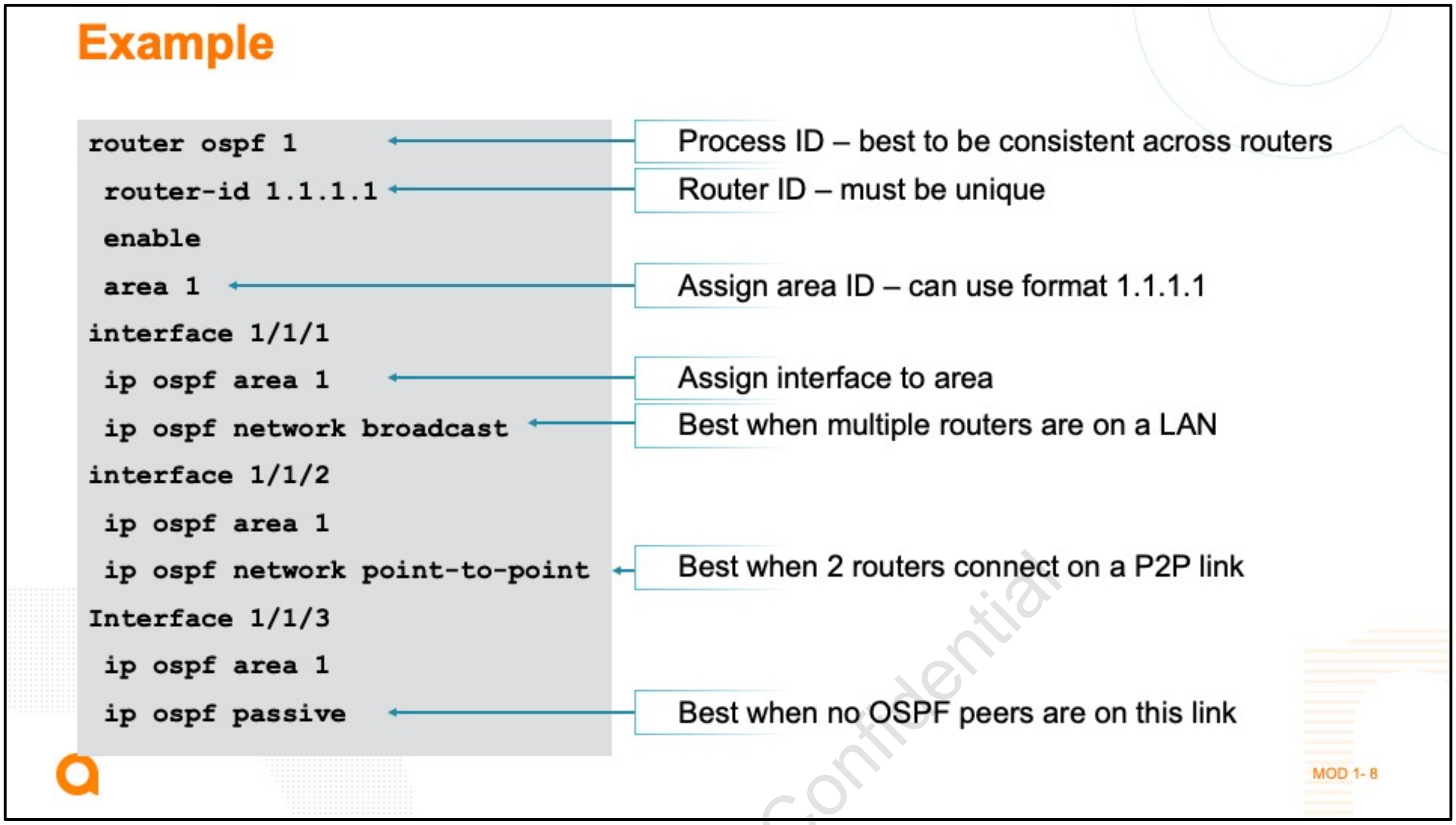
IP address 10.255.0.1/24, the router advertises that it has a link to 10.255.0.0/24.

VLANs that need to be advertised in OSPF include:

- Those that connect routers (like VLAN 255 on Access-1 and Core-1)
- Those that support endpoints (like VLAN 10 on Core-1)
- If you are only enabling OSPF on the VLAN because you want to advertise its subnet, and you do not want the router to form adjacencies with other routers on the VLAN, configure the VLAN as a passive OSPF interface.

- When you enable OSPF on a VLAN, you also specify the VLAN's OSPF area. The area ID must match on all OSPF routers connected on the same VLAN or subnet.

Aruba Training-Confidential



Example

```
router ospf 1
 router-id 1.1.1.1
 enable
 area 1
interface 1/1/1
 ip ospf area 1
 ip ospf network broadcast
interface 1/1/2
 ip ospf area 1
 ip ospf network point-to-point
Interface 1/1/3
 ip ospf area 1
 ip ospf passive
```

Process ID – best to be consistent across routers

Router ID – must be unique

Assign area ID – can use format 1.1.1.1

Assign interface to area

Best when multiple routers are on a LAN

Best when 2 routers connect on a P2P link

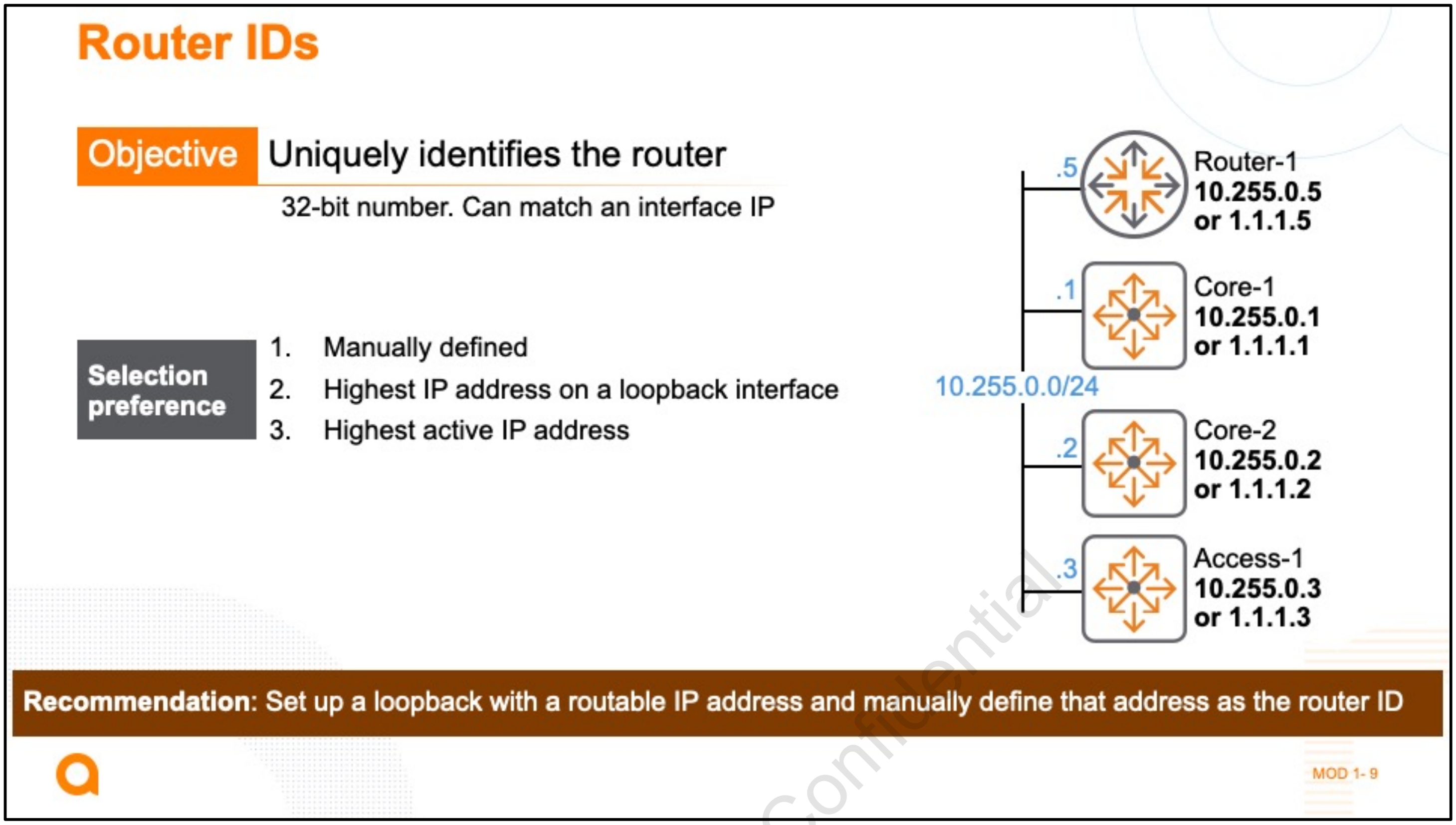
Best when no OSPF peers are on this link

MOD 1- 8

Aruba Training-Confidential

The figure shows a basic OSPF configuration. After you complete these steps, your routers automatically create a topology of your network and learn routes to every subnet within it. If a link fails, the routers inform each other of the change and learn new routes.

By default, OSPF is enabled when you create the routing process the first time. If you disable it with the disable command, use the corresponding enable command to re-enable it. This allows you to temporarily turn off OSPF without losing your OSPF configuration.



Router IDs

Objective Uniquely identifies the router

32-bit number. Can match an interface IP

Selection preference

1. Manually defined
2. Highest IP address on a loopback interface
3. Highest active IP address

10.255.0.0/24
.5
Router-1
10.255.0.5
or 1.1.1.5
.1
Core-1
10.255.0.1
or 1.1.1.1
.2
Core-2
10.255.0.2
or 1.1.1.2
.3
Access-1
10.255.0.3
or 1.1.1.3

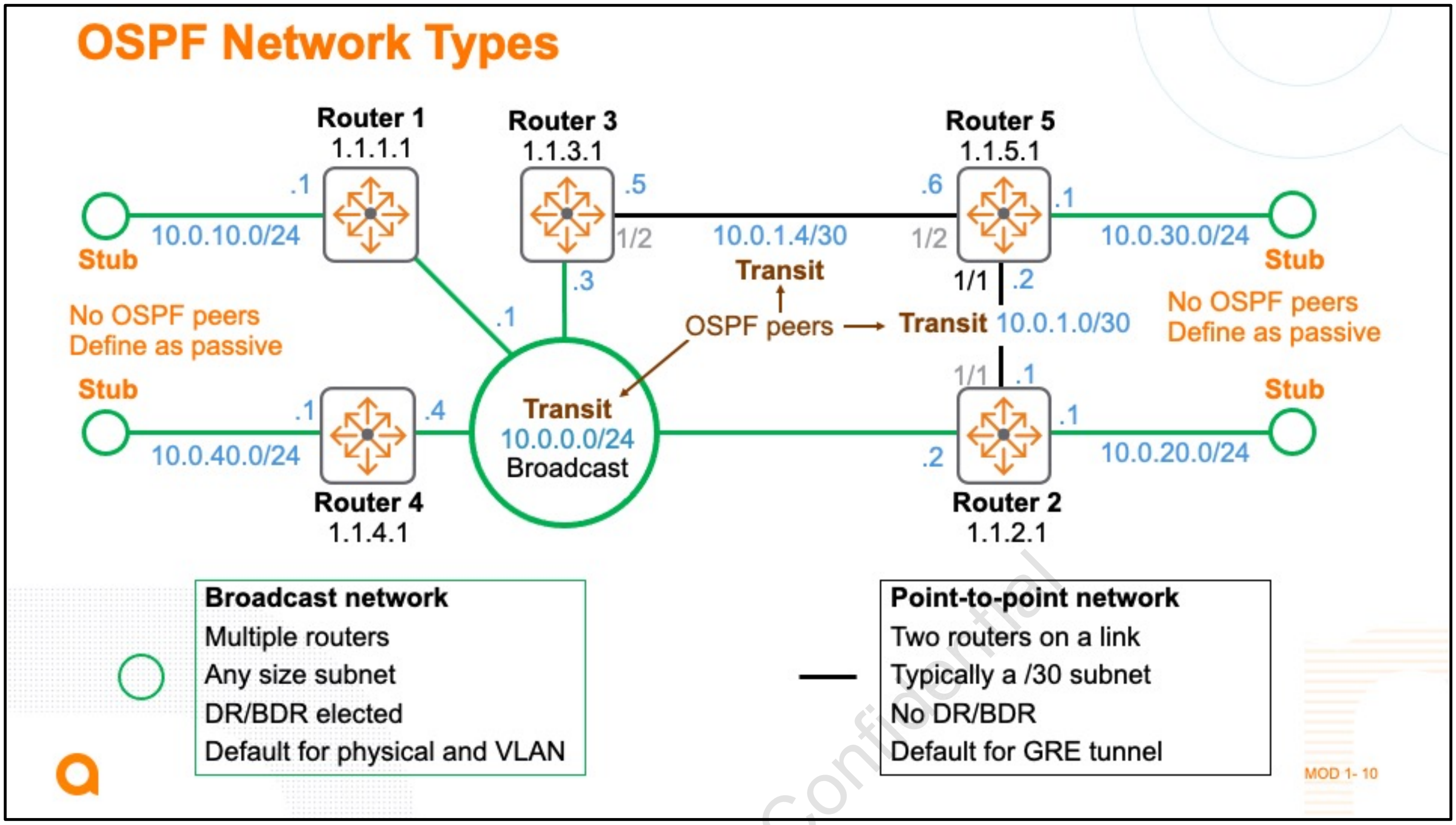
Recommendation: Set up a loopback with a routable IP address and manually define that address as the router ID

MOD 1- 9

Aruba Training-Confidential

The router ID uniquely identifies the router within the OSPF system. This ID has the format of an IP address, and it can be an IP address on one of the routing device's interfaces. However, you can also set an ID manually, and this ID does not necessarily need to match an IP address on the router. The figure shows a very simple topology in which routers connect on a single shared subnet, and they each have a router ID that matches their IP address on that subnet. You could also define a logical loopback interface on each router, and set the router ID to match that loopback IP address. This ensures that the associated IP address is always reachable, even if some of the router's links fail.)

Note: It is best practice to set up a loopback interface with the router ID and including the loopback interface in an OSPF area. This ensures that the admin can always ping the address, even if links' statuses change, which greatly facilitates troubleshooting OSPF problems.



OSPF Network Types

Broadcast network
Multiple routers
Any size subnet
DR/BDR elected
Default for physical and VLAN

Point-to-point network
Two routers on a link
Typically a /30 subnet
No DR/BDR
Default for GRE tunnel

MOD 1- 10

Aruba Training-Confidential

OSPF routers run the protocol on layer 3 interfaces, or networks. This includes VLAN interfaces, physical interfaces set to routed mode, link aggregations set to routed mode, and loopback interfaces. AOS-CX switches support two types of networks: broadcast and point-to-point. Ethernet is a broadcast protocol, and a VLAN is a common example of a broadcast network.

You can connect OSPF routers to the VLAN in some sort of redundant tree topology, directly to each other, or through Layer 2 switches.

The figure shows an example of a Layer 3 connection with a circle, to which all routers connect. This “circle” is likely a Layer 2 switch - the Layer 2 connections are not shown. In a broadcast network, a designated router (DR) is elected to establish adjacencies with other routers on the link (subnet). DRs also receive and flood link state updates (LSUs). A backup DR (BDR) is also elected to take over if the DR fails. OSPF routers multicast messages to the DR and BDR at 224.0.0.6. The purpose of DR/BDR election is efficiency. Non-DR routers need not communicate with several other routers on the link – only the DR/BDR.

When you directly connect two OSPF routing switches, it is generally preferable to set up a point-to-point network. This network is typically associated with a /30 subnet. On AOS-CX switches, GRE tunnel interfaces are point-to-point by default, and loopback interfaces are a special loopback type. It could also include VLAN interfaces, but only the two routers belong to the associated VLAN. But most interfaces are broadcast networks by default, including physical route-only ports and VLAN interfaces. However, you can set route only ports and link aggregations to point-to-point, assuming that the routed link connects directly to a routed link on another OSPF device. You can even set VLAN interfaces to point-to- point if they follow the rules. (Remember, though, that a VSX LAG between a VSX pair running OSPF and another OSPF router should be a broadcast network because the network has three routers: each VSX switch and the connected router.) Since a point-to-point network has only two routers, it does not require the DR and BDR roles.

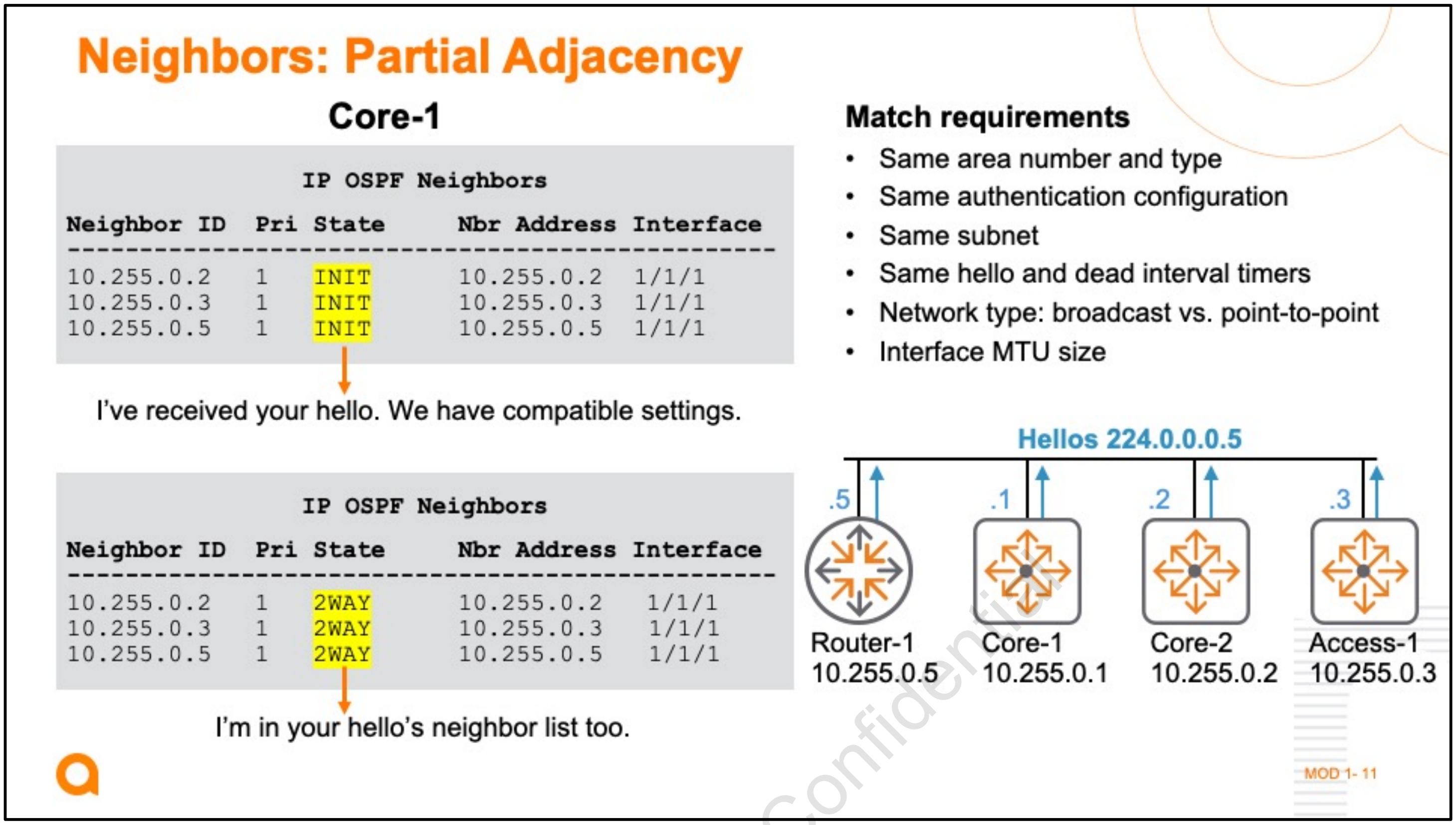
OSPF also distinguishes between transit and stub networks. A transit network is any network on which a router discovers at least one neighbor; it could be a broadcast or point-to-point network. A stub network is any network on which the router does not detect another OSPF device. The reason that the router runs OSPF on the stub network is to advertise the network to other routers in the AS. Often a stub network is a VLAN interface for which the OSPF router is the default gateway.

Note: Two routers (even a VSX core) on same subnet, with passive interfaces, also results in stub network.

You can define the passive option for interfaces that will not connect to other OSPF peers, and/or to ensure that the interface will not form an OSPF adjacency should any OSPF neighbors be connected to that interface. This can slightly reduce OSPF overhead on the link, and improves security by ensuring that rogue routers will not connect to your system - especially if used along with OSPF passwords.

Note: OSPF point-to-point networks converge faster than OSPF broadcast networks and therefore are more preferred when applicable.

Aruba Training-Confidential



Neighbors: Partial Adjacency

Core-1

```
              IP OSPF Neighbors

Neighbor ID  Pri State      Nbr Address Interface
-------------------------------------------------
10.255.0.2    1   INIT      10.255.0.2  1/1/1
10.255.0.3    1   INIT      10.255.0.3  1/1/1
10.255.0.5    1   INIT      10.255.0.5  1/1/1
```

I've received your hello. We have compatible settings.

```
              IP OSPF Neighbors

Neighbor ID  Pri State      Nbr Address Interface
-------------------------------------------------
10.255.0.2    1   2WAY      10.255.0.2   1/1/1
10.255.0.3    1   2WAY      10.255.0.3   1/1/1
10.255.0.5    1   2WAY      10.255.0.5   1/1/1
```

I'm in your hello's neighbor list too.

Match requirements

- Same area number and type
- Same authentication configuration
- Same subnet
- Same hello and dead interval timers
- Network type: broadcast vs. point-to-point
- Interface MTU size

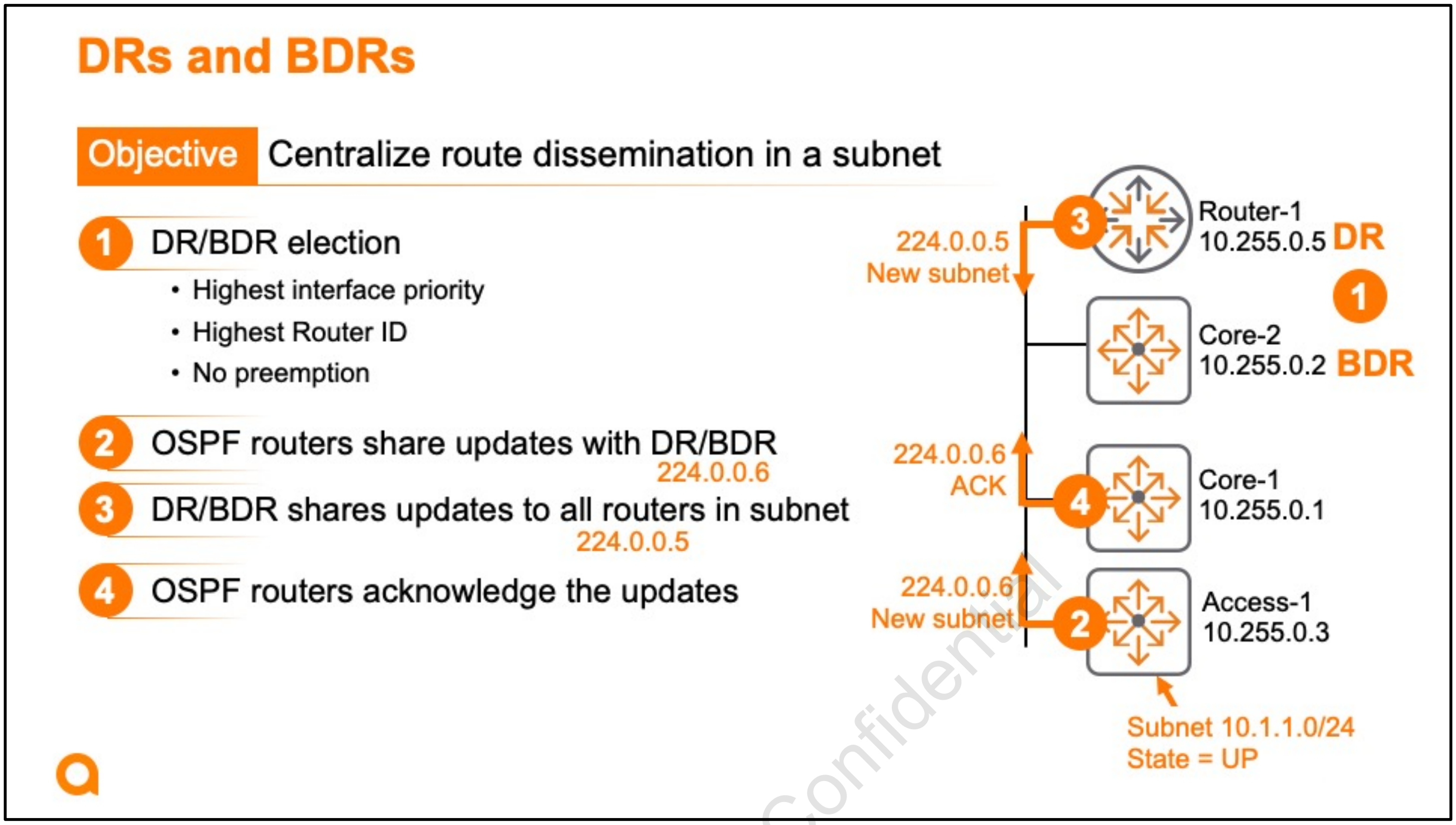
Aruba Training-Confidential

Now you will review how an OSPF router discovers its neighbors. The router sends hellos on each OSPF interface to multicast address 224.0.0.5 – the "all OSPF routers" address. Therefore, all other OSPF routers in the same subnet will hear the hello.

An OSPF interface is a Layer 3 interface with an IP address on which OSPF is enabled. On AOS-CX switches, the Layer 3 interface is a VLAN. The hello includes information about the router's OSPF settings on the interface, and a list of neighbors discovered on the interface. The neighbor list is empty until the router receives hellos from neighbors.

A receiving router checks the settings in the hello to make sure that they match its own settings on the interface. The figure shows the settings that must match for the router to accept the hello. Note that these settings might be mismatched initially, or configuration changes might cause these settings to change after neighbors already have an established relationship. Routers continue to transmit hellos periodically, so the neighbor will receive the hello with the mismatched settings and remove the neighbor relationship. If these do not match, the OSPF routers will fail to form an adjacency and will be stuck in an INIT state.

But if a peer's settings match, the router adds it as an OSPF neighbor in the INIT state. It also adds the neighbor's router ID in the neighbor list included in its hellos. When an OSPF router receives a hello with its own ID in this list, it knows that it has bidirectional connectivity with the neighbor. It moves the neighbor to 2WAY state.



DRs and BDRs

Objective Centralize route dissemination in a subnet

1 DR/BDR election
- Highest interface priority
- Highest Router ID
- No preemption

2 OSPF routers share updates with DR/BDR 224.0.0.6

3 DR/BDR shares updates to all routers in subnet 224.0.0.5

4 OSPF routers acknowledge the updates

Aruba Training-Confidential

An OSPF router will not form adjacencies to just any router. Instead, a client/server design is implemented in OSPF on each broadcast segment. For each multi-access broadcast segment, such as Ethernet, there is a Designated Router (DR) and a Backup Designated Router (BDR) as well as other OSPF routers, called DROTHERs. As an example, if you have 10 VLANs/subnets in your switched area, you'll have 10 DRs and 10 BDRs. Remember, point-to-point links only have two routers, so there is no advantage nor need for DR/BDRs.

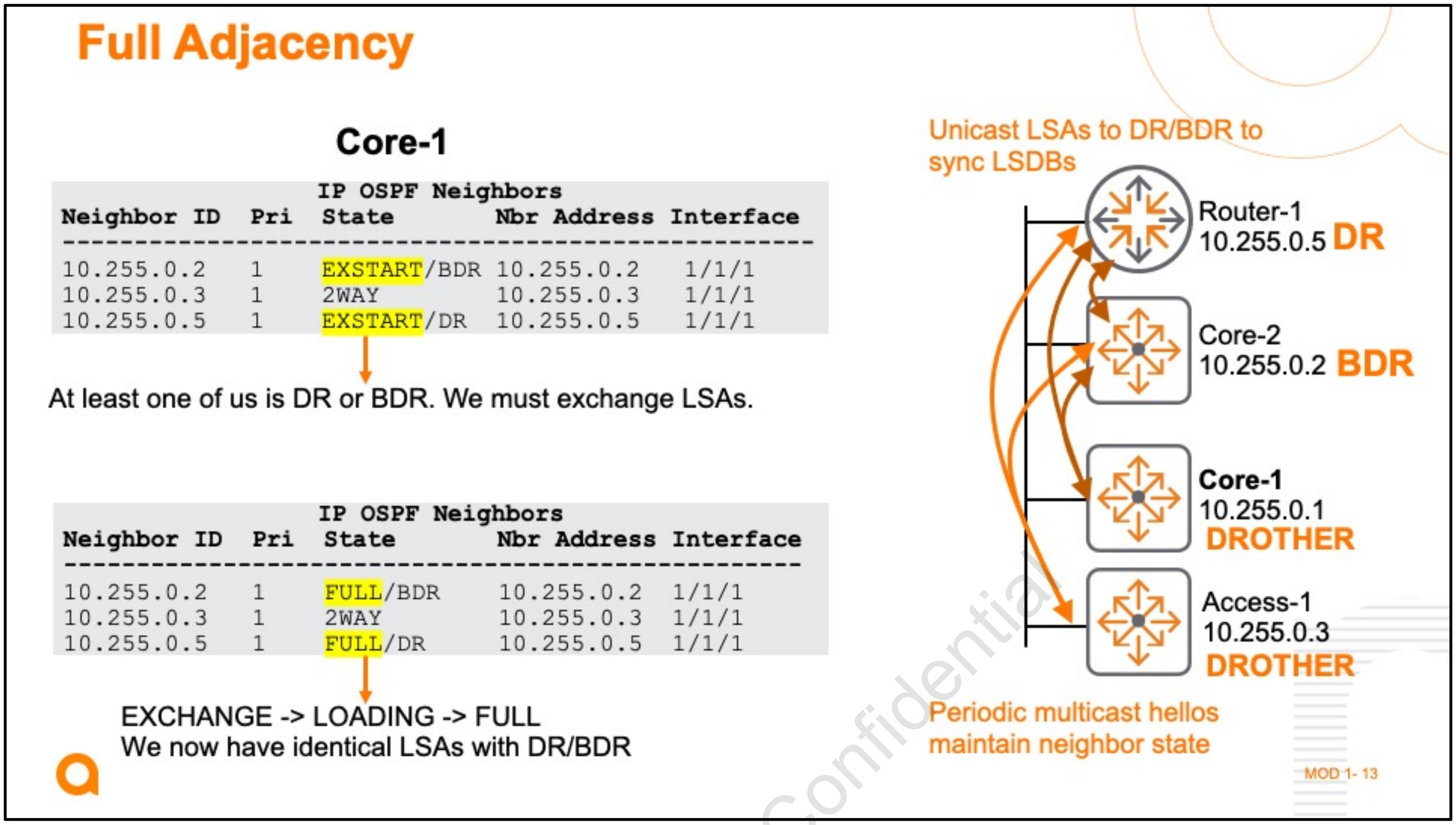
The OSPF router with the highest priority becomes the DR for the segment. If there is a tie, the router with the highest router ID (not IP address on the segment) becomes the DR. By default, all routers have a priority of 1 (priorities can range from 0 to 255—it's an 8-bit value). If the DR fails, the BDR is promoted to DR and another router is elected as the BDR. If a router does not become a DR or BDR, it is referred to as a designated router-other, or commonly DROTHER, for short. There is no preemption of the DR/BDR role. If a DR/BDR exist and a router with a higher router ID joins the segment, it will not preempt the existing DR/BDR. Election only occurs if there is no DR or BDR on a segment.

Note: This feature was specifically designed based on the hardware (or lack thereof) capabilities of routers in the 1980s. Today, this is a moot point and you should not be concerned which of your routers play which role.

When an OSPF router comes up, it forms adjacencies with the DR and the BDR on each multi-access segment to which it is connected; if it is connected to three segments, it will form three sets of adjacencies. Any exchange of routing information is between these DR/BDR routers and the other OSPF neighbors on a segment (and vice versa). An OSPF router talks to a DR using the IP multicast address of 224.0.0.6. The DR and the BDR talk to all OSPF routers using the 224.0.0.5 multicast IP address.

OSPF routers use link state advertisements (LSAs) to communicate with each other. One type of LSA is a hello, which is used to form neighbor relationships and as a keep-alive function. Hellos are generated every ten seconds. When sharing link information (directly connected routes), links are sent to the DR (224.0.0.6) and the DR disseminates this to everyone else (224.0.0.5) on the segment. On point-to-point links, since no DR/BDR is used, all OSPF packets are addressed to 224.0.0.5.

Aruba Training-Confidential



Full Adjacency

Core-1

```
                    IP OSPF Neighbors
Neighbor ID  Pri  State          Nbr Address  Interface
---------------------------------------------------------
10.255.0.2   1    EXSTART/BDR 10.255.0.2    1/1/1
10.255.0.3   1    2WAY        10.255.0.3    1/1/1
10.255.0.5   1    EXSTART/DR  10.255.0.5    1/1/1
```

At least one of us is DR or BDR. We must exchange LSAs.

```
                    IP OSPF Neighbors
Neighbor ID  Pri  State          Nbr Address  Interface
-------------------------------------------------------
10.255.0.2   1    FULL/BDR      10.255.0.2  1/1/1
10.255.0.3   1    2WAY          10.255.0.3  1/1/1
10.255.0.5   1    FULL/DR       10.255.0.5  1/1/1
```

EXCHANGE -> LOADING -> FULL
We now have identical LSAs with DR/BDR

Aruba Training-Confidential

The DR and BDR form a Full adjacency with all other routers. Or you could say that all DROTHER routers from a full adjacency with the DR and BDR. DROTHER routes only reach the 2WAY state between each other.

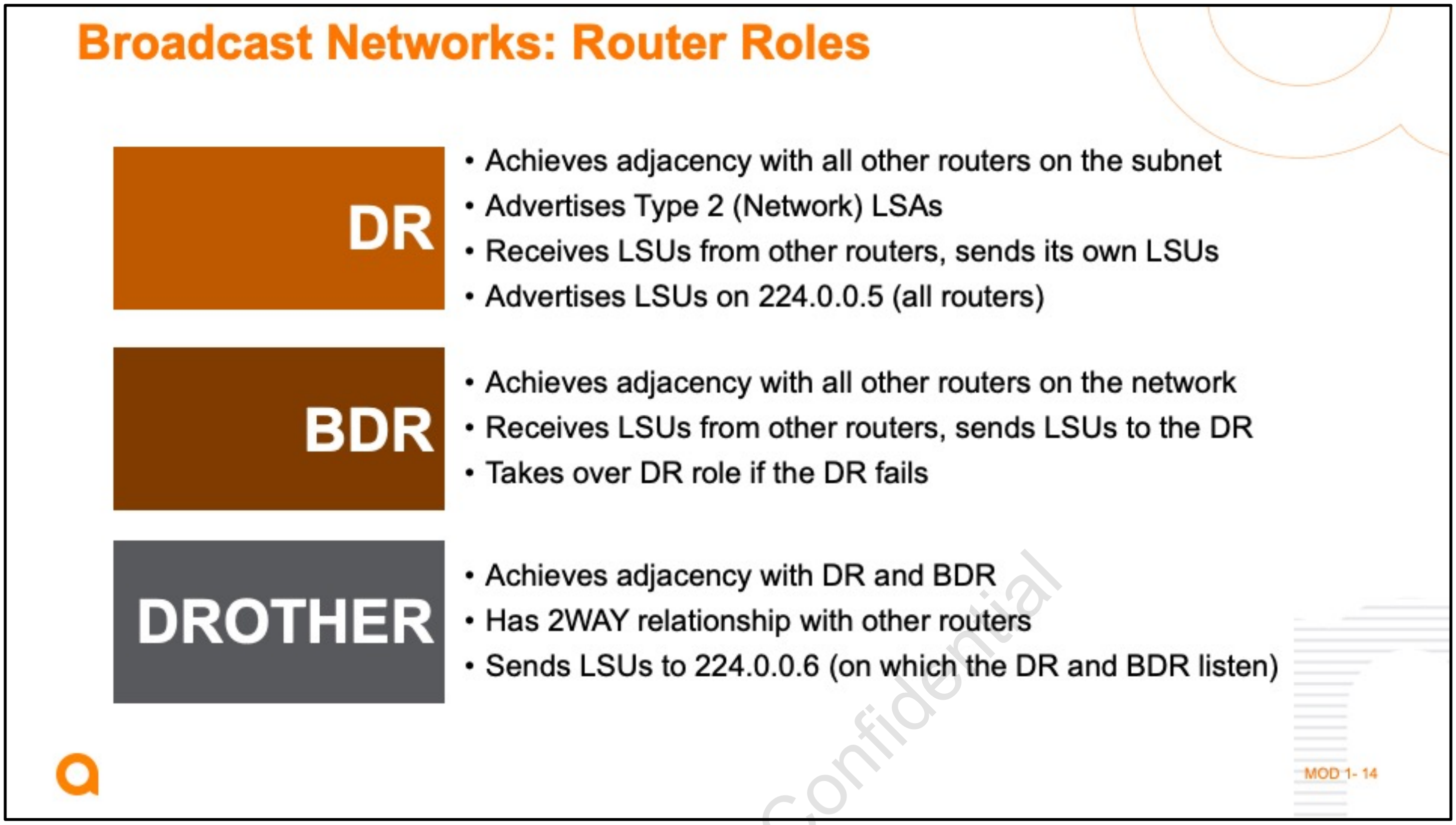
To achieve full adjacency, OSPF neighbors send unicasts to describe their databases to each other and exchange LSAs until their link state databases (LSDBs) match. In this way, information about the complete topology spreads from neighbor to neighbor throughout the network.

OSPF routers must establish full adjacency with:

The single neighbor on a point-to-point interface

The DR and BDR of a broadcast or non-broadcast multi-access (NBMA) network

Periodic hellos (every 10 seconds, by default) maintain the 2WAY or FULL state. If the OSPF router misses four hellos from a neighbor, it moves the neighbor to down. You will learn about how OSPF routers detect that neighbors are down in more detail later.



Broadcast Networks: Router Roles

DR

- Achieves adjacency with all other routers on the subnet
- Advertises Type 2 (Network) LSAs
- Receives LSUs from other routers, sends its own LSUs
- Advertises LSUs on 224.0.0.5 (all routers)

BDR

- Achieves adjacency with all other routers on the network
- Receives LSUs from other routers, sends LSUs to the DR
- Takes over DR role if the DR fails

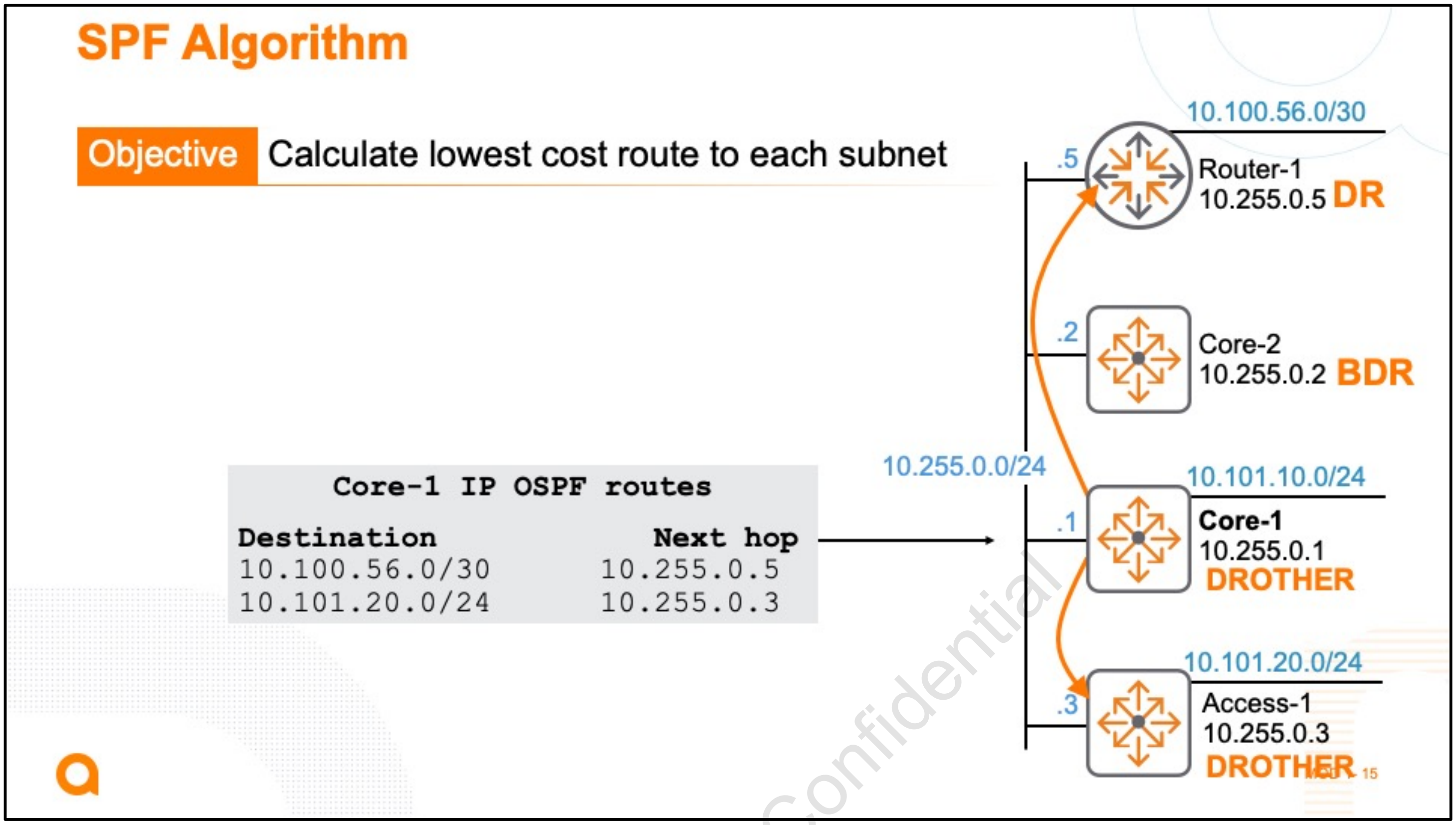
DROTHER

- Achieves adjacency with DR and BDR
- Has 2WAY relationship with other routers
- Sends LSUs to 224.0.0.6 (on which the DR and BDR listen)

MOD 1- 14

The figure summarizes the router roles on a broadcast network. Take a moment to study these roles.

Aruba Training-Confidential



SPF Algorithm

Objective Calculate lowest cost route to each subnet

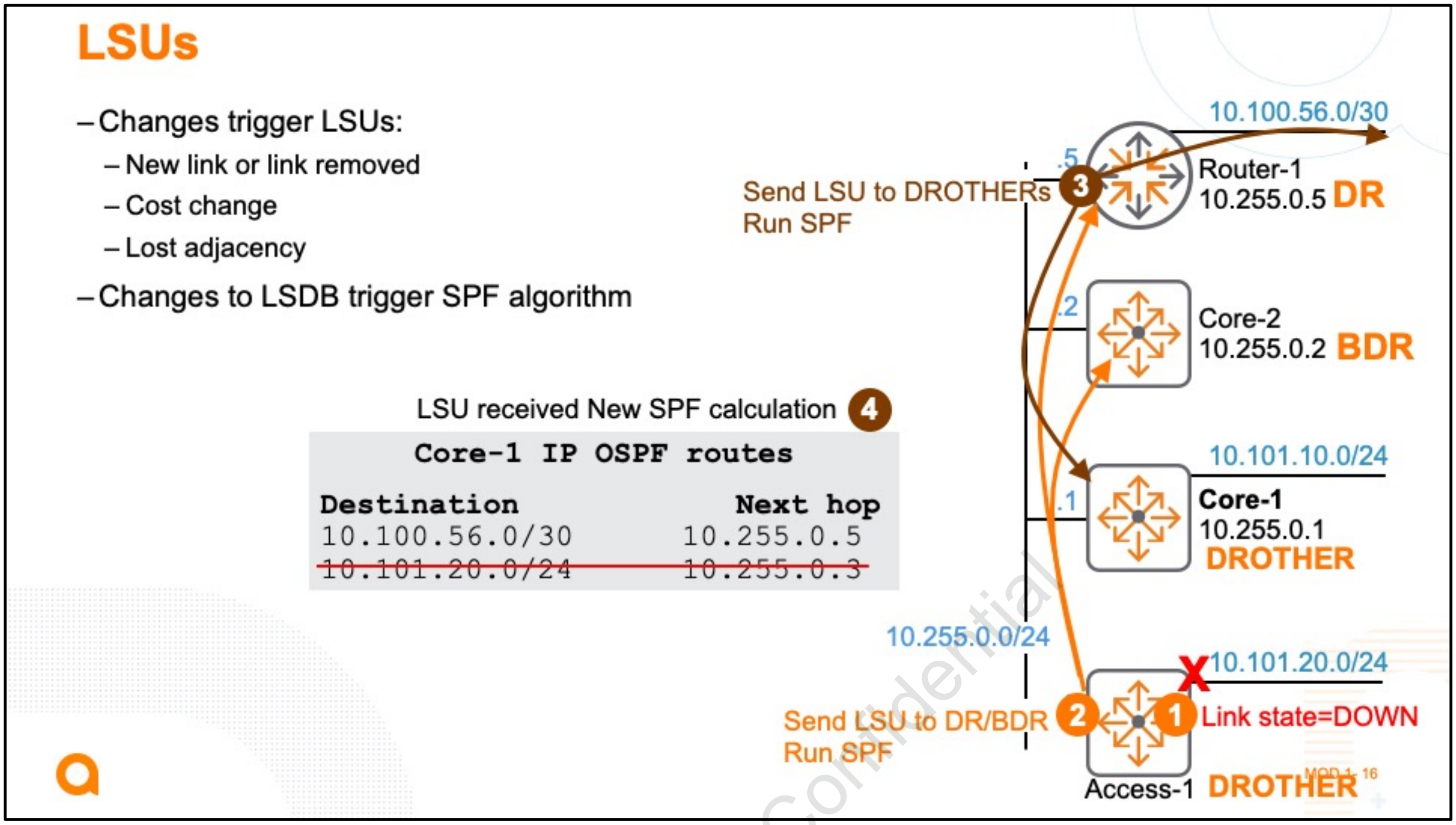
Core-1 IP OSPF routes

Destination	Next hop
10.100.56.0/30	10.255.0.5
10.101.20.0/24	10.255.0.3

Aruba Training-Confidential

Within an OSPF area, the Type 1 and Type 2 LSAs give each router enough information to create a topology (sometimes called a directed graph) a bit like the one in the figure. That is, they know which routers connect to other routers by the links that they share, and they know which routers offer gateways to which stub networks, which are networks advertised by only that router.

Each router then calculates the best route to each network—each advertised link—based on this topology. Specifically, they run the shortest path first (SPF), or Dijkstra algorithm, which determines the lowest cost path between the router and the network. In the figure, Core-1 has learned that it has a route to 10.100.56.0/30 through next hop 10.255.0.5 This is the actual Router-1 IP address on the transit network, not the router ID, although they are the same in this instance.



LSUs

- Changes trigger LSUs:
  - New link or link removed
  - Cost change
  - Lost adjacency
- Changes to LSDB trigger SPF algorithm

LSU received New SPF calculation 4

Core-1 IP OSPF routes

Destination	Next hop
10.100.56.0/30	10.255.0.5
~~10.101.20.0/24~~	~~10.255.0.3~~

Aruba Training-Confidential

LSAs have an age, and they age out of the database when the LSA reaches the max age (standard 1800 seconds). Routers periodically re-advertise the LSAs that they have generated to reset the age and maintain them. If a router needs to withdraw one of its LSAs, it can flush it from the system by re-advertising the LSA with an age set to the max age.

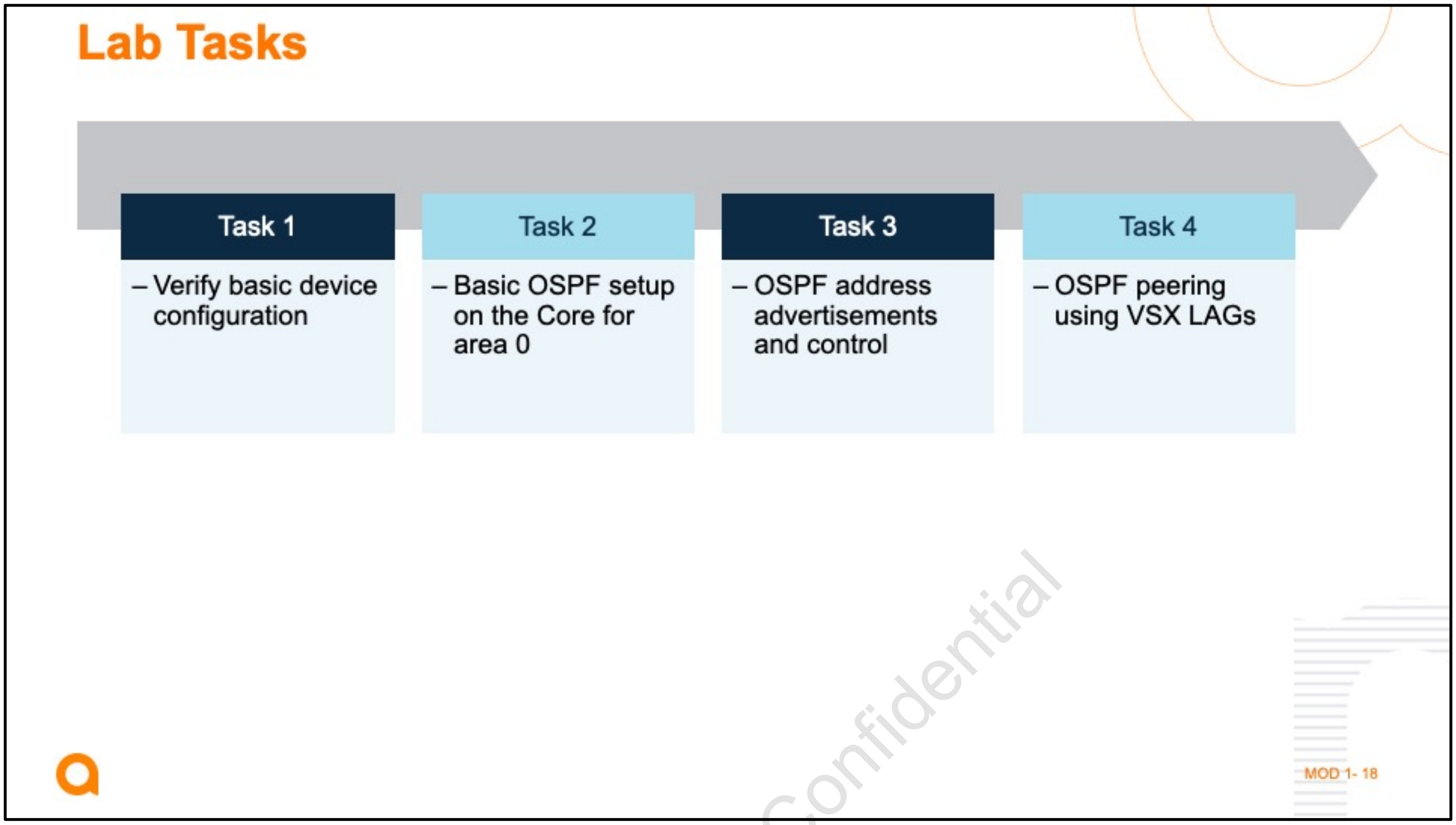
But an LSA aging out is not the typical way that routers learn about changes to the topology; this would be far too slow. Instead, if a router needs to change information in one of its LSA, it sends a Link State Update (LSU) and waits for an acknowledgement. Suppose Access-1's 10.101.20.0/24 subnet changes state – from UP to DOWN. It sends this information in a Type 1 (router) LSU, and the runs the SPF algorithm to calculate its new best routes and then adjust its route table accordingly.

The DR receives this update, floods the LSU out all other interfaces, and to DROTHERs on the 10.255.0.0/24 subnet. DROTHERs like Core-1 receive this update, flood it out any other links they may be connected to, and runs SPF to calculate their new best paths. In the figure, Core-1 has removed the 10.101.20.0/24 subnet from its route table. In a more complex and redundant topology—such as the ones you will explore later—the router might find a new next hop for the route.

Lab Activity

Lab 6.1

Aruba Training-C



Lab Tasks

Task 1

– Verify basic device configuration

Task 2

– Basic OSPF setup on the Core for area 0

Task 3

– OSPF address advertisements and control

Task 4

– OSPF peering using VSX LAGs

MOD 1- 18

The figure provides a brief review of lab tasks. Please see your lab guide for details.

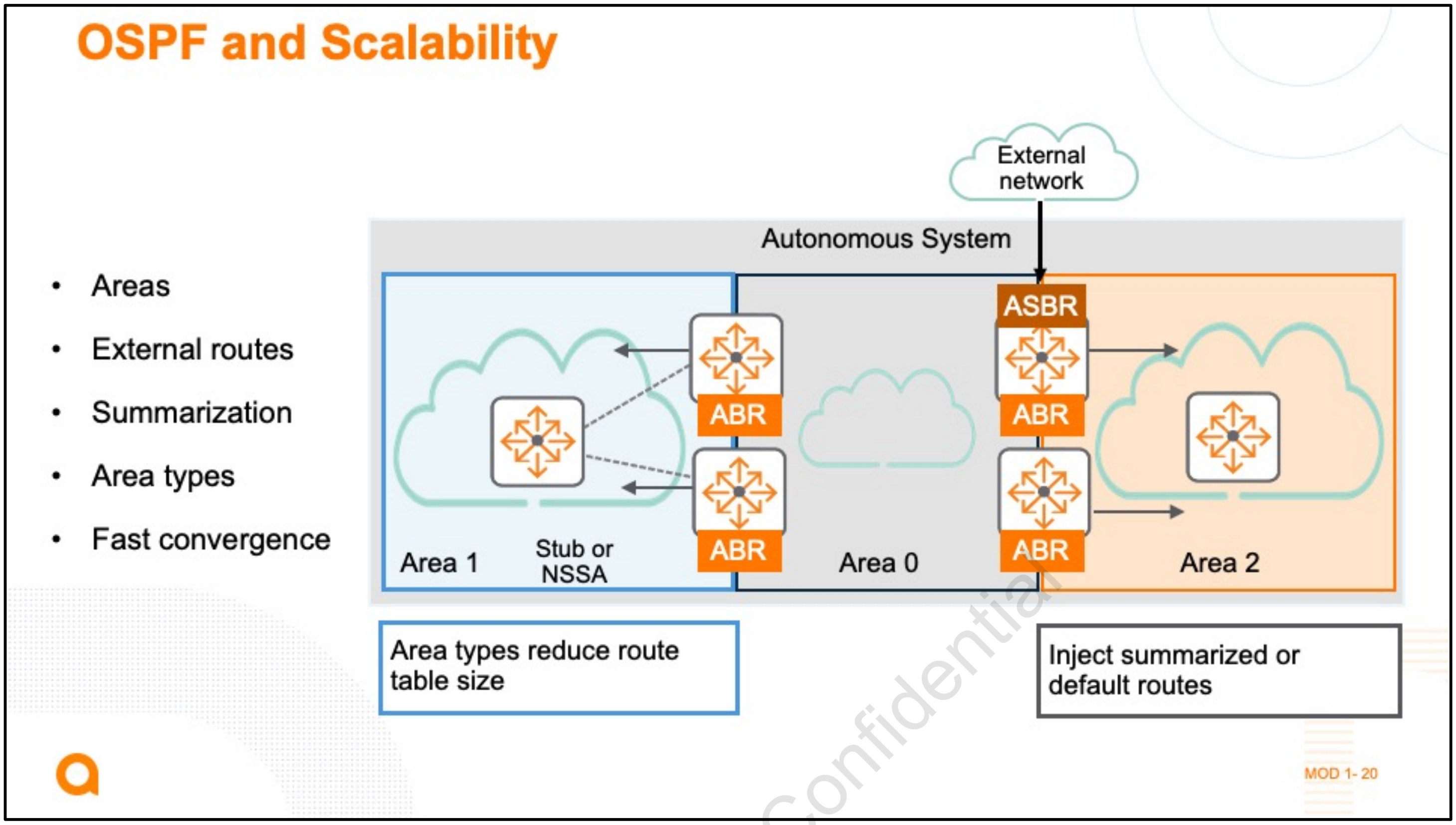
Aruba Training-Confidential

Multi-Area OSPF

MOD 1- 19

You will now learn about dividing OSPF autonomous systems into multiple areas. In addition to learning how to establish multiple areas, you will learn why you would want to do so. You will explore the differences between intra-areas OSPF routes and inter-area OSPF routes, as well as the LSAs that help to propagate routing information across the area boundaries.

Aruba Training-Confidential



OSPF and Scalability

- Areas
- External routes
- Summarization
- Area types
- Fast convergence

Area types reduce route table size

Inject summarized or default routes

MOD 1- 20

Most of the rest of this module will discuss the scalability features in OSPF, including the following:

Areas and their components, including Area Border Routers (ABRs) and the type of LSAs they generate between areas (LSA Type 3)

External routes and their components, including Autonomous System Boundary Routers (ASBRs) and the type of LSAs that are generated (LSA Type 4 and 5)

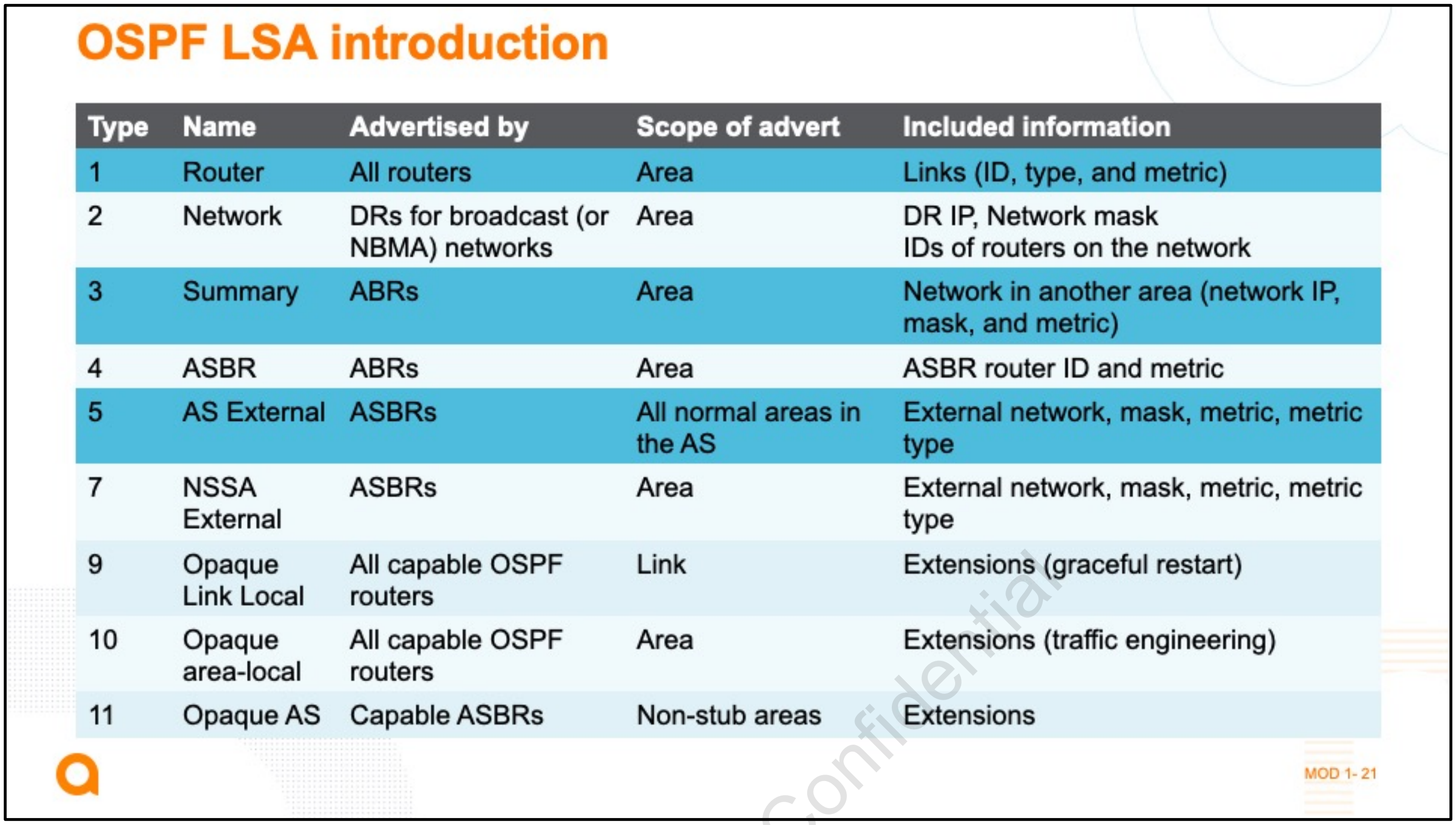
Summarization, including how ABRs can summarize and filter routes to reduce the size of the link state database in non-backbone areas

Area types, including stubby, totally stubby, not-so-stubby (NSSA) and totally NSSA areas to reduce the number of routes in a non-backbone area

Fast convergence, including features like Bi-Directional Forwarding Detection (BFD), VSX operations with OSPF, and OSPF graceful restart

This section will discuss implementing multiple areas in OSPF.

Aruba Training-Confidential



OSPF LSA introduction

Type	Name	Advertised by	Scope of advert	Included information
1	Router	All routers	Area	Links (ID, type, and metric)
2	Network	DRs for broadcast (or NBMA) networks	Area	DR IP, Network mask IDs of routers on the network
3	Summary	ABRs	Area	Network in another area (network IP, mask, and metric)
4	ASBR	ABRs	Area	ASBR router ID and metric
5	AS External	ASBRs	All normal areas in the AS	External network, mask, metric, metric type
7	NSSA External	ASBRs	Area	External network, mask, metric, metric type
9	Opaque Link Local	All capable OSPF routers	Link	Extensions (graceful restart)
10	Opaque area-local	All capable OSPF routers	Area	Extensions (traffic engineering)
11	Opaque AS	Capable ASBRs	Non-stub areas	Extensions

MOD 1- 21

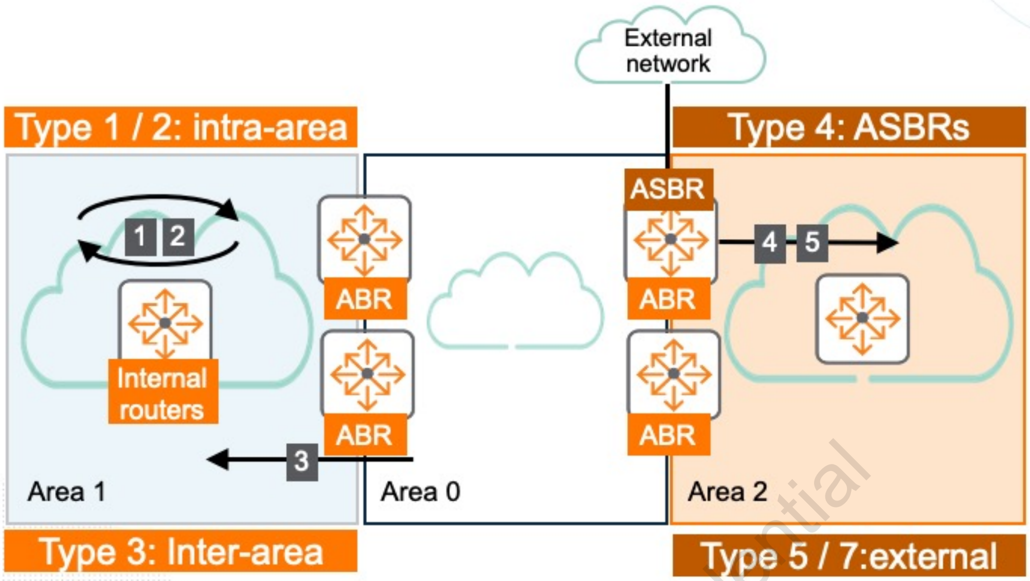
Aruba Training-Confidential

The chart summarizes key information for various LSA types. This table does not show Type 6 LSAs because they advertise multicast groups for Multicast OSPF (MOSPF), a protocol that routers rarely use and that AOS-CX switches do not support.

Recall that Type 1 LSAs are router advertisements – the router advertises information about itself. Type 2 LSAs are Network advertisement – routers advertise their connected subnets. You will learn about LSA types 3 to 7 in the rest of this module.

Other LSA types include extended or opaque LSAs, which enable features such as OSPF graceful restart and traffic engineering. Routers only advertise opaque LSAs to neighbors that indicate that they are opaque capable during adjacency establishment. You will learn about graceful restart later in this module. Other opaque LSAs are beyond the scope of this course. Not included in the table is the Type 8 LSA, which has special functions in OSPFv3, used for IPv6 routing. OSPFv3 is also beyond the scope of this course.

Types and Usage

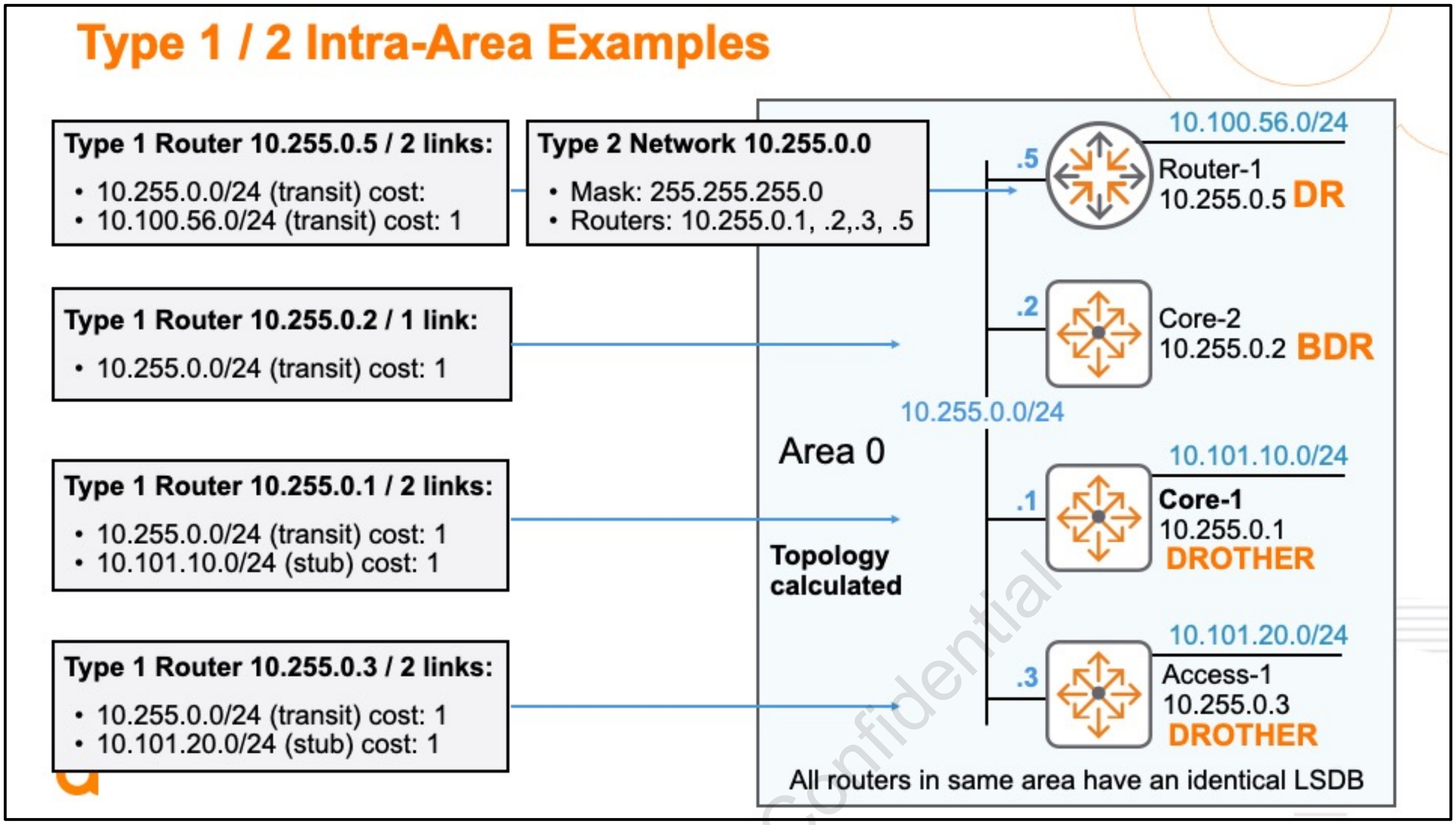


MOD 1- 22

Aruba Training-Confidential

The figure illustrates where you'll find the common LSA types and who advertises them.

You see Type 1 and 2 LSAs inside an area. ABRs inject Type 3 summary LSAs into an area. ASBRs send type 5 external LSAs, and sometimes Type 7 LSAs, while Type 4 LSAs are information about those ASBRs.



Type 1 / 2 Intra-Area Examples

Type 1 Router 10.255.0.5 / 2 links:

- 10.255.0.0/24 (transit) cost:
- 10.100.56.0/24 (transit) cost: 1

Type 2 Network 10.255.0.0

- Mask: 255.255.255.0
- Routers: 10.255.0.1, .2,.3, .5

Type 1 Router 10.255.0.2 / 1 link:

- 10.255.0.0/24 (transit) cost: 1

Type 1 Router 10.255.0.1 / 2 links:

- 10.255.0.0/24 (transit) cost: 1
- 10.101.10.0/24 (stub) cost: 1

Type 1 Router 10.255.0.3 / 2 links:

- 10.255.0.0/24 (transit) cost: 1
- 10.101.20.0/24 (stub) cost: 1

All routers in same area have an identical LSDB

Aruba Training-Confidential

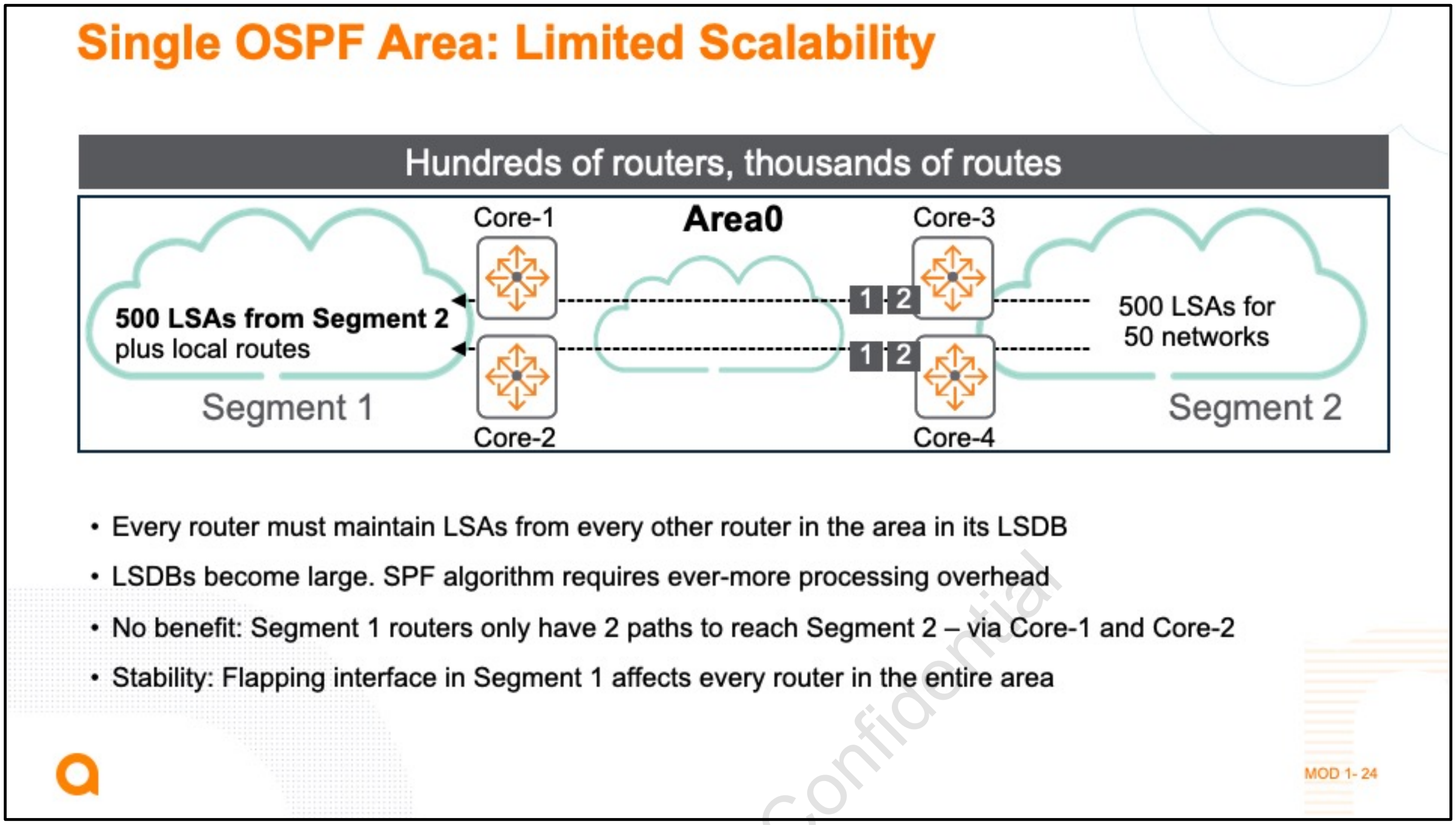
Let’s review what is in the LSDB and how routers use that information to create a topology of the OSPF area.

An OSPF router generates a Type 1 Router LSA, with information about its own OSPF networks, or links, and the cost of using those links. The figure gives a summarized view of just some of the critical information in the LSA; the real LSA has many fields.

The DR on a broadcast or NBMA transit network also generates a Network, or Type 2, LSA with information about all the routers on that network. This allows routers throughout the system to determine which routers are available on a network. A stub network, which is a network with just one OSPF router on it, does not require a Type 2 LSA. The Type 1 LSA of the single router on that network is enough to advertise the stub network’s location.

LSAs show the cost to reach a link - higher bandwidth links having a lower cost. However, you can set the cost as preferred. The router labels itself as the advertising router for these LSAs, using its router ID.

The designated router (DR) in the transit broadcast network has also sent out network LSAs for the network indicating which routers connect to it.



Single OSPF Area: Limited Scalability

- Every router must maintain LSAs from every other router in the area in its LSDB
- LSDBs become large. SPF algorithm requires ever-more processing overhead
- No benefit: Segment 1 routers only have 2 paths to reach Segment 2 – via Core-1 and Core-2
- Stability: Flapping interface in Segment 1 affects every router in the entire area

Aruba Training-Confidential

The figure shows a single area OSPF network with two segments. The segments might be a campus LAN and a data center, a campus LAN and a large branch office, buildings in a large three tier-campus topology, or any other segments within a larger network. Segment 1 and Segment 2 each represent many routers, which ultimately connect to two core routing switches. These core routers connect through a network backbone, which might also include more routers and routing switches. The figure only shows two segments for simplicity, but a large enterprise network might have several such segments.

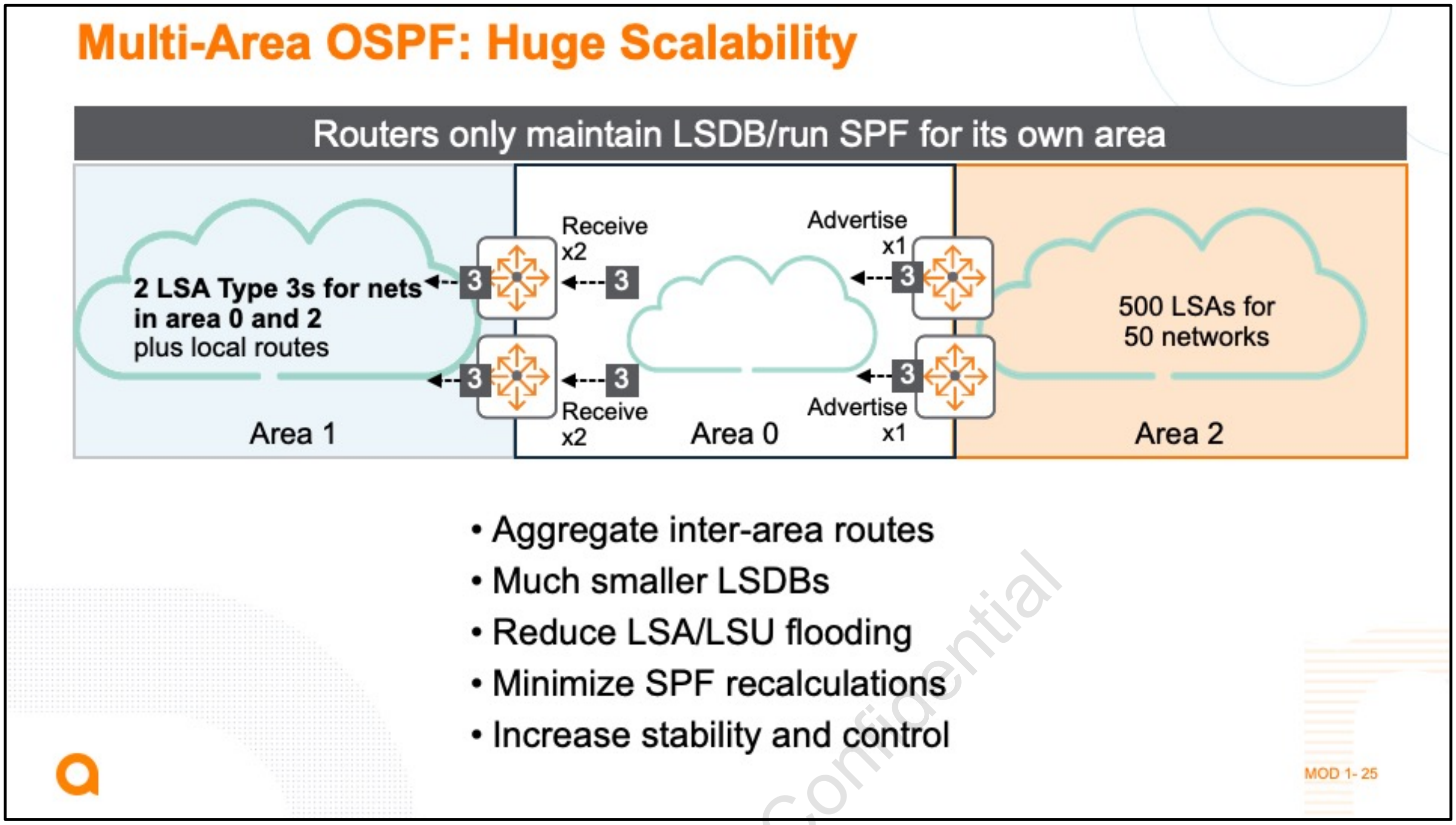
Every OSPF router in an area must maintain LSAs from every other router in the area in its LSDB. The top example shows how a single area implementation forces each of these Type 1 and 2 LSAs to propagate across the system. Plus, the figure only mentions LSAs associated with segment 2. Each OSPF router must also maintain LSAs for routers and networks in segment 1 and the backbone.

As a domain grows, the LSDB becomes excessively large and difficult to process. Routers must dedicate valuable processing resources to running the SPF algorithm. Often this resource consumption adds little value. As shown in the figure, networks often feature aggregation such that one segment connects to another through a few core devices. With only two exit points from segment 1 towards segment 2, the routing switches in segment 1 have no reason to maintain so much granular information about networks in segment 2.

In addition, suppose that two routing switches in segment 2 provide a connection to the subnet A. One of those switches has a flapping interface. It loses its connection to subnet A and sends an LSU withdrawing that subnet from its list of links. Every routing switch in the network backbone and segment 1 receives this LSU, adjusts the LSDB, and re-runs the Dijkstra algorithm on the topology. However, these routers still reach subnet A through the same core routing switches. When the flapping interface comes back up, the switch sends another LSU, which propagates throughout the entire network, causing every OSPF router to recalculate shortest paths, again without any significant change.

What is the solution?

Aruba Training-Confidential



Multi-Area OSPF: Huge Scalability

Routers only maintain LSDB/run SPF for its own area

- Aggregate inter-area routes
- Much smaller LSDBs
- Reduce LSA/LSU flooding
- Minimize SPF recalculations
- Increase stability and control

MOD 1- 25

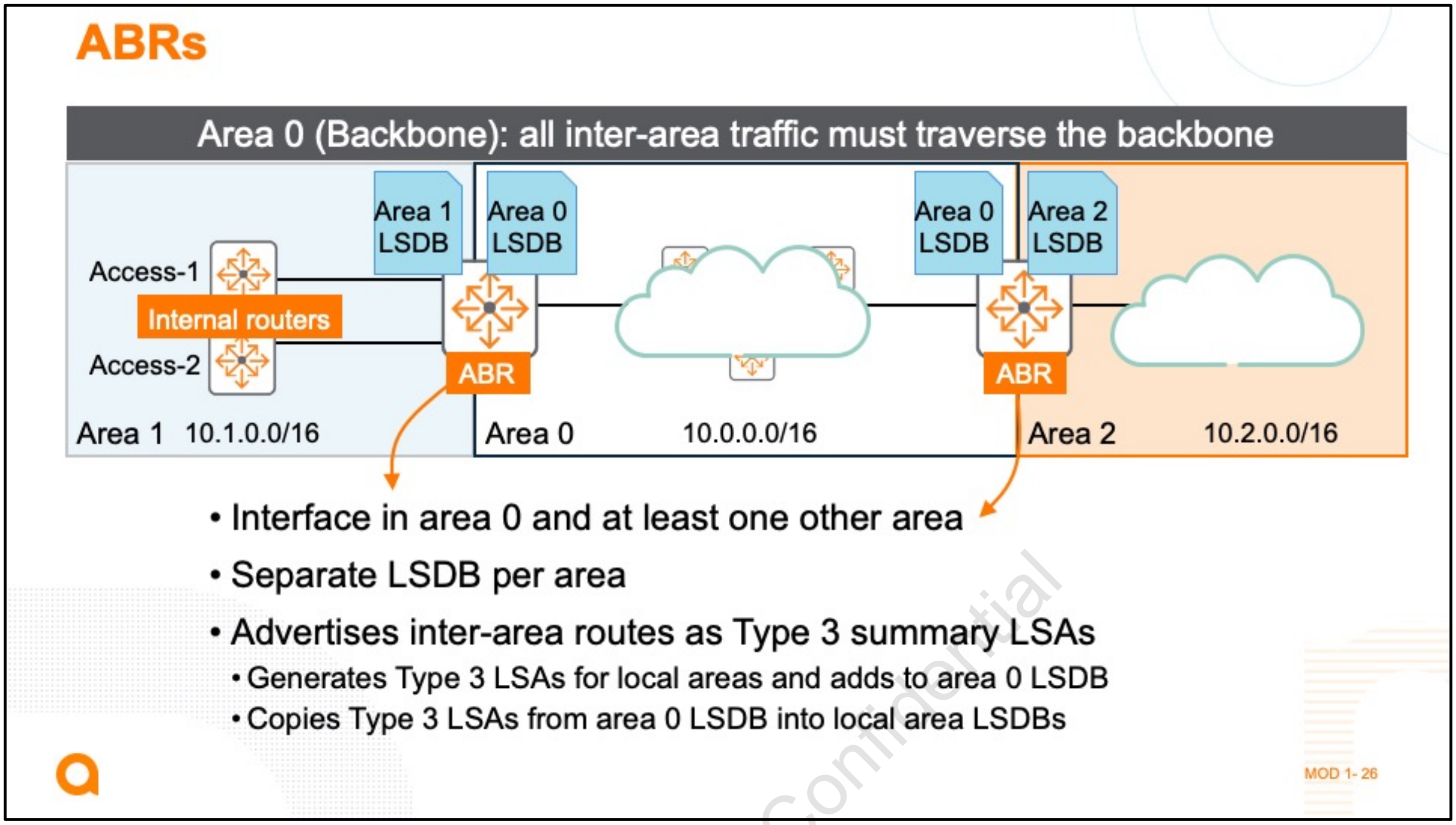
Aruba Training-Confidential

OSPF areas segment the domain into more stable and manageable pieces. Each OSPF routing device maintains the LSDB and runs the shortest path algorithm only for its own areas. OSPF ABRs forward routing information between areas. Type 1 and Type 2 LSAs do not flow across area boundaries. ABRs summarize this huge number of Type 1 and Type 2 LSAs into a few Type 3 summary LSAs.

This all serves to dramatically decrease LSDB size and improve stability - routers do not have to re-run the SPF algorithm as often, and when they do, it runs on smaller LSDB. Understand that ABRs automatically summaries LSAs – the advertisements about routes. They do not automatically summarize the routing tables themselves. If you had 100,000 routes in your routing table with a single area design, you will still have 100,000 entries after using areas.

However, now that you have ABRs you can configure route summarization, which can drastically reduce your routing table sizes. Routers in Area 1 may only have the few dozen or hundred routes in that area, plus a default route (0.0.0.0) to reach all other subnets via the ABRs. These benefits still apply even for powerful routers that could support very large areas.

This is especially true as relates to network stability. With ABRs and route summarization, you have effectively segmented your network into smaller failure domains. A flapping route in Area 1 has absolutely no effect on Area 0 or Area 2. This can be a big boost in network stability and ease of troubleshooting.



ABRs

Area 0 (Backbone): all inter-area traffic must traverse the backbone

- Interface in area 0 and at least one other area
- Separate LSDB per area
- Advertises inter-area routes as Type 3 summary LSAs
  - Generates Type 3 LSAs for local areas and adds to area 0 LSDB
  - Copies Type 3 LSAs from area 0 LSDB into local area LSDBs

MOD 1- 26

Aruba Training-Confidential

You learned how each OSPF area defines an isolated region beyond which Type 1 and Type 2 LSAs and shortest path topologies do not extend. ABRs receive these Type 1 and 2 LSAs from one area, summarize them into a Type 3 summary LSA, and inject them into another area.

Area 0 has a special role in multi-area networks – the backbone. All areas must connect directly to the Area 0 backbone, and all inter-area traffic must traverse the backbone. In the figure Area 1 reaches Area 2 via the backbone, and vice versa – always through an ABR. Each ABR must have at least one interface in area 0, and at least one interface in some other area. So, ABRs have interfaces in at least two areas, one of which is area 0. OSPF design requires all areas to connect to this backbone. This is nearly always a physical link, but there is a logical "virtual link" that is sometimes used.

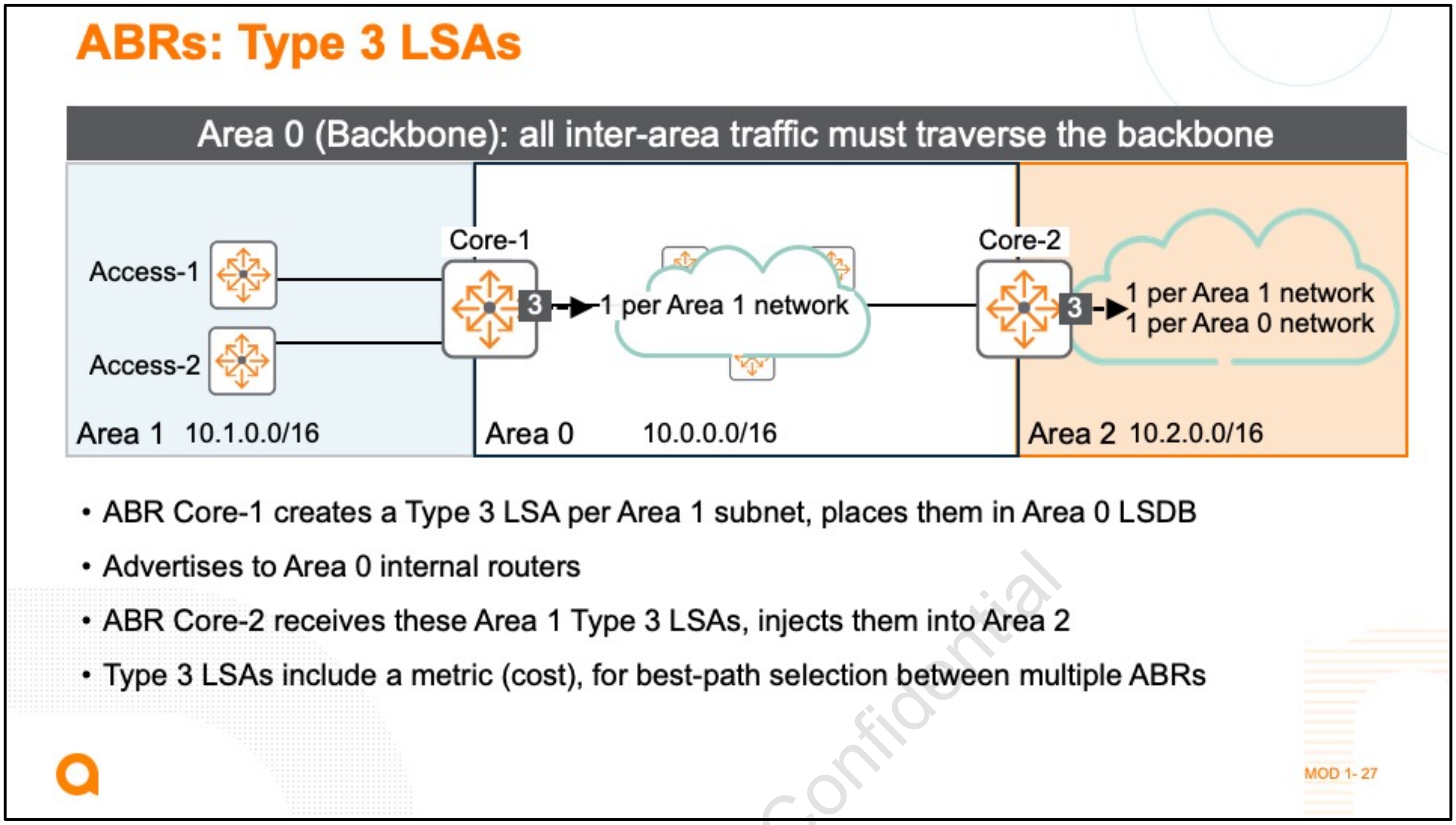
Each ABR maintains one LSDB per connected area. Think about it. All those Type 1 and 2 LSAs are synchronized among all routers in Area 1 – all Area 1 routers have an identical LSDB. This is true for all areas. These Type 1/2 routes are summarized into Type 3 LSAs, and injected into the other areas. So, Area 2 does not have any topology information for Area 1 (connectivity detail) but area 2 knows about all Area 1 networks. It simply knows that it can reach these subnets via any connected ABR.

In this example, Core-1 is an ABR. It synchronizes Type 1 and Type 2 LSAs with other OSPF routing devices in area 1. The figure shows just two other routing devices, Access-1 and Access-2, for simplicity. In the real world, the area could have dozens or hundreds of routers. Access-1 and Access-2 reside entirely within area 1 and are called internal routers. That connected ABR, Access-1, and Access-2 have identical area 1 LSDBs.

The ABR follows the same process in area 0, exchanging Type 1 and Type 2 LSAs and synchronizing its area 0 LSDB with routers such as Router-1 and Core-2.

In each area, the ABR also creates Type 3, or summary LSAs, which advertise routes that it knows in other areas. The sections below elaborate on how the ABR conducts this advertisement.

Aruba Training-Confidential



ABRs: Type 3 LSAs

- ABR Core-1 creates a Type 3 LSA per Area 1 subnet, places them in Area 0 LSDB
- Advertises to Area 0 internal routers
- ABR Core-2 receives these Area 1 Type 3 LSAs, injects them into Area 2
- Type 3 LSAs include a metric (cost), for best-path selection between multiple ABRs

MOD 1- 27

Aruba Training-Confidential

The Core-1 ABR creates a Type 3, or summary LSA, for each network in area 1. It places these Type 3 LSAs in its LSDB for area 0. The ABR can then advertise its routes to area 1 networks to other routers in area 0.

For example, Core-1 learned of an inter-area route to 10.1.20.0/24 from Access-1’s Type 1 LSA, contained in its area 1 LSDB. From that LSDB, Core-1 creates an intra-area route to 10.1.20.0/24 . Core-1 then generates a Type 3 LSA for 10.1.20.0/24 and places it in its area 0 LSDB. Other area 0 routers, such as Core-2, receive Core-1’s Type 3 LSA and add it to their area 0 LSDB. They can then calculate an inter-area route to 10.1.20.0/24. Core 2 also injects these Type 3 LSAs from Area 1 into Area 0.

In addition to the network address, the Type 3 LSA includes a metric (cost), which enables routers to select the best inter-area route when multiple ABRs advertise a Type 3 LSA with the same network ID.

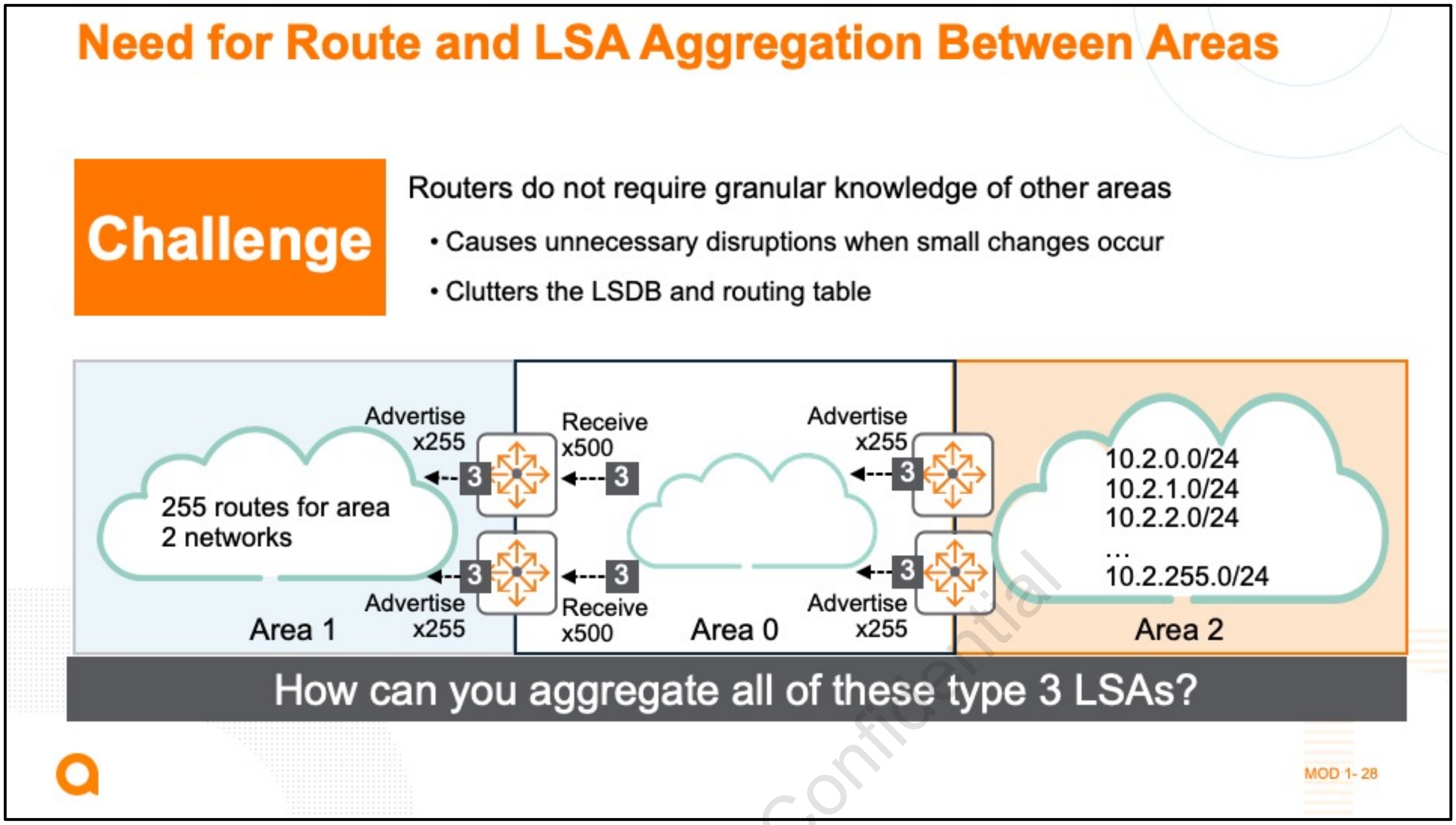
Of course, the same things happens in the other direction Core-1 learns creates Type 3 LSAs for the networks in Area 0 and injects them into Area 1.

And Core-2 has generated Type 3 LSAs for each network in area 2 and advertised them in area 0. Core-1 calculates an inter-area routes to each area 2 network, generates Type 3 LSAs for them, and injects them into area 0. They arrive at Core-1, which injects them into Area 1.

If Core-1 connected to another area, such as area 3, it would place Type 3 LSAs for each internal area 3 route in the area 0 LSDB and the area 1 LSDB. In other words, for each area, an ABR creates Type 3 LSAs with routes to other areas to which the ABR connects. Remember, though, that the inter-area routes always traverse area 0 even if the same ABR supports multiple non-backbone areas. That is, you cannot configure an OSPF router with area 1 and area 3 and expect the router to advertise area 3 networks in area 1 and area 1 networks in area 3. The router must have interfaces in area1, area 3, and area 0.

The figure shows how internal routers Access-1 and Access-2 receive the Type 3 LSAs for area 0 and area 2. Note that, although area 1 routers learn routes to area 2, the Type 3 LSAs themselves do not cross area boundaries. Instead Core-2, connected to area 2 and area 0, generates Type 3 LSAs for area 2 in area 0. It is the advertising router for these LSAs. Core-1 receives these LSAs in area 0 and generates new Type 3 LSAs for advertising these routes into area 1, and it is the advertising router for these new LSAs.

Aruba Training Confidential



Need for Route and LSA Aggregation Between Areas

Challenge

Routers do not require granular knowledge of other areas

- Causes unnecessary disruptions when small changes occur
- Clutters the LSDB and routing table

How can you aggregate all of these type 3 LSAs?

MOD 1- 28

Aruba Training-Confidential

Simply dividing an AS into multiple areas begins to decrease the size of OSPF routers' LSDBs and to reduce the complexity of SPF calculations. However, to achieve the true benefits of multiple areas, you should further reduce the number of LSAs advertised between areas.

In this example network, both area 1 and area 2 have only two ABR exit points toward the backbone and other areas. Internal area 1 routers and backbone routers might benefit from some level of detail in the information that they receive about area 2. For example, each ABR's Type 3 LSAs reflect the metrics of reaching the network in question, enabling routers in other areas to select the best ABR for reaching each network in area 2.

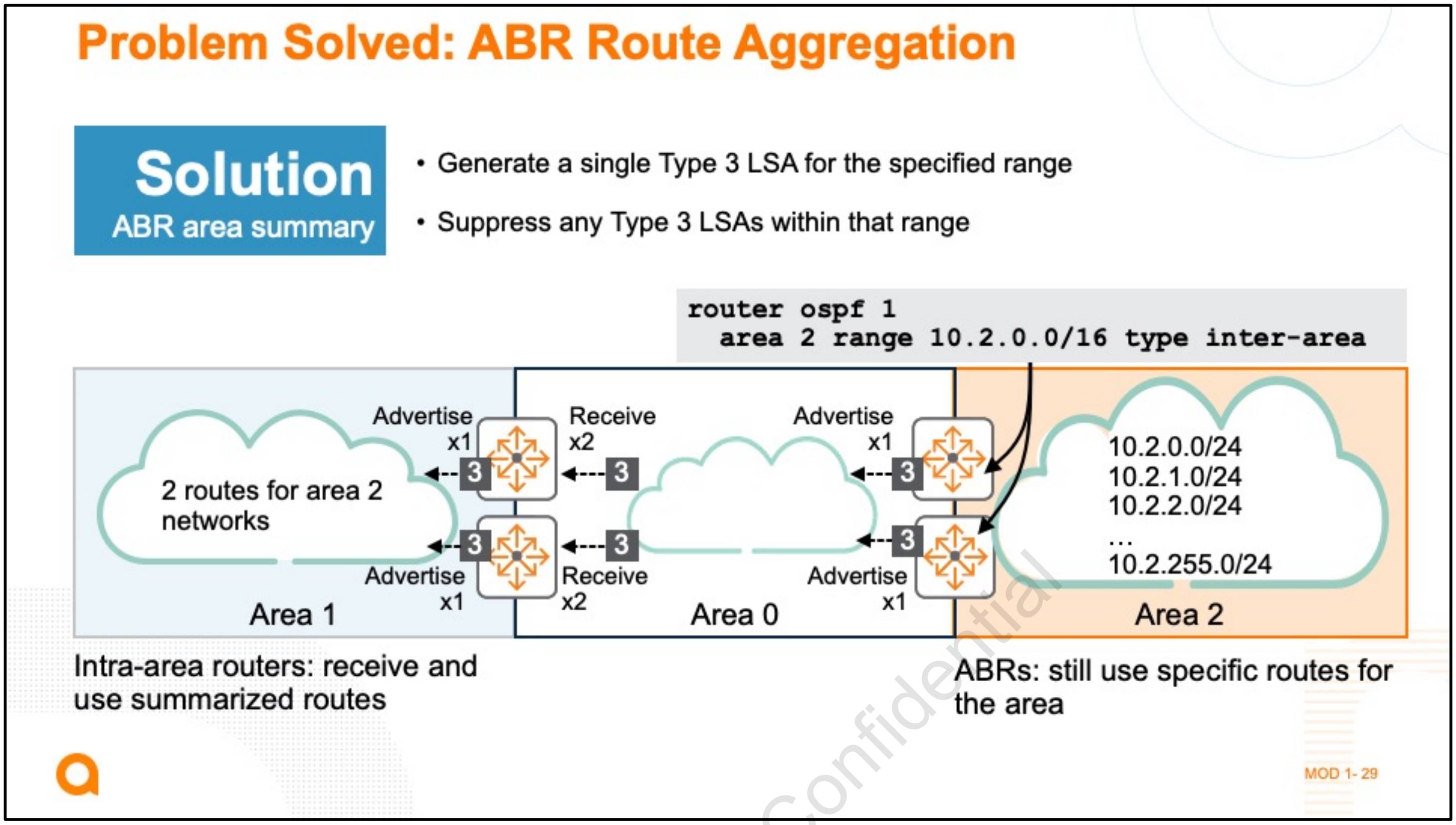
However, with only two entry points into area 2, routers (and, in particular, internal area 1 routers) are probably receiving too much granularity about each subnet. This information clutters the routers' LSDB and routing tables.

In addition, small changes in the area 2 topology could flood LSUs through the backbone and area 1, causing routers to re-run the SPF algorithm. For example, network 10.2.40.0/24 in area 2 becomes unavailable. The two area 2 ABRs withdraw their Type 3 LSA for that specific network in area 0, and then the two area 1 ABRs do the same in area 1. However, the area 1 routers have no other path to this network. It might be best to simply not inform routers beyond area 2 of the small change. If they continue to forward traffic for this network, the area 2 ABRs simply drop it.

You should aggregate these Type 3 LSAs to hide the unnecessary information and to smooth over small changes, thereby gaining the true benefits of a multiple area OSPF configuration.

Note: Although this example focuses on aggregating advertisements for area 2 networks, you would aggregate summaries of area 1 or area 0 networks for similar reasons.

Aruba Training-Confidential



Problem Solved: ABR Route Aggregation

Solution
ABR area summary

- Generate a single Type 3 LSA for the specified range
- Suppress any Type 3 LSAs within that range

```
router ospf 1
  area 2 range 10.2.0.0/16 type inter-area
```

Intra-area routers: receive and use summarized routes

ABRs: still use specific routes for the area

Aruba Training-Confidential

You can configure an ABR to aggregate its summary routes. Rather than generate a Type 3 LSA for every network in an area, it generates a few Type 3 LSAs, each of which includes many networks in its range. In other words, the Type 3 LSA advertises a network with a shorter prefix length such as a /20 or a /16 subnet that includes many /24 subnets.

You configure area summaries on the ABR that connects to the area to be aggregated. In the figure, you configure area summaries for area 2 on the two ABRs connected to Area 2. These ABRs then advertise the summaries to other backbone routers and ABRs in area 0. The other ABRs can then copy them into their own areas. In this example, the aggregated summary 10.2.0.0/16 shows up as a route on ABR A and on internal routing switches in area 1.

Similarly, you could configure the ABRs connected to Area 1 to summarize Area 1 routes and inject them into Area 0. All ABRs connect to Area 0, so you can configure them all to summarize area 0 routes and inject them into other area.

You see the drastic reduction in routes. Before summarization, Area 1 had 255 entries for Area 2. Now it has 2 entries.

The figure shows the syntax to configure this aggregation.

When you enter the area x range command, the ABR automatically suppresses the generation of individual Type 3 LSAs for subnets within the specified network. However, it continues to advertise individual Type 3 LSAs for networks that do not fall into the range. Thus, it is best to assign contiguous IP addresses within an area – it ensures easy route aggregation.

The ABRs for area 1 now receive one Type 3 LSA with the aggregated summary from each of the two ABRs that connect to area 2. Each area 1 ABR selects the best LSA for the summary range and creates an inter-area route to 10.2.0.0/16 through the best next-hop.

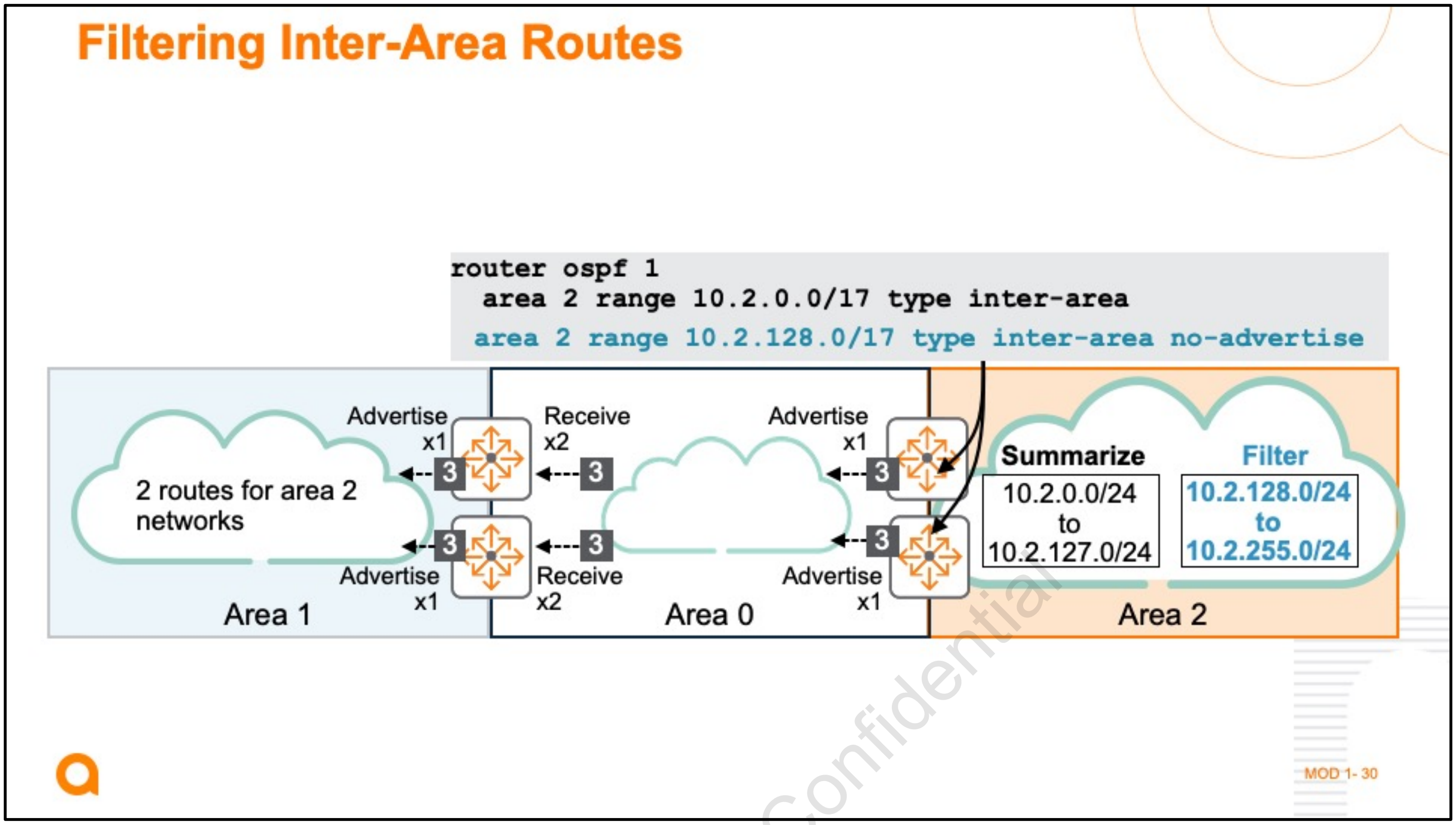
Each area 1 ABR then generates an area 1 Type 3 LSA to advertise this inter-area route to routers in area 1. The internal area 1 routers now have two LSAs for all area 2 networks, and they learn routes to only the one aggregated network (up to two routes if the ABRs advertise equal metrics and the internal routers support Equal-Cost Multi-Path [ECMP]).

Aruba Training-Confidential

Note: After you configure area summaries on an area's ABR, you can see the summaries in the LSDB and routing tables for routers not connected to that area.

You typically want to create as few area summaries as possible, but this is not always possible. Areas might have grown organically and without advance planning, creating non-contiguous ranges or ranges that do not fit neatly within a CIDR subnet. For example, area 2 might include networks in the 10.2.0.0-10.2.255.255 range as well as 10.3.0.0-10.3.23.255. Area 3 includes networks in the 10.3.0.0-10.3.19.255 range and the 10.3.30.0-10.3.39.255 range. You cannot use simply use 10.2.0.0/16 for the area 2 summary and 10.3.0.0/16 for the area 3 summary, due to the non-contiguous and overlapping ranges. Instead, you must use several area summaries for area 2 and several summaries for area 3.

Important: Examples throughout this module use the /16 prefix length for summary ranges to illustrate the concept in a simple way. When you create a range for your area, choose an appropriate range. For example, your area might include 16 /24 subnets, 10.1.0.0/24 to 10.1.15.0/24. This is equal to the 10.1.0.0/20 network, and you should specify that in the area <ID> range command.



Filtering Inter-Area Routes

```
router ospf 1
  area 2 range 10.2.0.0/17 type inter-area
  area 2 range 10.2.128.0/17 type inter-area no-advertise
```

MOD 1- 30

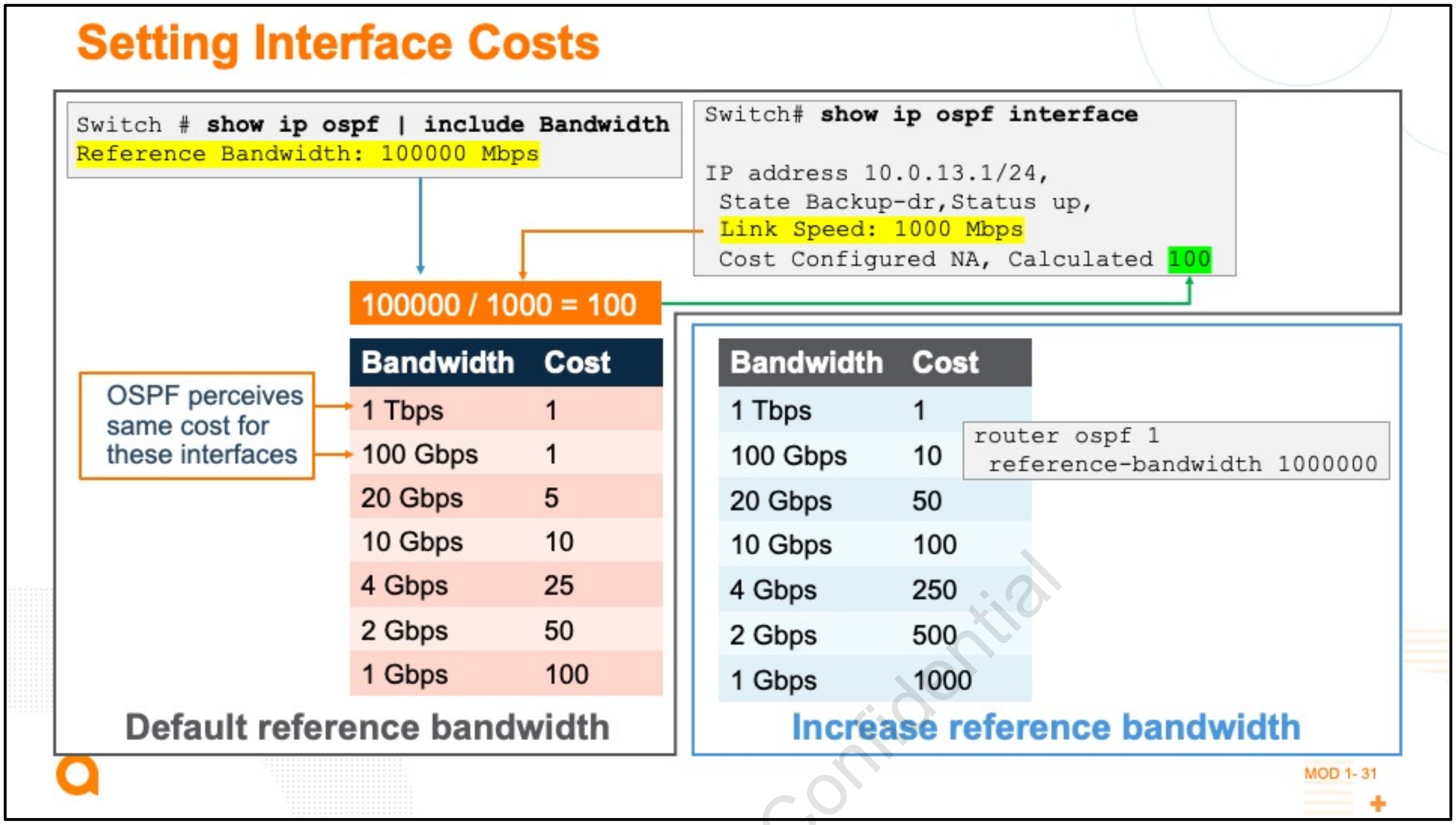
Remember that ABRs advertise all routes from one area to another, unless you configure an area summary for a specific range. You can also configure an area summary with the special no-advertise option, which prevents networks from being advertised outside of their local area.

Suppose your company has a strict security policy for one of the sites in OSPF area 2. The policies allow only local access to high-security subnets 10.2.128.0/24 through 10.2.255.0/24 (10.2.128/17). Devices in other areas should still be able to reach 10.2.0.0/24 through 10.2.127.0/24 (10.2.0.0/17). Of course, a firewall can filter out prohibited traffic, but, in this case, you want to prevent any routing of traffic from other areas to the restricted networks.

Aruba Training-Confidential

Therefore, you configure two area summaries on the area 2 ABRs, one for 10.2.0.0/17 and one for 10.2.128.0/17. ABRs advertise the first but not the second, as shown in the figure. If you later wanted to re-enable advertising the filtered subnets, you could simply re-enter the command, but without the no-advertise option. Simply removing the non-advertised summary would enable the advertisement of individual routes without aggregation.

Note that the filtering occurs at the area 2 boundary; you cannot apply the filtering to some areas and not others. For example, you cannot use this configuration to block inter-area routes to the area 2 subnets in area 1 but allow them in area 0.



Setting Interface Costs

```
Switch # show ip ospf | include Bandwidth
Reference Bandwidth: 100000 Mbps
```

```
Switch# show ip ospf interface

IP address 10.0.13.1/24,
 State Backup-dr,Status up,
 Link Speed: 1000 Mbps
 Cost Configured NA, Calculated 100
```

100000 / 1000 = 100

OSPF perceives same cost for these interfaces

Bandwidth	Cost
1 Tbps	1
100 Gbps	1
20 Gbps	5
10 Gbps	10
4 Gbps	25
2 Gbps	50
1 Gbps	100

Default reference bandwidth

Bandwidth	Cost
1 Tbps	1
100 Gbps	10
20 Gbps	50
10 Gbps	100
4 Gbps	250
2 Gbps	500
1 Gbps	1000

```
router ospf 1
 reference-bandwidth 1000000
```

Increase reference bandwidth

MOD 1-31

Aruba Training-Confidential

OSPF path selection is based on link cost, which is based on the reference bandwidth, which defaults to 100000. This is shown in the figure, with the output of the show ip ospf command. AOS-CX devices divide this reference bandwidth by interface link speed in Mbps to determine cost. So, the default reference of 100000 Mbps / 1000 Mbps (1 Gbps) = 100. This is proven in the output of show ip ospf interface.

One problem with this default is shown in the figure. The default reference of 100000 Mbps / 100000 Mbps (100 Gbps) = 1. OSPF cannot use fractions, and so all interfaces faster than 100 Gbps will also have a cost of 1. Thus, OSPF will perceive a 100 Gbps path and a 1 Tbps path as being equal cost. If you have no links faster than 100 Gbps, this is not an issue.

| If you DO have links faster than 100 Gbps, then you should change the reference bandwidth to a higher value. In the figure, you have changed the reference bandwidth to 1,000,000. Thus, 1 Tbps links have a cost of 1, while 100 Gbps links have a cost of 10, and so on - as shown in the figure. OSPF can now distinguish between these fast links.

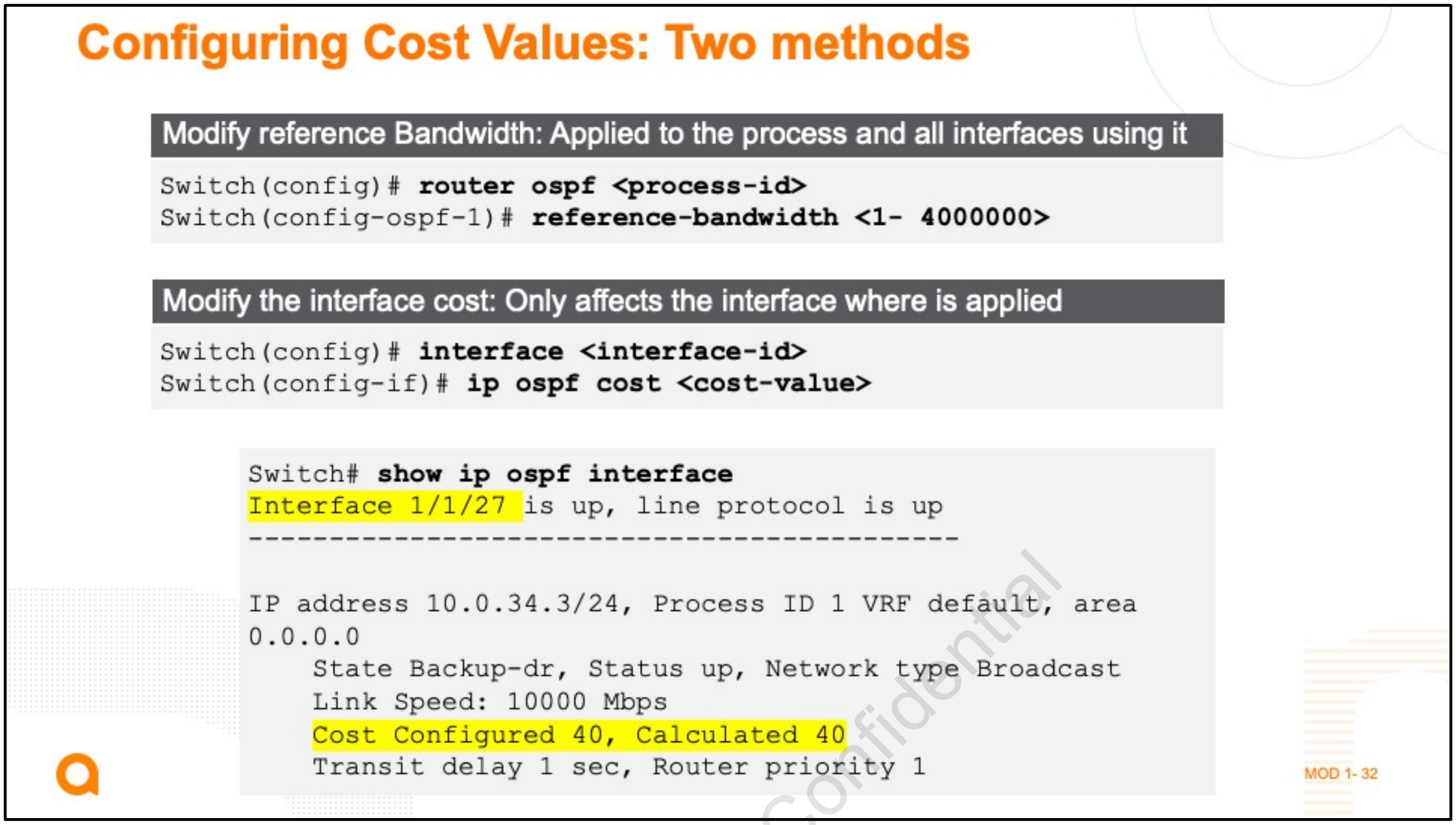
Another issue is that by default, AOS-CX VLAN interfaces have a link speed of 1000 Mbps (1 Gbps), and so have a cost of 100. Because every VLAN interface has the same cost, route selection becomes a matter of hop count, rather than true lowest cost path, if you are using layer 3 VLAN interfaces. If you modify the reference bandwidth to 1,000,000 as shown, then all VLAN interfaces would have a cost of 1000.

Changing the reference bandwidth does not mitigate the issue of VLAN interfaces having the same cost. If this is an issue, you can change the individual cost values on a per-interface basis:

- Interface vlan41
- ip ospf cost <1-65535>

- As you can see in the generic example above, you can set each individual interface to an OSPF cost value of your choice – any value between 1 and 65535 inclusive.

Aruba Training-Confidential



Configuring Cost Values: Two methods

Modify reference Bandwidth: Applied to the process and all interfaces using it

```
Switch(config)# router ospf <process-id>
Switch(config-ospf-1)# reference-bandwidth <1- 4000000>
```

Modify the interface cost: Only affects the interface where is applied

```
Switch(config)# interface <interface-id>
Switch(config-if)# ip ospf cost <cost-value>
```

```
Switch# show ip ospf interface
Interface 1/1/27 is up, line protocol is up
--------------------------------------------

IP address 10.0.34.3/24, Process ID 1 VRF default, area
0.0.0.0
    State Backup-dr, Status up, Network type Broadcast
    Link Speed: 10000 Mbps
    Cost Configured 40, Calculated 40
    Transit delay 1 sec, Router priority 1
```

MOD 1- 32

Aruba Training-Confidential

Preferably, you want to set the interface costs globally for the OSPF process by changing the bandwidth reference with the following configuration:

switch(config)# router ospf <process-id>

switch(config-ospf-1)# reference-bandwidth <Mbps>

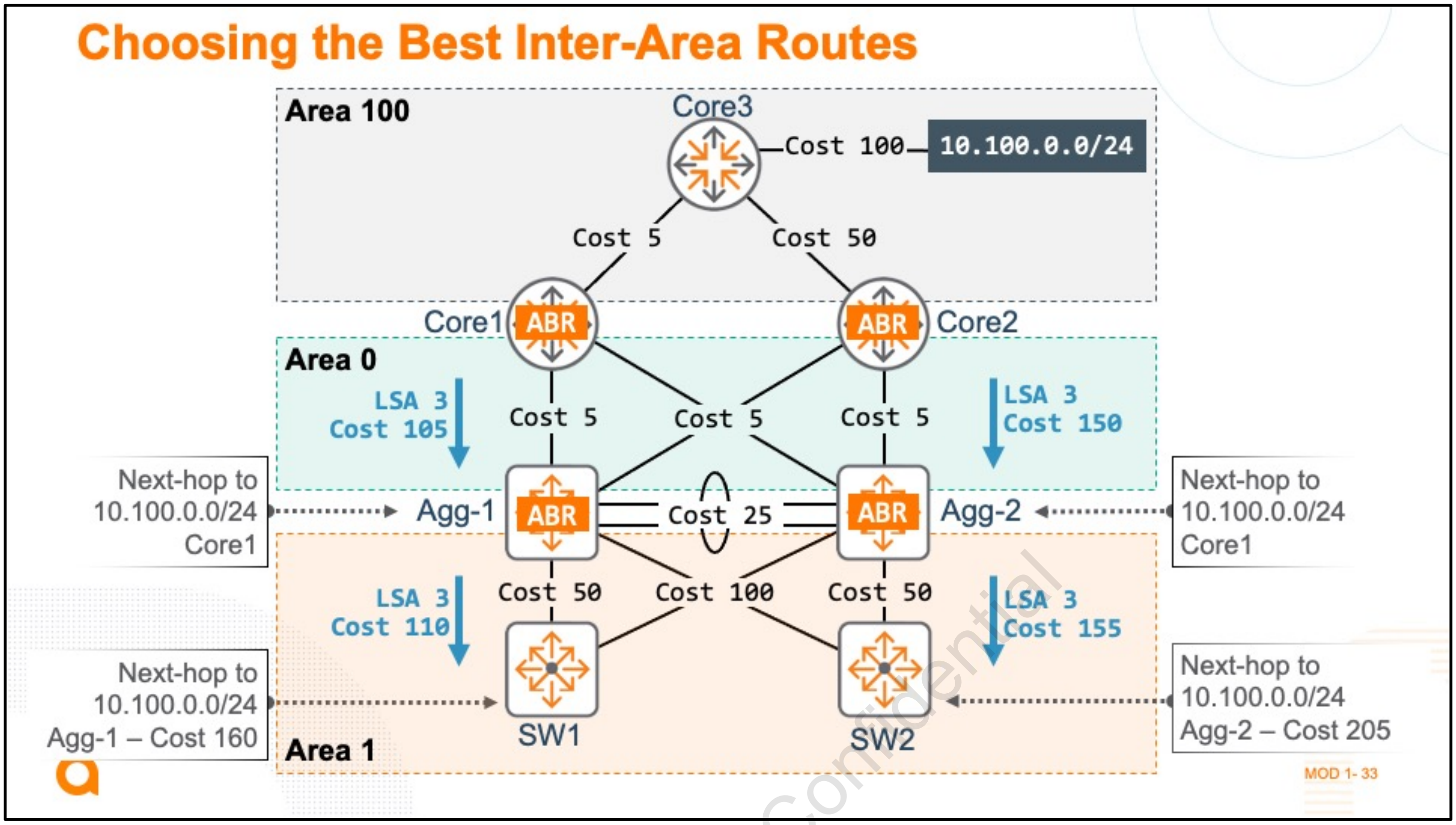
If the OSPFv2 interface cost is not explicitly set. then the cost of all the OSPFv2 interfaces is recalculated based on the reference bandwidth and link speed of the interface. For VLAN interfaces the link speed value is taken as 1 Gbps, if the OSPFv2 interface cost is not explicitly set. The bandwidth range is 1-4000000 Mbps, where the default is 100000 Mbps.

You can use this command to set the cost on a layer 3 interface, which will override the reference calculation:

Switch(config)# interface <interface-ID>

Switch(config-if)# ip ospf cost <cost>

When this command is applied the router no longer considers the cost formula – it simply uses the manual value that was entered.



Choosing the Best Inter-Area Routes

Aruba Training-Confidential

There may be times when you want multiple ABRs for an area, making it more important for you to understand inter-area route selection. Recall that the SPF algorithm only operates within areas - routers do not create a topology of the complete AS, then calculate inter-area routes from that. Instead, OSPF routers use the SPF algorithm to calculate a path to each ABR in its area. It considers the ABR as a gateway to the networks that the ABR advertises in Type 3 LSAs.

This discussion assumes that area 1 and area 100 are normal, stub, or NSSA areas, but not Totally stub or Totally stub NSSAs. The Totally stub areas do not receive Type 3 LSAs, except a default route, so the routers in those areas cannot choose best paths for inter-area traffic based on anything except the cost in the default route.

A Type 3 LSA advertises a route to the network. The routers that receive these advertisements do not create a topology of the networks in other areas. Instead they treat the ABR as a gateway to all the networks which the ABR advertises in Type 3 LSAs. That is, those networks are leaves hanging off the ABR in the shortest path tree.

Each router determines the shortest path to the ABR or ABRs advertising Type 3 LSAs to it. It uses Dijkstra's (SPF) algorithm to calculate this path as normal for inter-area routes – run for that router's area. The next hop for reaching the ABR becomes the next hop for all inter-area routes advertised by that ABR in Type 3 LSAs.

When the area includes multiple ABRs, internal routers learn multiple inter-area routes for that network. These internal routers use the metric to choose the best path. The metric for each inter-area route consists of:

Metric advertised in the Type 3 LSA, which the ABR copies from its metric for the target route

Metric for reaching the advertising ABR, from the area's shortest path topology

Because each ABR advertises the cost of its own route to the network, cost increments across the complete system. First consider a situation without route aggregation.

Core1 advertises 10.100.0.0/24 as a Type 3 LSA with metric 105 in area 0, and Core2 advertises this route with metric 150. Agg-1 adds 5 to each of these metrics because that is its cost for reaching each of the routers. It selects Core1 as offering the best route and adds an inter-area route with metric 110 to its OSPF table. When Agg-1 generates an LSA for 10.100.0.0/24 in area 1, it advertises metric 110. SW1 adds its metric for reaching Agg-1 to 110, assigning the inter-area route to 10.100.0.0/24 a metric of 160.

Inter-area routing in OSPF combines the best aspects of link-state and distance-vector routing. SW1, for example, uses link-state routing to determine its shortest path to its two ABRs (Agg-1 and Agg-2). It can then use those same two paths for all routes to other areas without considering the link states of networks in those areas. Even though SW1 only knows a small part of the topology, traffic that it routes toward other areas still receives the benefits of link-state routing across the complete path. For example, SW1 sends traffic to 10.100.0.0/24 on its shortest path to Agg-1. Agg-1, in turn, routes the traffic using the shortest path to Core1, which it has calculated from its detailed area 0 topology. Finally, Core1 routes the traffic using the best path calculated by the detailed area 100 topology. In each area, routers maintain just the local segment of the complete path for inter-area traffic.

Also consider how the area boundaries enhances stability. When an OSPF router receives an update for a Type 3 LSA, it does not run the complete

Aruba Training Confidential

shortest path algorithm. Instead it adjusts the networks behind each ABR, as indicated by the update, and adds a new inter-area route, removes an inter-area route, or selects a different inter-area route (due to a cost change), as appropriate. The router does not, on the other hand, need to re-calculate the path to the ABR, which is the only segment of the path with which it is concerned.

Some loss of information is introduced by route aggregation at the area boundaries; however, the simplicity is typically worth this loss. For example, in this topology, network administrators might configure an aggregated range of 10.100.0.0/16 for area 100 on Core1 and Core2. They can assign cost 105 to this range on Core1 but cost 150 to the range on Core2. Agg-1 can still choose the route through Core1 as the lowest cost route. In this example, Agg-1 and Agg-2 have the same cost to reach Core1, so they advertise the same cost for the 10.100.0.0/16 route into area 1. However, Agg-2, for example, had a better cost path to Core1, it would advertise a lower cost route, and internal routers in area 1 would direct their area 100 traffic to that ABR.

You could also use multiple area summaries when an area has multiple ABRs, and you want to prefer one ABR for one segment of the area but another ABR for another segment. This approach is more complicated, however. You must also take care to configuring matching summaries on each ABR (although the costs would not match). Otherwise, you can accidentally introduce routing loops.

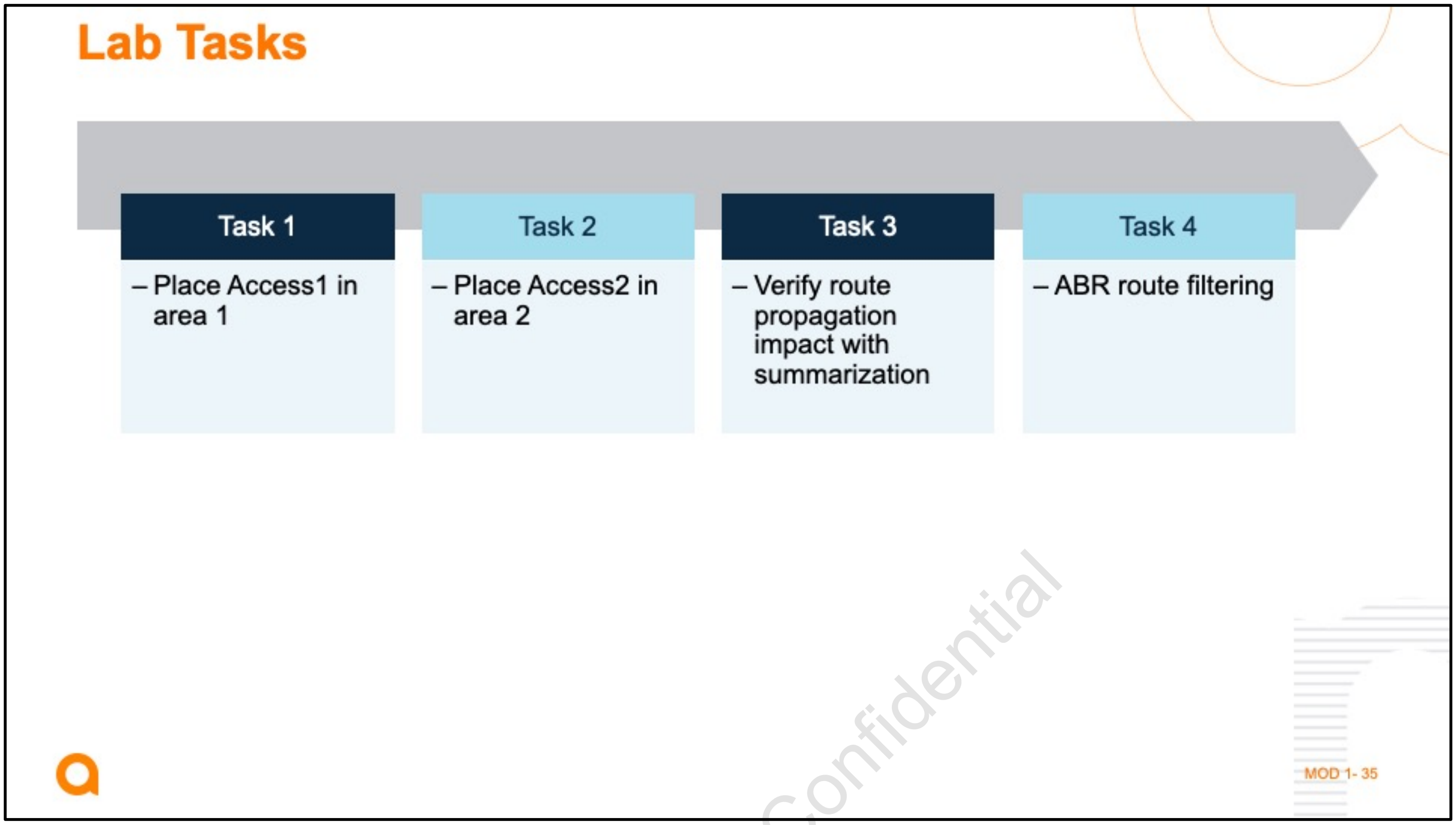
Aruba Training-Confidential

Lab Activity

Lab 6.2

Its time for another lab activity.

Aruba Training-C



Lab Tasks

Task 1
- Place Access1 in area 1

Task 2
- Place Access2 in area 2

Task 3
- Verify route propagation impact with summarization

Task 4
- ABR route filtering

MOD 1- 35

The figure provides a brief review of lab tasks. Please see your lab guide for details.

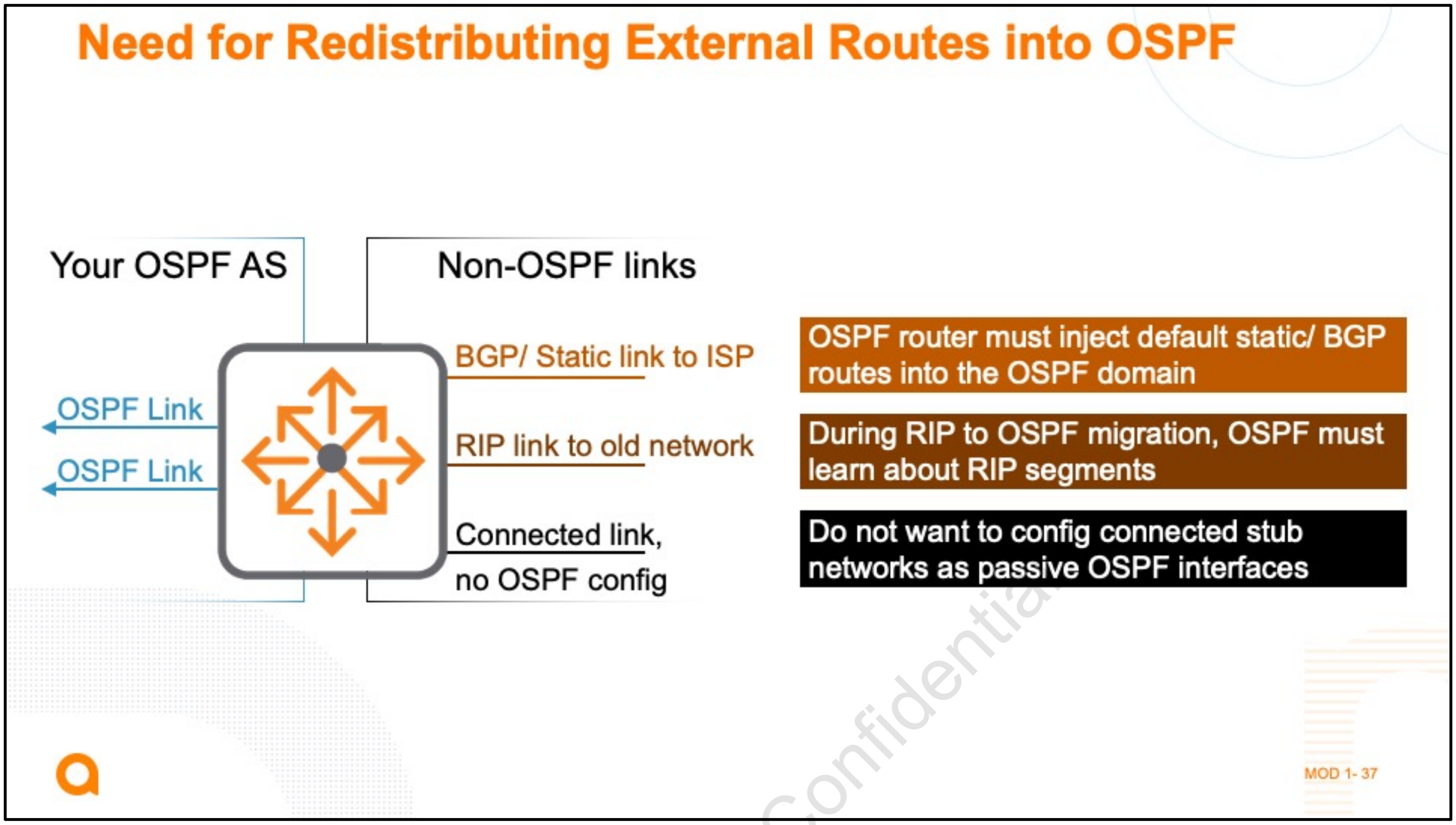
Aruba Training-Confidential

Route Redistribution

MOD 1- 36

Here the focus is on connecting an OSPF autonomous system (AS) to a foreign, external network, which may use a different routing protocol, like BGP. You learn to redistribute these routes into OSPF, and related concepts of advertisement and path selection.

Aruba Training-Confidential



Need for Redistributing External Routes into OSPF

Your OSPF AS

Non-OSPF links

OSPF Link

OSPF Link

BGP/ Static link to ISP

RIP link to old network

Connected link, no OSPF config

OSPF router must inject default static/ BGP routes into the OSPF domain

During RIP to OSPF migration, OSPF must learn about RIP segments

Do not want to config connected stub networks as passive OSPF interfaces

MOD 1- 37

Aruba Training-Confidential

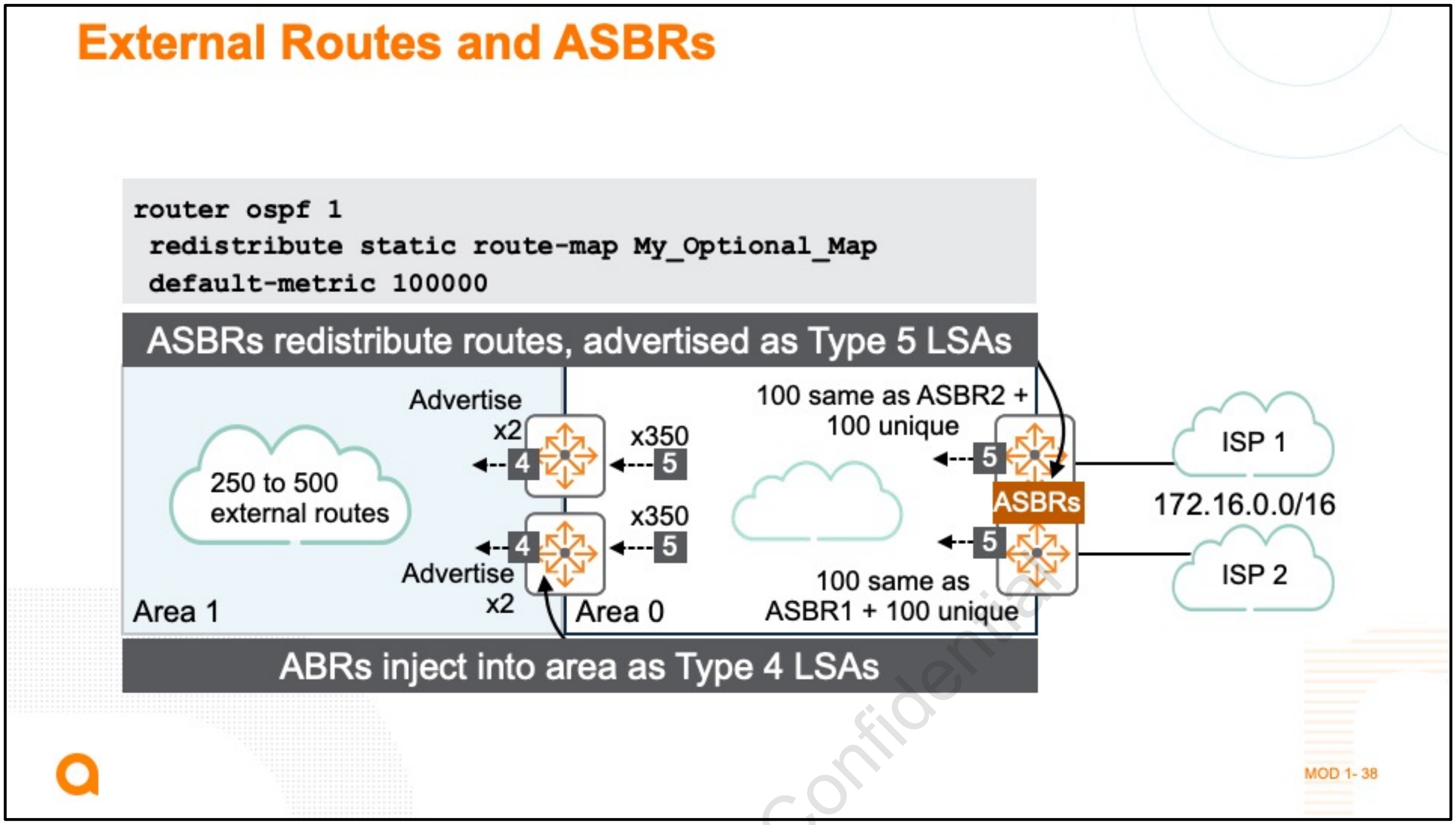
As a link state protocol, OSPF routers advertise network interfaces that run OSPF. But sometimes an OSPF router has other routes in its IP routing table, which must be advertised to OSPF neighbors.

Perhaps one or more routers connect to an ISP with static default or BGP routes. You must inject these routes into your OSPF domain.

Or suppose you are migrating your network from RIP to OSPF. During the upgrade, some segments run RIP. You must inject these routes into the already-migrated OSPF AS, and the old RIP segments must know about the OSPF networks.

For security and efficiency reasons, you might choose not to enable OSPF on a stub network. So you redistribute connected routes into OSPF , to be advertised to the rest of your AS.

Use route redistribution to connect OSPF to any foreign routing protocol.



External Routes and ASBRs

```
router ospf 1
 redistribute static route-map My_Optional_Map
 default-metric 100000
```

ASBRs redistribute routes, advertised as Type 5 LSAs

ABRs inject into area as Type 4 LSAs

The router that redistributes external routes is an Autonomous System Border Router (ASBR) because it communicates routing information between the OSPF AS and other routing processes. A redistributed route becomes an external OSPF route, sometimes referred to as an AS external route. All routes calculated from Type 1, Type 2, and Type 3 LSAs, including both intra-area and inter-area routes, are internal OSPF routes.

ASBRs advertise external routes in Type 5 LSAs. The LSA includes the network address for the redistributed route, the router ID of the advertising ASBR, and the metric, which you assign when you redistribute the route into OSPF.

OSPF routers propagate the Type 5 LSA in LSUs throughout the entire AS, making Type 5 LSAs different from Type 1, Type 2, and Type 3 LSAs. (Remember: although Type 3 LSAs advertise inter-area routes, each area has its own set of Type 3 LSAs.)

To calculate an external route from the Type 5 LSA, a router refers to the LSA's advertising ASBR. The router has calculated a shortest path to each ASBR. The forwarding interface and next hop for that path become the forwarding interface and next hop for all external routes advertised by that ASBR.

Aruba Training-Confidential

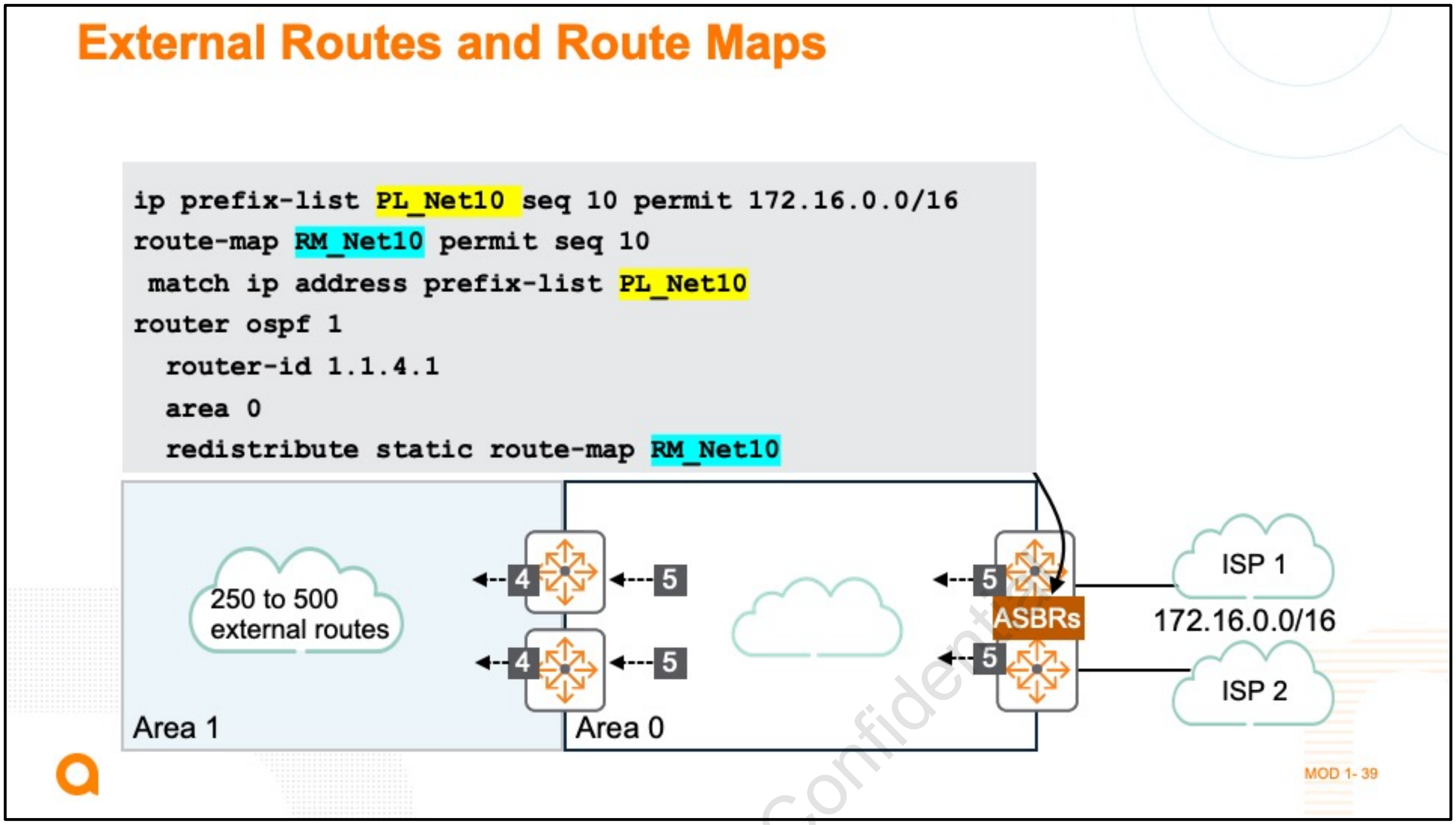
In other words, routers do not add the external networks to the directed graph for the shortest path algorithm. Instead they add them to a list of networks for which the advertising ASBR is the gateway. Therefore, changes to the particular external networks advertised by an ASBR do not affect the area topology.

The figure shows configuration syntax for ASBRs to redistribute static routes into OSPF. You could also specify connected (directly connected links not running OSPF) or BGP. The route map syntax is optional.

By default, all routes of the specified type are redistributed unless you employ route maps. When you redistribute routes into OSPF, they are assigned a seed metric to them: a beginning cost value. If you specify a default metric with the default-metric command, that would be applied to redistributed routes; if you want to have different metrics for different external routes, then use a route map to accomplish this task.

Note: Intra-area routes are always preferred over inter-area. Inter-area routes are always preferred over external routes. These rules are applied before the cost evaluation.

Aruba Training Confidential



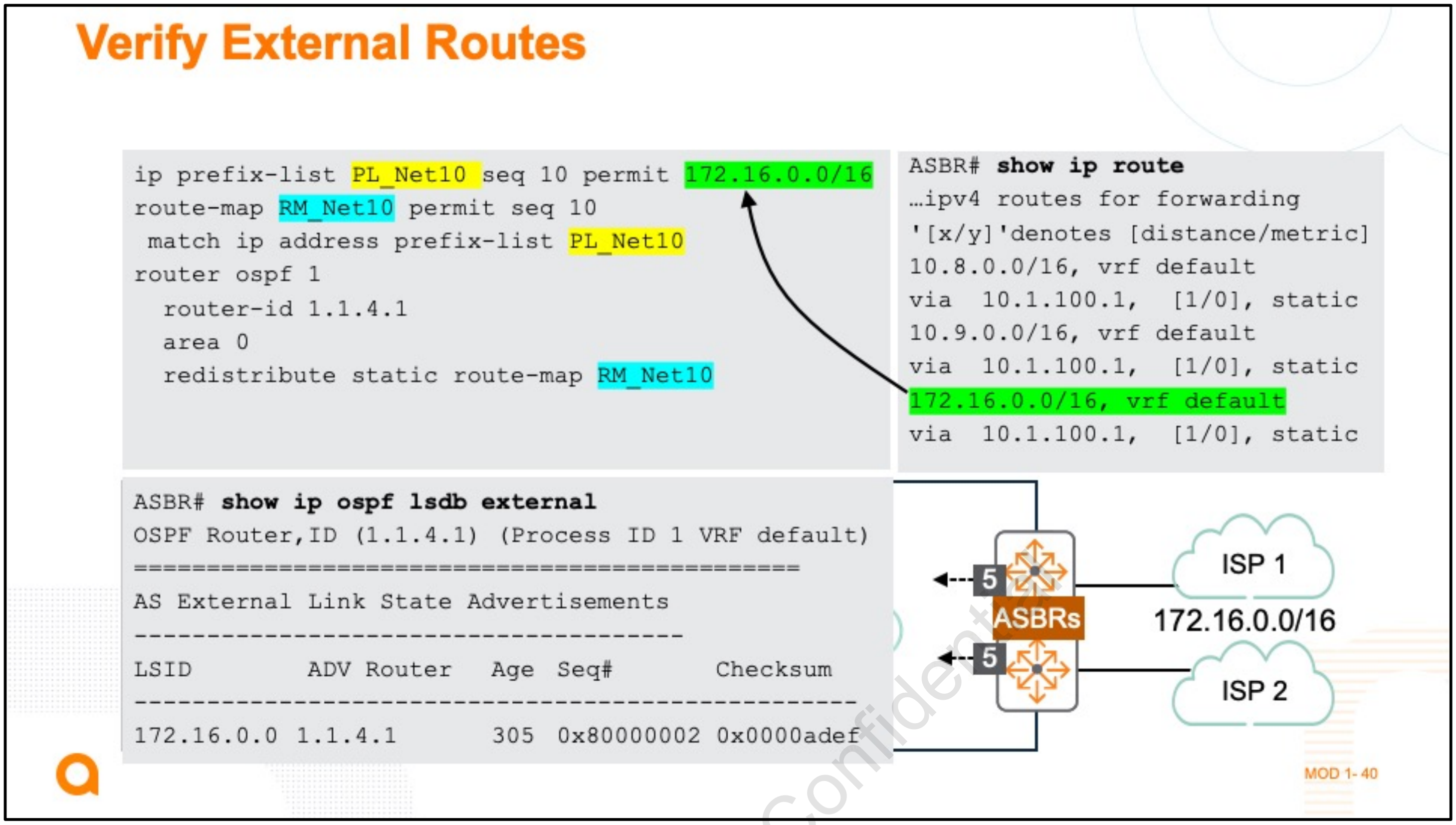
External Routes and Route Maps

```
ip prefix-list PL_Net10 seq 10 permit 172.16.0.0/16
route-map RM_Net10 permit seq 10
 match ip address prefix-list PL_Net10
router ospf 1
  router-id 1.1.4.1
  area 0
  redistribute static route-map RM_Net10
```

By default, when you redistribute a certain type of route into OSPF, the AOS-CX switch redistributes all routes of that type. However, you can filter the routes that are redistributed at the ASBR by applying a route map to the redistribute command, as shown in the figure.

Aruba Training-Confidential

The route map refers to an IP prefix list that specifies the networks to be redistributed into OSPF. In this example, the ASBR has several static routes, but you only want to redistribute 10.10.0.0/16 into OSPF. Optionally you can use various set parameters with match criteria, to modify how routes are advertised. You will learn more about this later in this module.



Verify External Routes

```
ip prefix-list PL_Net10 seq 10 permit 172.16.0.0/16
route-map RM_Net10 permit seq 10
 match ip address prefix-list PL_Net10
router ospf 1
  router-id 1.1.4.1
  area 0
  redistribute static route-map RM_Net10
```

```
ASBR# show ip route
…ipv4 routes for forwarding
'[x/y]'denotes [distance/metric]
10.8.0.0/16, vrf default
via  10.1.100.1,  [1/0], static
10.9.0.0/16, vrf default
via  10.1.100.1,  [1/0], static
172.16.0.0/16, vrf default
via  10.1.100.1,  [1/0], static
```

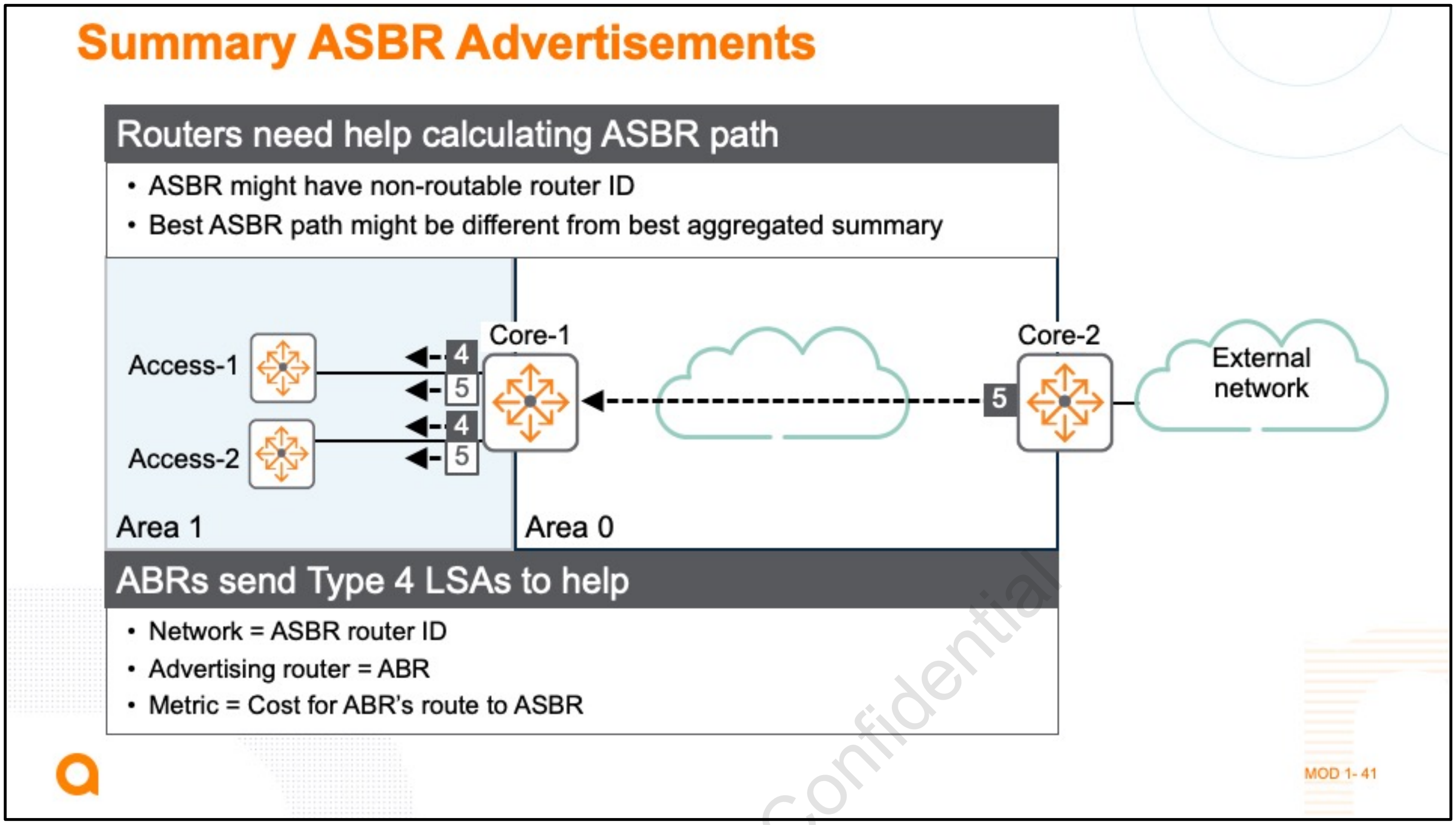
```
ASBR# show ip ospf lsdb external
OSPF Router,ID (1.1.4.1) (Process ID 1 VRF default)
==============================================
AS External Link State Advertisements
--------------------------------------
LSID         ADV Router   Age  Seq#        Checksum
----------------------------------------------------
172.16.0.0 1.1.4.1       305  0x80000002 0x0000adef
```

MOD 1- 40

Aruba Training-Confidential

The figure shows the ASBR routing table. Notice that the last static route matches the prefix list, as highlighted in green in the figure

The figure also shows the LSDB entry for the redistributed route. Note that on the ASBR, the local static route, because its administrative distance is 1 is placed in its local routing table; however, other adjacent OSPF routers will receive the LSBD redistributed route and place that in their routing table if it happens to have the best distance and metric.



Summary ASBR Advertisements

Routers need help calculating ASBR path

- ASBR might have non-routable router ID
- Best ASBR path might be different from best aggregated summary

ABRs send Type 4 LSAs to help

- Network = ASBR router ID
- Advertising router = ABR
- Metric = Cost for ABR's route to ASBR

MOD 1- 41

Aruba Training-Confidential

When a router receives a Type 5 LSA from an ASBR in the same area, it can use the area's topology to calculate the shortest path to that ASBR, the ASBR route. The next hop in this route becomes the next hop for all external networks for which that ASBR is the gateway.

Routers in different areas from the ASBR also receive the Type 5 LSA, but they need help determining how to reach the ASBR: Type 4 LSAs are used to denote the ASBRs themselves. The Type 5 LSA indicates the advertising ASBR by router ID. You have learned how to make router IDs routable. However, because OSPF does not require routable router IDs, it must provide another mechanism for informing routers in other areas how to reach the ASBR.

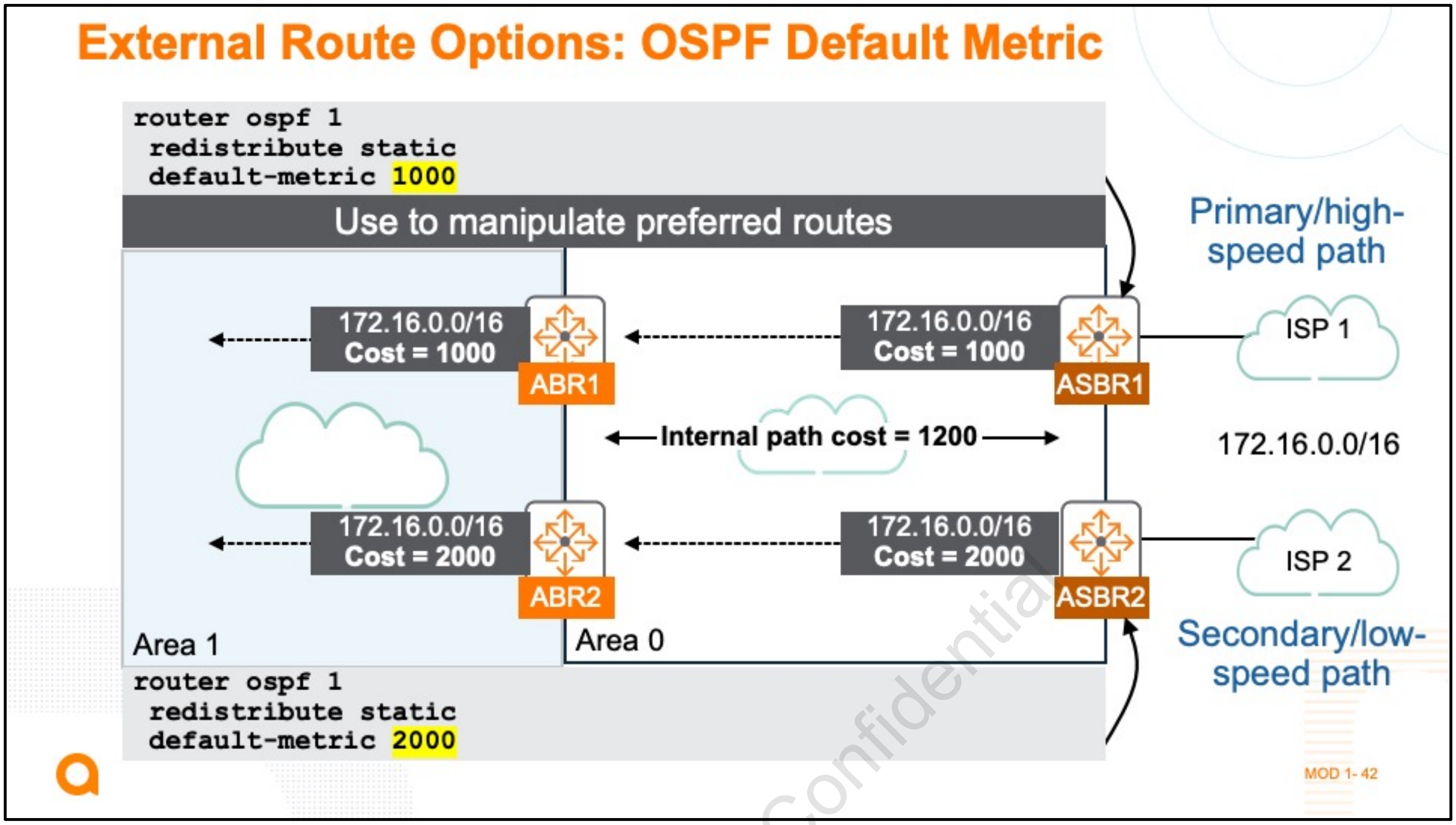
Also, even if OSPF did require routable router IDs, an OSPF router in another area might not want to use its typical inter-area routes to reach the ASBR. In areas with multiple ABRs, the ABR with the lowest cost for the broader area range might not provide the best path to the ASBR within the area.

OSPF defines the Type 4 summary LSA so that the routers in other areas can learn the best ABR through which to reach particular ASBRs. For each area, an ABR generates one Type 4 LSA for each ASBR that it can reach in another area. As the metric for this LSA, the ABR specifies the cost for its shortest path to the ASBR. For example, in the figure, Core-1 would specify its cost for reaching Core-2.

Internal routers that receive the Type 4 LSAs use them to select the best path to each ASBR in other areas. In this example, the choice is simple because area 1 has only one ABR.

The figure shows an external OSPF route learned by an internal area 1 router from the ASBR in area 0. You can see both the Type 4 LSAs (about the ASBR), and the Type 5 LSAs (about routes the advertised by the ASBR).

Aruba Training-Confidential



External Route Options: OSPF Default Metric

```
router ospf 1
 redistribute static
 default-metric 1000
```

Use to manipulate preferred routes

```
router ospf 1
 redistribute static
 default-metric 2000
```

MOD 1- 42

Aruba Training-Confidential

Consider how OSPF routers choose the best external route in networks with redundant ASBRs, as well as redundant paths to the ASBRs. Multiple ASBRs might send Type 5 LSAs for the same network. The Type 5 LSA includes a metric and metric type, helping other routers to choose the best ASBR for reaching that network.

The ASBR that advertises the Type 5 LSA chooses the metric type and sets the default metric, often called the seed cost or seed metric.

In the figure, ASBR1 is configured with a seed metric of 1000, while ASBR2 has seed metric = 2000. Thus, all routers will prefer to use ASBR1 to reach external networks, only using ASBR2 if the path via ASBR1 fails.

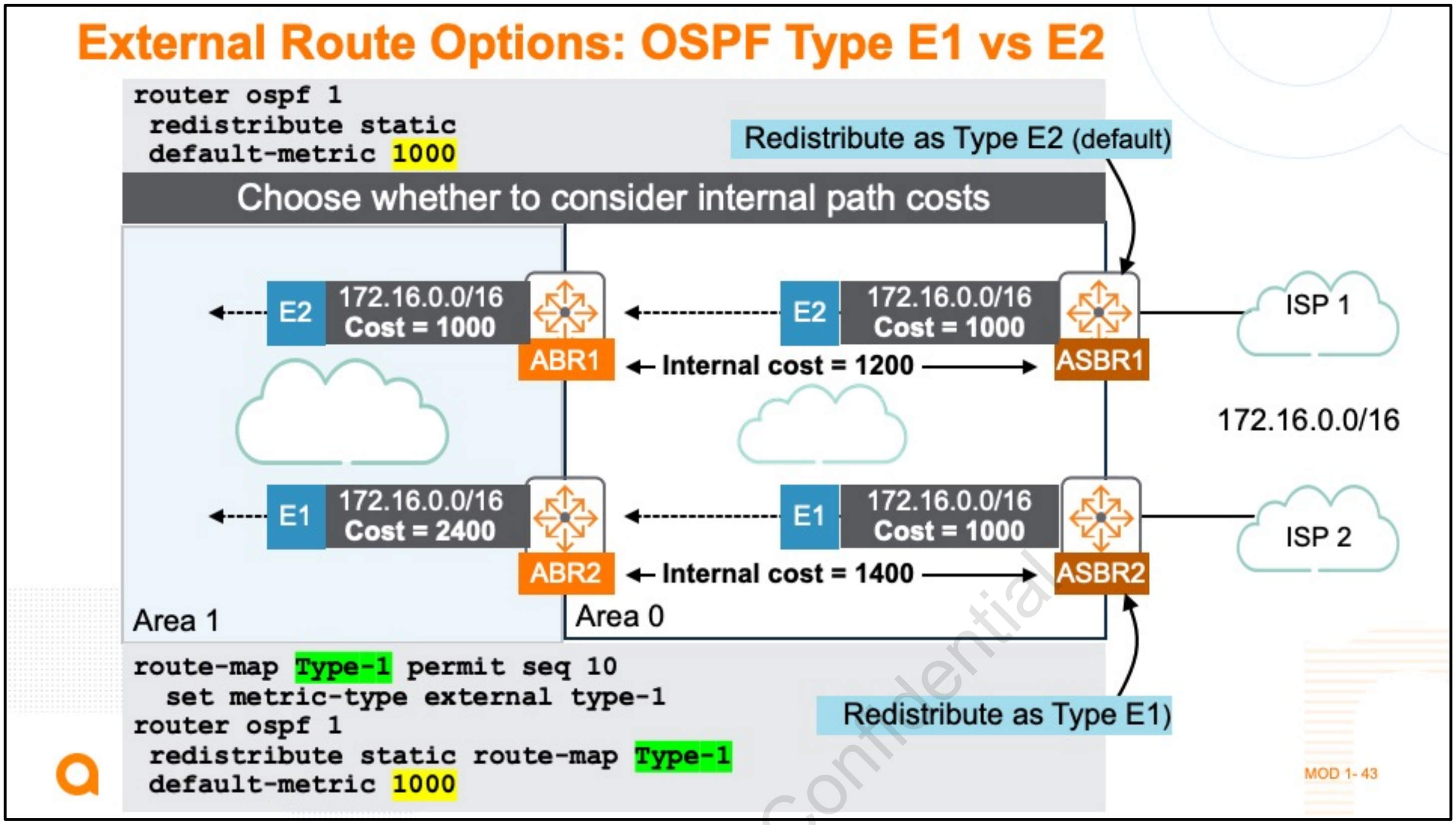
Perhaps ISP1 is your primary, high-speed link path to 172.16.0.0/15 subnets, while ISP2 is a lower-cost, slower link – only to be used in emergency/outage situations.

Or perhaps you know that ISP1 is simply a better, faster path, with lower latency and a better Service Level Agreement (SLA). You could also use this technique for load sharing, where ISP 1 is the primary path for 172.16.0.0/16 subnets, while ISP 2 is the primary path for 172.17.0.0/16 subnets.

Did you notice something strange about the example? ASBR1 and 2 advertise their default or seed metric of 1000 and 2000 respectively. This advertisement traverses Area 0 to reach ABR1 and ABR2, with a cost of 1200. However, ABR1 and 2 did not add this cost before they advertised routes toward Area 1 internal routers! They still advertise the same metric advertised by the ASBRs, without considering their cost to reach the ASBRs. Why is that?

The answer relates to how ASBRs inject external routes into OSPF as either External Type 1 (E1) or External Type 2 (E2) routes.

Aruba Training-Confidential



External Route Options: OSPF Type E1 vs E2

```
router ospf 1
 redistribute static
 default-metric 1000
```

Redistribute as Type E2 (default)

Choose whether to consider internal path costs

```
route-map Type-1 permit seq 10
  set metric-type external type-1
router ospf 1
 redistribute static route-map Type-1
 default-metric 1000
```

Redistribute as Type E1)

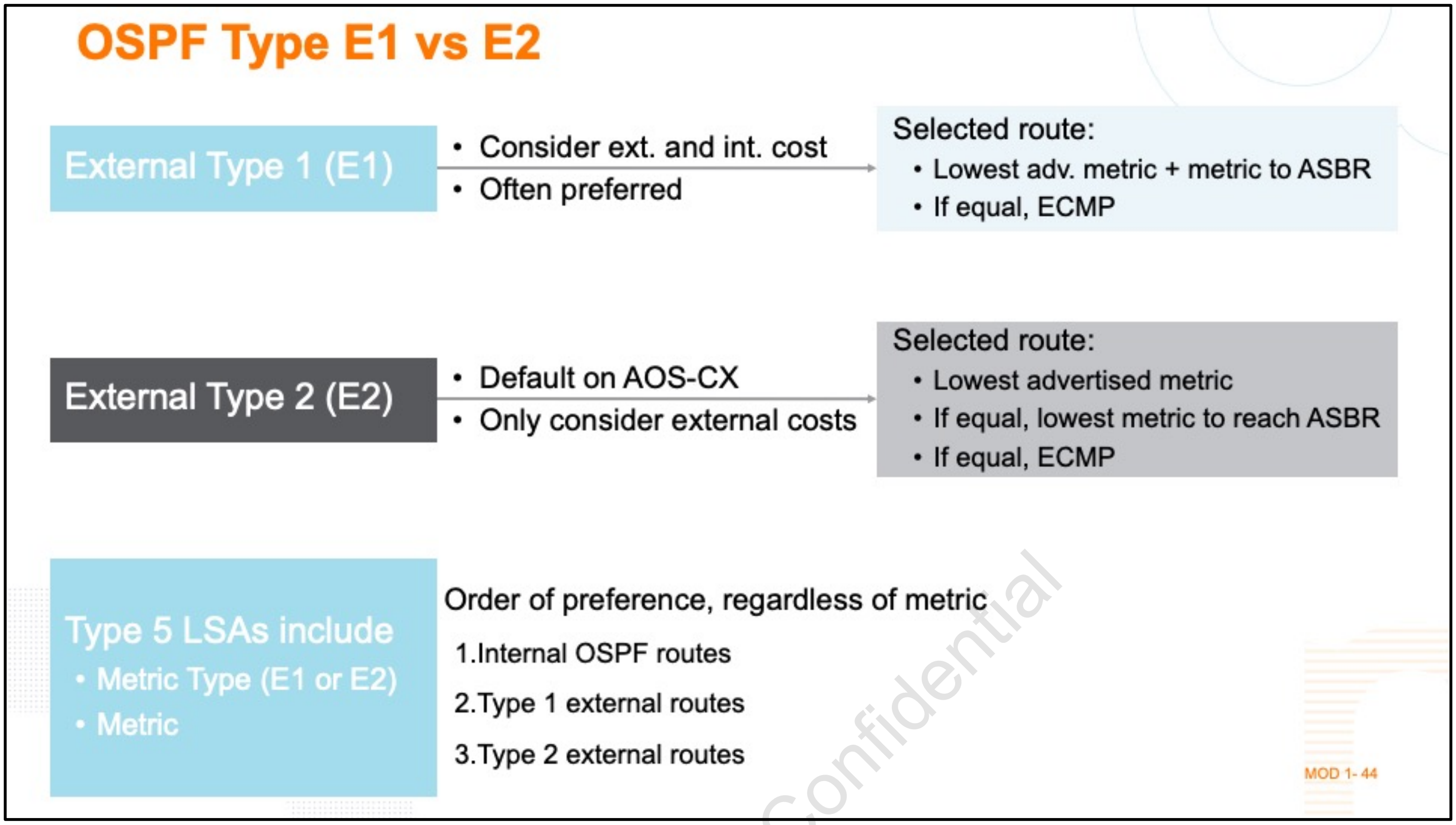
MOD 1- 43

Aruba Training-Confidential

In the figure, you have modified the ASBR configurations. ASBR1 remains unchanged. You have not explicitly configured ASBR1 to inject external routes as OSPF External Type 1 (E1) or Type 2 (E2). Thus, ASBR1 uses the default is of E2, where internal OSPF path costs are not considered. ASBR1 advertised its default metric as 1000, and it costs ABR1 1200 to reach ASBR1, but ABR1 does not consider this cost. It simply advertised the route with a cost of 1000.

You configured ASBR2 with a route map, with no match criteria. This means that all known external routes are to be advertised . The set statement indicates they are all to be advertised as type E1, where internal paths ARE considered. ASBR2 advertises 172.16.0.0/16 with a cost of 1000. ABRs has a cost of 1400 to reach ASBR2, and so it advertised 172.16.0.0/16 with a cost of 2400.

Note: You can use route maps to change the metric on a per-route basis if needed. Set the metric to a higher value than what is used internally in your OSPF network. So, if the highest native, internal OSPF path cost is 10,000, its good to set the seed metric to, say, 20,000.



OSPF Type E1 vs E2

External Type 1 (E1)

- Consider ext. and int. cost
- Often preferred

Selected route:

- Lowest adv. metric + metric to ASBR
- If equal, ECMP

External Type 2 (E2)

- Default on AOS-CX
- Only consider external costs

Selected route:

- Lowest advertised metric
- If equal, lowest metric to reach ASBR
- If equal, ECMP

Type 5 LSAs include

- Metric Type (E1 or E2)
- Metric

Order of preference, regardless of metric

1. Internal OSPF routes
2. Type 1 external routes
3. Type 2 external routes

MOD 1- 44

Aruba Training-Confidential

The figure summarizes key aspects of E1 and E2 routes.

With External Type 1 routes, both external and internal path costs are considered. This is often preferred. Routers select the external route through the ASBR with the lowest total cost. If a tie occurs, ECMP lets the router add multiple routes.

External Type 2 is the default on AOS-CX and most other devices. Only external path costs are considered – use the ASBR with the lowest advertised metric. However, if multiple LSAs tie for the lowest value, the metric for reaching the ASBR is a tie breaker. If a tie occurs again, the router uses Equal Cost Multi-Path (ECMP). If

You may wonder why internal path costs would not be considered. It is because when OSPF was developed, it was assumed that internal path costs where all relatively high-speed, low-cost links (10mbps at the time or slower). This was insignificant as compared to the very low-bandwidth high-cost external WAN links – likely operating at 1.544 Mbps or slower. Since the internal path costs were so insignificant, why consider them?

Today, this is less true, and even if it is, it is often best for OSPF routes to use Type E1 routes and consider internal path costs. You typically want routers to take the best path to reach the ASBR, and this can only truly happen if you configure ASBRs inject external routes as Type

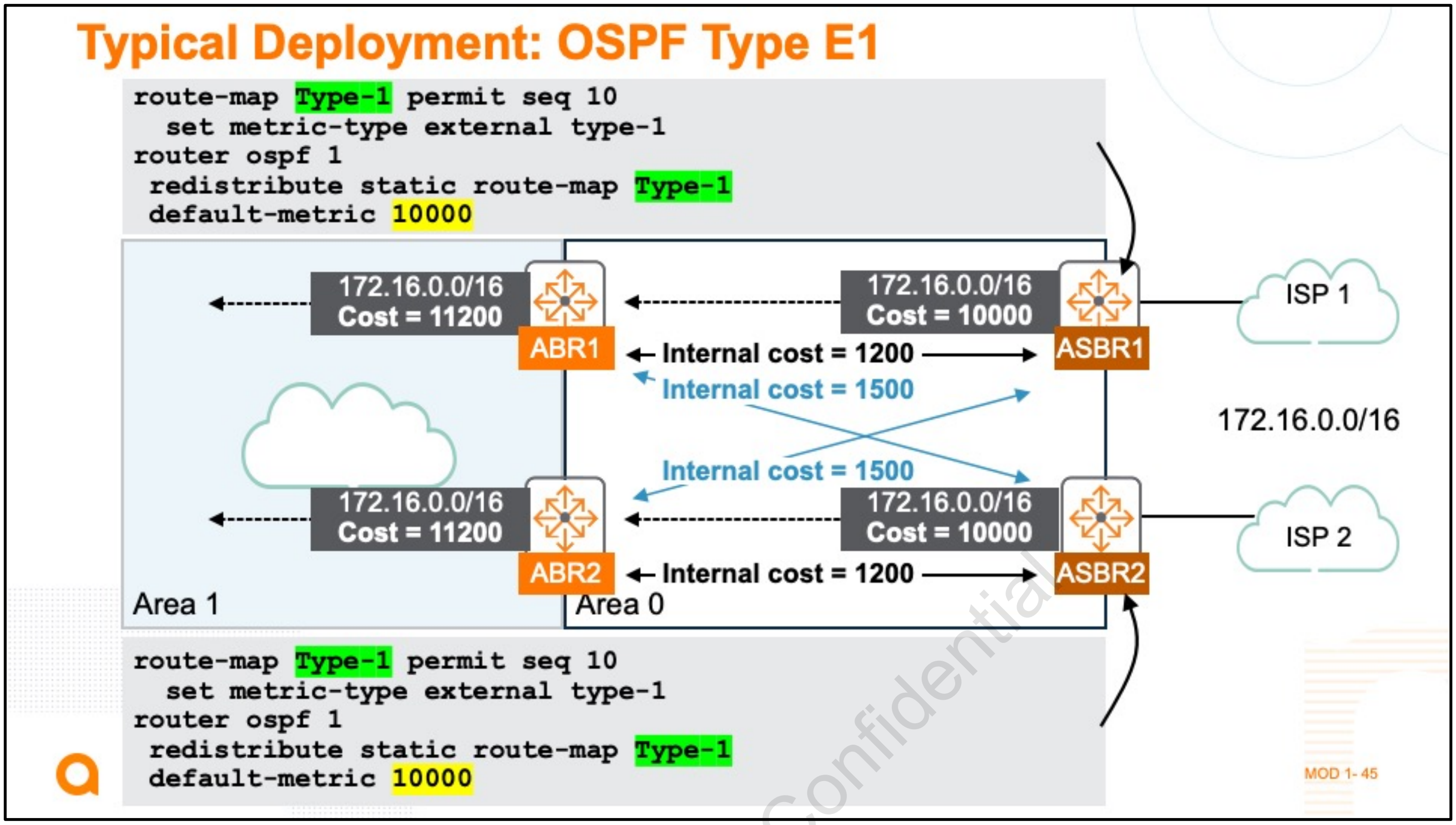
E1.

So, every external route advertised in a Type 5 LSA has two important properties: the metric-type, which is an “external” type 1 or type 2, and the metric.

By default, OSPF prefers routes based on the following order:
Internal OSPF routes
External type 1 routes
External type 2 routes

Let’s look at how a typical, modern network might be deployed.

Aruba Training-Confidential



Typical Deployment: OSPF Type E1

```
route-map Type-1 permit seq 10
  set metric-type external type-1
router ospf 1
 redistribute static route-map Type-1
 default-metric 10000
```

```
route-map Type-1 permit seq 10
  set metric-type external type-1
router ospf 1
 redistribute static route-map Type-1
 default-metric 10000
```

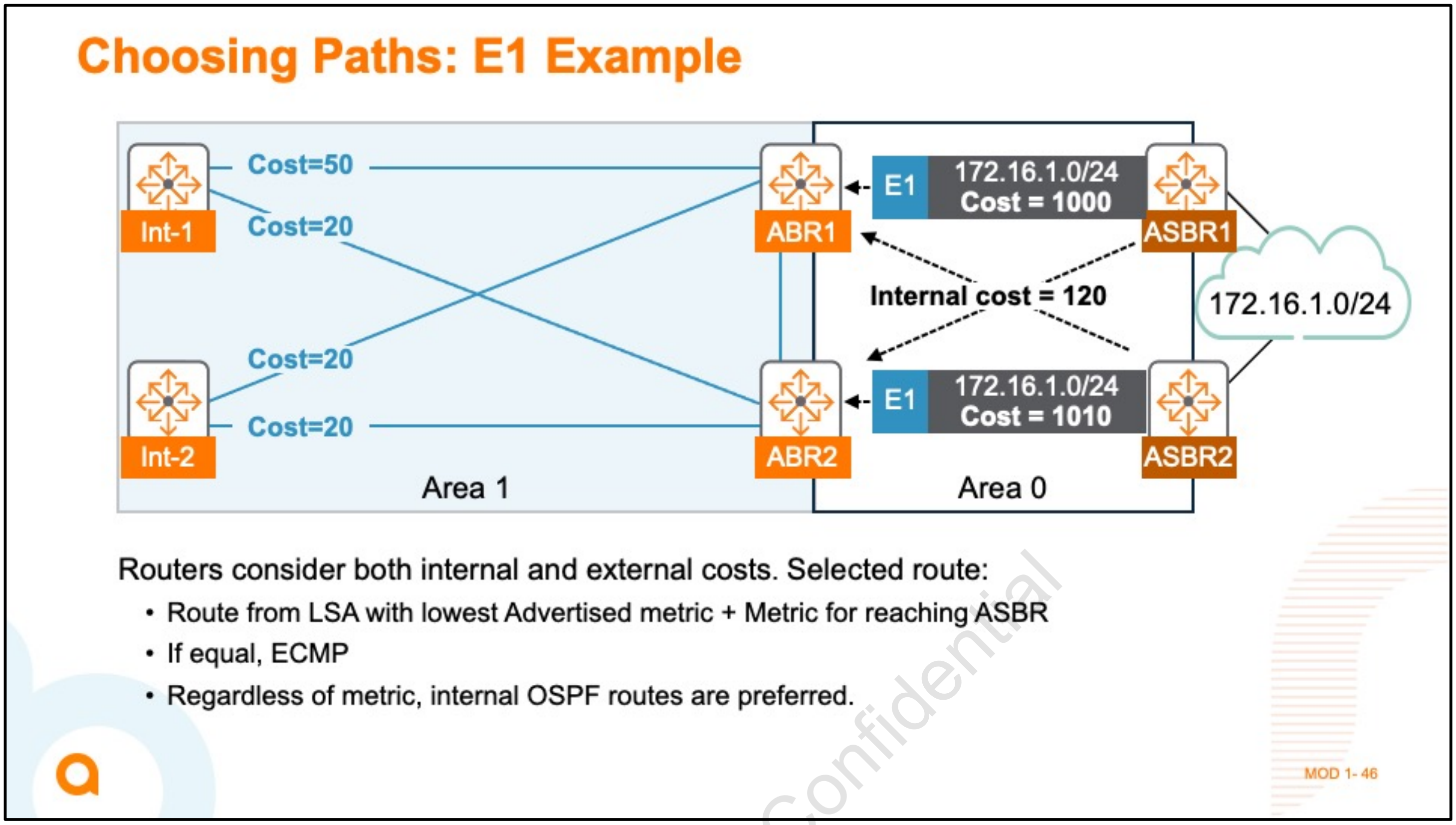
MOD 1- 45

Aruba Training-Confidential

The figure shows a common deployment scenario, where both ASBRs inject external routes as type E1, with the same default metric.

ABR1 has an internal cost of 1200 to reach ASBR1, and 1500 to reach ASBR2. Thus, total cost = 11200 to reach 172.16.0..0/16 via ASBR1, and 11500 to reach those subnets via ASBR2. It maintains both paths in its LSDB, but only the best path in its route table. ABR1 then advertises this best path to Area 1 internal routers.

ASBR2 and ABR2 have the same story. This is a nice solution for many deployments. All paths are available, but only the best paths are used. In case of failure anywhere in the path, routers will converge on the alternate path with little to no downtime.



Choosing Paths: E1 Example

Routers consider both internal and external costs. Selected route:

- Route from LSA with lowest Advertised metric + Metric for reaching ASBR
- If equal, ECMP
- Regardless of metric, internal OSPF routes are preferred.

MOD 1- 46

Aruba Training-Confidential

The figure shows a scenario, which you can use to solidify your understanding of how an OSPF router chooses between type 1 external routes to the same network.

Recall these facts:

The router uses the Type 5 LSA to calculate the external route, setting the correct forwarding interface and next hop for the path to the advertising ASBR.

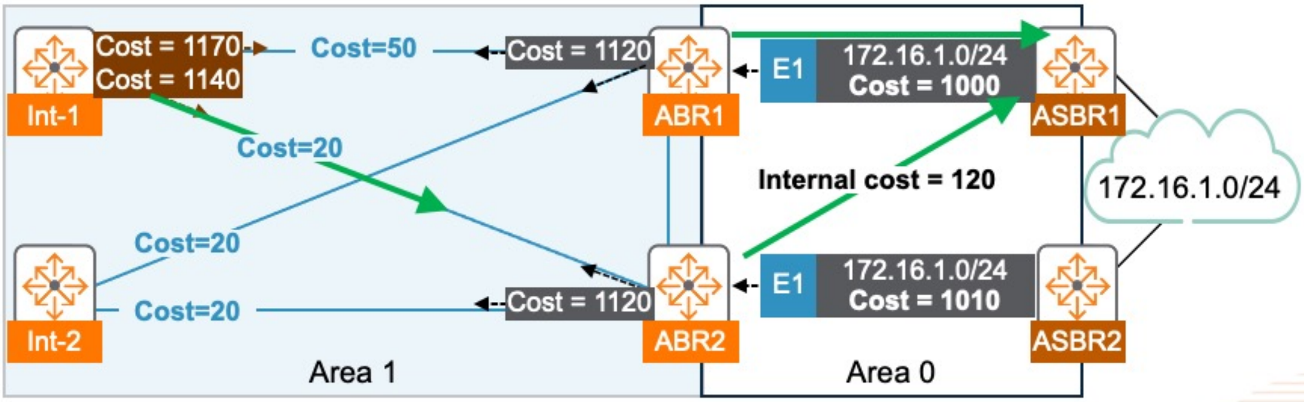
The router sets the external route's metric equal to:
The metric advertised in the Type 5 LSA + The metric for the forwarding path to the advertising ASBR

When the OSPF routing table already includes an external (type 2) OSPF route to this network, the router selects the route with the lower metric.

If both routes have the same metric, the router installs multiple routes in the OSPF routing table. (Otherwise, it uses a mechanism such as router ID to break the tie.)

Now study the figure. All routers support ECMP. Try not to look at the material below this figure and see if you can answer this question. Which ASBR does Internal router Int-1 select to reach the external 172.16.1.0/24?

Choosing Paths: E1 Example Answer



Routers consider both internal and external costs. Selected route:

- Route from LSA with lowest Advertised metric + Metric for reaching ASBR
- If equal, ECMP
- Regardless of metric, internal OSPF routes are preferred.

MOD 1- 47

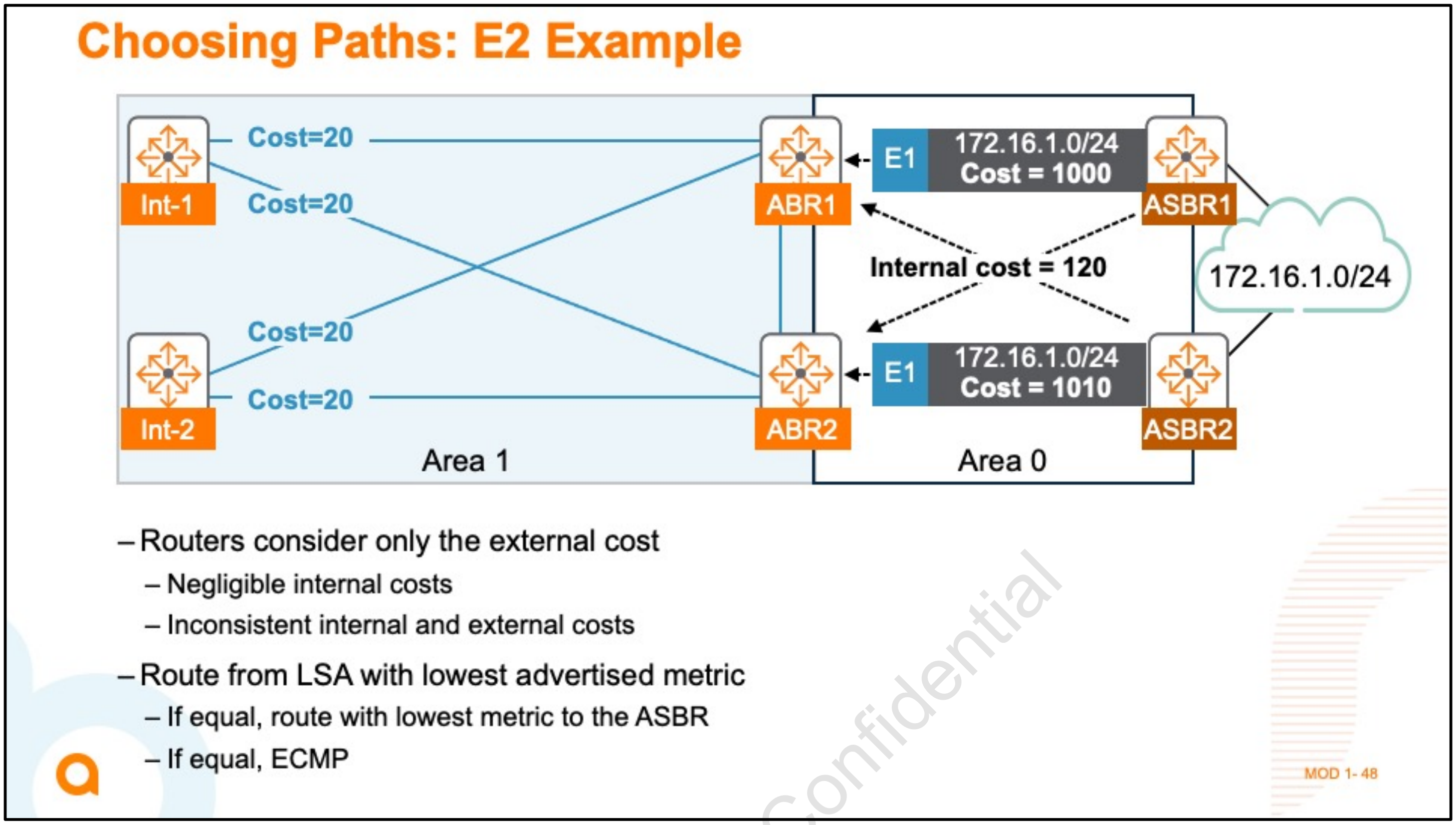
Aruba Training-Confidential

Which ASBR does Internal router Int-1 select to reach the external 172.16.1.0/24?

The figure shows the answer to this question. To figure it out, you must know that ABR1 learns of 172.16.1.0/24 from ASBR1, with a cost of 1000; and from ASBR2 with cost = 1010. The cost to reach both ASBRs is 120, so actual cost for ABR1 is 1120 via ASBR1 and 1130 via ASBR2. For ASBR1, the path via ASBR1 is best. The same holds true for ABR2 – it learns of 172.16.1.0/24 from both ASBRs, with ASBR1 being the lowest-cost path.

This is advertised into Area 1.

So, Int-1 learns of 172.16.1.0/24 via both ABR1 and 2, both with a best path of 1120. However, the cost to reach ABR1 is 50, for a total cost of 1170, while to reach ABR2 is 20, for a total cost of 1140. This is Int-1’s best path



Choosing Paths: E2 Example

- Routers consider only the external cost
  - Negligible internal costs
  - Inconsistent internal and external costs
- Route from LSA with lowest advertised metric
  - If equal, route with lowest metric to the ASBR
  - If equal, ECMP

MOD 1- 48

Aruba Training-Confidential

The figure shows a scenario where ASBRs advertise Type E2 routes. Keep the following in mind.

Type 5 LSAs use type E1 and type E2 metrics. When routers examine a type E2 metric, which is the default, they consider only the metric advertised by the ASBR. OSPF follows this high-level process to choose the best external route:

The router uses the Type 5 LSA to calculate the external route, setting the correct forwarding interface and next hop for the shortest path to the advertising ASBR.

- If the router is in the same area as the ASBR, it knows the correct forwarding interface and next hop from the area’s shortest path topology. Equal-Cost Multi-Path (ECMP) routing might apply.
- If the router is in a different area from the ASBR, it calculates a metric for the ASBR route by incrementing the metric in the Type 4 LSA with the metric for reaching the advertising ABR. If multiple ABRs advertise a Type 4 LSA to the same ASBR, the router chooses the ASBR route with the lower total cost.
- A router might have multiple equal cost paths to an ASBR. Multiple ABRs might offer the router an equal-cost path to the ASBR, and the router has a different next hop for each ABR. Or the router might have multiple equal-cost paths to one ABR.
- The best path might not match the IP route to the ASBR’s IP address in the routing table. For example, the ABRs in the figure above advertise the same cost to the ASBRs. However, in Figure 6-26, router A and router B might advertise the same cost for the area 0

summary. However, ABR-B advertises a better Type 4 LSA for ASBR-E. Therefore, the internal area 1 routers will use their best path to router B to calculate the path to the ASBR, rather than load-balance over the link to both ABRs.

The router sets the external route's metric equal to the metric advertised in the Type 5 LSA. The advertising ASBR added this metric when it redistributed the route into OSPF.

When the OSPF routing table already includes an external (type 2) OSPF route to this network, the router selects the route with the lower metric.

If both routes have the same metric, the router examines the metric for the forwarding path to the ASBR. It selects the route associated with the ASBR to which it has a lower cost path.

If forwarding paths to the ASBRs have equal costs and the router supports ECMP, the router installs multiple routes in the OSPF routing table. (Otherwise, it uses another mechanism such as the router ID to break the tie.)

Aruba Training-Confidential

When all ASBRs use type 2 metrics, an OSPF router will always select the external route through the ASBR that advertises the lowest metric for it, even if the router is closer to another ASBR that advertises the same route. However, if multiple ASBRs advertise the route with the same metric, other routers do take the cost of reaching the ASBR into account.

The reasoning behind this behavior is that the cost of reaching the ASBR through the OSPF AS may be negligible compared to the cost of reaching the external network. In addition, the ASBR might not be assigning metrics to the external routes that correspond to the metrics used in the OSPF AS. In other words, a metric of 10 for the external route might not mean the same thing that a 10 metric on an OSPF link means, in terms of how strongly the company prefers using a particular path.

Note: This figure examines the selection process when OSPF is choosing between identical external OSPF routes. Usually, a network advertised in a Type 5 LSA should not be associated with an internal OSPF route, but a misconfiguration could cause this situation to occur. In that case, the OSPF always prefers internal (intra-area and then inter-area) routes.

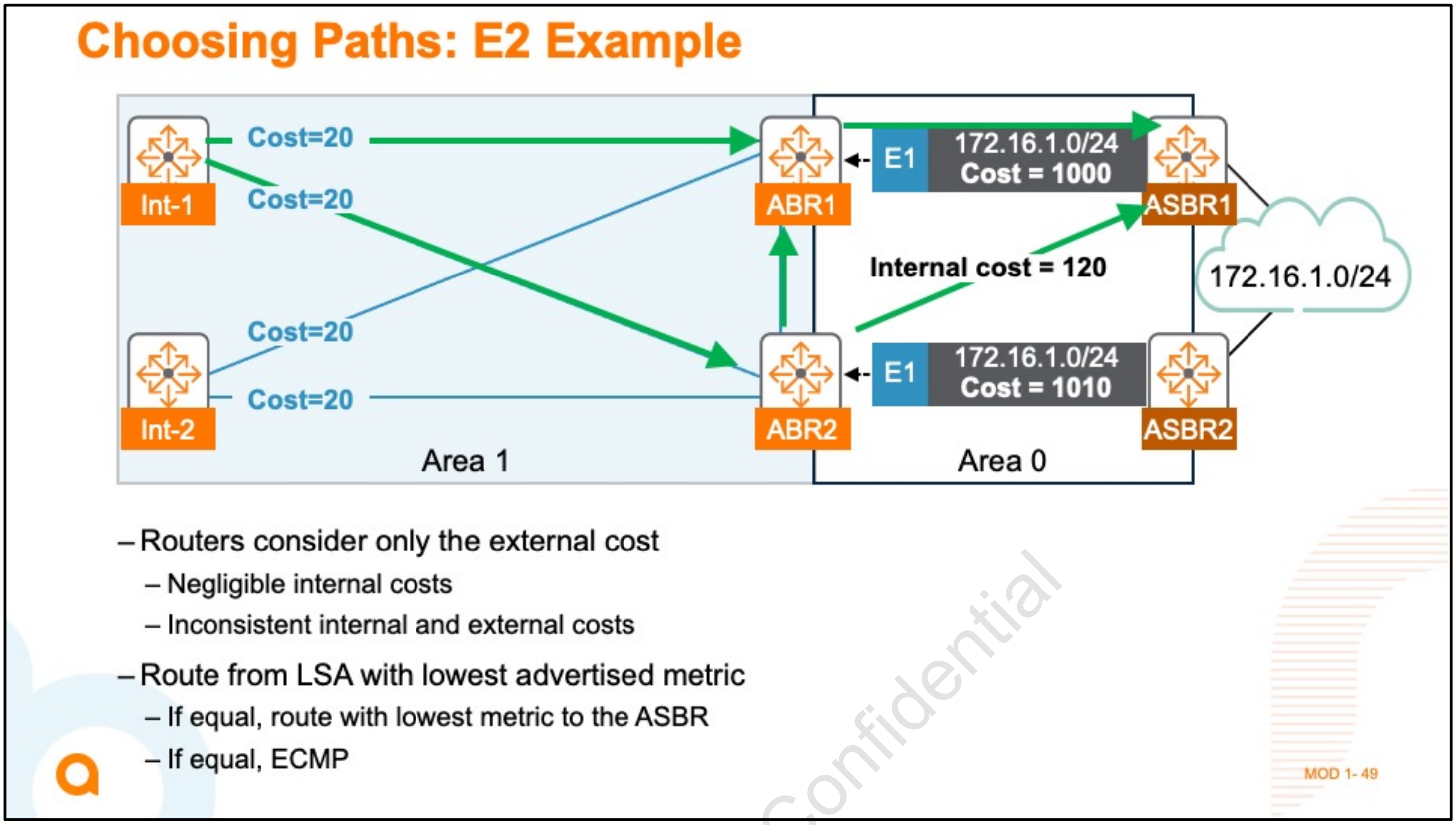
And, of course, the router uses administrative distance to choose between the OSPF route and routes learned through other means.

Discussion questions

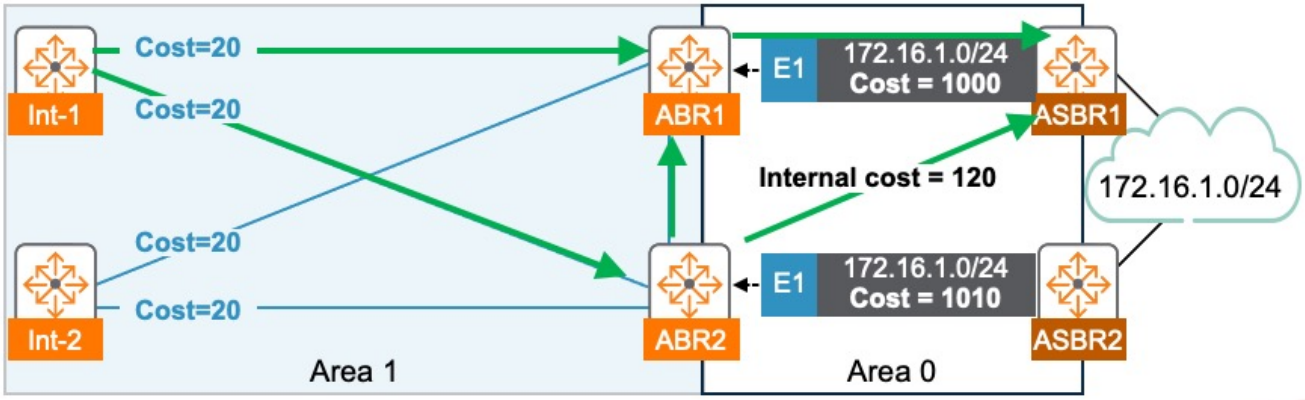
Which ASBR does router Int-1 select to reach the external route, 172.16.1.0/24?

What is the next hop in Int-1’s route table to reach 172.16.1.0/24? (This router supports ECMP.)

Aruba Training-Confidential



Choosing Paths: E2 Example



- Routers consider only the external cost
  - Negligible internal costs
  - Inconsistent internal and external costs
- Route from LSA with lowest advertised metric
  - If equal, route with lowest metric to the ASBR
  - If equal, ECMP

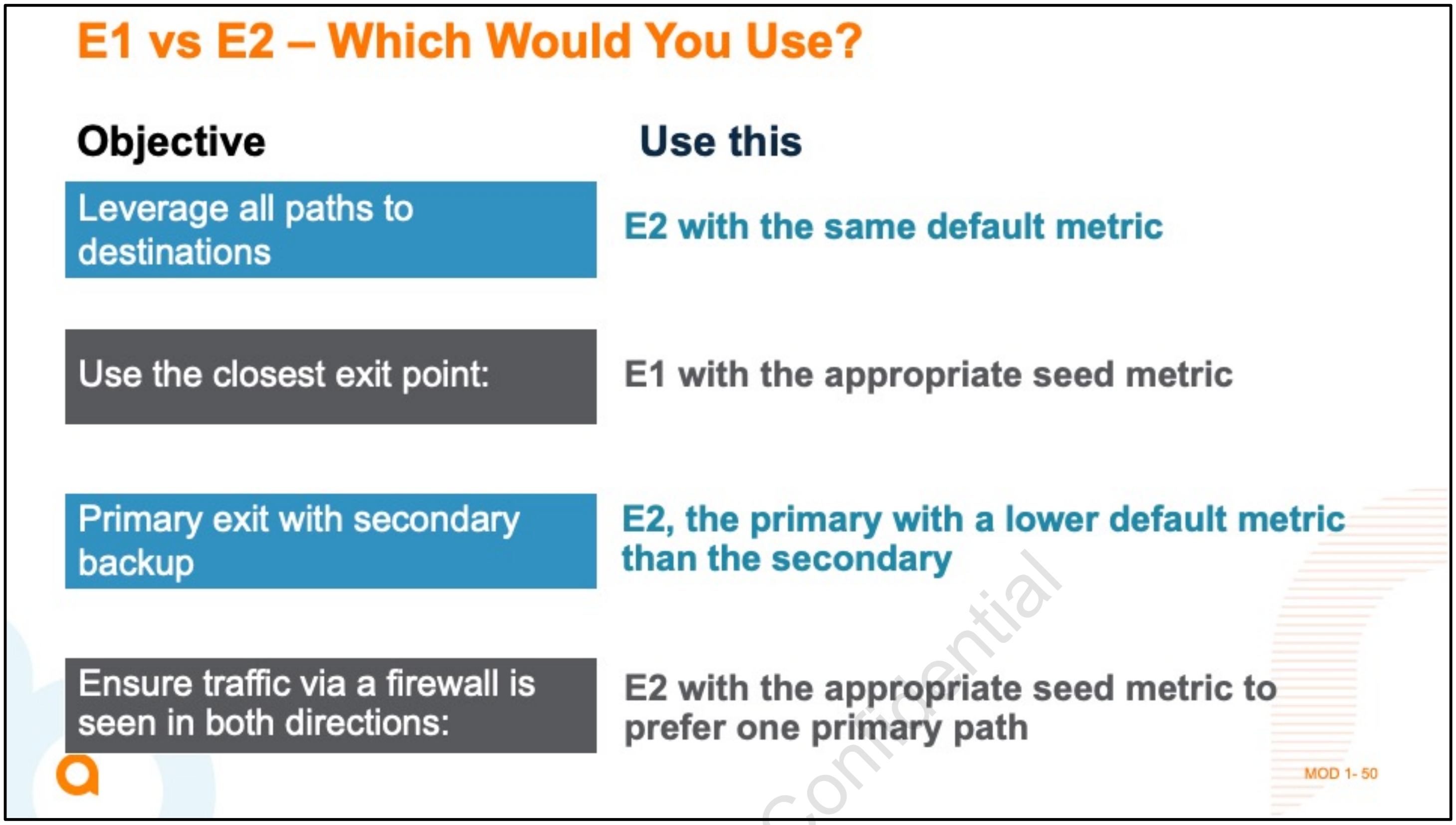
Aruba Training-Confidential

Which ASBR does router Int-1 select to reach the external route, 172.16.1.0/24?

The figure shows the paths that Int-1 would take in this scenario – all equal cost paths.

What is the next hop in Int-1's route table to reach 172.16.1.0/24? (This router supports ECMP.)

Both ABR1 and ABR2 would therefore appear in the router Int-1's route table as next hops.



E1 vs E2 – Which Would You Use?

Objective	Use this
Leverage all paths to destinations	E2 with the same default metric
Use the closest exit point:	E1 with the appropriate seed metric
Primary exit with secondary backup	E2, the primary with a lower default metric than the secondary
Ensure traffic via a firewall is seen in both directions:	E2 with the appropriate seed metric to prefer one primary path

MOD 1- 50

The figure helps determine which LSA 5 external type(s) you should use based on the policy you are trying to implement.

Aruba Training-Confidential

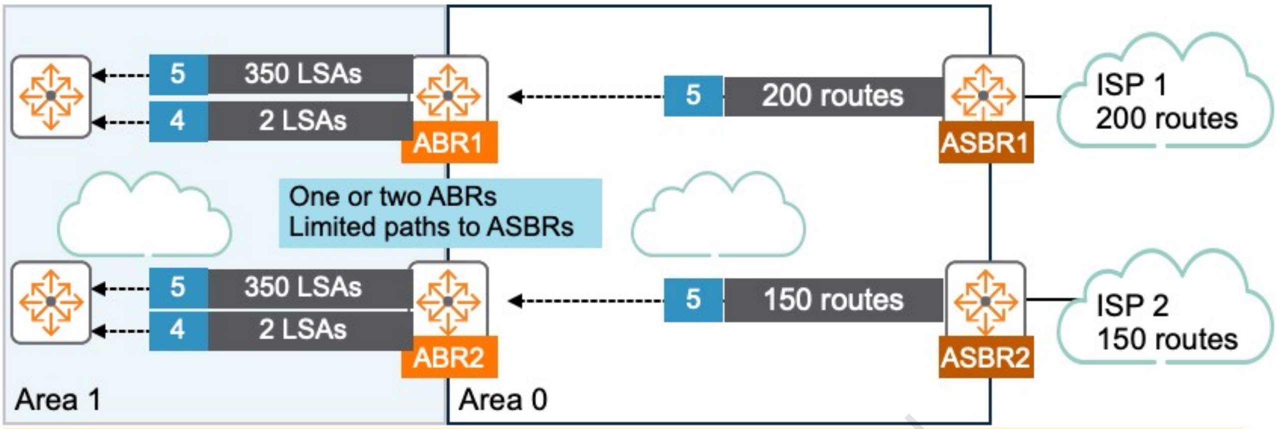
OSPF Area Types

MOD 1- 51

You will now explore the use of area types to more easily deal with external routes in OSPF.

Aruba Training-Confidential

Need for Eliminating Excessive External Routes



Examine each non-backbone area: Does it really need all external routes?

If no ASBRs, all external traffic must traverse the ABRs

If only 1 ABR, then only 1 path out of the Area – just use a default route

Even if 2 ABRs, default routes from ABRs may suffice

Aruba Training-Confidential

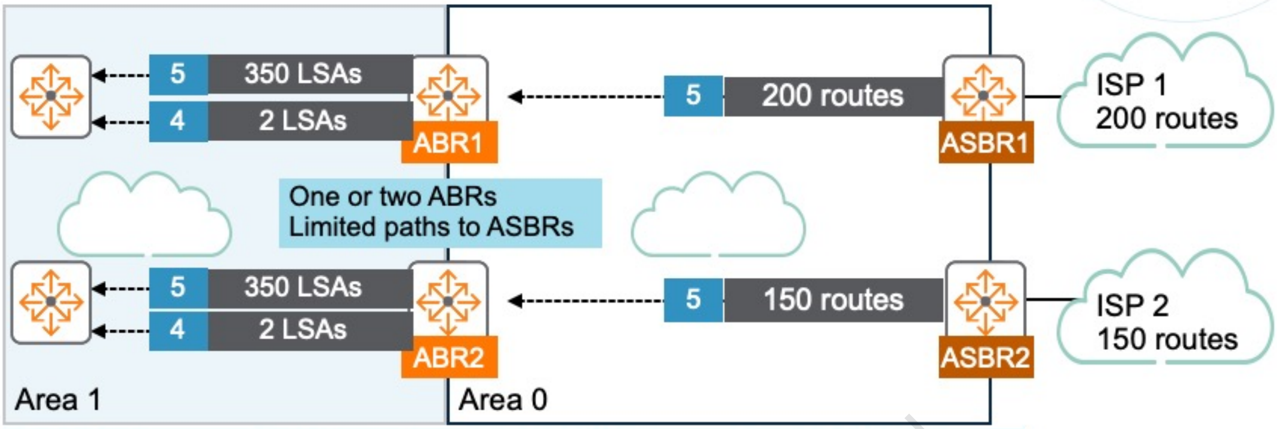
Look at this scenario, where 200 routes from ISP 1 and 150 routes from ISP 2 are injected into Area 0 as Type 5 LSAs. The ABRs then advertise these into Area 1, increasing route table size by 350 routes, along with associated memory and CPU usage. The internal routers in some OSPF areas rarely require all external routes.

Remember the primary benefit of individual Type 4 and Type 5 LSAs. Type 5 LSAs help routers choose the best ASBR for reaching an external network. If that ASBR is in another area, Type 4 LSAs further help the routers choose the best ABR to reach that ASBR, and perhaps to select the best ASBR. You should examine each non-backbone area and determine whether routers in the area truly need to choose between paths for particular external networks.

First, consider whether the area includes any ASBRs. If not, then all external traffic must traverse the ABRs. If there's only one ABR, there's only one path for external traffic. A default route through the ABR provides the same benefit as 350 individual external routes.

If the area has multiple ABRs, internal routers might benefit from choices for external paths. But ABRs often offer roughly equivalent paths to external networks, so internal routers can simply receive a default route from both ABRs.

Special Area Types

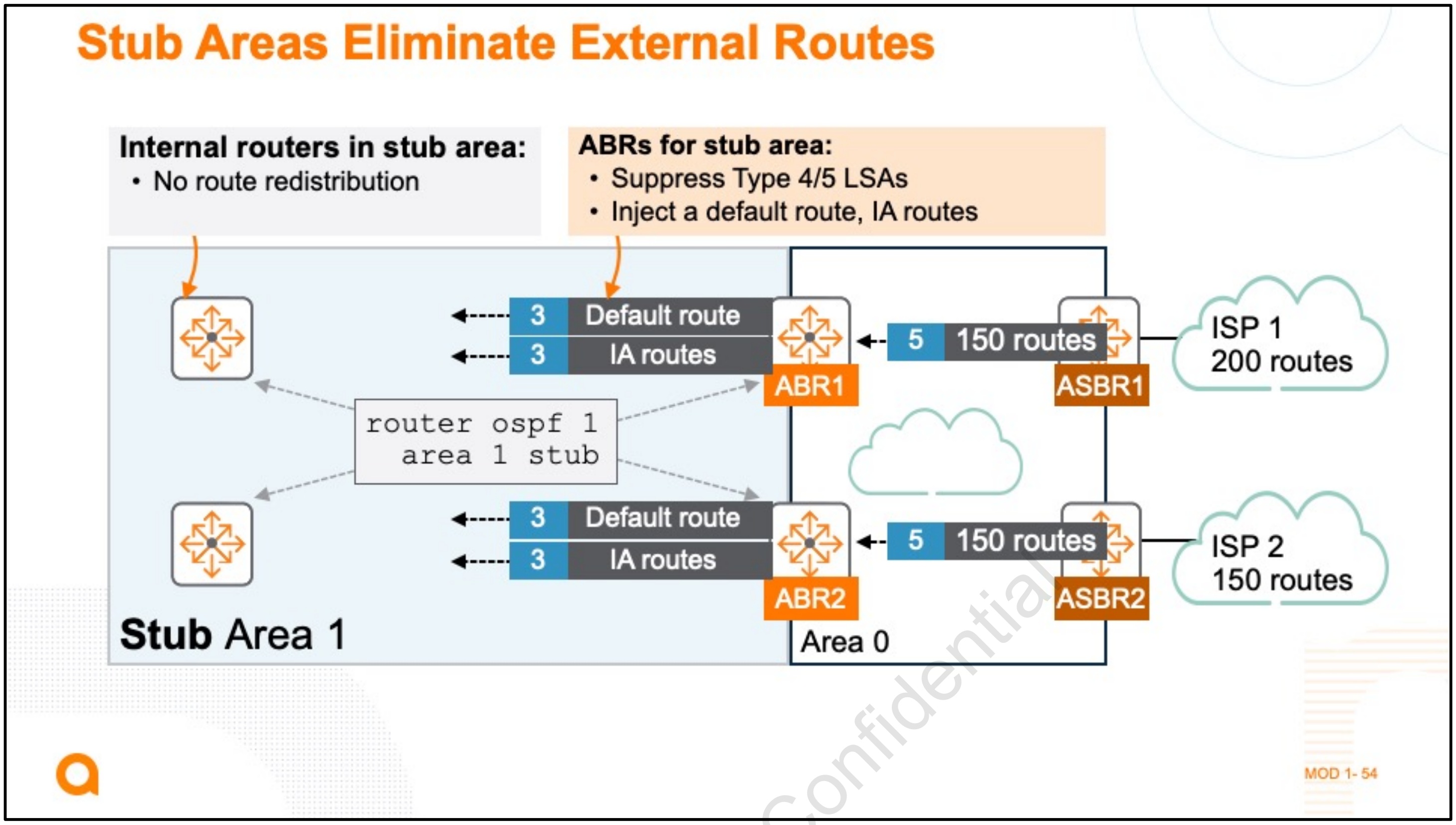


Area type	Intra-area routes	Type 3 LSAs Inter-area routes	Type 5 LSAs External routes
Normal	☑	☑	☑
Stub	☑	☑ + default route	
Totally stub	☑	Default only	

MOD 1- 53

Aruba Training-Confidential

Often an area only has one or two ABR exit points - internal routers will choose the same next hop or hops for all external routes. This is when you might choose to filter the Type 5 LSAs at the area boundary. You do so by configuring the area as a special type: stub or totally stubby. These area types reduce the type 3, 4 and/or 5 LSAs by only advertising a default route.



Stub Areas Eliminate External Routes

Internal routers in stub area:

- No route redistribution

ABRs for stub area:

- Suppress Type 4/5 LSAs
- Inject a default route, IA routes

```
router ospf 1
  area 1 stub
```

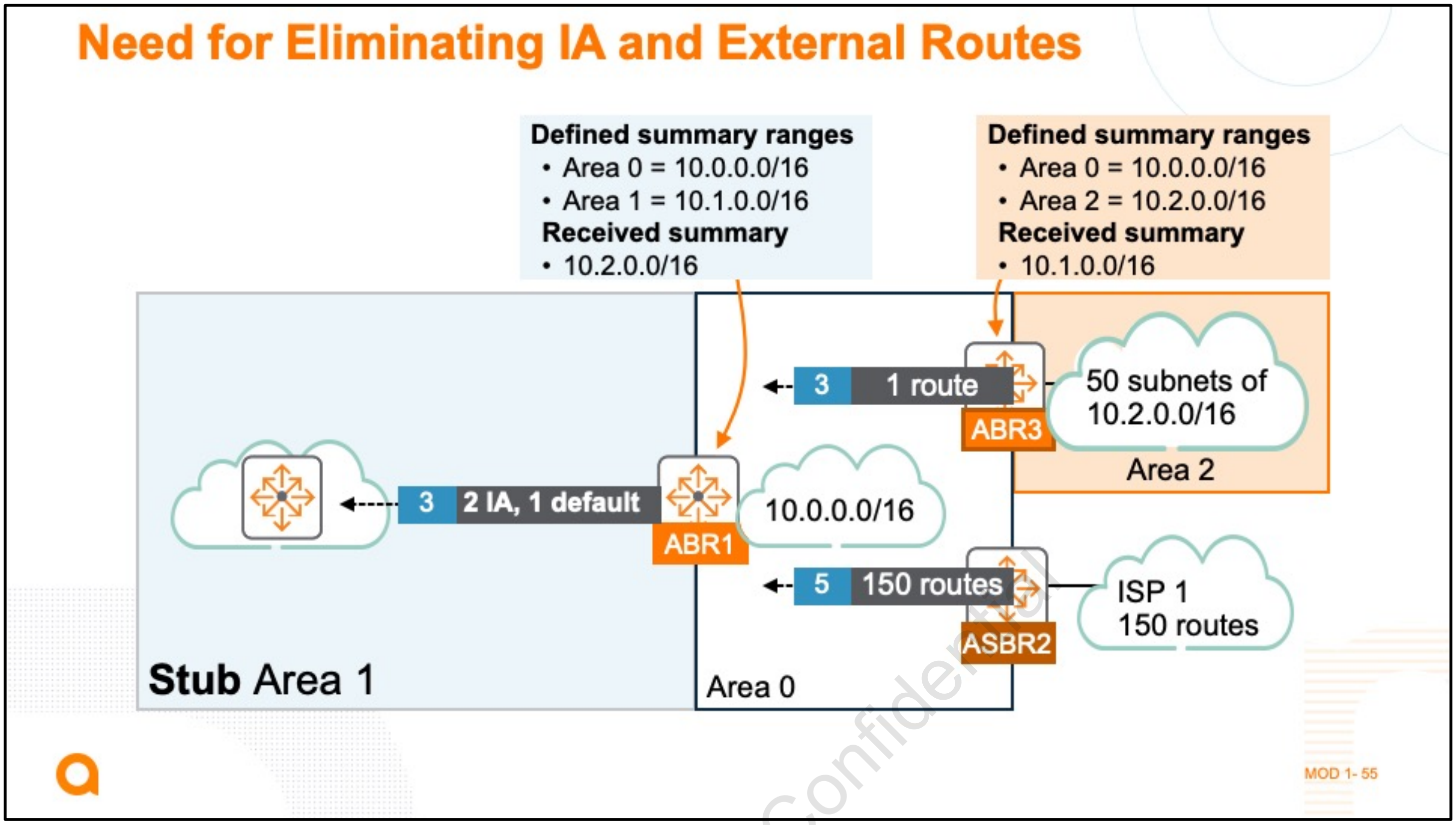
Aruba Training-Confidential

To eliminate external routes and advertisements from an area, define the area as a stub area. ABRs for stub areas do not forward Type 4 or Type 5 LSAs into those areas, and internal routers in those areas do not generate or accept them. You must define the stub setting on every OSPF router in the area – ABRs and internal routers. Use the syntax shown in the example. Routers with mis-matched area types will not become neighbors –routing is broken.

When you configure an ABR area as a stub, it automatically generates a default route and injects it into the stub area as a Type 3 LSA. As usual, the ABR assigns a metric to the default route, and you can manipulate metrics to set up one ABR as a preferred exit for external traffic.

Notice that Type 3 LSAs for inter-area (IA) routes are injected into the area, just like a normal area. Internal stub area routers receive summary LSAs for other areas through ABRs; the summary LSAs include individual or aggregated routes as configured on the ABRs. Therefore, internal routers in the stub area can continue to select the best path for inter-area traffic, based on the cost of the summary routes advertised by the ABRs.

In larger networks, you might also want to prevent these IA routes from being injected into the area.



Need for Eliminating IA and External Routes

Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 1 = 10.1.0.0/16

Received summary
- 10.2.0.0/16

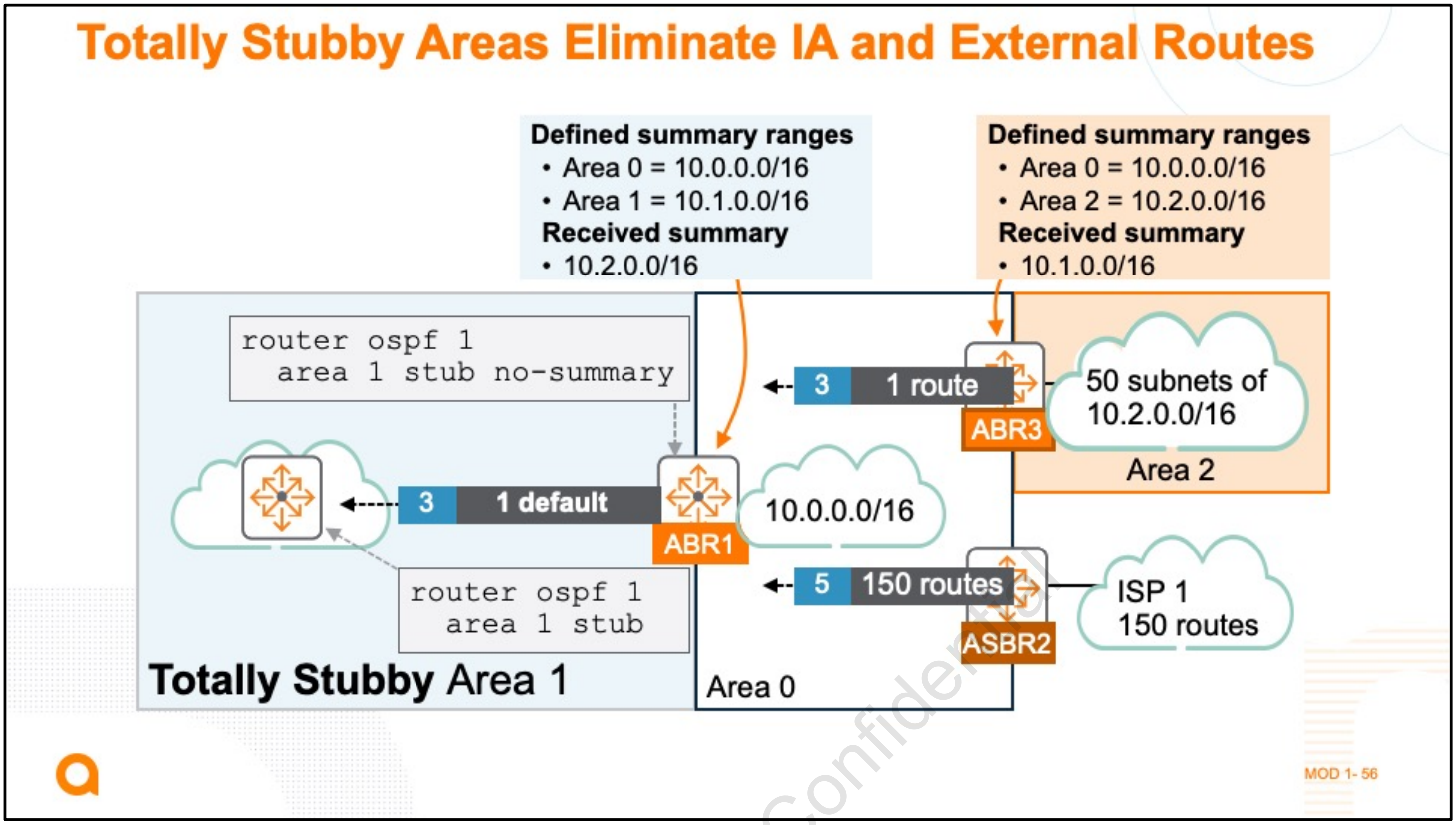
Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 2 = 10.2.0.0/16

Received summary
- 10.1.0.0/16

The figure shows a scenario where area 1 is a stub. All traffic destined to it terminates there. The area has only one ABR exit for all non-local traffic, including traffic destined to the backbone, to other OSPF areas, and to the ISP. In such an environment, the internal area 1 routers do not gain a benefit from multiple inter-area routes.

Aruba Training-Confidential

You can configure the OSPF system to suppress all inter-area routes in this area in addition to all external routes.



Totally Stubby Areas Eliminate IA and External Routes

Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 1 = 10.1.0.0/16

Received summary
- 10.2.0.0/16

Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 2 = 10.2.0.0/16

Received summary
- 10.1.0.0/16

```
router ospf 1
  area 1 stub no-summary
```

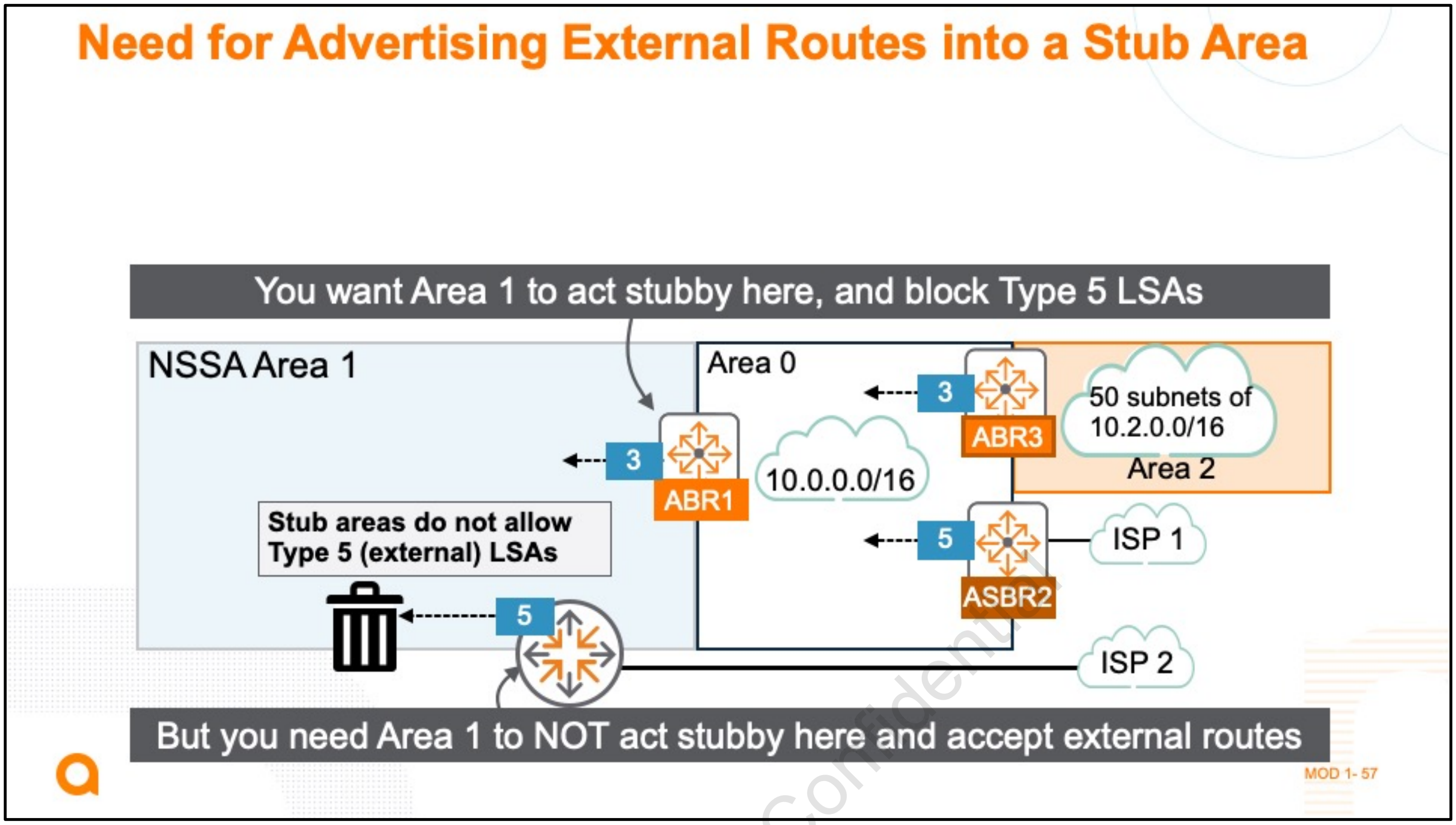
```
router ospf 1
  area 1 stub
```

Aruba Training-Confidential

You define a totally stub area by configuring the ABR or ABRs connected to that area to suppress summaries. Use the OSPF configuration shown in the figure on all ABRs that connect to the area. Configure internal, non-ABR routers in the area with the regular area 1 stub command.

The no-summary option transforms a stub area to a totally stub area – it suppresses all external route advertisements as normal for a stub area. It also prevents the ABR from generating non-aggregated Inter-Area (IA) summary routes for this area. For example, the ABR in the figure would not generate area 1 summary LSAs for known routes in area 0 and area 2. Instead, the internal routers in area 1 use the default route to reach the other areas.

The number of routes eliminated may seem fairly trivial in this example, but in more complex networks, totally stubby areas can greatly reduce the size of LSDBs and route tables.



Need for Advertising External Routes into a Stub Area

You want Area 1 to act stubby here, and block Type 5 LSAs

Stub areas do not allow Type 5 (external) LSAs

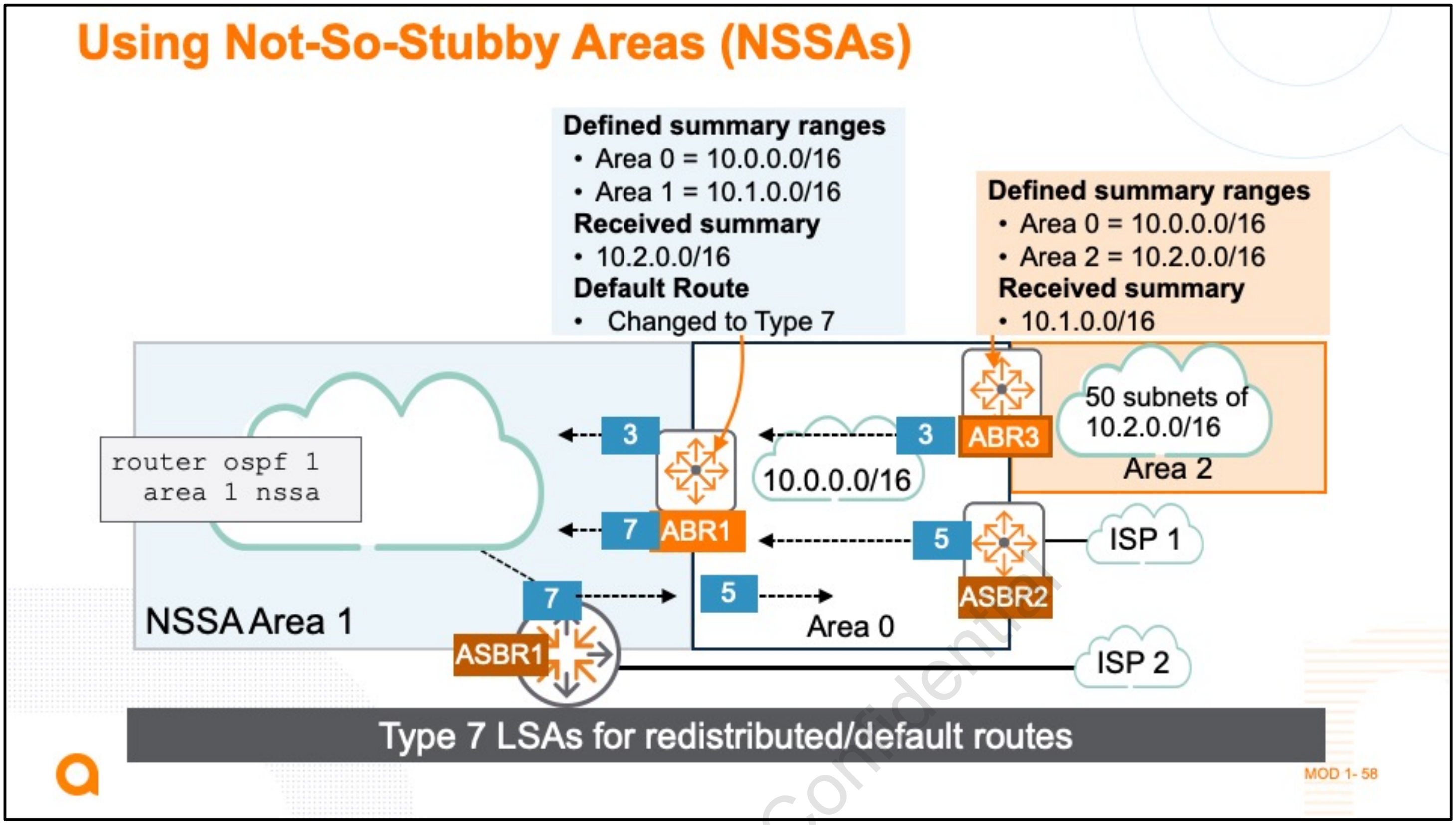
But you need Area 1 to NOT act stubby here and accept external routes

You may need to create a stub or totally stub area, to reduce table sizes and improve stability. Remember, these area types do not allow Type 5 LSAs. This means that you cannot have routers in a stub area connected to external networks, doing route redistribution, and thus creating Type 5 LSAs.

Aruba Training-Confidential

Sometimes you need this functionality. You want the area disallow all Type 5 areas like a stub, and perhaps all Type 3's, like a totally stubby area. However, you need this area to be connected to external subnets that use a foreign routing protocol (static, BGP, etc.). And remember, any redistributed route is defined as an external route. Therefore, you would encounter a similar issue if you decided to redistribute connected routes into OSPF rather than configure the associated layer 3 interfaces as passive OSPF interfaces.

You need the area to act stubby in one sense, and not so stubby in another sense.



Using Not-So-Stubby Areas (NSSAs)

Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 1 = 10.1.0.0/16

Received summary
- 10.2.0.0/16

Default Route
- Changed to Type 7

Defined summary ranges
- Area 0 = 10.0.0.0/16
- Area 2 = 10.2.0.0/16

Received summary
- 10.1.0.0/16

```
router ospf 1
  area 1 nssa
```

Type 7 LSAs for redistributed/default routes

MOD 1- 58

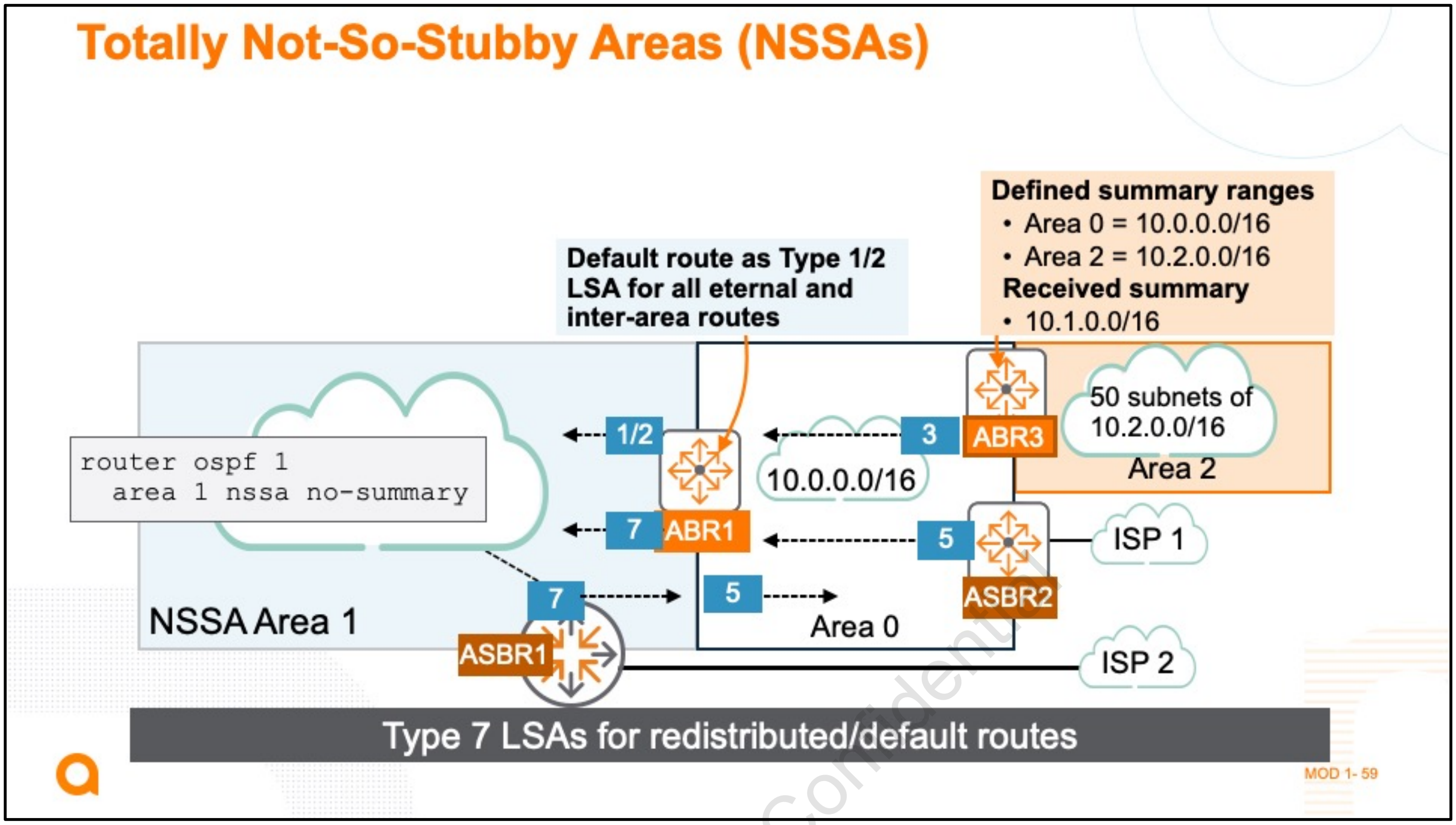
Aruba Training-Confidential

A not-so-stubby area (NSSA) furnishes the solution for this problem. You must define this area type on every OSPF router in the area, as shown in the figure.

An NSSA acts like a stub area. However, the area permits external routes that are advertised in OSPF by a local ASBR. These external routes are carried in Type 7 LSAs.

Of course, you must configure the local ASBR (ASBR1 in the figure) to redistribute routes, advertise a default route, or both, as you learned in the section on configuring ASBRs and external routes. Like other ASBRs, the NSSA ASBR assigns a metric type to the external routes that it injects into OSPF, which you can adjust to influence route selection. For NSSA external routes, these are designated as N1 and N2, instead of the normal E1 and E2 metric types.

The ABR translates the Type 7 LSAs (N1 and N2) into area 0 as Type 5 LSAs (E1 and E2) that it appears to originate. The ABR still filters Type 5 LSAs from other areas and does not advertise them in this area. It also continues to advertise a default route in the NSSA, but as a Type 7 LSA rather than a Type 3. If the NSSA's ASBR advertises a default route as well, you must adjust metrics such that routers choose the intended routes.



Totally Not-So-Stubby Areas (NSSAs)

Defined summary ranges

- Area 0 = 10.0.0.0/16
- Area 2 = 10.2.0.0/16

Received summary

- 10.1.0.0/16

Default route as Type 1/2 LSA for all eternal and inter-area routes

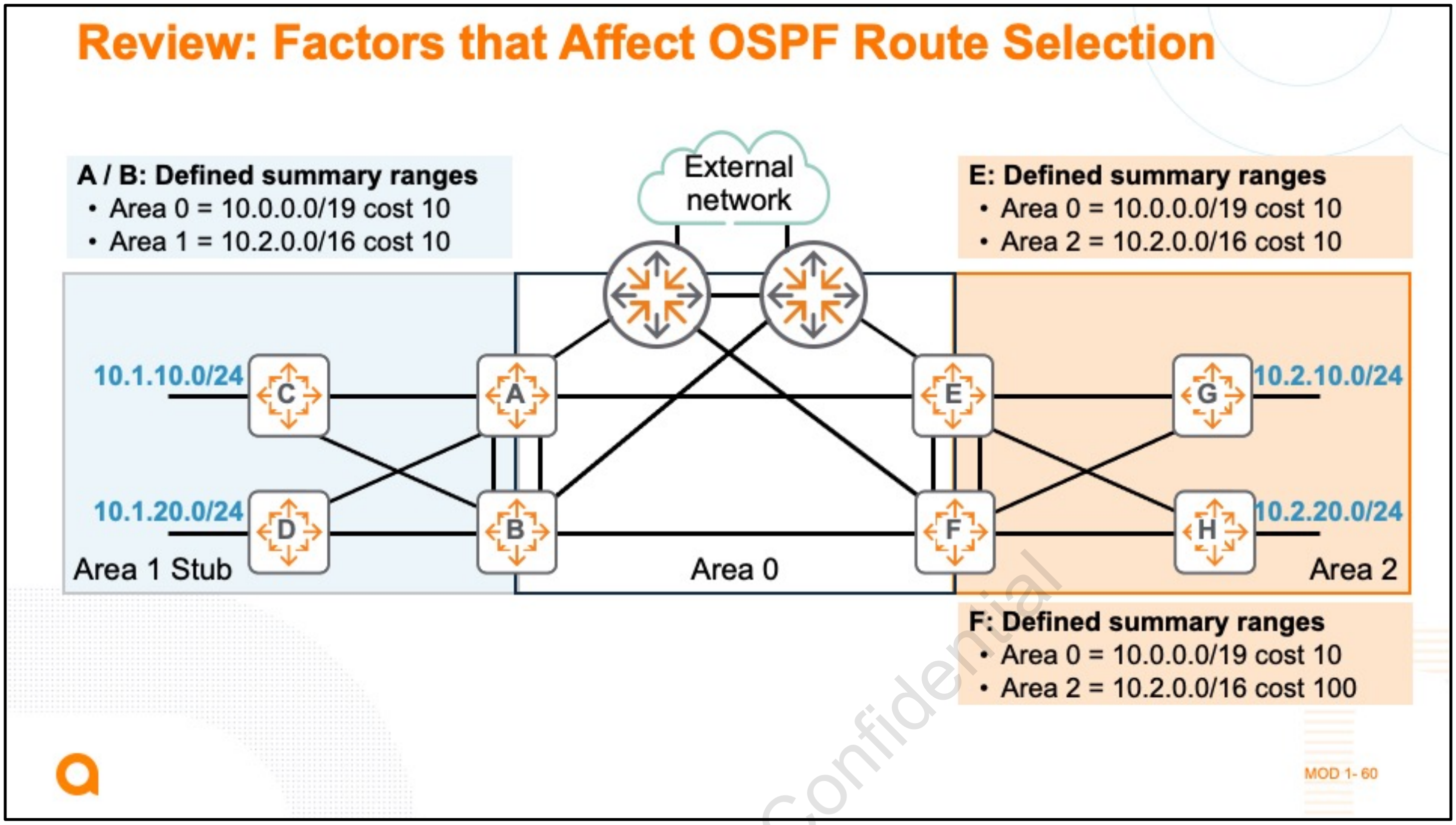
```
router ospf 1
  area 1 nssa no-summary
```

Type 7 LSAs for redistributed/default routes

Aruba Training-Confidential

You can also set up an NSSA area as a totally not-so-stubby area, similar to a totally stubby area. You only need to add the no-summary option to the area configuration, as shown in the figure.

As a totally NSSA, the ABR prevents all external or inter-area routes from entering the area. Instead, the ABR injects a default route into the area for all external and inter-area routes.



Review: Factors that Affect OSPF Route Selection

A / B: Defined summary ranges
- Area 0 = 10.0.0.0/19 cost 10
- Area 1 = 10.2.0.0/16 cost 10

E: Defined summary ranges
- Area 0 = 10.0.0.0/19 cost 10
- Area 2 = 10.2.0.0/16 cost 10

F: Defined summary ranges
- Area 0 = 10.0.0.0/19 cost 10
- Area 2 = 10.2.0.0/16 cost 100

Aruba Training-Confidential

What factors affect Router C’s next hop for 10.1.20.0/24?

Consider Router C, an internal router in area 1. Area 1 is a stub area with two ABRs, Router A and Router B. For intra-area destinations like 10.1.20.0/24, link costs alone determine the best path and selected next hop. In other words, if the cost of the C-to-B link + the B-to-D link is less than the cost of the C-to-A + the A-to-D link, C will use B as the next hop. If the costs are the same, C will learn two routes and use a hash to assign conversations to each (ECMP).

What factors affect Router C’s next hop for 10.2.20.0/24?

Link costs also play a large factor in determining how C forwards inter-area traffic such as that destined to 10.2.20.0/24. Remember that C uses the 10.2.0.0/19 route for all traffic destined to area 2 (10.2.0.0 to 10.2.31.255). You must consider the link costs from C to its ABRs and then from those ABRs to the area 2 ABRs. Then you add the cost of the 10.2.0.0/19 summary advertised by the area 2 ABR. In other words, the cost configured for the summary range on the area 2 ABR influences the path for the inter-area traffic destined to that area, but a router’s cost for reaching that ABR is also taken into account.

What factors affect Router C’s next hop for external destinations?

Because this is a stub area, Router C uses a default route to reach external destinations. It chooses the default route based on the metric advertised by the ABR and the cost for reaching the ABR. The next hop is the next hop in the best path to the advertising ABR. So, the next hop for the default route is affected by link costs within area 1 and the cost set in the area 1 stub command on the ABRs.

What factors affect Router A's next hop for external destinations?

More factors affect Router A's choice of next hop for routes to external destinations. The metric type (1 or 2) advertised by the ASBRs plays a role, as does the advertised metric itself. Link costs within area 0 can also affect the choice. If ASBRs use type 1 LSAs, these costs are incremented with the advertised metric when Router A chooses the best of two external LSAs for the same destination. Even if ASBRs use type 2 LSAs, the area 0 costs will still affect Router A's best path to the ASBR that advertises the best LSA.

If you change Area 1 to Totally Stubby, what factors affect how Router C forwards inter-area traffic?

As a final review, pretend that you have changed area 0 to a totally stub network. In this case, Router C uses its default route for both inter-area and external OSPF routes. Therefore, the ways to influence how Router C forwards inter-area traffic are to manipulate the link costs within area 1 and the cost in the area 1 stub command on the ABRs.

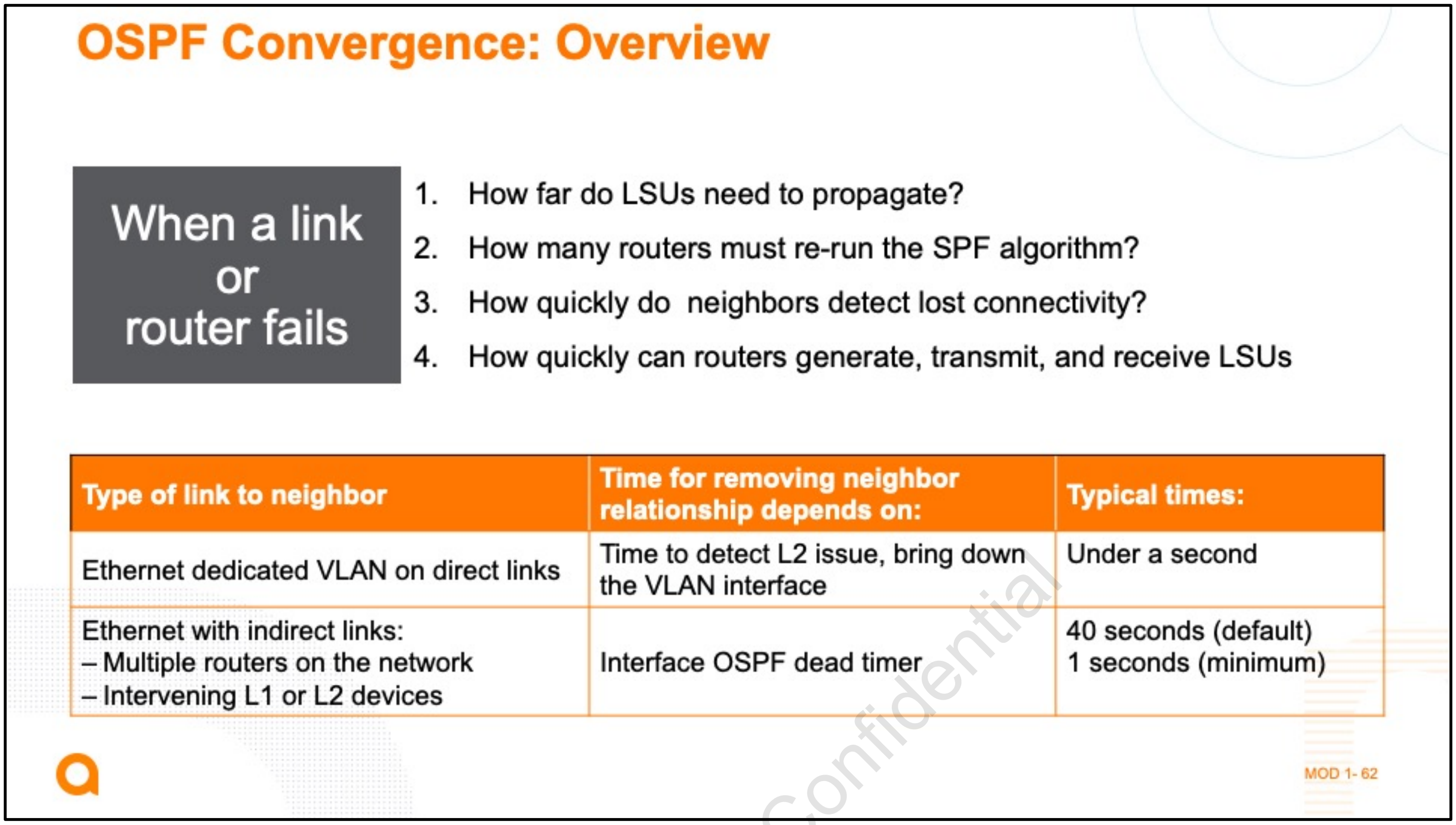
Aruba Training-Confidential

OSPF Redundancy

MOD 1- 61

Let’s look at OSPF redundancy.

Aruba Training-Confidential



OSPF Convergence: Overview

When a link or router fails

1. How far do LSUs need to propagate?
2. How many routers must re-run the SPF algorithm?
3. How quickly do neighbors detect lost connectivity?
4. How quickly can routers generate, transmit, and receive LSUs

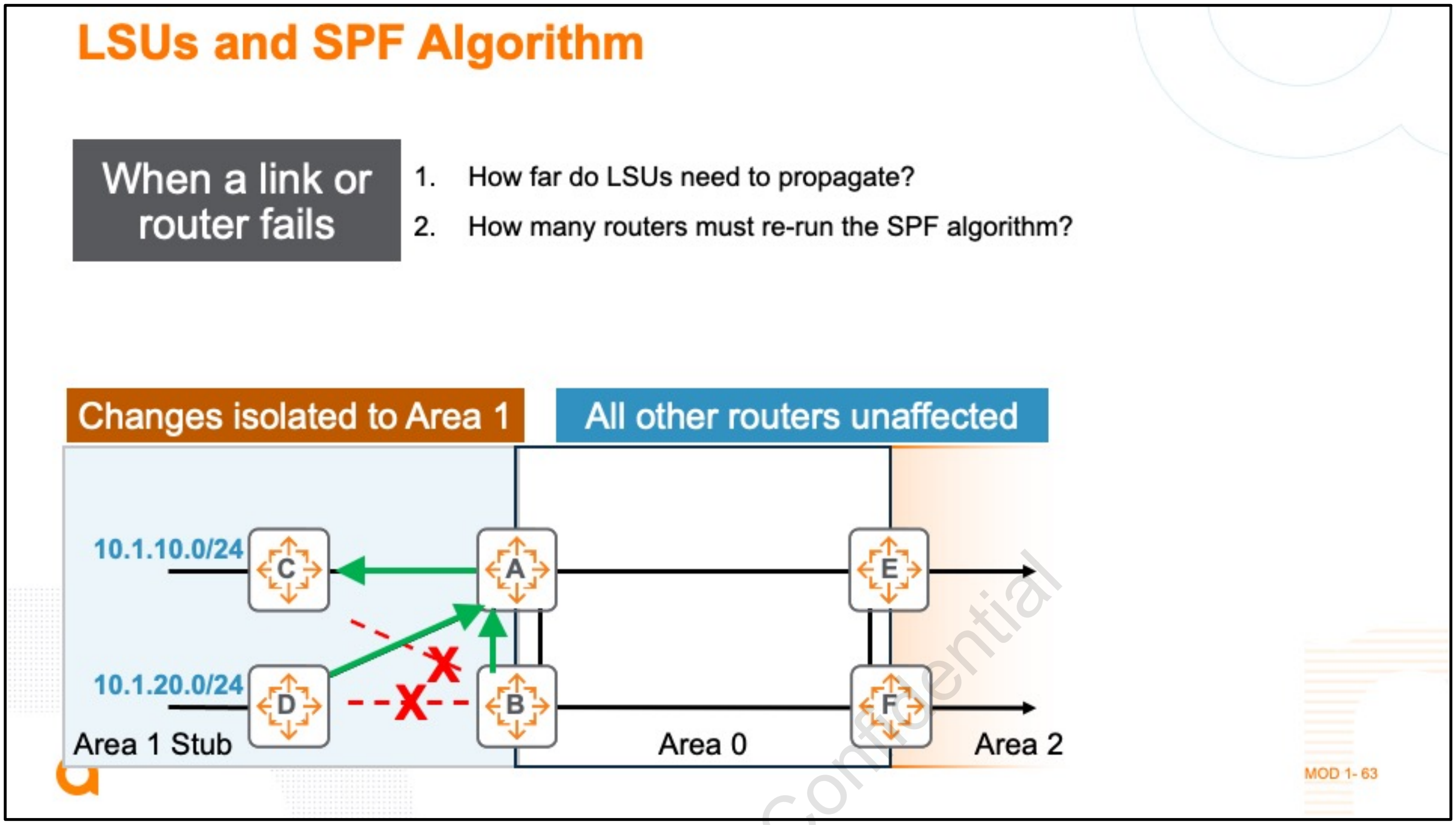
Type of link to neighbor	Time for removing neighbor relationship depends on:	Typical times:
Ethernet dedicated VLAN on direct links	Time to detect L2 issue, bring down the VLAN interface	Under a second
Ethernet with indirect links: – Multiple routers on the network – Intervening L1 or L2 devices	Interface OSPF dead timer	40 seconds (default) 1 seconds (minimum)

MOD 1- 62

In a topology with redundant links, OSPF automatically handles reconvergence if a link fails. You should consider how your design and configuration can affect reconvergence time. The figure summarizes factors that affect whether connectivity is disrupted and for how long.

Take a moment to review this material before you explore the first two questions.

Aruba Training-Confidential



LSUs and SPF Algorithm

When a link or router fails

1. How far do LSUs need to propagate?
2. How many routers must re-run the SPF algorithm?

Changes isolated to Area 1

All other routers unaffected

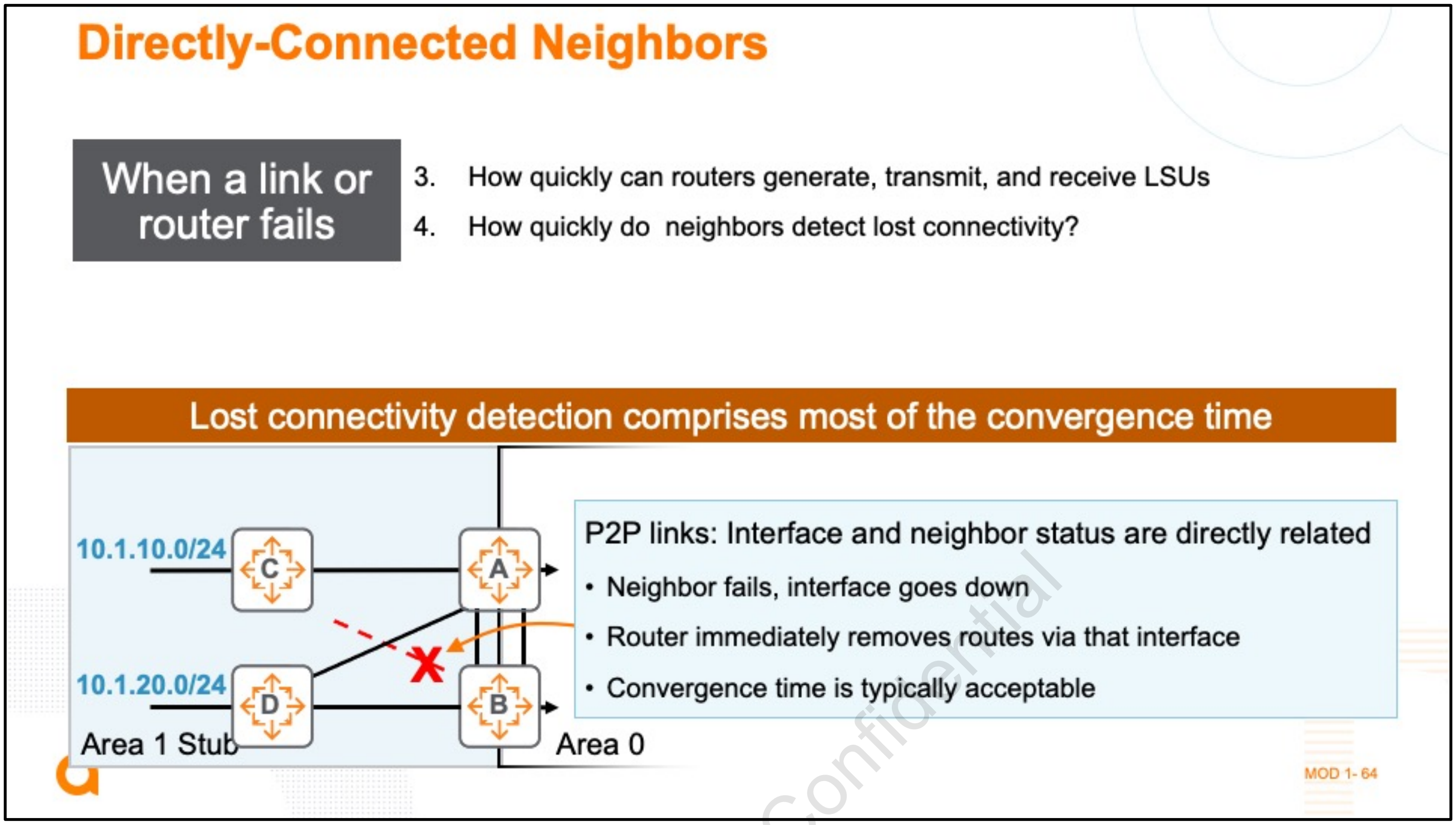
MOD 1- 63

For the first two questions, consider how what you have learned about multi-area designs help to stabilize the system and enhance convergence.

Area borders act as boundaries for changes, and often LSUs generated by a change within the area will not prompt LSUs for other areas. In addition, only changes within the OSPF area cause a router to re-run the full SPF algorithm. Updates to summary, summary-ASBR, and external LSAs require just a partial re-run. The router simply adds a new network to the list of networks reachable behind the ABR, removes a network from the list, or adjusts the cost for reaching the network. The router does not need to re-calculate and run the shortest path algorithm because it cares only about the segment of the inter-area path between itself and the proper ABR for reaching the inter-area network. Those paths have not changed.

In the figure, the link between routers C and B is down, as is the link between routers D and B. Routers in area 1 learn this information and converge on alternate paths. Router D and B route all traffic to 10.1.10.0/24 through Router A. However, Router A and Router B still have a route to this network, so they do not withdraw its Type 3 LSA for the subnet. It might need to update the cost, but usually it will not even need to do that because it is aggregating the area 1 summaries anyway. As far as routers in area 0 and area 2 are concerned nothing has changed. As you see, changes and any attendant disruptions are isolated to area 1.

Aruba Training-Confidential



Directly-Connected Neighbors

When a link or router fails

3. How quickly can routers generate, transmit, and receive LSUs
4. How quickly do neighbors detect lost connectivity?

Lost connectivity detection comprises most of the convergence time

P2P links: Interface and neighbor status are directly related

- Neighbor fails, interface goes down
- Router immediately removes routes via that interface
- Convergence time is typically acceptable

MOD 1- 64

Aruba Training-Confidential

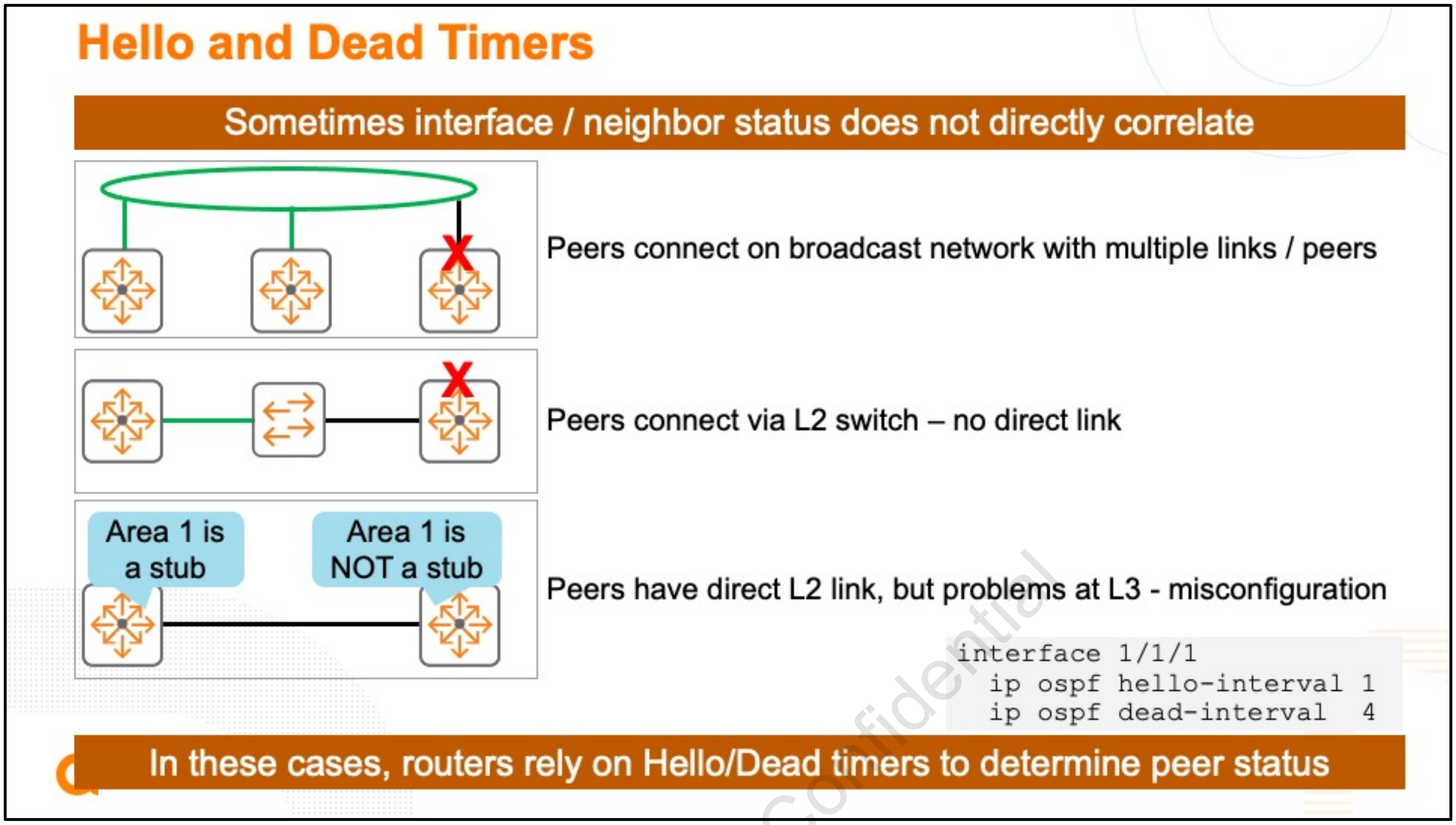
Now consider how quickly routers can generate, transmit, and receive LSUs. The default timers often provide sufficiently speedy LSU generation and propagation. When very fast convergence is required, you can adjust OSPF timers like the LSA generation interval. However, in most cases, you should leave the intervals at their defaults; tuning them is beyond the scope of this course.

In a well-designed OSPF AS with proper area boundaries, the time to detect failed neighbor relationships often determines how seamlessly OSPF routes fail over.

Suppose an OSPF router loses its next-hop connection to one of its routes. Either a link failed, or the neighbor as a whole is down. To remove the invalid route and use an alternate, the router must first detect the primary path failure. This often takes up the bulk of the time required for routes to failover.

When two neighbors have a point-to-point connection on a direct link, the status of the interface correlates to the status of the neighbor. If the neighbor fails, the interface with the direct connection goes down. The router can immediately remove any routes through that interface. Similarly when a VLAN interface is dedicated to a single physical link, the failure of that link brings the VLAN interface down. The routing switch can then remove any routes through that interface. In either case, the time to detect the Link Layer problem affects the time for detecting the failed neighbor relationship. This time varies depending on the media and is acceptable for many companies' needs.

However, sometimes the status of an OSPF interface and the status of an OSPF neighbor do not directly correlate.



Hello and Dead Timers

Sometimes interface / neighbor status does not directly correlate

Peers connect on broadcast network with multiple links / peers

Peers connect via L2 switch – no direct link

Peers have direct L2 link, but problems at L3 - misconfiguration

```
interface 1/1/1
  ip ospf hello-interval 1
  ip ospf dead-interval  4
```

In these cases, routers rely on Hello/Dead timers to determine peer status

Aruba Training-Confidential

Examples relevant to AOS-CX switches include:

The neighbors or peers connect on a broadcast network such as a VLAN but the VLAN has multiple links, and possibly multiple neighbors, associated with it. In the figure, a router is down (red X), but the other router's links remain up (green).

The neighbors are connected via a non-routing Layer 2 switch, and so do not have a direct Layer 2 connection. In the figure a router is down (red X), but the other routers connection to the switch remains up (green).

The neighbors have a direct Layer 2 connection, but problems occur at Layer 3. For example, a misconfiguration has prevented one of the neighbors from implementing OSPF on the network. In the figure, one router in the area is configured as a stub, but the other is not. You might go back to the first section of this module and review the criteria in OSPF hello packets that must match for neighbors to form an adjacency.

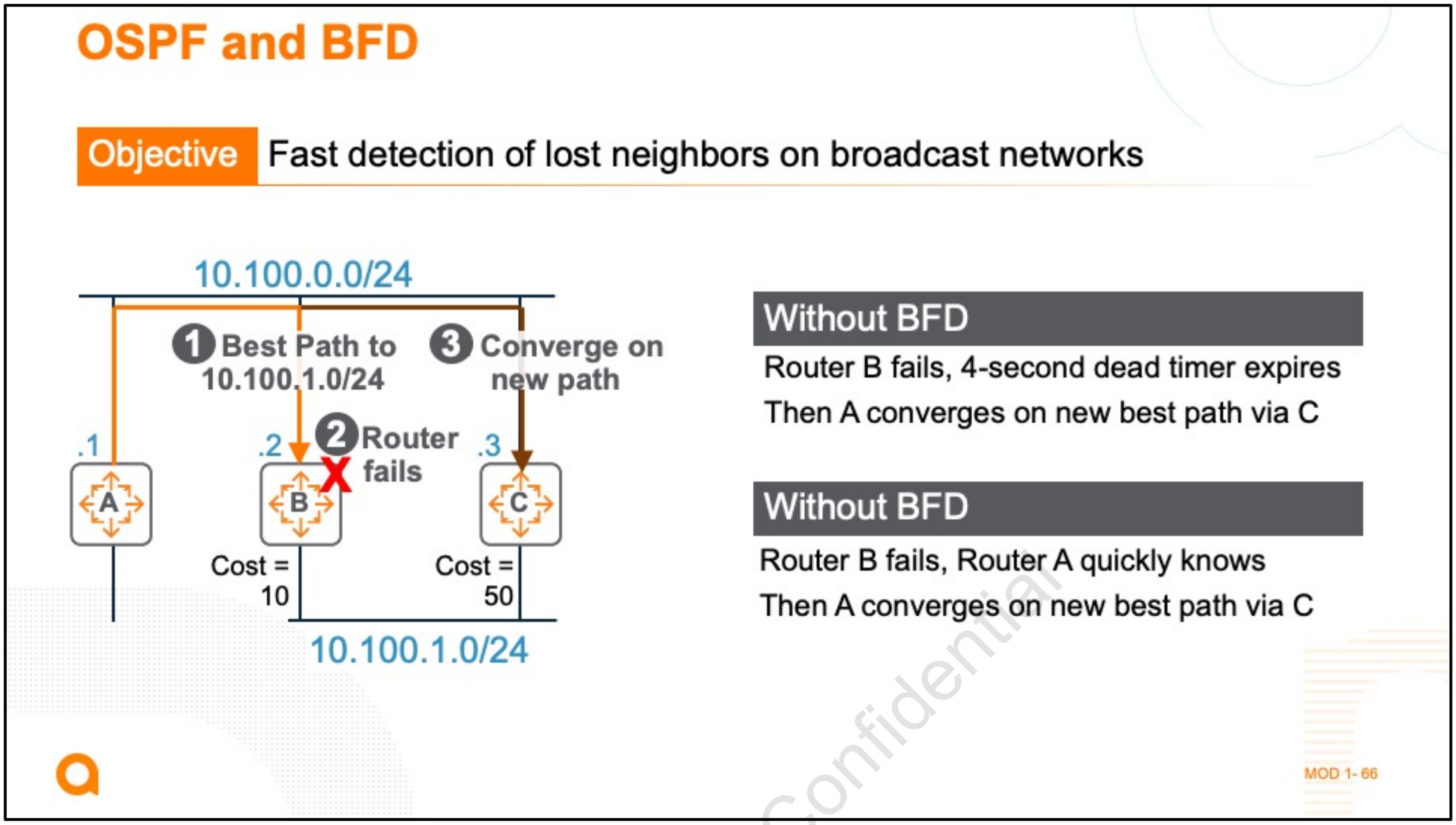
In the cases above, OSPF routers must rely on timers to determine the neighbor status. If the dead time expires without a hello from the neighbor, the router considers the neighbor down and removes routes using that neighbor as a next hop from the routing table. The OSPF dead timer is always four times the hello timer. Because the default hello timer is 10 seconds, the default dead timer is 40 seconds, which is far too long for contemporary networks.

You can adjust the hello and dead timers, but the lowest standard hello timer is 1 second, making the lowest dead timer 4 seconds, still too long for many contemporary networks. See the figure for syntax.

Note: Take care when setting the dead interval timer. The dead timer, by default, is 4 times the hello timer. Typically, it is not recommended to set the hello and dead interval timer to 1 second each, since this could cause inadvertent missed hellos.

We need faster convergence mechanisms than these old hello/dead timers.

Aruba Training-Confidential



OSPF and BFD

Objective Fast detection of lost neighbors on broadcast networks

Without BFD

Router B fails, 4-second dead timer expires

Then A converges on new best path via C

Without BFD

Router B fails, Router A quickly knows

Then A converges on new best path via C

MOD 1- 66

Aruba Training-Confidential

There are several ways you can help OSPF respond quickly and gracefully in the event of topology changes such as failed link.

Bidirectional Forwarding Detection (BFD) tests the connectivity between two IP addresses in a BFD session. BFD reports to OSPF when connectivity is lost. Routers use that information to take appropriate actions, depending on functions tied to BFD.

When you enable BFD on an OSPF layer 3 interface, BFD assumes responsibility for testing neighbor connectivity. If the BFD session fails, the router removes the BFD neighbor relationship, and so can remove all LSAs through that neighbor and update routing paths. The interface still sends out OSPF hellos, but BFD will generally detect any problems first.

In the figure, three routers connect on subnet 10.100.0.0/24. Routers B and C also connect to network 10.100.1.0/24. Router A uses Router B as the next hop for 10.100.1.0/24 because Router B advertised the lowest cost.

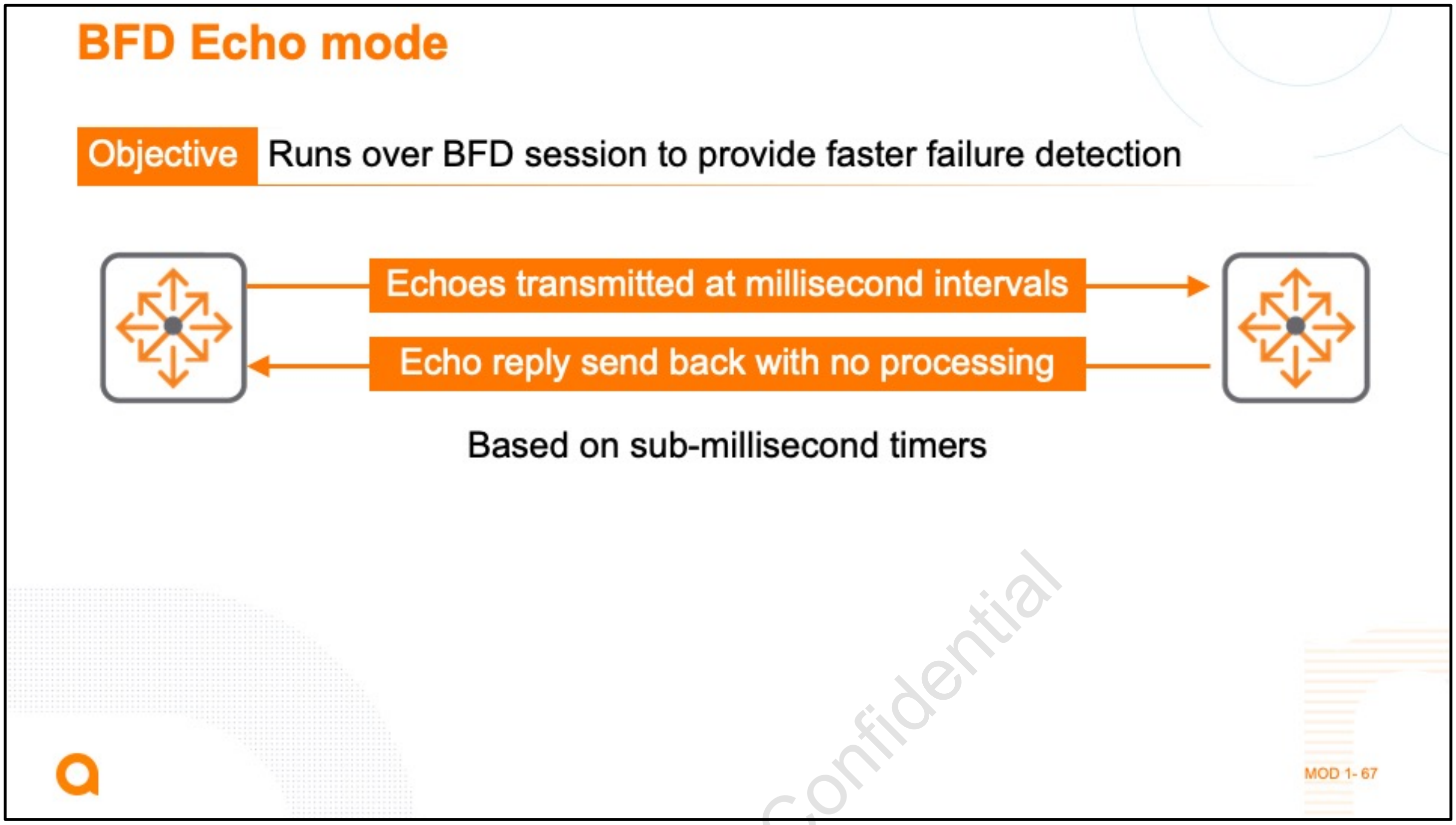
If BFD is not in use and Router B fails or its link on 10.100.0.0/24 fails, Router A continues sending traffic to 10.100.1.0/24 through 10.100.0.2 until its neighbor relationship times out. This could be 40 seconds if the dead timer is at its default value, or as fast as 4 seconds if you lower the dead timer to the recommended safe minimum. Because 10.100.0.2 is no longer up, this traffic does not reach its destination.

With BFD, Router A quickly detects that 10.100.0.2 is no longer available – typically this is less than 1 second. BFD removes the neighbor relationship, causing Router A to lose adjacency with Router B. Router A can then immediately start using Router C as the next hop to 10.100.1.0/24.

Note: If an OSPF neighbor does not support BFD, the switch continues to try to establish a BFD session with that neighbor. It lists the session as down, but, because the session was never up, the down state does not cause the switch to tear down the neighbor relationship.

What makes BFD so much faster?

Aruba Training-Confidential



BFD Echo mode

Objective Runs over BFD session to provide faster failure detection

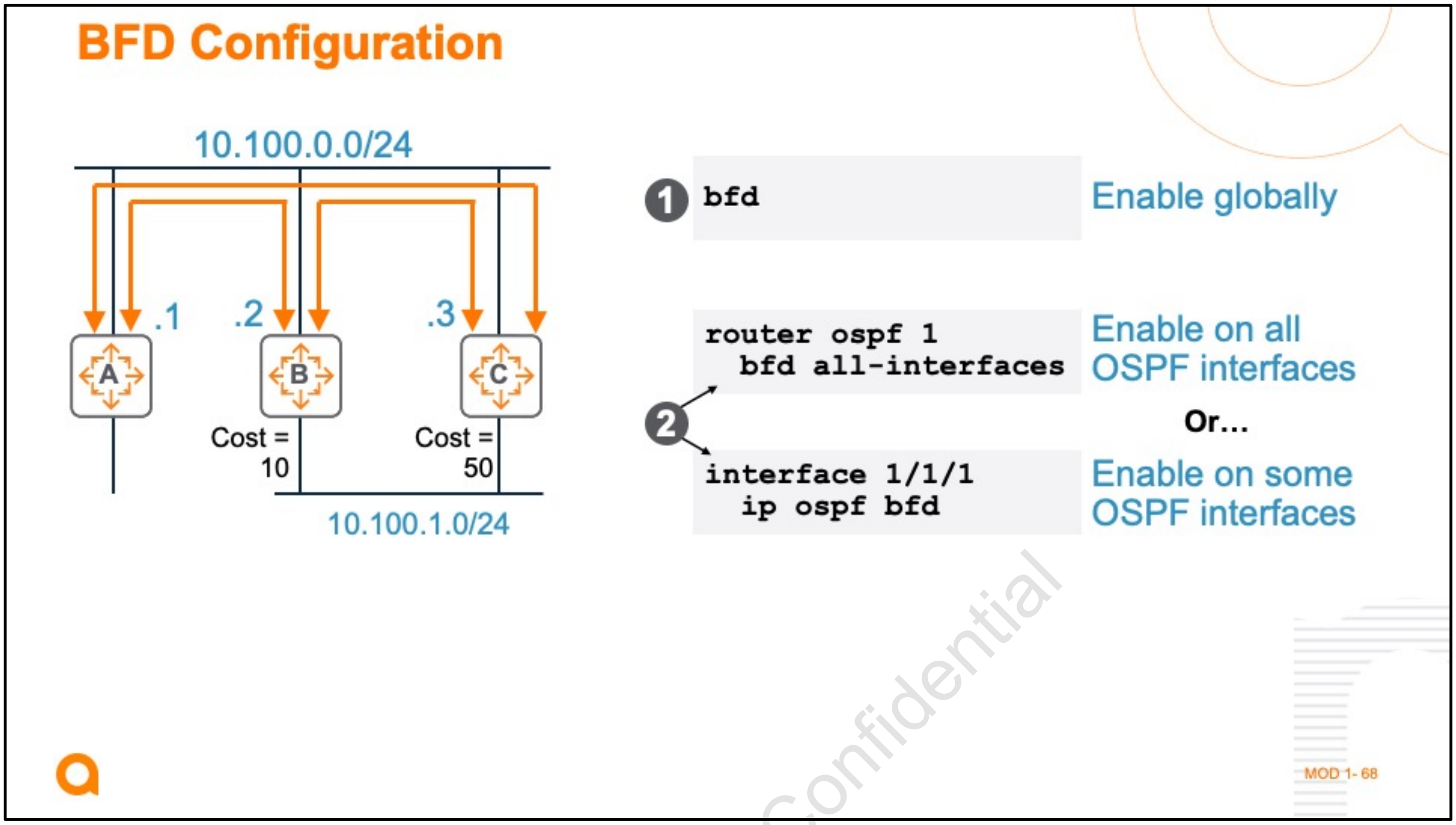
Echoes transmitted at millisecond intervals

Echo reply send back with no processing

Based on sub-millisecond timers

When you configure BFD, echo mode is enabled by default. In Echo mode, routers periodically send BFD echo packets. The peer router returns the received BFD echo packets back without processing them. If the sender does not receive BFD echo packet from the peer within the specified interval, the session is considered down. Because peers respond to echo packets without processing them, detection is FAST.

Aruba Training-Confidential



BFD Configuration

1
```
bfd
```
Enable globally

2
```
router ospf 1
  bfd all-interfaces
```
Enable on all OSPF interfaces

Or...

```
interface 1/1/1
  ip ospf bfd
```
Enable on some OSPF interfaces

MOD 1- 68

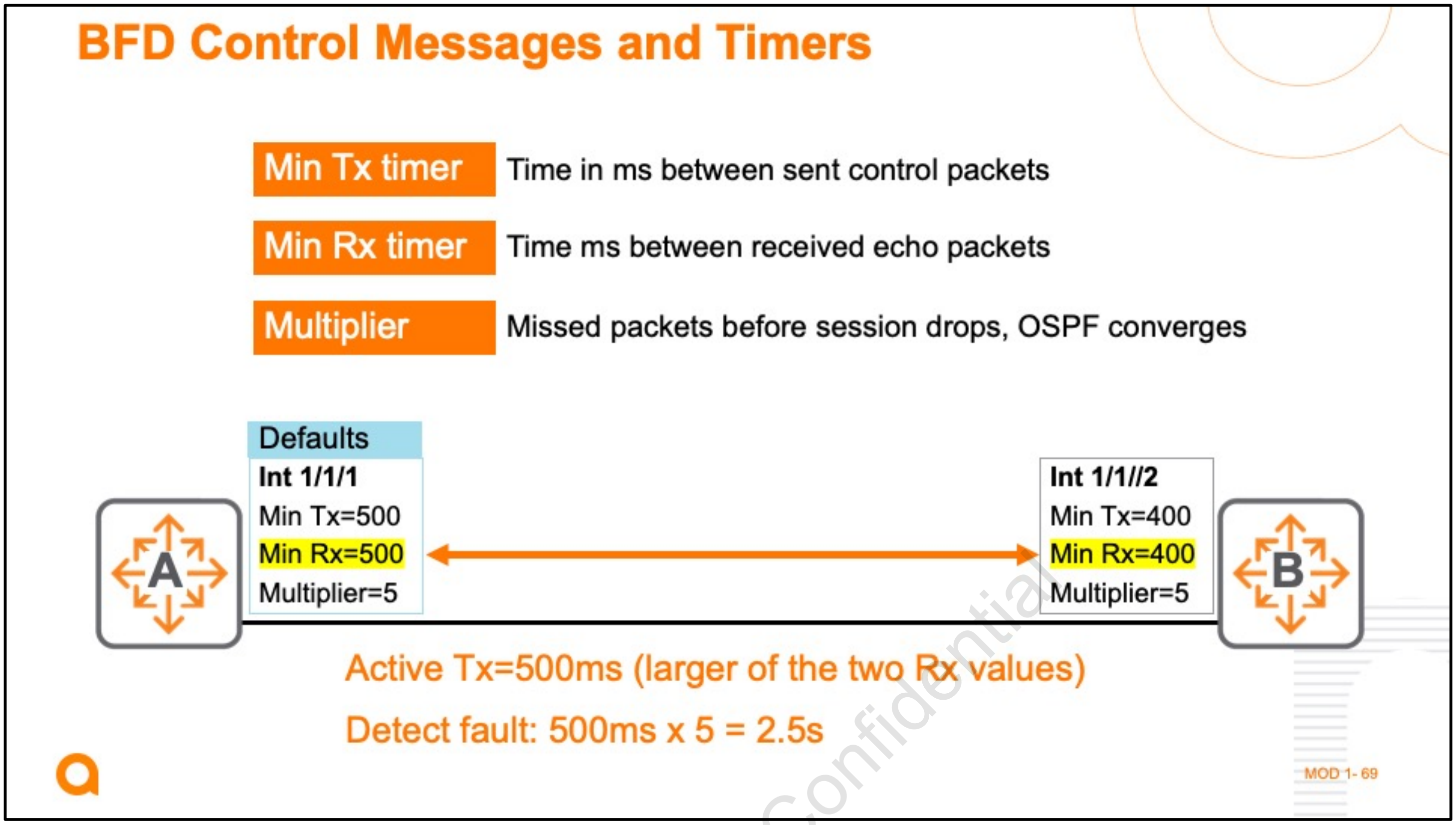
Aruba Training-Confidential

First enable BFD globally, from global config mode, as shown in the figure. When you enable BFD, it automatically enables in echo mode. There is a bfd echo disable command, which disables echo mode, but this is not recommended. Use the no form of the command to re-enabled.

Next you can either enable OSPF BFD on all interfaces in the OSPF context. Or you can enter the desired Layer 3 interface context, and enable BFD on that interface specifically, as shown in the figure

Because AOS-CX switches use active mode, the interface automatically attempts to establish a BFD session with each OSPF neighbor on that interface. It transmits a unicast BFD control packets on UDP port 3784 to each neighbor. As you see in the figure, a BFD session is point-to-point. In this example, the 10.100.0.0/24 network has three OSPF routers, each of which implements OSPF BFD. Each router attempts to establish a session with each of its two neighbors for six total BFD sessions on the network.

To establish a BFD session, the neighbors simply determine that both support BFD and negotiate timers.



BFD Control Messages and Timers

Min Tx timer	Time in ms between sent control packets
Min Rx timer	Time ms between received echo packets
Multiplier	Missed packets before session drops, OSPF converges

Active Tx=500ms (larger of the two Rx values)

Detect fault: 500ms x 5 = 2.5s

MOD 1- 69

Aruba Training-Confidential

Each neighbor interface has three settings that control convergence time.

Minimum Transmit (Tx) timer: Minimum time between sent echo (control) packets. Default = 500 milliseconds (ms)

Minimum Receive (Rx) timer: Minimum time between expected echo packets received from peer. Default = 500ms

Detection Multiplier: number of consecutive missed packets received before declaring the session dead and informing OSPF to converge on an alternate path

For most deployments, the default values should not be changed. If you must change these settings, use the syntax shown below.

```
(config-if)# bfd min-transmit-interval=500
(config-if)# bfd-min-echo-receive-interval=500
(config-if)# bfd detect-multiplier=5
```

In the figure, Router A has been left at the default values, while Router B’s timers have been lowered to 400ms. Interestingly, this has no effect on convergence times.

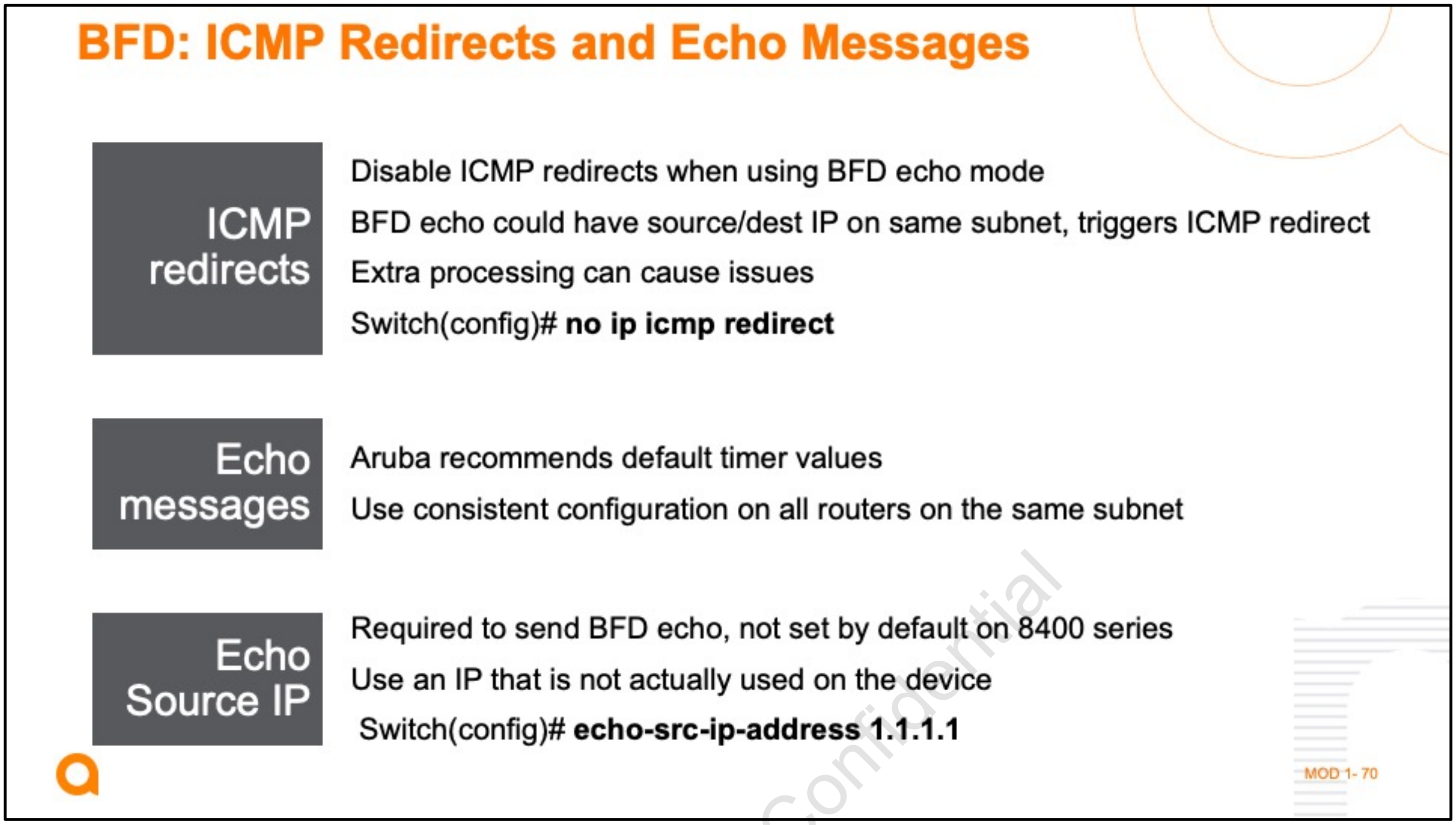
The two routers share echo/control packets and compare values. The actual “Active Tx” timer used is the larger of the local peer setting vs the remote peer settings. Thus, the actual transmit interval is negotiated to 500ms. Of course, this scenario is only to help you understand how timers work. You should be consistent with your configurations, and typically leave timers at default values.

The multiplier is 5, and so 5 consecutive packets missed, at 500ms apart = 2.5 seconds before the BFD session is torn down, which causes the OSPF neighbor relationship to be declared down - even if the dead timer has not yet expired. You can lower the multiplier to 4 or 3 if you have a reliable link that loses little traffic. If you set the Min Tx and Min Rx timers to the minimum value (1 second), do not set the detect multiplier lower than 3. Understand that lowering these values can be dangerous – especially if you have links that occasionally become congested. This could cause routers to converge when links are down, and reduce overall system stability.

Aruba Training Confidential

The active Tx timer and detect time on the two neighbors in a BFD session do not have to match. However, they often do.

When the switch uses the control message timeout alone to determine when a session fails, it operates in asynchronous mode. In this mode, BFD does not provide sub-second detection of issues, which is often what you want from BFD. Therefore, you should usually enable OSPF BFD to operate in echo mode.



BFD: ICMP Redirects and Echo Messages

ICMP redirects

Disable ICMP redirects when using BFD echo mode

BFD echo could have source/dest IP on same subnet, triggers ICMP redirect

Extra processing can cause issues

Switch(config)# **no ip icmp redirect**

Echo messages

Aruba recommends default timer values

Use consistent configuration on all routers on the same subnet

Echo Source IP

Required to send BFD echo, not set by default on 8400 series

Use an IP that is not actually used on the device

Switch(config)# **echo-src-ip-address 1.1.1.1**

MOD 1- 70

Aruba Training-Confidential

ICMP redirects

It is also best practice to disable ICMP redirects on the switches that use BFD in echo mode. Before you enable OSPF BFD, you should disable ICMP redirects on the AOS-CX switch. In some cases, the echo could have a source and destination on the same subnet, which would usually trigger the switch to send an ICMP redirect. The extra processing can cause issues on the switch. Disabling ICMP redirects prevents these issues.

```
Switch(config)# no ip icmp redirect
```

Echo messages

If you want, you can adjust the BFD timers, but Aruba recommends the default values (see the AOS-CX Command-Line Interface Guide for more details on how to do this.

You should set up BFD in the same way on all the OSPF routers connected on the interface.

In echo mode, control messages still maintain the BFD session, but echo messages, which are transmitted at subsecond intervals, allow BFD to respond to issues much more quickly. The interface implementing BFD sends an echo message to the other side of the session. The other side of the session simply echoes the message back without processing it. If the interface does not receive its echo message back, BFD takes down the session and the OSPF neighbor relationship. (The switches still exchange control messages in echo mode, but these become less important for determining when the session has failed.)

Echo source IP address

By default, AOS-CX switches support echoing echo messages back to their BFD neighbors at a 500 ms interval. However, they will not send echo messages themselves until you set a global BFD echo source for the packets, which is not set by default on the 8400s. Until you set the address, the BFD session will have echo mode disabled. You should choose a unique “dummy” IP address that is not used on this device.

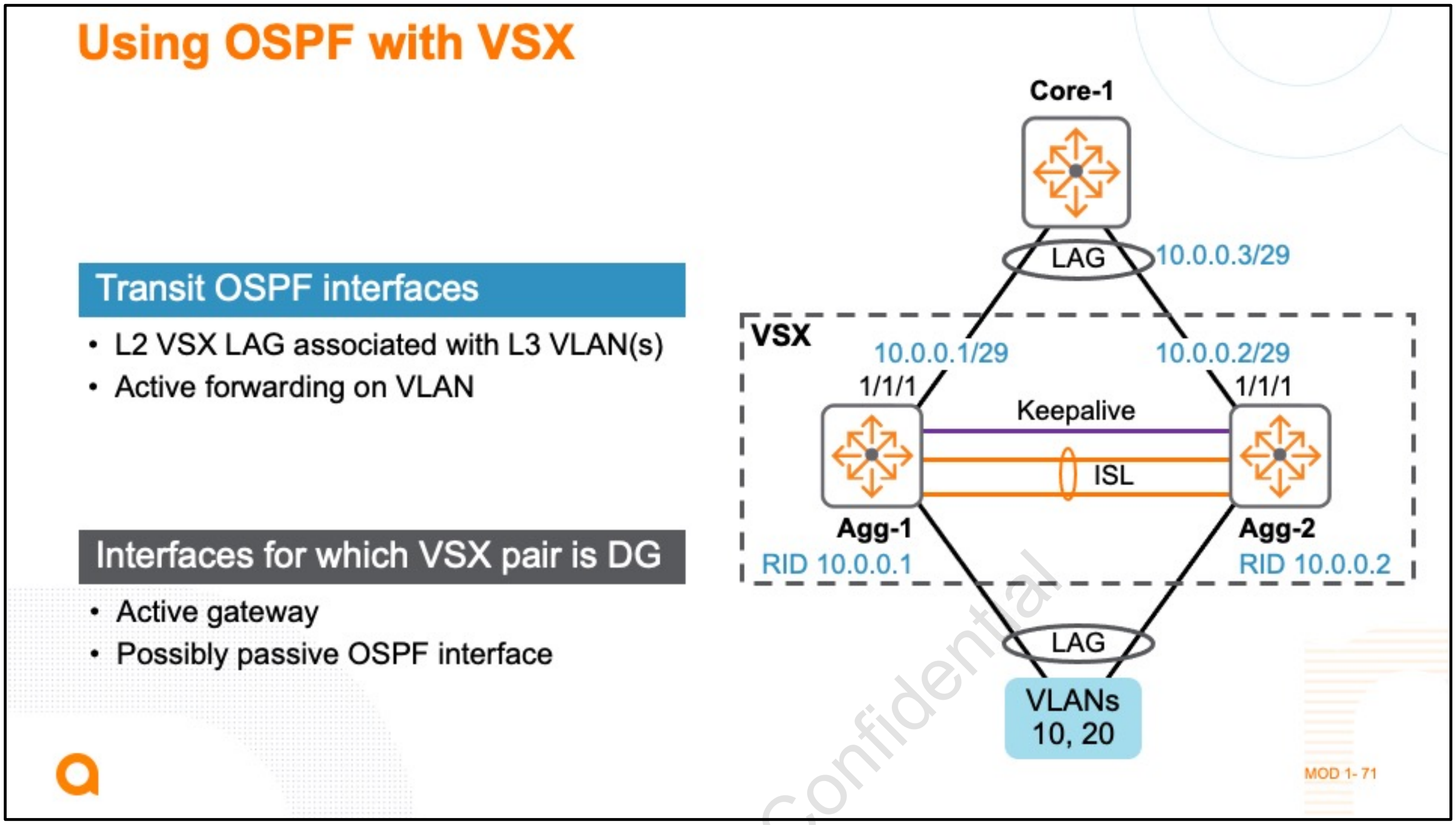
On the 8400 switches, you should next set a global echo mode IP address, which the switch includes as the source IP address for the echo packets in all sessions. Remember that the destination is the switch’s own IP address, so the source is just a “dummy” address. It should be an IP address on a subnet you’re not using on the switch or the network. You cannot set this address on 8320 or lower switches.

```
Switch(config)# echo-src-ip-address <IPV4-ADDR>
```

This sets the source IPv4 address for BFD echo packets. This address is used in all echo sessions. The source IP address must not be on the same network segment as any switch interface, otherwise a large number of ICMP redirect packets may be sent by the remote device, causing network congestion.

The switch uses this as the source for the echo packets. The destination is the switch’s own IP address, which allows the neighbor to send message back without processing it. Setting the source IP address to a different subnet ensures that the neighbor does not have to process the packet and send ICMP redirects.

Aruba Training-Confidential



Using OSPF with VSX

Transit OSPF interfaces

- L2 VSX LAG associated with L3 VLAN(s)
- Active forwarding on VLAN

Interfaces for which VSX pair is DG

- Active gateway
- Possibly passive OSPF interface

Aruba Training-Confidential

Each VSX member runs OSPF separately. In the figure, aggregation layer switches Agg-1 and Agg-2 are a VSX pair. They have a VSX LAG to Core-1, and each switch has an IP address on the Layer 3 VLAN interface assigned to the LAG. If you were using multiple VRFs, the LAG would carry multiple VLANs; but in this example, the LAG has just one VLAN.

Core-1 learns ECMP routes to VLAN 10 and 20 subnets through both Agg-1 and Agg-2. Because the ECMP next hop selection algorithm will not necessarily match the link aggregation selection algorithm, the core switch might send traffic through next hop 10.0.0.1 but on the link to Agg-2. The active forwarding, which is set per-VLAN, enables Agg-2 in this situation to route the traffic itself even though the traffic was destined to 10.0.0.1. You should typically enable active forwarding on every OSPF interface that is a transit network.

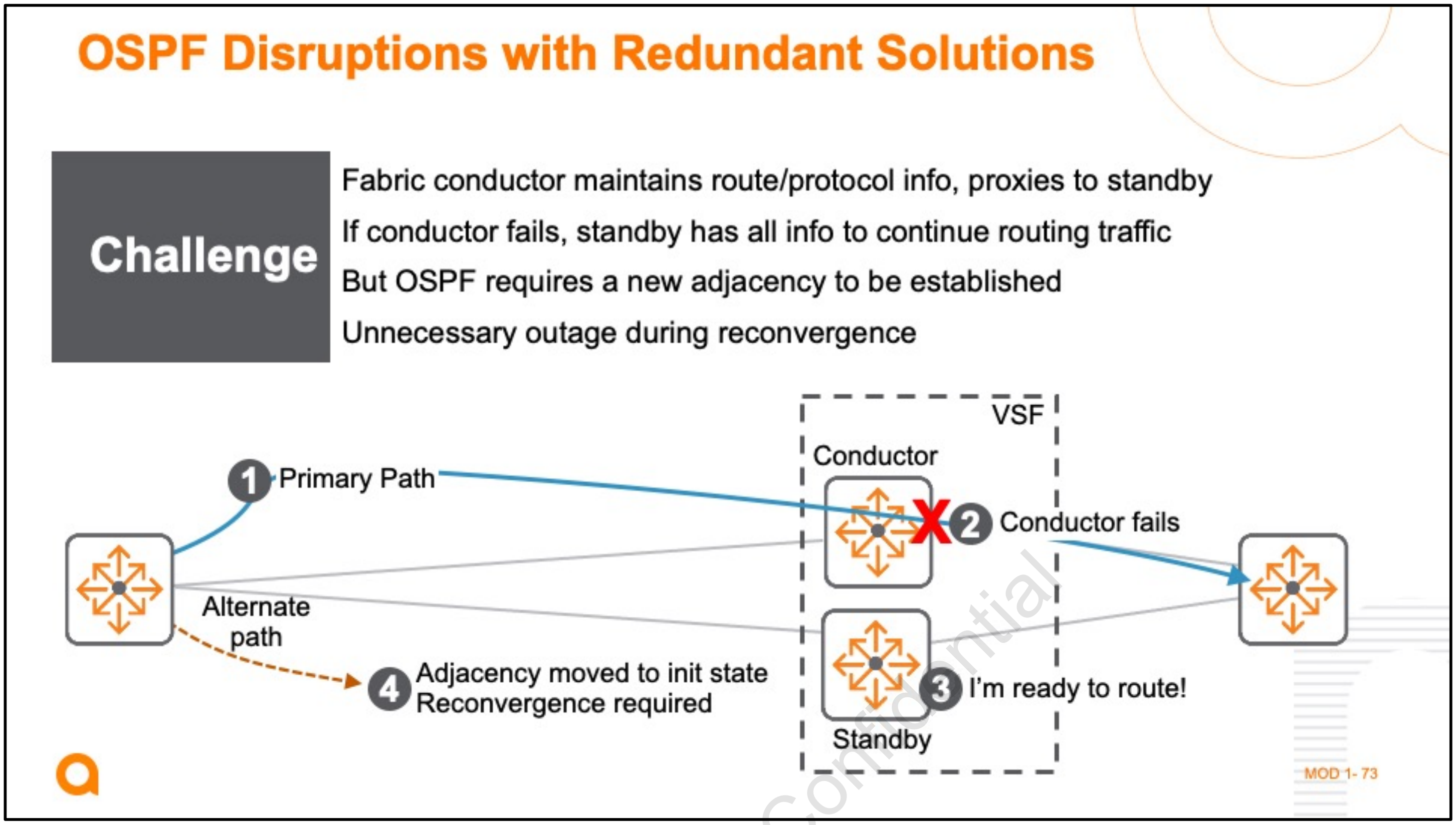
The VSX pair is probably the default gateway for subnets such as those associated with VLANs 10 and 20 . You should typically set up the active gateway feature on those VLANs. Recall that this feature enables the VSX switches to act as active-active default gateways for a subnet. You might run OSPF on these subnets, but in passive mode - VSX peers need not become OSPF neighbors on all subnets. Alternatively, you could redistribute connected routes into OSPF.

Additional OSPF Features

MOD 1- 72

Now you look at several additional features – perhaps needed in certain situations.

Aruba Training-Confidential



OSPF Disruptions with Redundant Solutions

Challenge

Fabric conductor maintains route/protocol info, proxies to standby

If conductor fails, standby has all info to continue routing traffic

But OSPF requires a new adjacency to be established

Unnecessary outage during reconvergence

Aruba Training-Confidential

Consider a scenario where some OSPF routers must also implement other protocols for high availability:

Virtual Switching Framework (VSF) (preferred when possible on the 6300s)

Virtual Router Redundancy Protocol (VRRP)

VSX

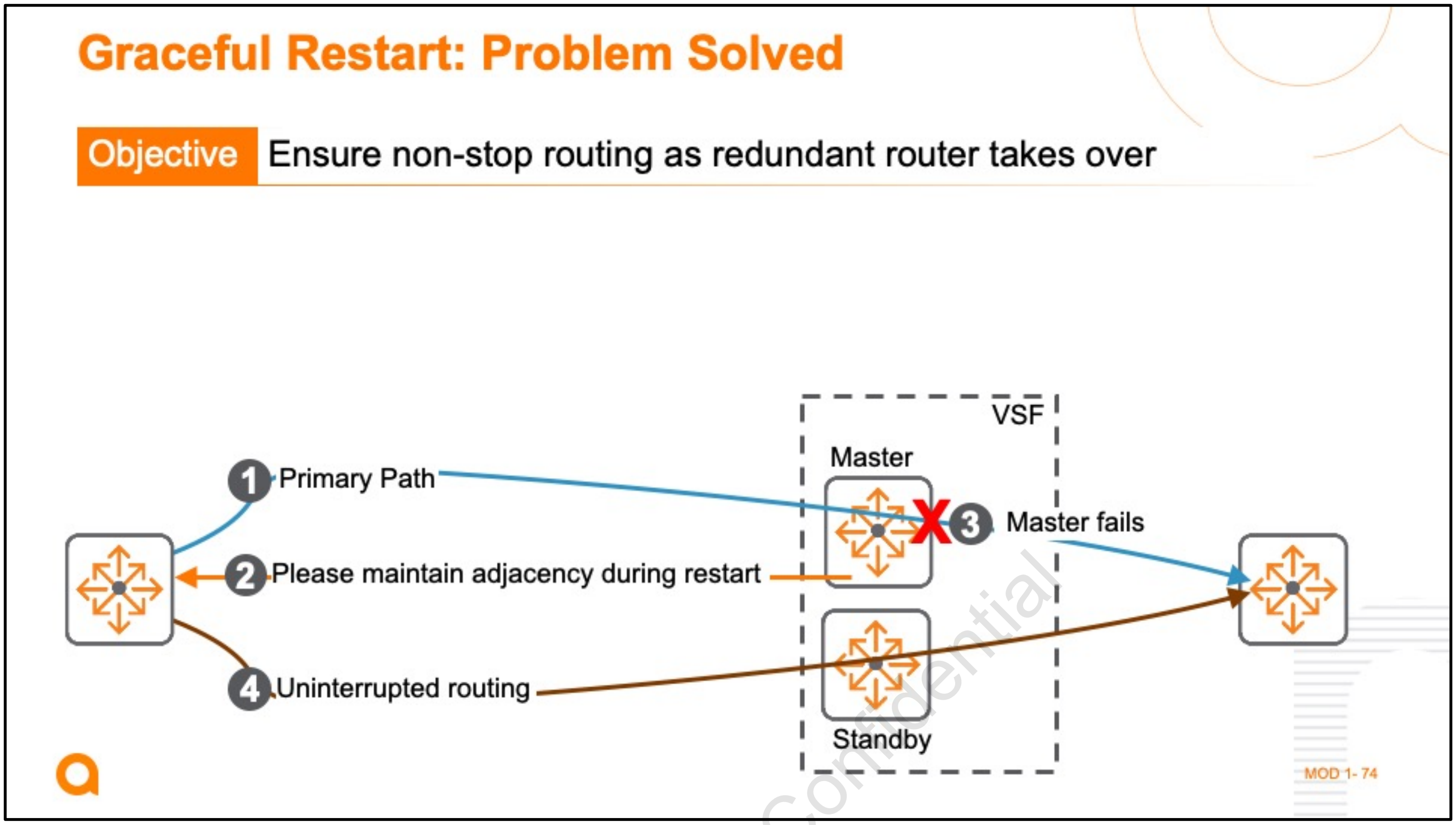
However, even though default gateway functions are maintained, failure of the VSF master or VRRP master could still disrupt routing and connectivity. OSPF graceful restart ensures non-stop routing as the standby member takes over as the master. This is illustrated in the figure.

The VSF fabric master maintains routing tables and protocol information in its control plane. In a chassis switch, the master's primary management module maintains this information. The master proxies this information to other members' control planes, but if the master fails, the routing processes must restart on the new master.

The VSF members have routes proxied to the Forwarding Information Bases (FIBs) in their control planes (or the control planes of their LPUs, in the case of chassis-based switches). They can continue to route traffic with these stale routes until the new master builds the routes in its own control plane—and, when the VSF fabric is using static routes or Routing Information Protocol (RIP) routes, they do seamlessly.

However, standard OSPF does not allow OSPF neighbors to entirely resynchronize the LSDBs without tearing down the old adjacency relationship. Therefore, when the new master tries to re-establish adjacency, the VSF fabric's neighbors take down the neighbor relationship and send out LSUs as if the fabric were no longer available. Routes reconverge away from VSF fabric—even though the fabric is perfectly capable of routing the traffic—causing unnecessary disruption to the traffic flow and possible interruption to service.

Aruba Training-Confidential



Graceful Restart: Problem Solved

Objective Ensure non-stop routing as redundant router takes over

VSF

Master

Standby

1 Primary Path

2 Please maintain adjacency during restart

3 Master fails

4 Uninterrupted routing

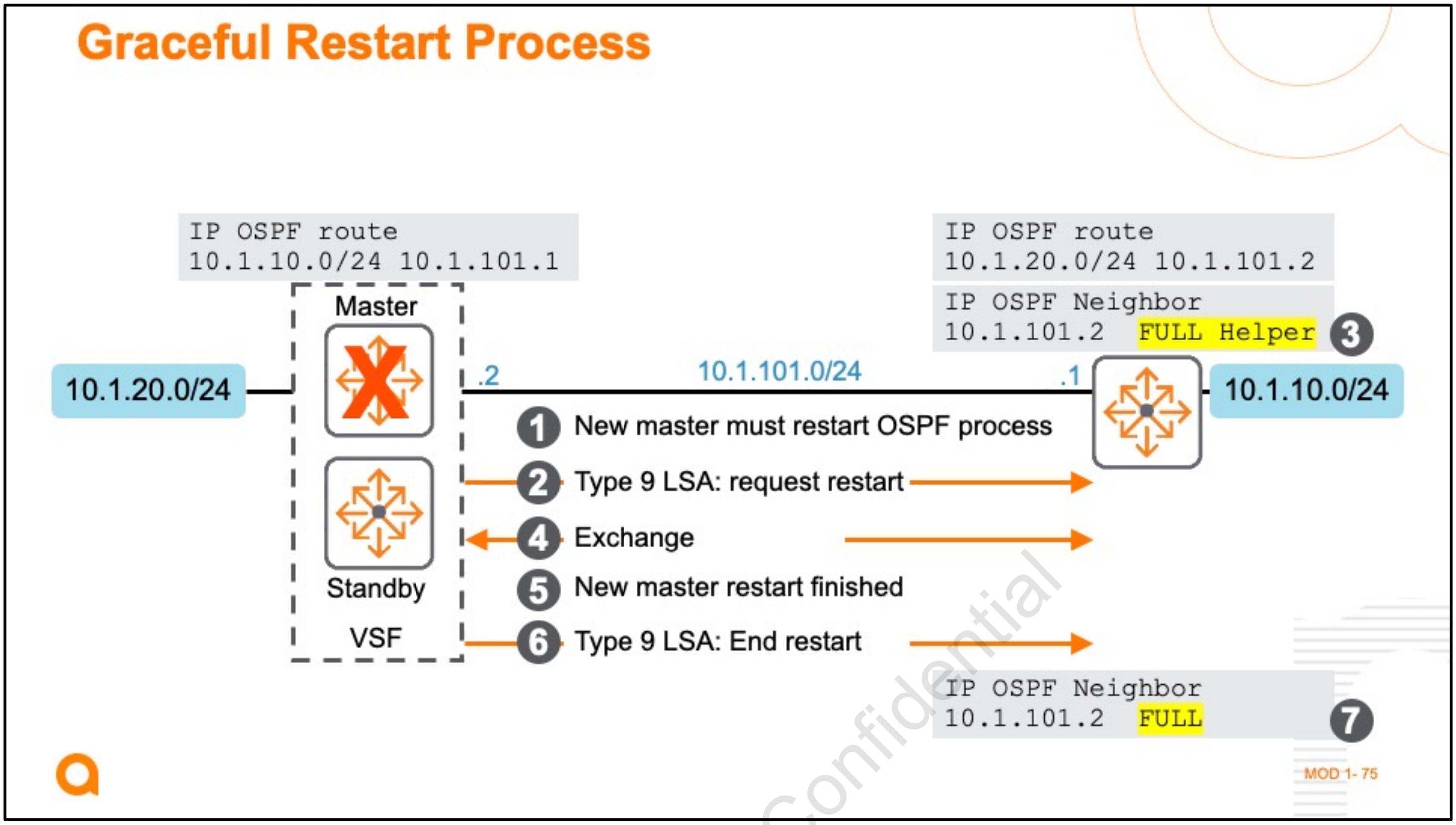
MOD 1- 74

Aruba Training-Confidential

OSPF graceful restart enables the VSF fabric to ask its neighbors to maintain the neighbor and/or adjacency relationship undisturbed while the new master restarts the OSPF process. Thus, traffic continues toward its destination undisturbed.

An OSPF router with redundant management modules should also implement OSPF graceful restart. The failover of the primary management module to the standby module is analogous to the failover of the VSF master to the standby member. Although the OSPF router or routing switch can use its stale routes to forward traffic, it must also send OSPF graceful restart requests to inform its neighbors not to drop its LSAs while the new management module rebuilds the adjacencies.

Note that OSPF graceful restart also helps to maintain routing system stability in any circumstance in which an OSPF router must restart its OSPF processes, including when a network administrator restarts the process for maintenance purposes (such as applying configuring a new router ID).



Graceful Restart Process

IP OSPF route
10.1.10.0/24 10.1.101.1

IP OSPF route
10.1.20.0/24 10.1.101.2

IP OSPF Neighbor
10.1.101.2 FULL Helper 3

1 New master must restart OSPF process

2 Type 9 LSA: request restart

4 Exchange

5 New master restart finished

6 Type 9 LSA: End restart

IP OSPF Neighbor
10.1.101.2 FULL 7

MOD 1- 75

Aruba Training-Confidential

When an OSPF router restarts the process—for example, because a main management module fails and a redundant one takes over—it must re-establish neighbor relationships. This triggers neighbor routers to clear the router and its LSAs out of the topology. Routing is disrupted across the network until the router finishes restarting the process. AOS-CX switches support OSPF graceful restart, which lets an OSPF router restart the process without causing disruptions for traffic.

Here's the IETF standard graceful restart process:

The GR Restarter is a router that uses graceful restart. It determines the need to restart its OSPF process.

The GR Restarter sends a Type 9 opaque LSA (grace LSA) to neighbors, asking them to move to graceful helper mode. The LSA includes:

- A grace period timer, which specifies how long the neighbors stay in graceful helper mode
- A graceful restart reason:
  - 0 = Unknown
  - 1 = Software restart
  - 2 = Software upgrade
  - 3 = Failover from an active to a standby management module, or, for VSF, failover from master to standby
- The GR Restarter's IP address

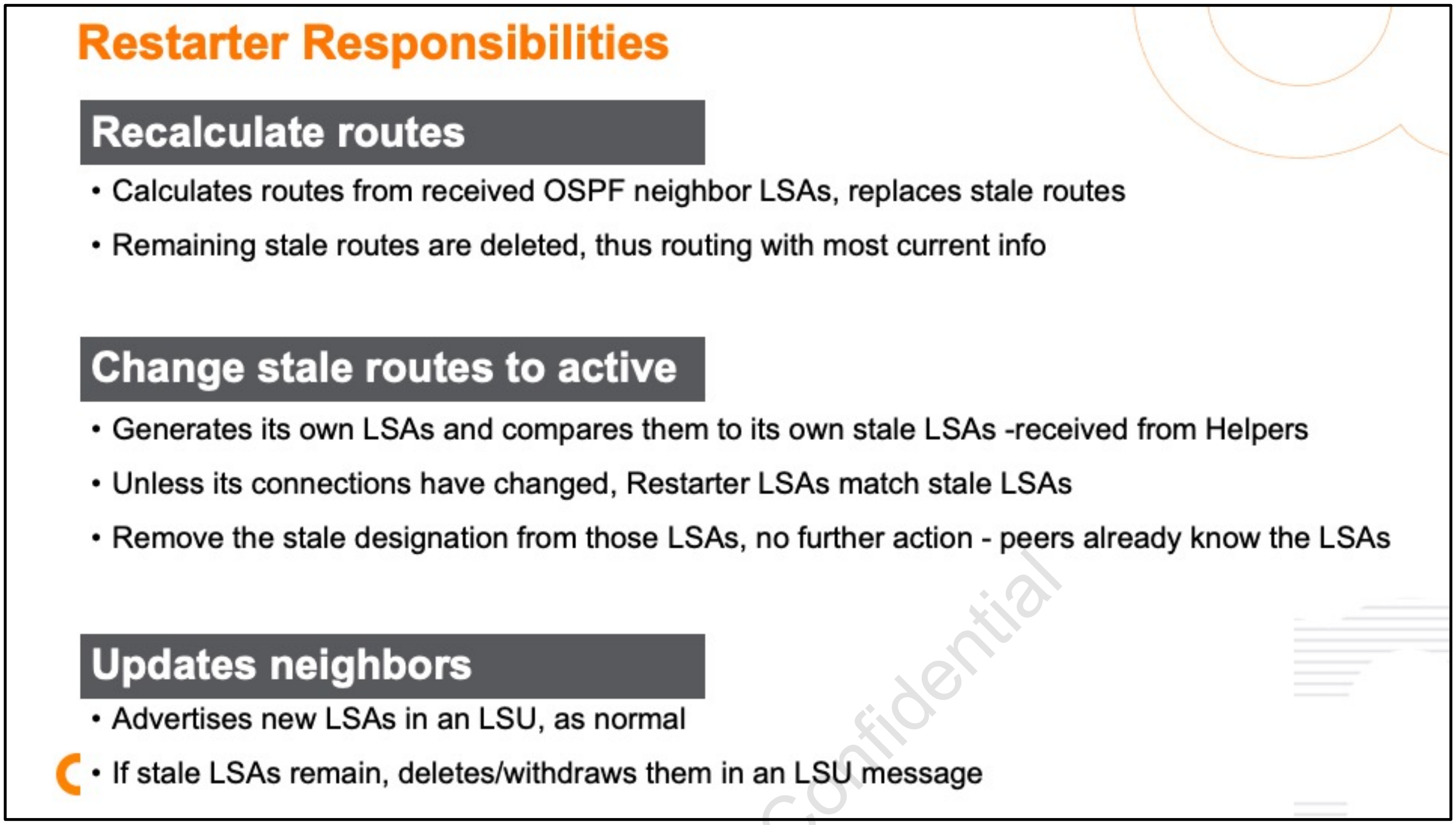
Neighbors must be configured as IETF standard GR Helpers. They place the GR Restarter in graceful restart mode. In this mode, GR Helpers act as if they have full adjacency with the GR Restarter, regardless of what is occurring. Thus, Helpers continue routing traffic to the Restarter and maintain LSAs from it, trusting that the Restarter can truly continue to route traffic with its stale routes.

During this time, the GR Restarter is establishing neighbor relationships and adjacencies with the GR Helpers. As always during this process, the switch obtains LSAs advertised by OSPF routing devices throughout the area. Some of these LSAs were advertised by itself (or the previous VSF master); the Restarter labels those LSAs as stale.

After the GR Restarter establishes an adjacency with all neighbors, it sends a grace LSA, asking Helpers to leave helper mode; i.e., grace period time is set to 0.

The GR Helpers leave helper mode. However, because they have achieved adjacency with the GR Restarter, they continue to route traffic to it without interruption.

Aruba Training-Confidential



Restarter Responsibilities

Recalculate routes

- Calculates routes from received OSPF neighbor LSAs, replaces stale routes
- Remaining stale routes are deleted, thus routing with most current info

Change stale routes to active

- Generates its own LSAs and compares them to its own stale LSAs -received from Helpers
- Unless its connections have changed, Restarter LSAs match stale LSAs
- Remove the stale designation from those LSAs, no further action - peers already know the LSAs

Updates neighbors

- Advertises new LSAs in an LSU, as normal
- If stale LSAs remain, deletes/withdraws them in an LSU message

Aruba Training-Confidential

GR Restarter has continued routing traffic with stale routes. However, it must ensure the truth of what it asked the GR Helpers to trust: it knows the correct routes, and the topology that it advertised before the restart remains accurate. During and after the process of establishing adjacency, the GR Restarter recalculates routes, changes stale routes to active routes, and determines whether any updates are necessary:

The GR Restarter calculates routes from the LSAs that it received from OSPF neighbors. The new routes replace the stale routes. If any stale routes remain, they are deleted as no longer valid. The GR Restarter now knows that it is routing traffic with the most up-to-date information.

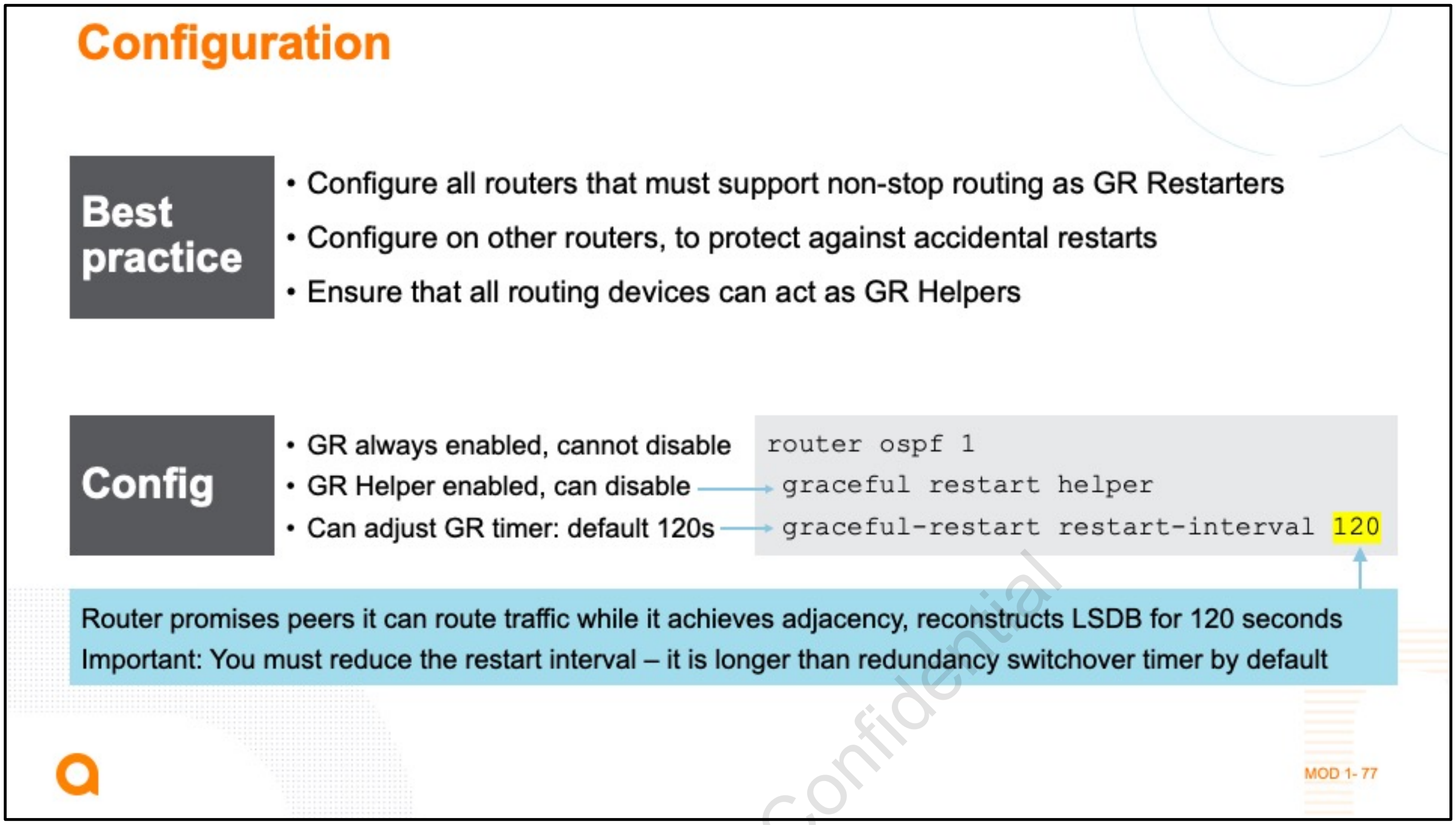
The GR Restarter generates its own LSAs and compares them to its own stale LSAs, which it received from the Helpers. Unless its connections have changed, the GR Restarter's LSAs will match its stale LSAs. The GR Restarter removes the stale designation from those LSAs and takes no further action (its neighbors already know the LSAs).

If the GR Restarter has generated any new LSAs, it advertises them to neighbors as normal in an LSU message. Similarly, if any stale LSAs remain (no matching current LSA), the GR Restarter deletes those LSAs and withdraws them in an LSU message.

The GR Restarter and the GR Helpers now know the proper LSAs, network topology, and routes. Throughout the process, all of the OSPF routing switches continued routing traffic with stale routes, only removing the routes if they were demonstrated to be truly invalid. Because the stale routes typically prove to be valid, the routing switches typically transition seamlessly from routing with stale routes to routing with active routes.

Note: The OSPF graceful restart standard dictates that GR Helpers must leave helper mode if the topology changes in a way that would affect the GR Restarter. That is, the GR Helper generates or receives an LSA that it would normally send to the GR Restarter. This provision prevents routing loops.

Aruba Training-Confidential



Configuration

Best practice

- Configure all routers that must support non-stop routing as GR Restarters
- Configure on other routers, to protect against accidental restarts
- Ensure that all routing devices can act as GR Helpers

Config

- GR always enabled, cannot disable
- GR Helper enabled, can disable
- Can adjust GR timer: default 120s

```
router ospf 1
  graceful restart helper
  graceful-restart restart-interval 120
```

Router promises peers it can route traffic while it achieves adjacency, reconstructs LSDB for 120 seconds
Important: You must reduce the restart interval – it is longer than redundancy switchover timer by default

MOD 1- 77

Aruba Training-Confidential

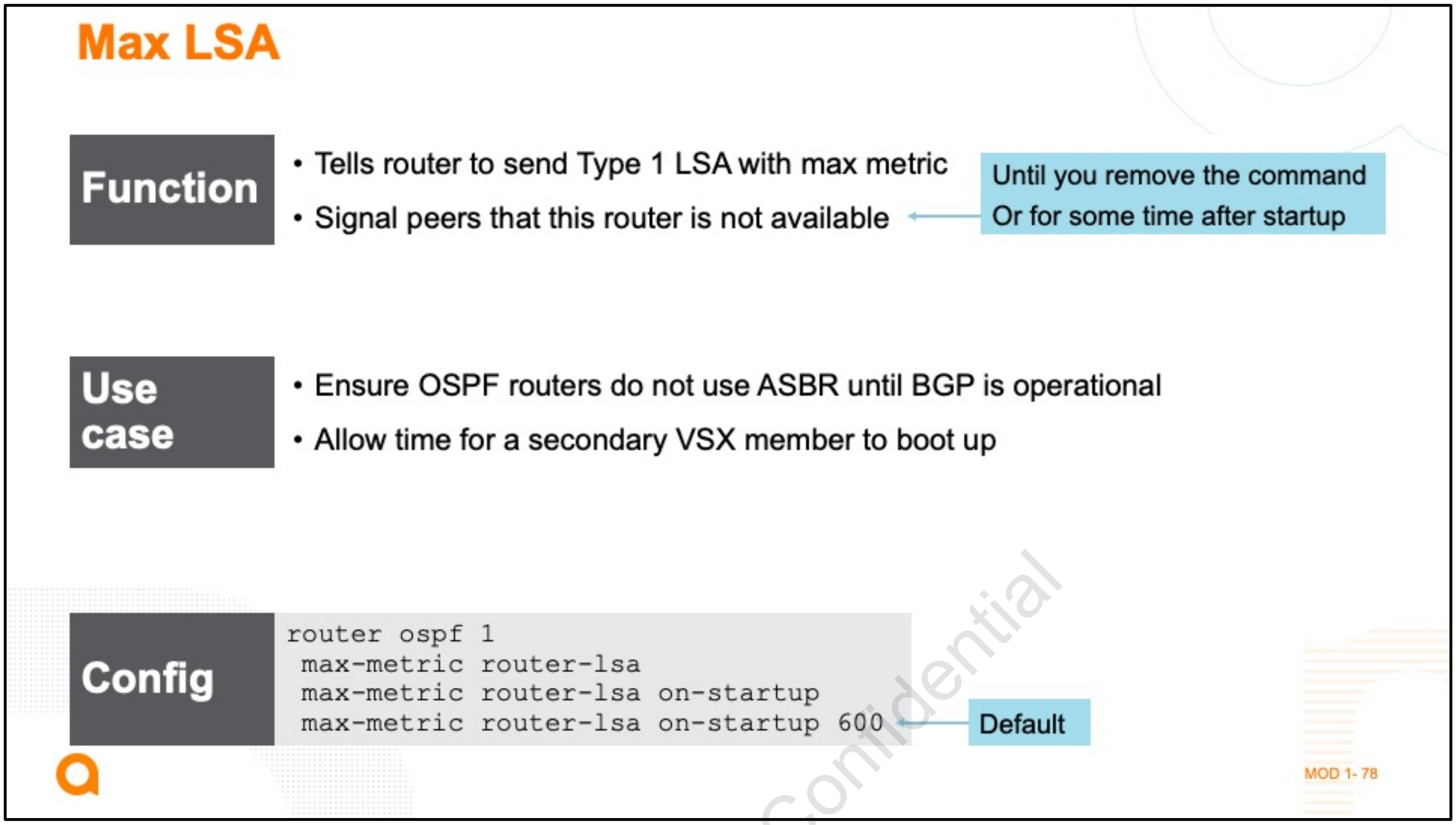
OSPF graceful restart deployment is rather simple. Configure VSF fabrics, and any routing devices that must support non-stop routing during management module failover, as GR Restarters. As a best practice, you could also enable the restarter function on other routing devices that support it, protecting the network in case an administrator accidentally restarts the OSPF process in a live environment.

All neighbors of the GR Restarters must act as GR Helpers. As a best practice, you should generally ensure that all OSPF routing devices in the AS support this function.

Configuration

Graceful restart itself is always enabled and cannot be disabled. Graceful restart helper is enabled by default, but can be disabled. However, you can adjust the graceful restart timer (defaults to 120 seconds) and enable or disable the helper function, as shown in the figure.

After this timer expires, the switch removes any stale routes—that is, routes not recalculated from the reconstructed LSDB. The restart interval dictates the amount of time that the switch promises its neighbors that it can continue to route traffic while it achieves adjacency and reconstructs the LSDB. Therefore, the rapid switchover time must exceed the restart interval. (For example, exceed the restart interval by 3 seconds.)



Max LSA

Function

- Tells router to send Type 1 LSA with max metric
- Signal peers that this router is not available

Until you remove the command
Or for some time after startup

Use case

- Ensure OSPF routers do not use ASBR until BGP is operational
- Allow time for a secondary VSX member to boot up

Config

```
router ospf 1
 max-metric router-lsa
 max-metric router-lsa on-startup
 max-metric router-lsa on-startup 600
```

Default

MOD 1- 78

Aruba Training-Confidential

Function

The max LSA feature is another feature than can smooth OSPF convergence. This feature tells the OSPF router to send out a Type 1 Router LSA with the maximum metric either permanently (until you remove the command) or for a period after the switch starts up. The max metric signals to other OSPF routers that the router is not available.

Use cases

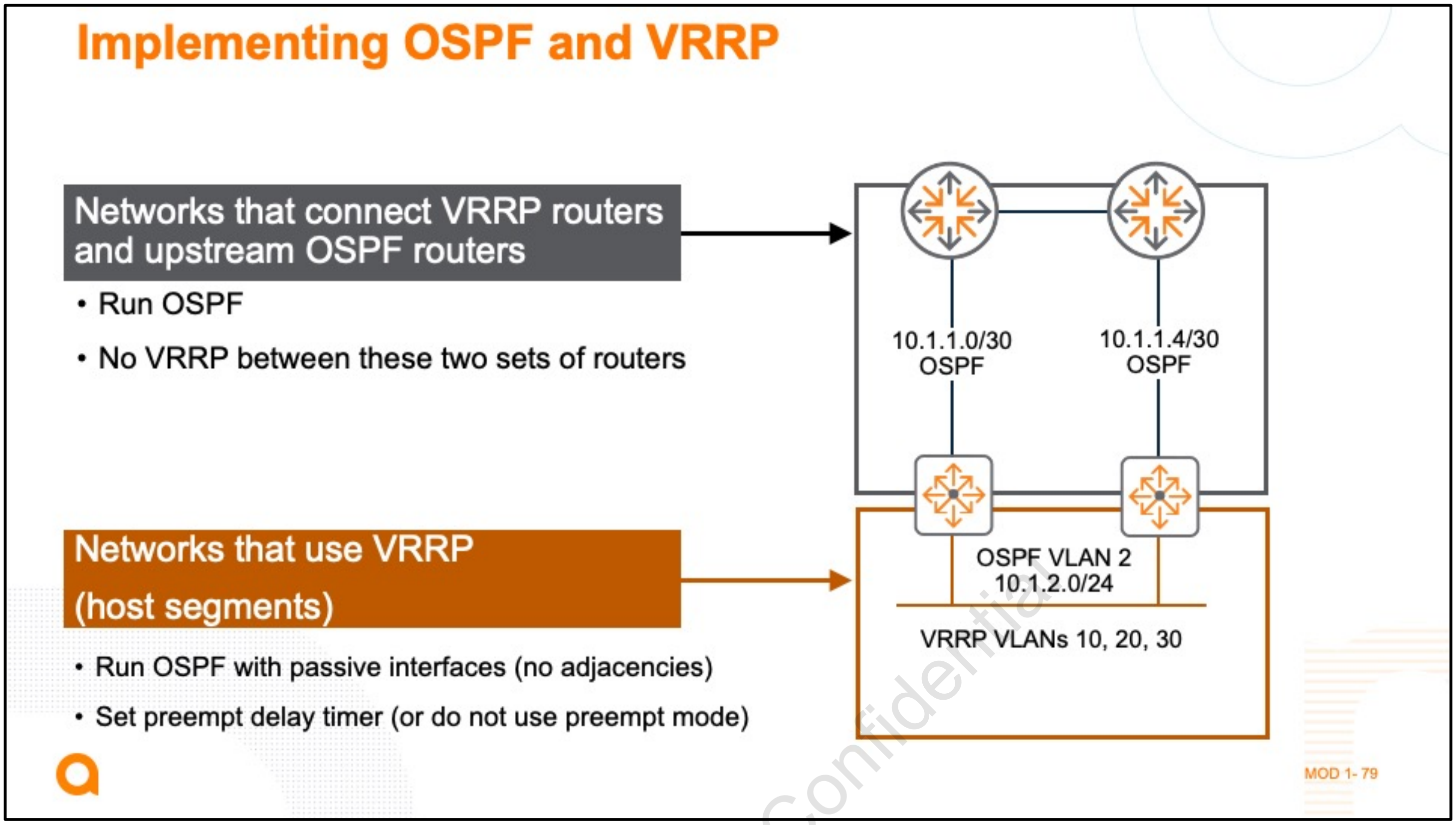
You would typically use this feature on an ASBR that redistributes routes from another protocol. You want to ensure that other OSPF routers do not send traffic through the ASBR until BGP is fully operational. BGP can take longer to load the full table than OSPF. Simply set the max LSA time period to a period a bit longer than you think it will take the ASBR to load all routing protocols.

Another use case relates to a secondary VSX member that requires time to boot up. The feature allows the secondary VSX member to learn upstream routes while the VSX uplink delay is keeping the VSX LAGs down. So when the VSX LAGs come up, the routing table is populated. Later, the secondary member will reduce the LSA metric, so it can also attract traffic from backbone for the local users.

Configuration

The figure shows the configuration for the MAX LSA feature:

The default on-startup delay is 600 seconds.



Implementing OSPF and VRRP

Networks that connect VRRP routers and upstream OSPF routers

- Run OSPF
- No VRRP between these two sets of routers

Networks that use VRRP (host segments)

- Run OSPF with passive interfaces (no adjacencies)
- Set preempt delay timer (or do not use preempt mode)

MOD 1- 79

Aruba Training-Confidential

As you have learned, you should use VSF to provide redundancy for the default router role whenever possible. VSF provides redundancy while allowing the VSF fabric to act as a single entity in the routing protocol, making implementing VSF with OSPF almost as easy as implementing OSPF without redundancy. However, sometimes customer requirements might force you to implement OSPF and VRRP. You must design the solution carefully to ensure that the OSPF and VRRP functions work will together.

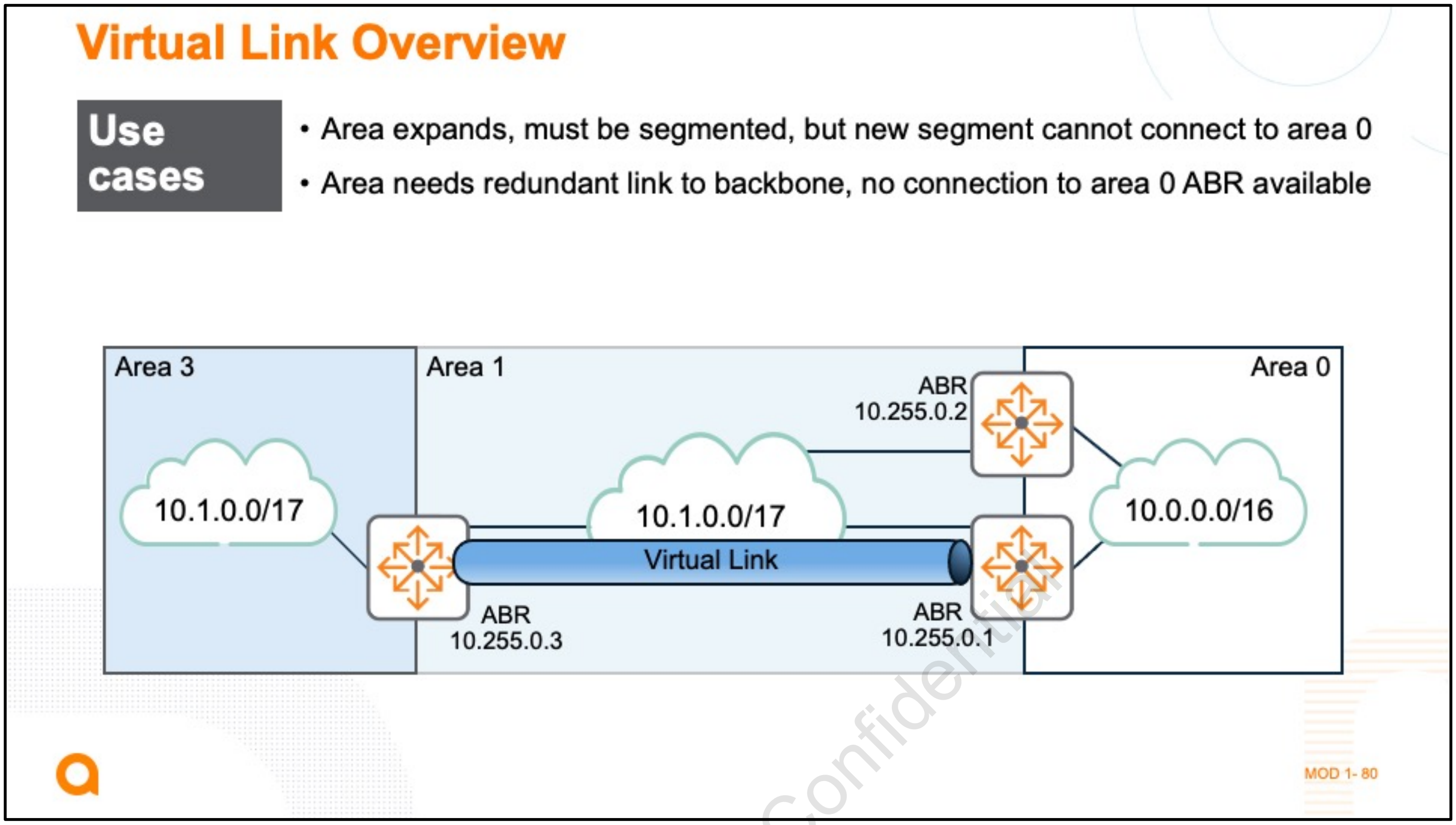
When a router implements VRRP and OSPF on the same interface, it uses its actual IP address to run OSPF. However, the protocol interoperations can cause complexities. To minimize these complexities, you should avoid running OSPF and VRRP on the same interfaces.

First, connect VRRP routers to upstream OSPF routers on their own networks and do not run VRRP on these networks. This design helps an upstream router to immediately determine when it loses connectivity to a VRRP router.

Second, do not establish OSPF adjacencies on the interfaces that run VRRP (where the hosts reside). Because VRRP provides redundancy for default gateway services, VRRP runs on directly connected networks, which are generally stub networks. You can redistribute the directly connected networks as external networks. (Make sure to set up the area as a normal area or an NSSA.) Aruba recommends that you configure the interfaces as passive OSPF interfaces.

Third, choose the VRRP design in which none of the VRRP routers own the virtual IP address. Otherwise, when the VRRP Owner fails and comes back up, it immediately takes over routing on the interfaces for which it owns the VRRP address. However, it has not yet established adjacencies with upstream OSPF routers, disrupting connectivity on those VLANs. By configuring a virtual IP address that none of the VRRP routers owns, you can prevent this problem. Either disable preempt mode or set a preempt mode delay time that is longer than the time required to establish adjacency with all neighbors and calculate the OSPF routes (generally, two or three minutes).

Aruba Training-Confidential



Virtual Link Overview

Use cases

- Area expands, must be segmented, but new segment cannot connect to area 0
- Area needs redundant link to backbone, no connection to area 0 ABR available

Aruba Training-Confidential

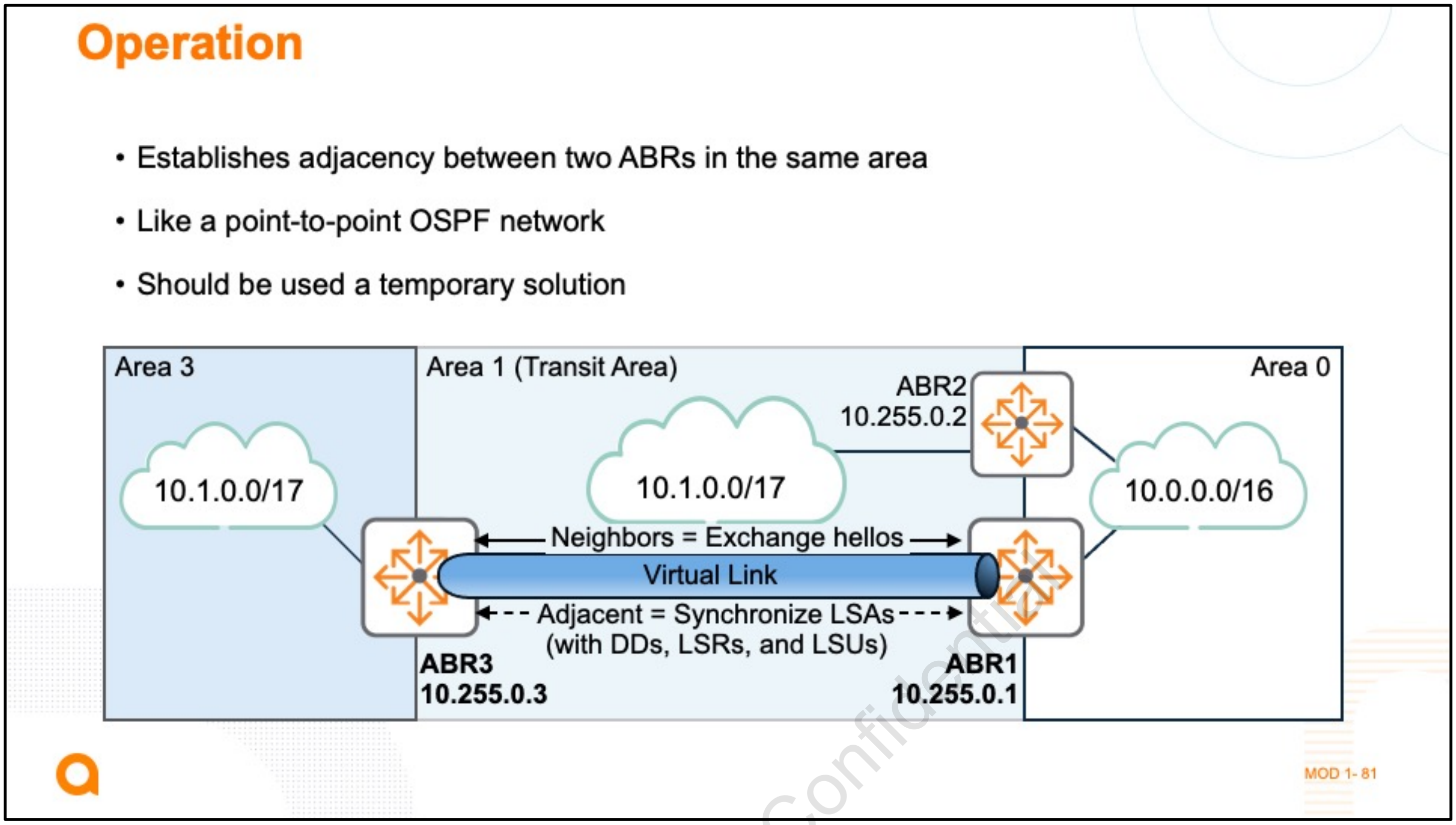
Well-designed OSPF systems should ensure that you can always connect new areas directly to the backbone. However, sometimes the company expands in unexpected ways, and you cannot meet this requirement.

The figure shows a scenario where the company must expand to new sites (10.1.128.0/17). Routers in area 1 are available for connecting to the new routers. However, area 1 has grown so large that you would prefer to establish the new sites in area 3. However, the rules of OSPF prevent you from creating the new area – all areas must connect to area 0, and all inter-area traffic must be via Area 0.

So, you create a virtual link between two ABRs in area 1.

Similar issues can arise when you need to add a redundant link between an area and the backbone area 0. As you have learned, you can create link aggregations between internal area routers and the ABR, which might be one router or, for even better redundancy, two physical routing switches in a VSF fabric. You can also connect internal area routers to two ABRs in area 0. You should choose one of these solutions whenever possible.

In this case, however, a second connection between a particular site and an ABR at the backbone is not available. You must establish the link through a transit site.



Operation

- Establishes adjacency between two ABRs in the same area
- Like a point-to-point OSPF network
- Should be used a temporary solution

Area 3

Area 1 (Transit Area)

Area 0

10.1.0.0/17

10.1.0.0/17

10.0.0.0/16

ABR2
10.255.0.2

Neighbors = Exchange hellos

Virtual Link

Adjacent = Synchronize LSAs
(with DDs, LSRs, and LSUs)

ABR3
10.255.0.3

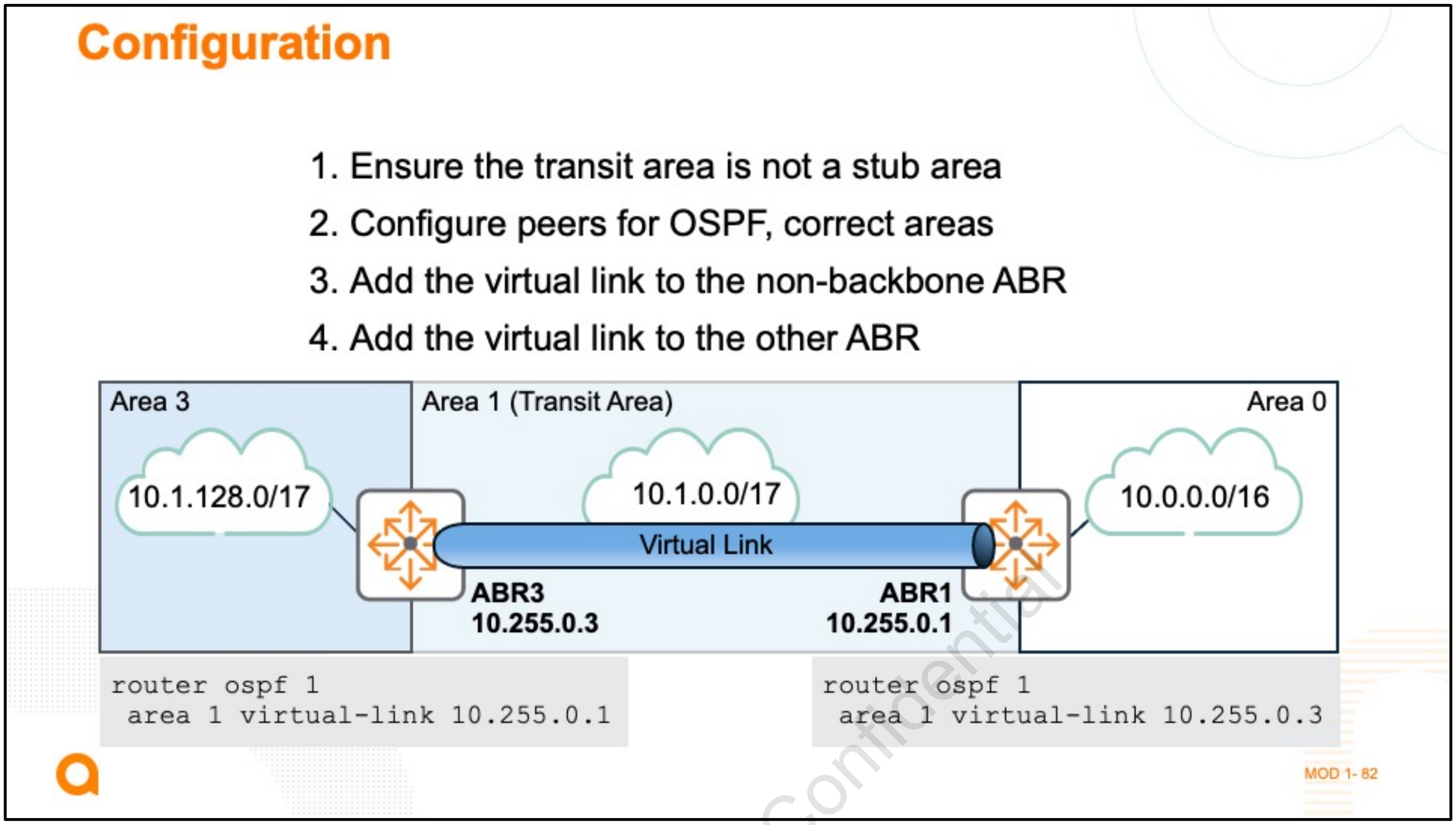
ABR1
10.255.0.1

MOD 1- 81

You establish the virtual link across a transit area (area 1, in this example). ABR3 connects Area 3 to Area 1, and ABR1 connects Area 0 to Area 1. You form a virtual link between these two ABRs in Area 1. The virtual link acts like a shared network interface, allowing the two OSPF routing devices to exchange hellos and establish adjacency. Logically, the virtual link acts as a point-to-point OSPF network on which both ABRs have an interface.

Note that you should not generally rely on virtual links as part of your design because they introduce complexity. They also negate some of the benefits of multi-area designs by leaking backbone routing information into non-backbone areas. Instead use virtual links as a patch when you have no other way to create the necessary link to the backbone at the current time.

Aruba Training-Confidential



Configuration

1. Ensure the transit area is not a stub area
2. Configure peers for OSPF, correct areas
3. Add the virtual link to the non-backbone ABR
4. Add the virtual link to the other ABR

```
router ospf 1
 area 1 virtual-link 10.255.0.1
```

```
router ospf 1
 area 1 virtual-link 10.255.0.3
```

MOD 1- 82

Aruba Training-Confidential

The slide indicates the process for adding the virtual link. Before completing these steps, you should configure the basic OSPF settings on all of the routers involved in the solution. The section below provides more detailed steps.

Make sure that the transit area is not defined as a stub area or NSSA.

Configure the virtual link on the ABR that does not connect to area 0.

- As mentioned above, this ABR should have its basic OSPF settings defined:
  - Router ID
  - The area without the direct connection to the backbone (area 3 in the example)
  - The transit area (area 1 in the example)
  - OSPF enabled on the appropriate interfaces in each area
  - Any other settings required for the OSPF solution (such as aggregated area summaries)
- In the transit area (area 1 in the example), configure a virtual link. Specify the router ID of the device at the other end of the link as shown in the figure.

Configure the virtual link on the ABR that connects to area 0.

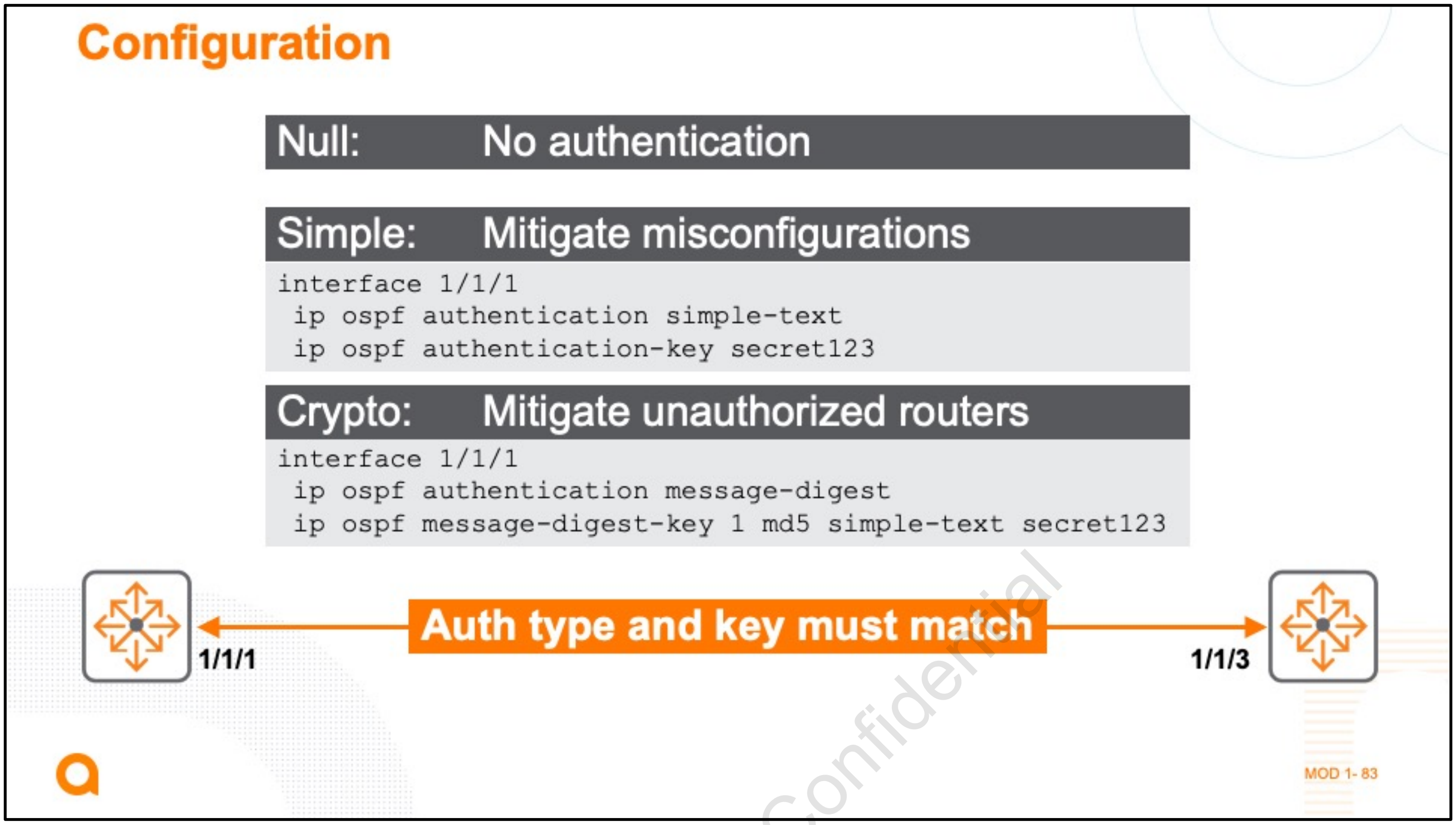
- Again, this ABR should have its basic OSPF settings defined:
  - Router ID
  - Area 0
  - The transit area (area 1 in the example)
  - OSPF enabled on the appropriate networks in each area

  - Any other settings required for the OSPF solution
- In the transit area, configure a virtual link, specifying the router ID of the device at the other end of the link, as shown in the figure.
  Note: Note that the OSPF routing devices have connectivity within the transit area. As always, if you want to be able to ping the router ID, use an IP address on loopback interface that runs OSPF.

You must check consistency for the typical settings required for two OSPF routers on a link to achieve adjacency. These settings include the OSPF timers. (The virtual link does not have a network associated with it, so you do not have to check consistency for its type.)

Note: There are other options you can configure for virtual links. See the AOS-CX IP Routing Guide for more detail.

Aruba Training-Confidential



Configuration

Null: No authentication

Simple: Mitigate misconfigurations

```
interface 1/1/1
 ip ospf authentication simple-text
 ip ospf authentication-key secret123
```

Crypto: Mitigate unauthorized routers

```
interface 1/1/1
 ip ospf authentication message-digest
 ip ospf message-digest-key 1 md5 simple-text secret123
```

Auth type and key must match

MOD 1- 83

Aruba Training-Confidential

You implement authentication for OSPF processes on router interfaces. OSPF includes authentication data in all OSPF packets sent, and checks this data in all received OSPF packets. Only OSPF routers with matching authentication type and keys can establish neighbor relationships.

The figure shows the three authentication types, as described below.

Null—The default authentication type, in which OSPF interfaces do not enforce authentication. The OSPF packets have an authentication field, but the field does not include data. An OSPF interface with null authentication does not examine the authentication field – it simply processes all valid OSPF packets.

Simple—The OSPF interface places the authentication key in every OSPF packet header. Other OSPF routers, examine this field and accept or reject packets accordingly.
The actual password or key is in plaintext, so anyone who can capture OSPF packets can learn the key. You should use this form of authentication only if you primarily aim to eliminate misconfigurations and routers inadvertently connecting on the wrong networks. You should not use it to secure your OSPF system from unauthorized routers and hackers.

Cryptographic (MD5)—The OSPF interface hashes the OSPF packet data and the key. It then places the hash at the end of the packet. The authentication field includes the key ID, a length for the hash, and a cryptographic sequence number; however, it does not include the key itself, protecting it from eavesdroppers.

A router receiving the OSPF packet hashes the packet with its secret key, identified by the key ID. If the hash matches, the router accepts the packet.

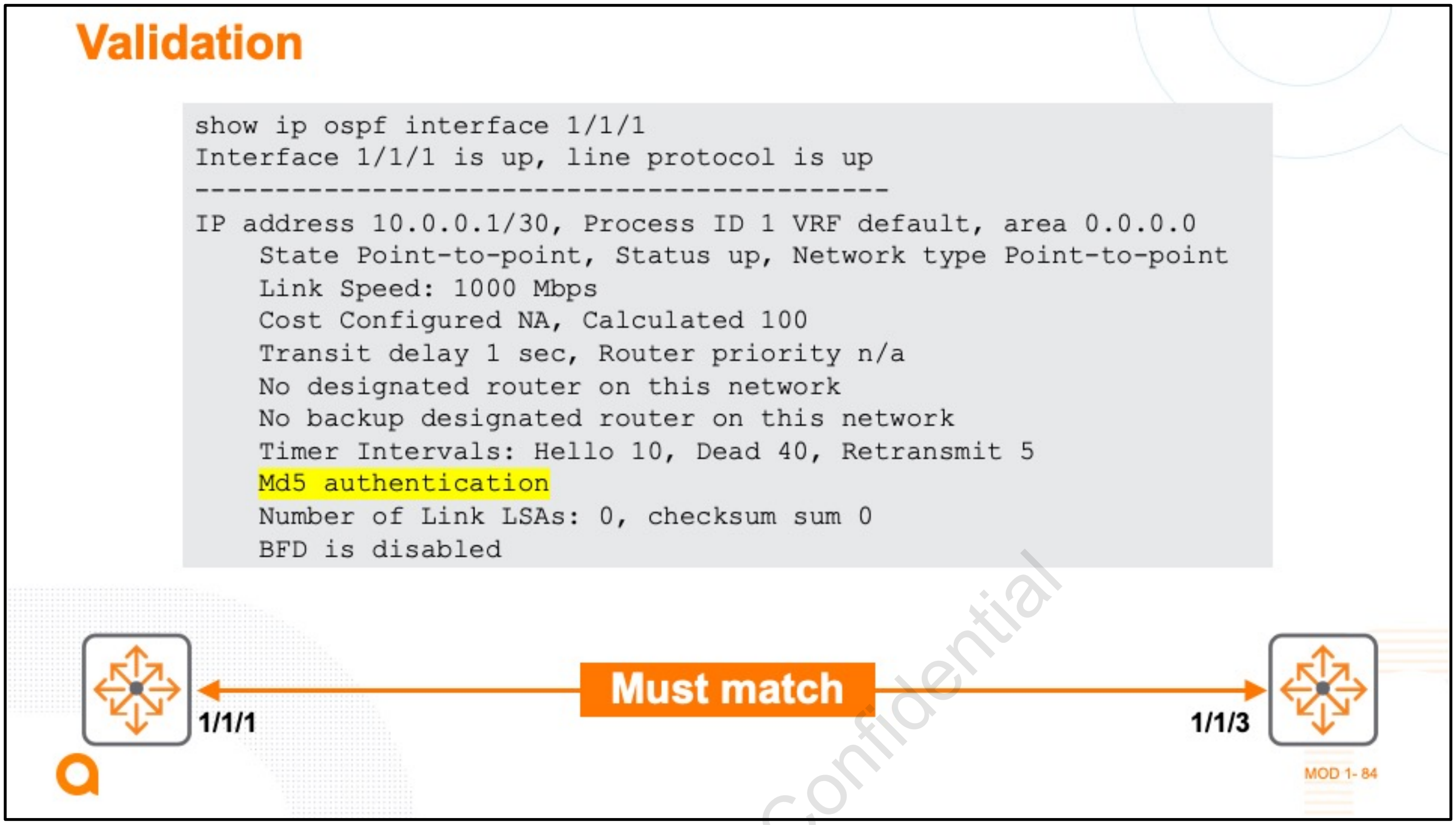
The cryptographic sequence number provides a degree of protection against replay attacks. An OSPF interface only accepts packets with a sequence number equal to or higher than the number for the most recent accepted OSPF packet from this neighbor. The sequence number might be a counter but is usually based on the system clock. In either case, the number always increases.

If a hacker obtains a legitimate OSPF packet and attempts to retransmit it later, the packet's cryptographic sequence number is no longer valid, and the other routers reject it. A hacker could replay a packet in the window before the legitimate neighbor sends a packet with a higher sequence, but this window is quite small.

Cryptographic authentication provides authenticity, ensuring that only OSPF interfaces configured with the same key can communicate. It does not provide privacy: the contents of the packet are not encrypted.

All Aruba routers and routing switches support MD5 for the cryptographic hashing algorithm.

Aruba Training-Confidential



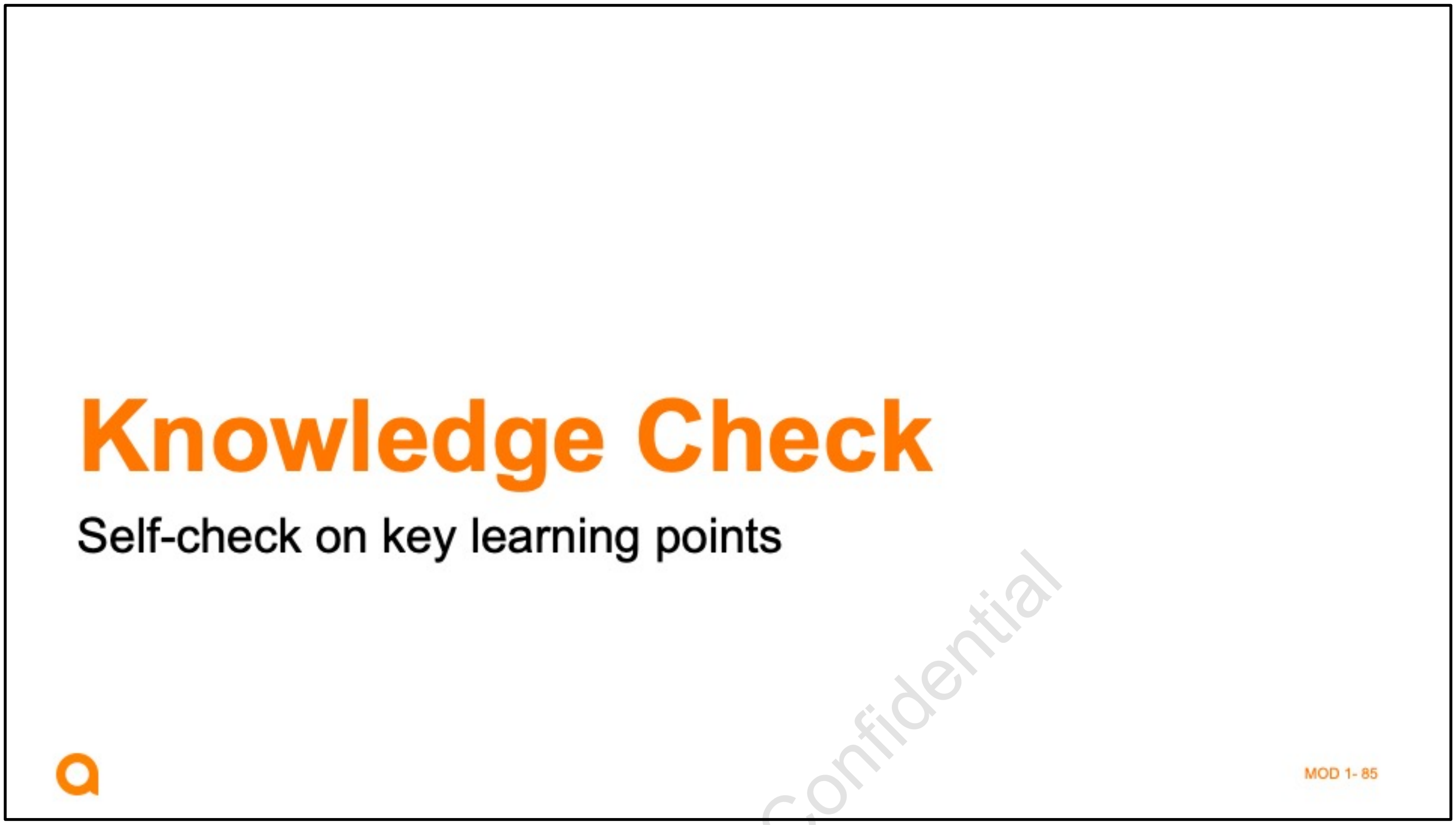
Validation

```
show ip ospf interface 1/1/1
Interface 1/1/1 is up, line protocol is up
-------------------------------------------
IP address 10.0.0.1/30, Process ID 1 VRF default, area 0.0.0.0
    State Point-to-point, Status up, Network type Point-to-point
    Link Speed: 1000 Mbps
    Cost Configured NA, Calculated 100
    Transit delay 1 sec, Router priority n/a
    No designated router on this network
    No backup designated router on this network
    Timer Intervals: Hello 10, Dead 40, Retransmit 5
    Md5 authentication
    Number of Link LSAs: 0, checksum sum 0
    BFD is disabled
```

MOD 1- 84

You can verify if key authentication is enabled for an interface as shown in the figure.

Aruba Training-Confidential

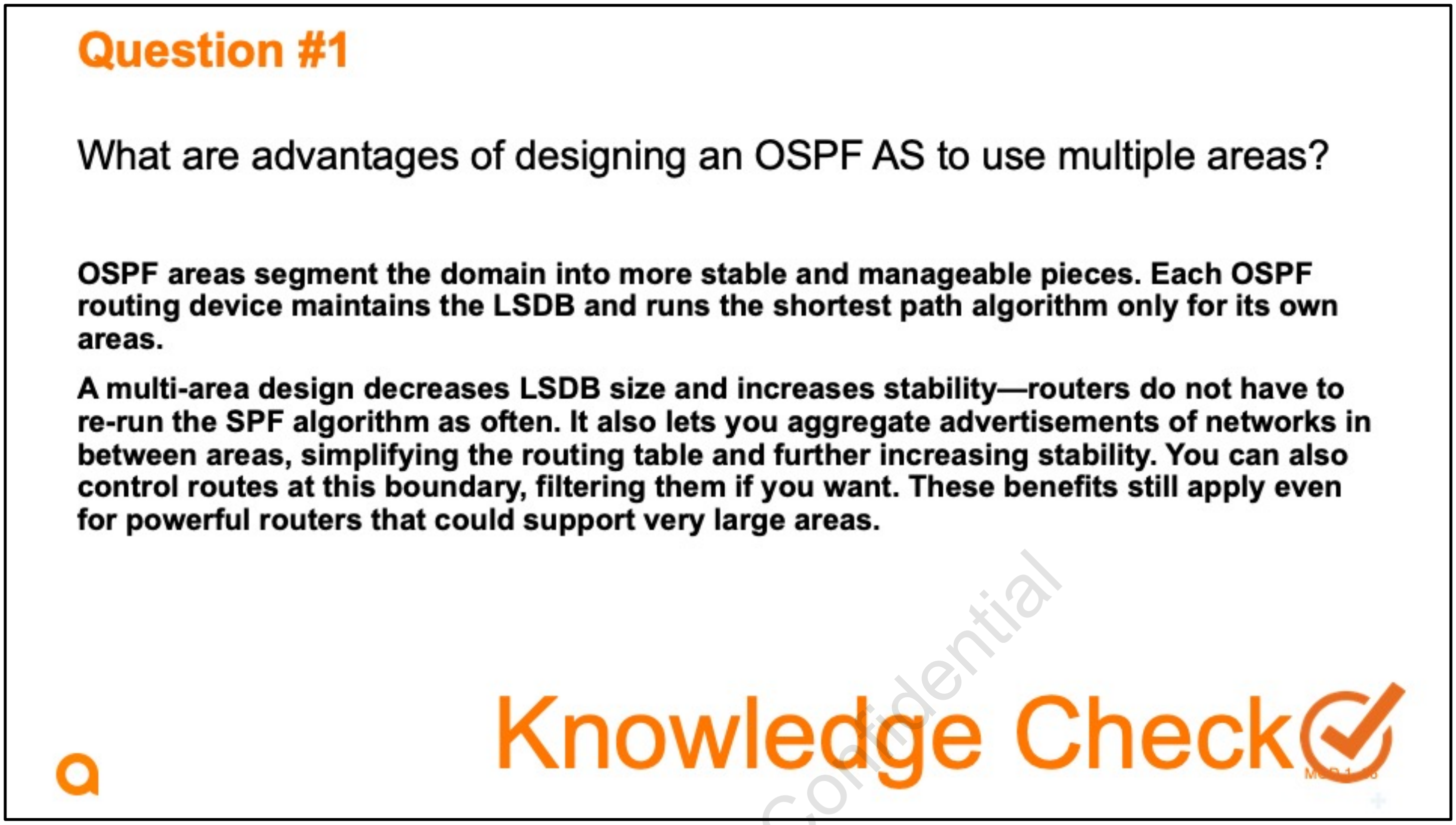


Knowledge Check

Self-check on key learning points

MOD 1- 85

Aruba Training-Confidential



Question #1

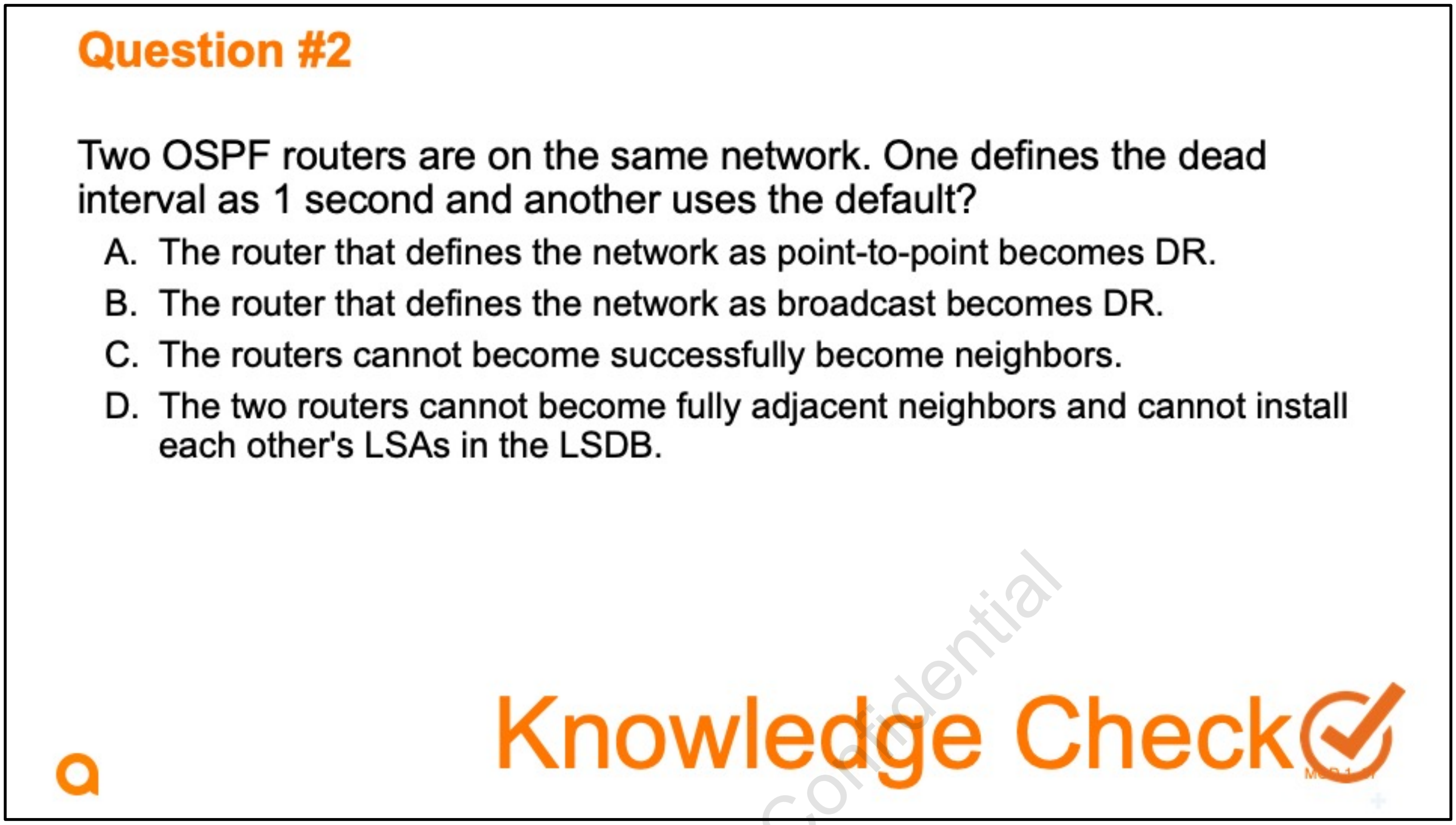
What are advantages of designing an OSPF AS to use multiple areas?

OSPF areas segment the domain into more stable and manageable pieces. Each OSPF routing device maintains the LSDB and runs the shortest path algorithm only for its own areas.

A multi-area design decreases LSDB size and increases stability—routers do not have to re-run the SPF algorithm as often. It also lets you aggregate advertisements of networks in between areas, simplifying the routing table and further increasing stability. You can also control routes at this boundary, filtering them if you want. These benefits still apply even for powerful routers that could support very large areas.

Knowledge Check

Aruba Training-Confidential



Question #2

Two OSPF routers are on the same network. One defines the dead interval as 1 second and another uses the default?

A. The router that defines the network as point-to-point becomes DR.

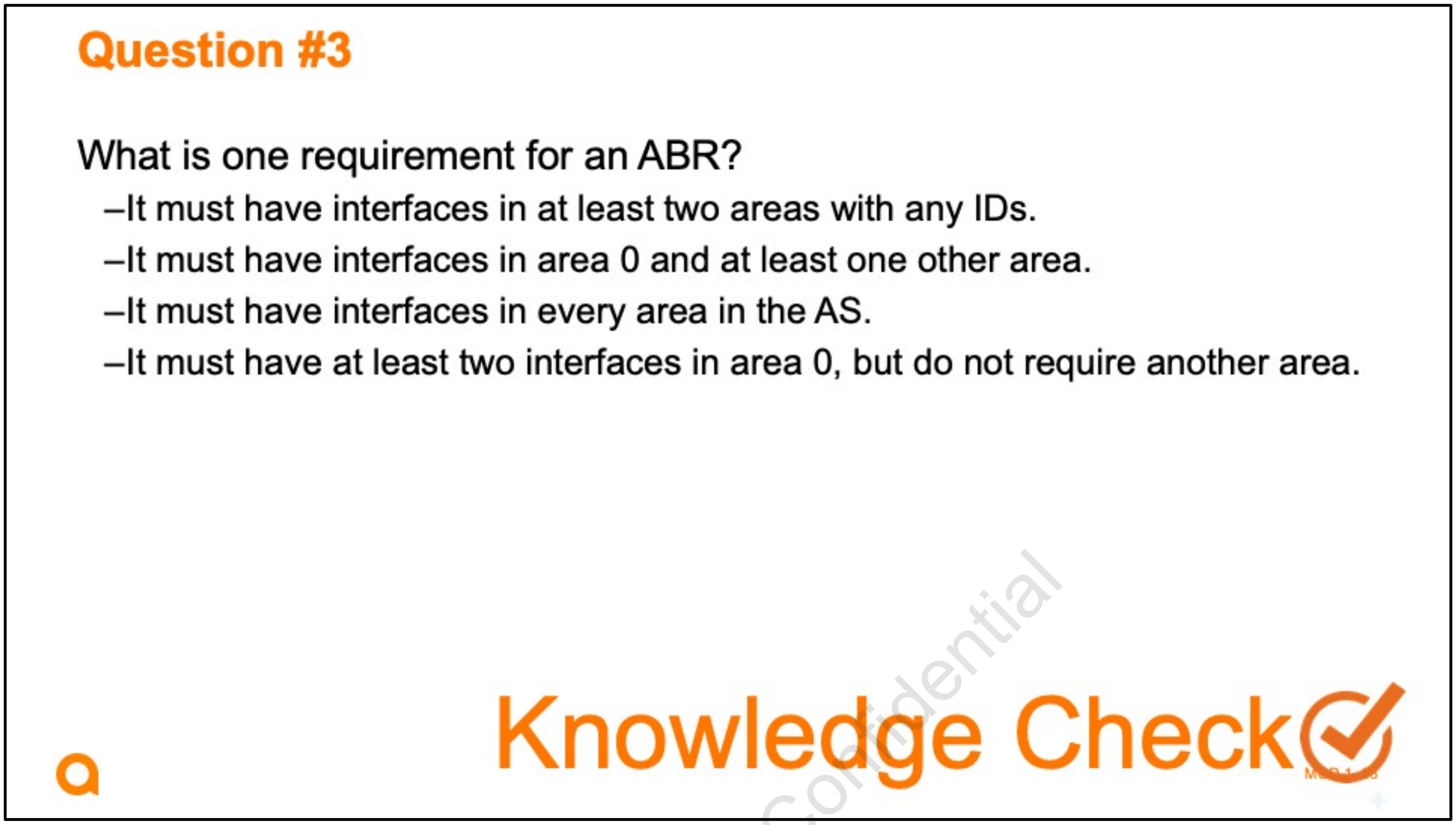
B. The router that defines the network as broadcast becomes DR.

C. The routers cannot become successfully become neighbors.

D. The two routers cannot become fully adjacent neighbors and cannot install each other's LSAs in the LSDB.

Knowledge Check

Aruba Training-Confidential



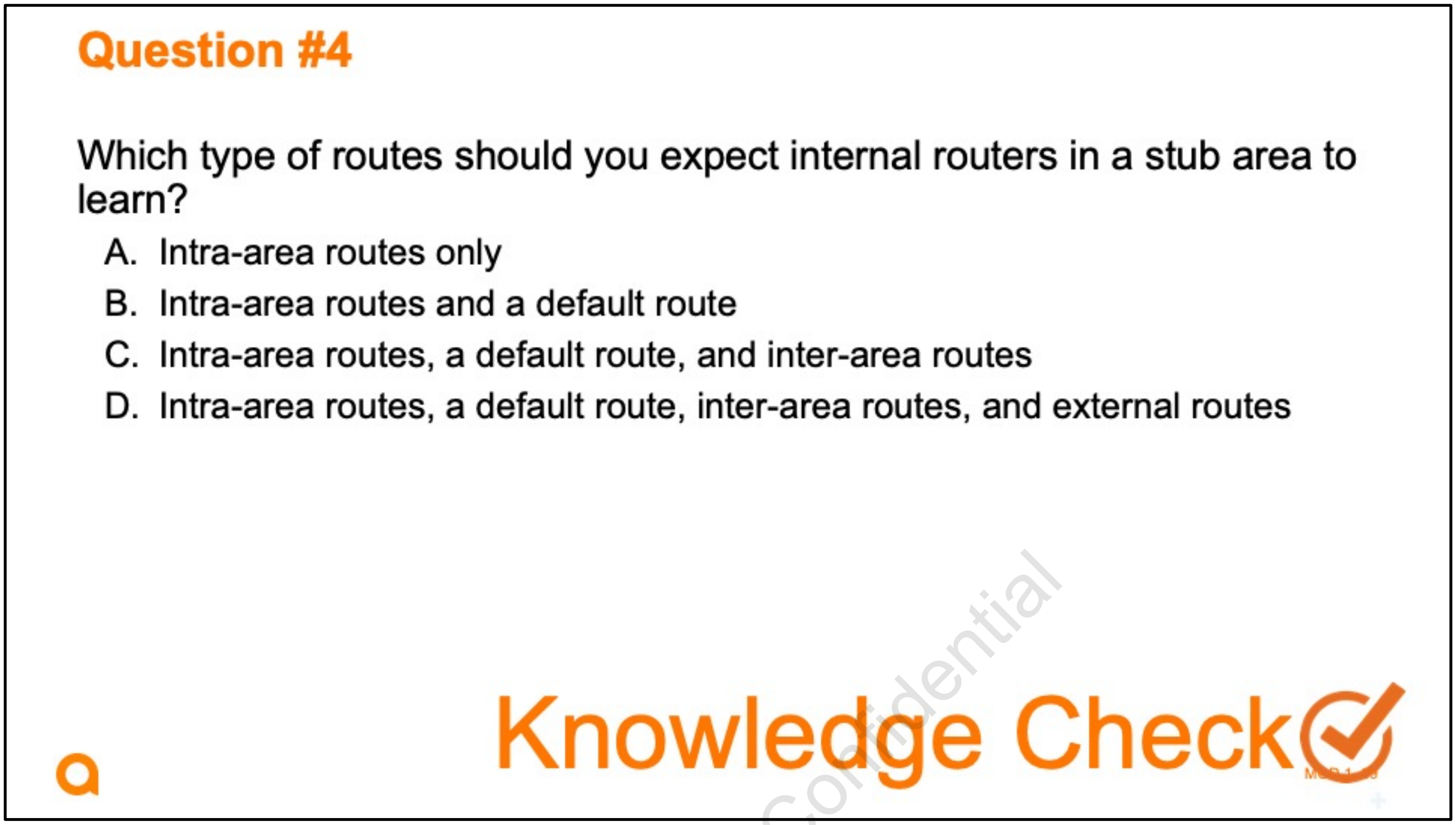
Question #3

What is one requirement for an ABR?

- –It must have interfaces in at least two areas with any IDs.
- –It must have interfaces in area 0 and at least one other area.
- –It must have interfaces in every area in the AS.
- –It must have at least two interfaces in area 0, but do not require another area.

Knowledge Check

Aruba Training-Confidential



Question #4

Which type of routes should you expect internal routers in a stub area to learn?

A. Intra-area routes only

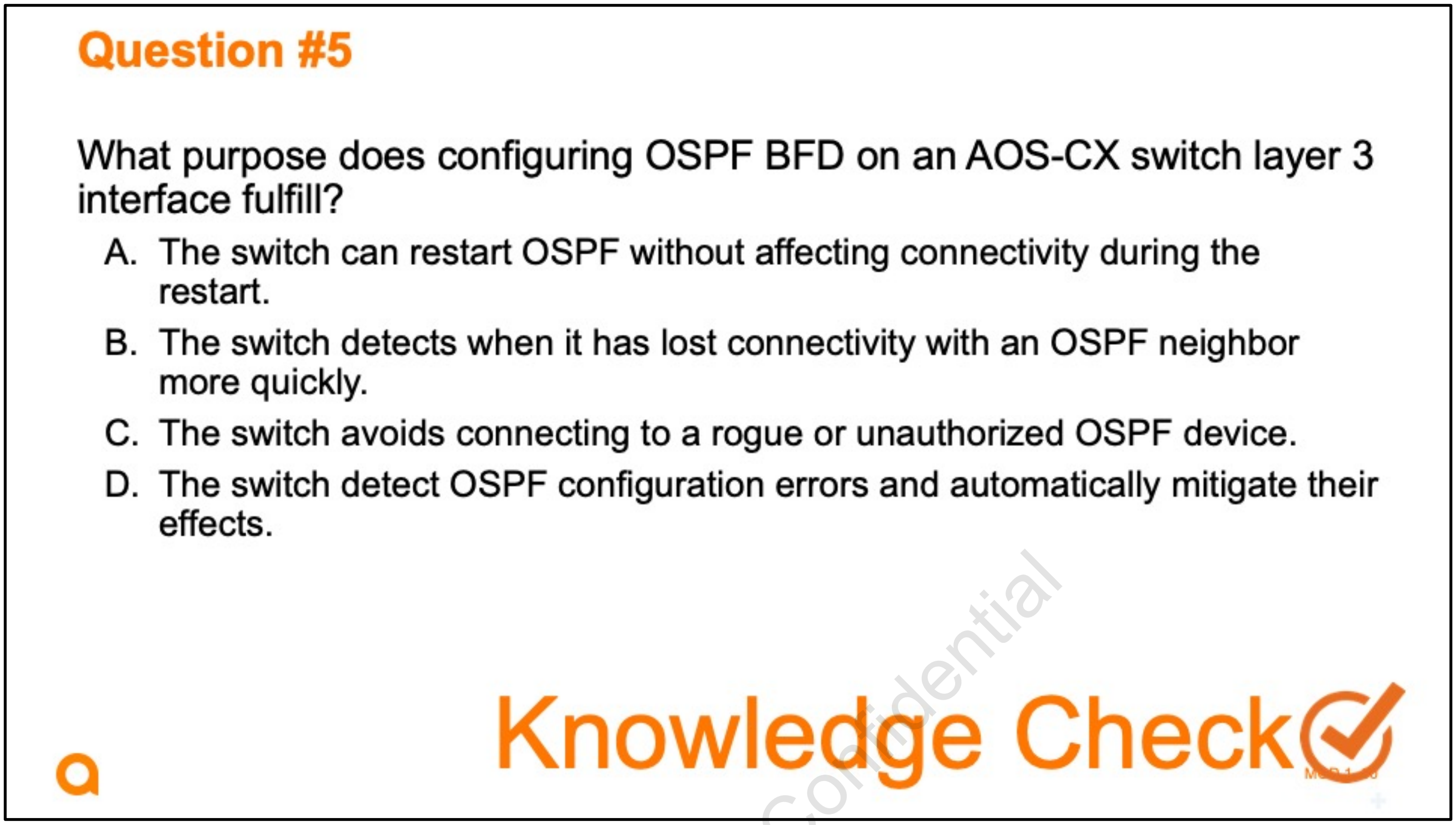
B. Intra-area routes and a default route

C. Intra-area routes, a default route, and inter-area routes

D. Intra-area routes, a default route, inter-area routes, and external routes

Knowledge Check

Aruba Training-Confidential



Question #5

What purpose does configuring OSPF BFD on an AOS-CX switch layer 3 interface fulfill?

A. The switch can restart OSPF without affecting connectivity during the restart.

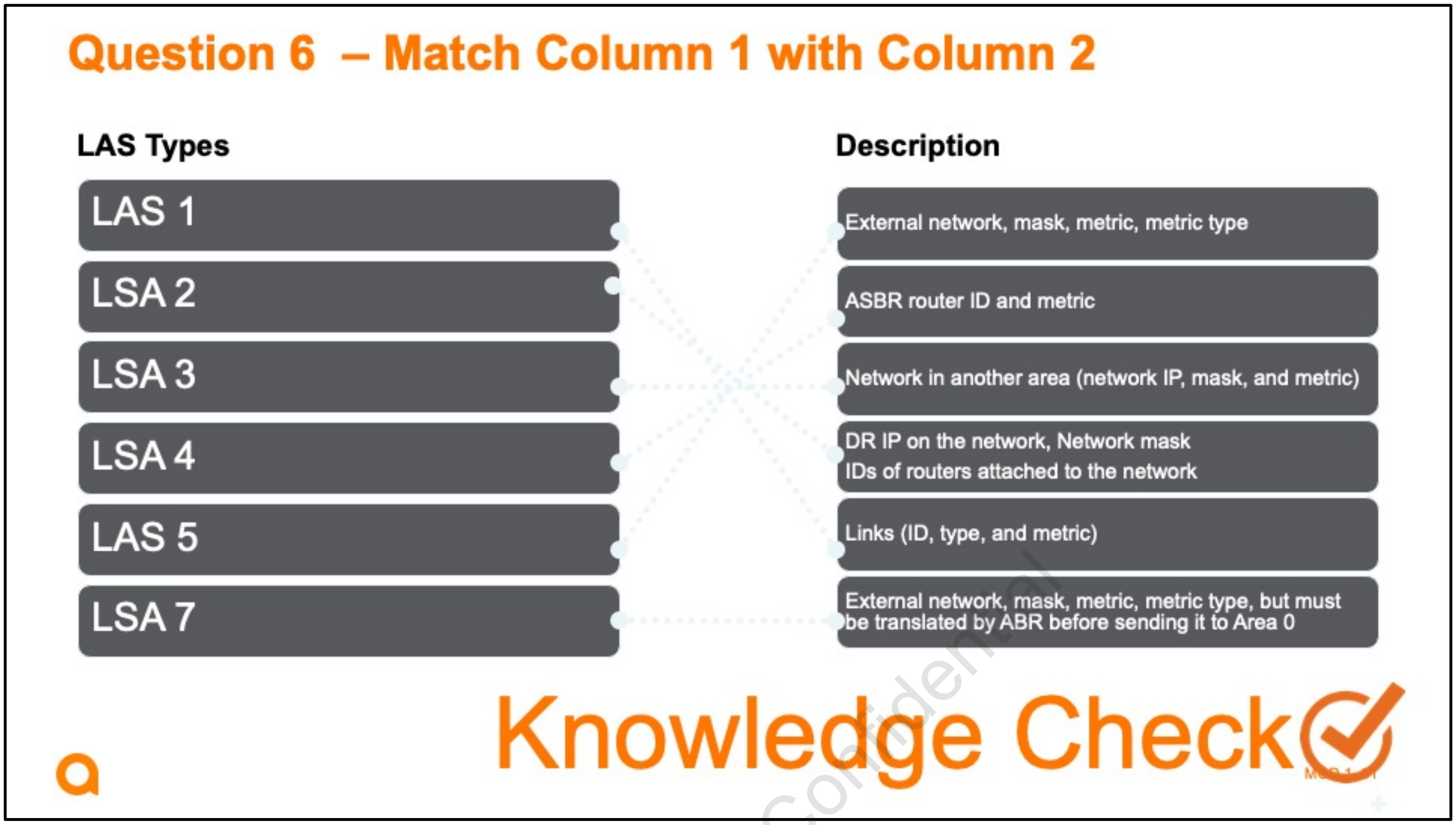
B. The switch detects when it has lost connectivity with an OSPF neighbor more quickly.

C. The switch avoids connecting to a rogue or unauthorized OSPF device.

D. The switch detect OSPF configuration errors and automatically mitigate their effects.

Knowledge Check

Aruba Training-Confidential



Question 6 – Match Column 1 with Column 2

LAS Types	Description
LAS 1	External network, mask, metric, metric type
LSA 2	ASBR router ID and metric
LSA 3	Network in another area (network IP, mask, and metric)
LSA 4	DR IP on the network, Network mask IDs of routers attached to the network
LAS 5	Links (ID, type, and metric)
LSA 7	External network, mask, metric, metric type, but must be translated by ABR before sending it to Area 0

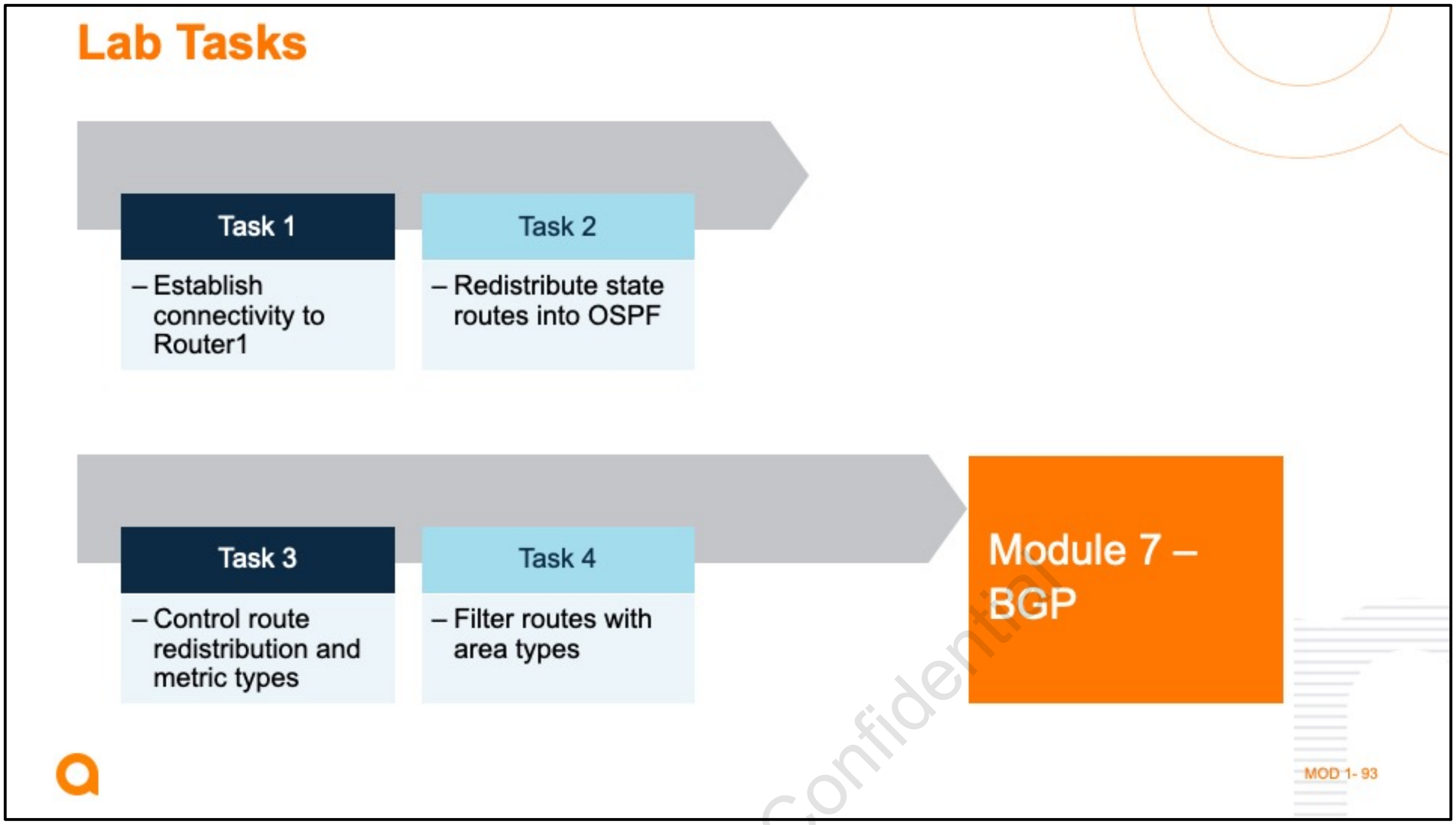
Knowledge Check

Aruba Training-Confidential

Lab Activity

Lab 6.3

Aruba Training-



Lab Tasks

Task 1

– Establish connectivity to Router1

Task 2

– Redistribute state routes into OSPF

Task 3

– Control route redistribution and metric types

Task 4

– Filter routes with area types

Module 7 – BGP

MOD 1- 93

The figure provides a brief review of lab tasks. Please see your lab guide for details.

Aruba Training-Confidential

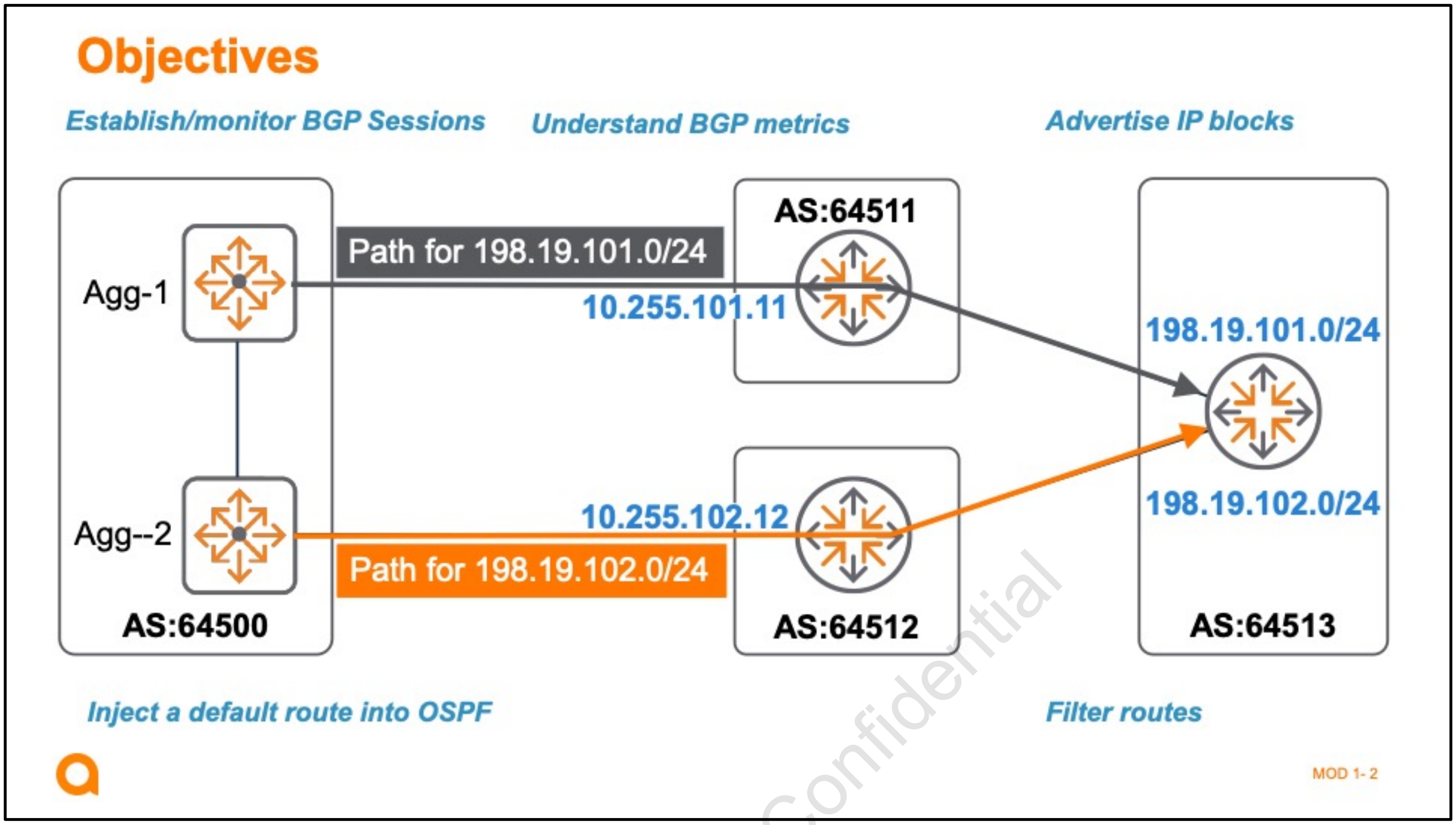
Rev # 20.21

BGP

Implementing ArubaOS-CX Switching

Hi team, and welcome back. This is Module 7 – BGP.

Aruba Training-Confidential



Objectives

Establish/monitor BGP Sessions

Understand BGP metrics

Advertise IP blocks

Inject a default route into OSPF

Filter routes

After completing this module, you should be able to:

Establish and monitor BGP sessions

Understand BGP route metrics

Advertise an IP block to multiple ISP routers

Configure a BGP router to advertise a default route in OSPF

Filter routes to prevent transit networks

Aruba Training-Confidential

Overview

- Intro to BGP
- Establish BGP Connections
- BGP Routing Metrics
- Receive/Advertise Routes
- Route Control
- Lab Activity

MOD 1- 3

This module introduces you to Border Gateway Protocol (BGP), a powerful routing protocol designed for Internet Service Providers (ISPs) and very large enterprise networks.

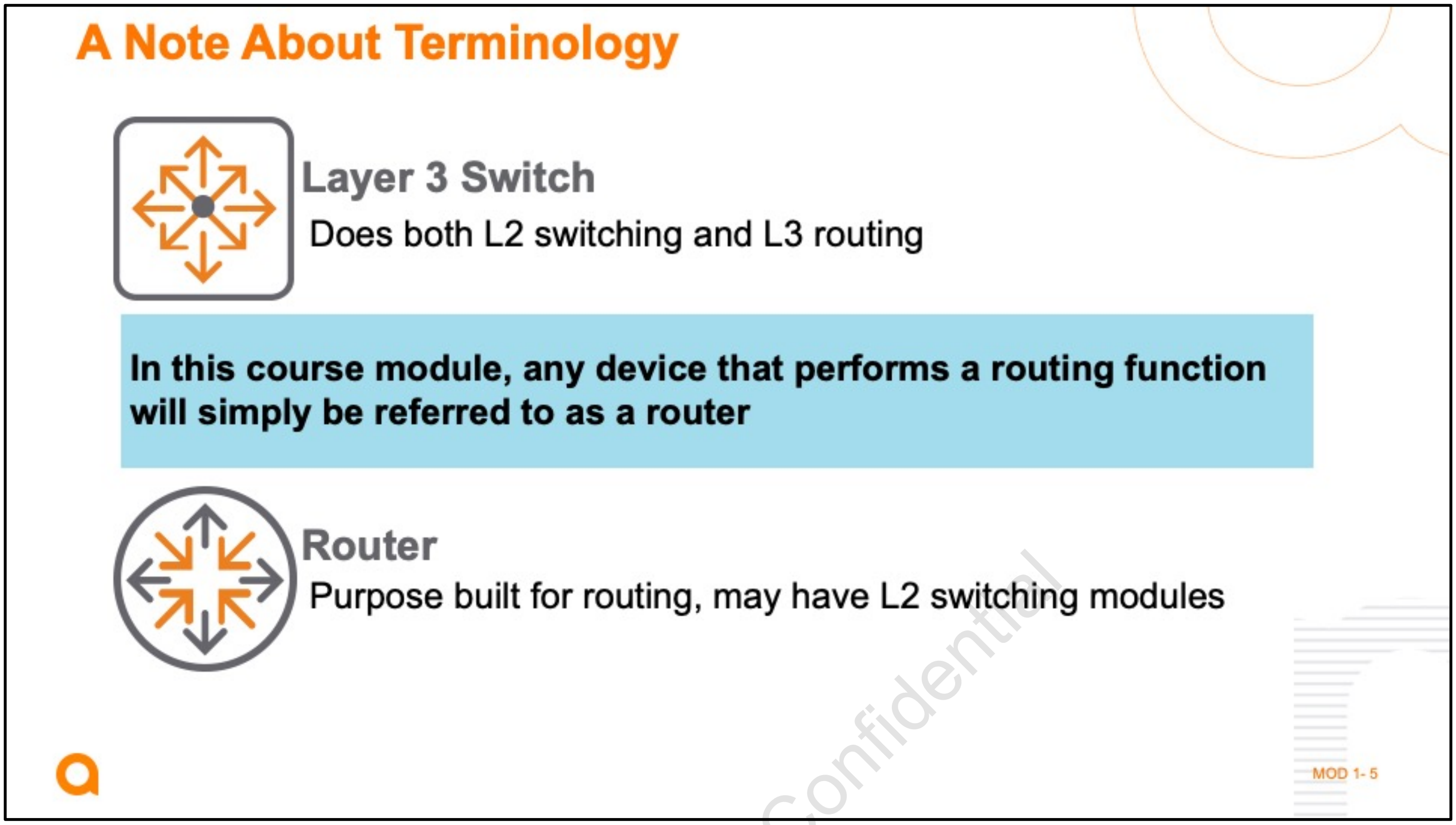
You'll learn to describe and deploy BGP in customer networks, along with strategies for making BGP meet the specific requirements – BGP basics, and use cases; setup neighbors and control advertisements. You'll learn how to influence route selection with route maps, and understand how route reflectors can improve BGP scalability. Of course, some lab activities ties it all together.

Aruba Training-Confidential

Intro to BGP

We begin with an intro.

Aruba Training-Confidential



A Note About Terminology

Layer 3 Switch

Does both L2 switching and L3 routing

In this course module, any device that performs a routing function will simply be referred to as a router

Router

Purpose built for routing, may have L2 switching modules

MOD 1- 5

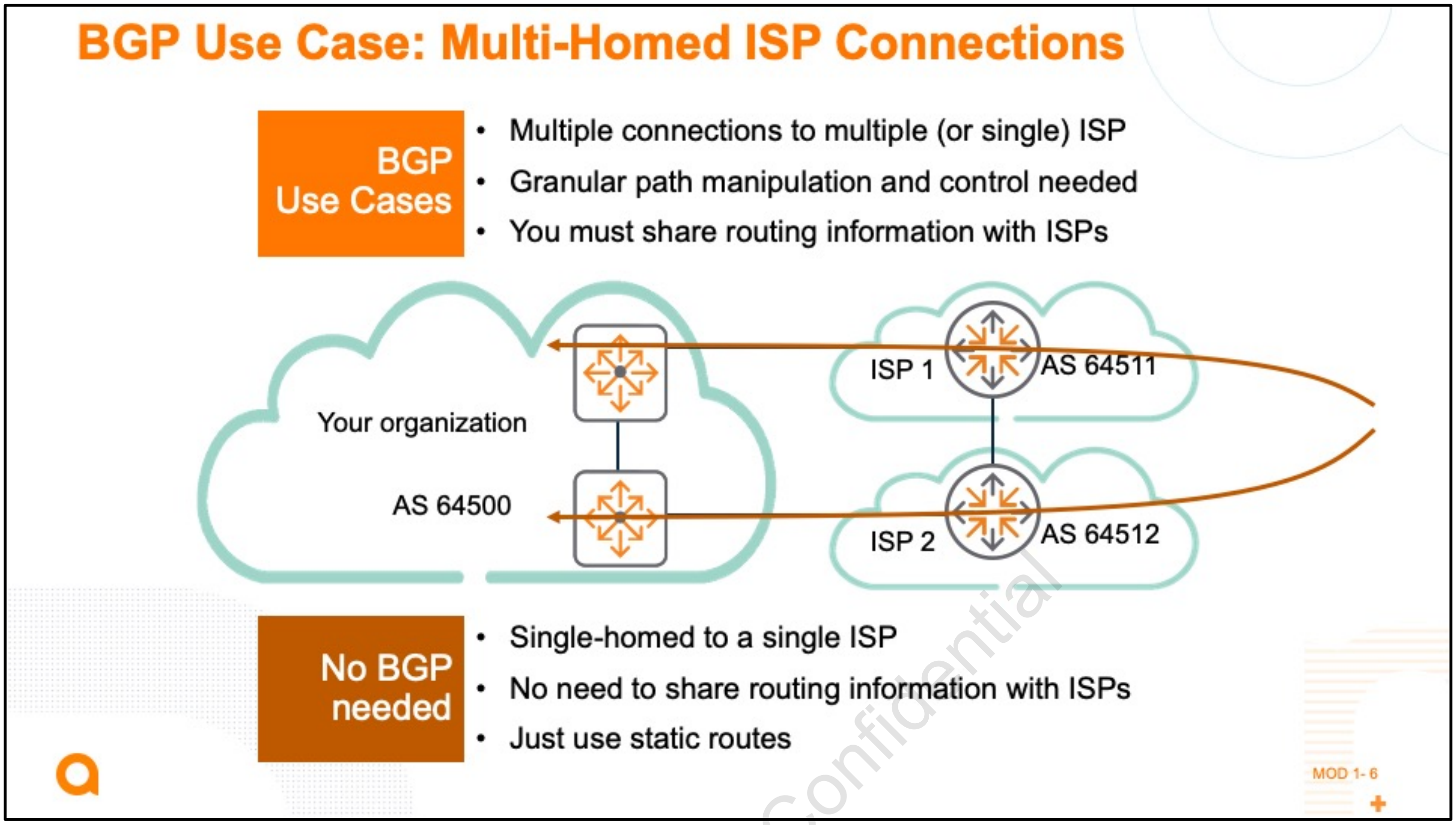
Aruba Training-Confidential

You may know that a Layer 3 switch is a device that can perform both Layer 2 switching functions and Layer 3 routing functions, depending on how you configure the device.

A router is a device that is purpose-built for routing, although some routers are modular, and may accommodate Layer 2 switching modules.

The point is, both devices perform a routing function. While performing these L3 functions, there is no significant functional difference between the two devices.

So, when discussing routing technology in this module, the term “router” is typically used to refer to any such device, unless there is some specific need to differentiate between the two.



BGP Use Case: Multi-Homed ISP Connections

BGP Use Cases

- Multiple connections to multiple (or single) ISP
- Granular path manipulation and control needed
- You must share routing information with ISPs

No BGP needed

- Single-homed to a single ISP
- No need to share routing information with ISPs
- Just use static routes

MOD 1- 6

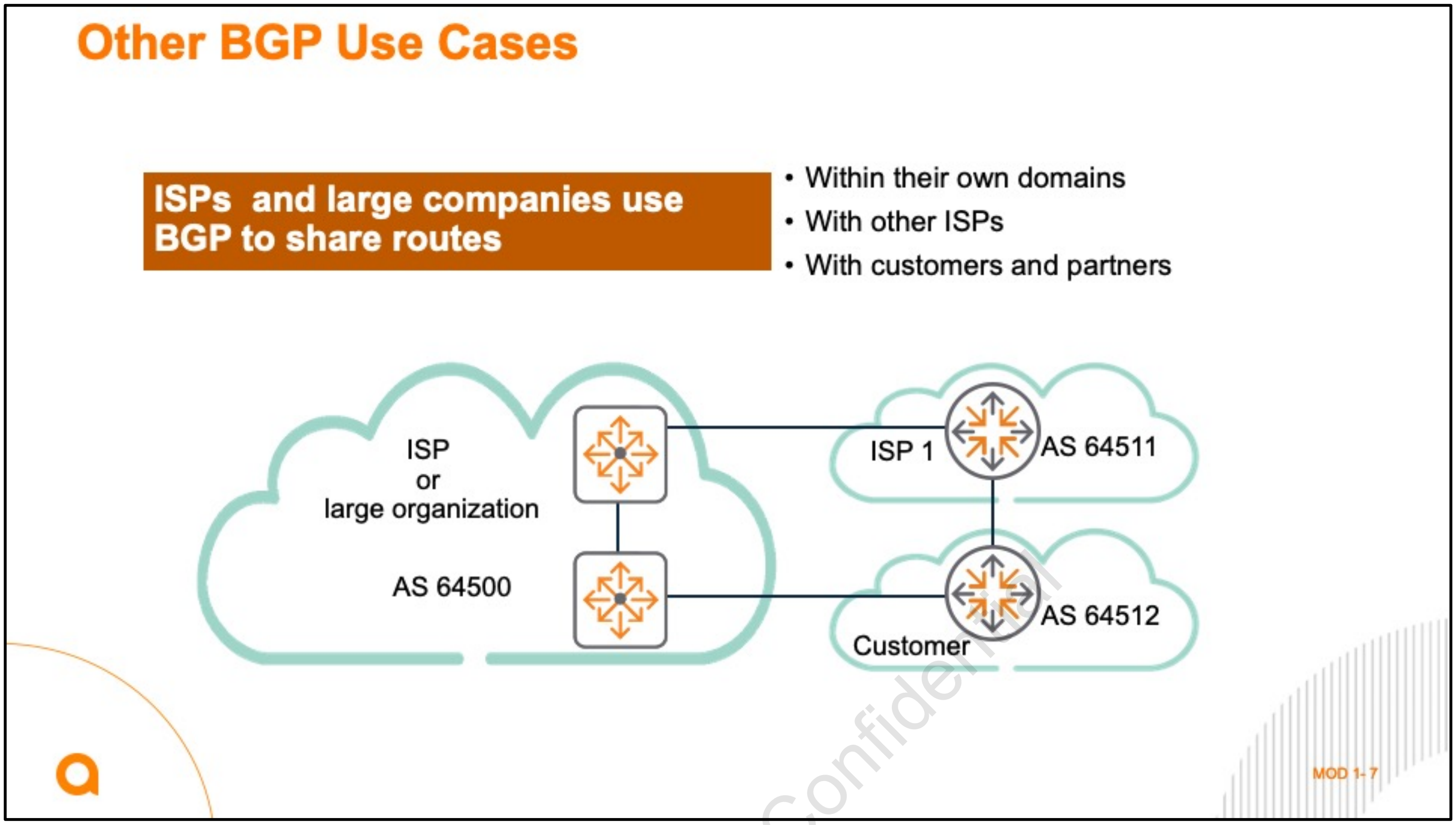
Aruba Training-Confidential

BGP is an exterior gateway protocol (EGP), which means that it communicates routes between different organizations, typically Internet Service Providers (ISPs). The figure illustrates BGP's primary use case - a company must communicate with multiple ISP routers. Another use case is when a large company connects to a single provider, and you need more granular path manipulation and control than is available with simple static routes. In other words, You only need to do so when the company specifically needs to communicate routing information with the ISP.

You also might need to advertise different IP blocks over different connections. This allows the ISPs to advertise your company's networks to the rest of the world, such that some paths may be best reached via ISP1, while others are best reached via ISP 2.

| In many cases, BGP is not required. If you only connect to a single ISP, and no granular path control or manipulation is needed, you can simply use static routes, reaching all internet sites via your ISP routers. The ISP handles announcing the block of public IP addresses assigned to the company.

The figure's dual-ISP example shows a different router connected to each ISP. This improves resiliency over a single connection but can be slightly more complex because it requires two BGP routers, which communicate with each other as well as the ISP routers.



Other BGP Use Cases

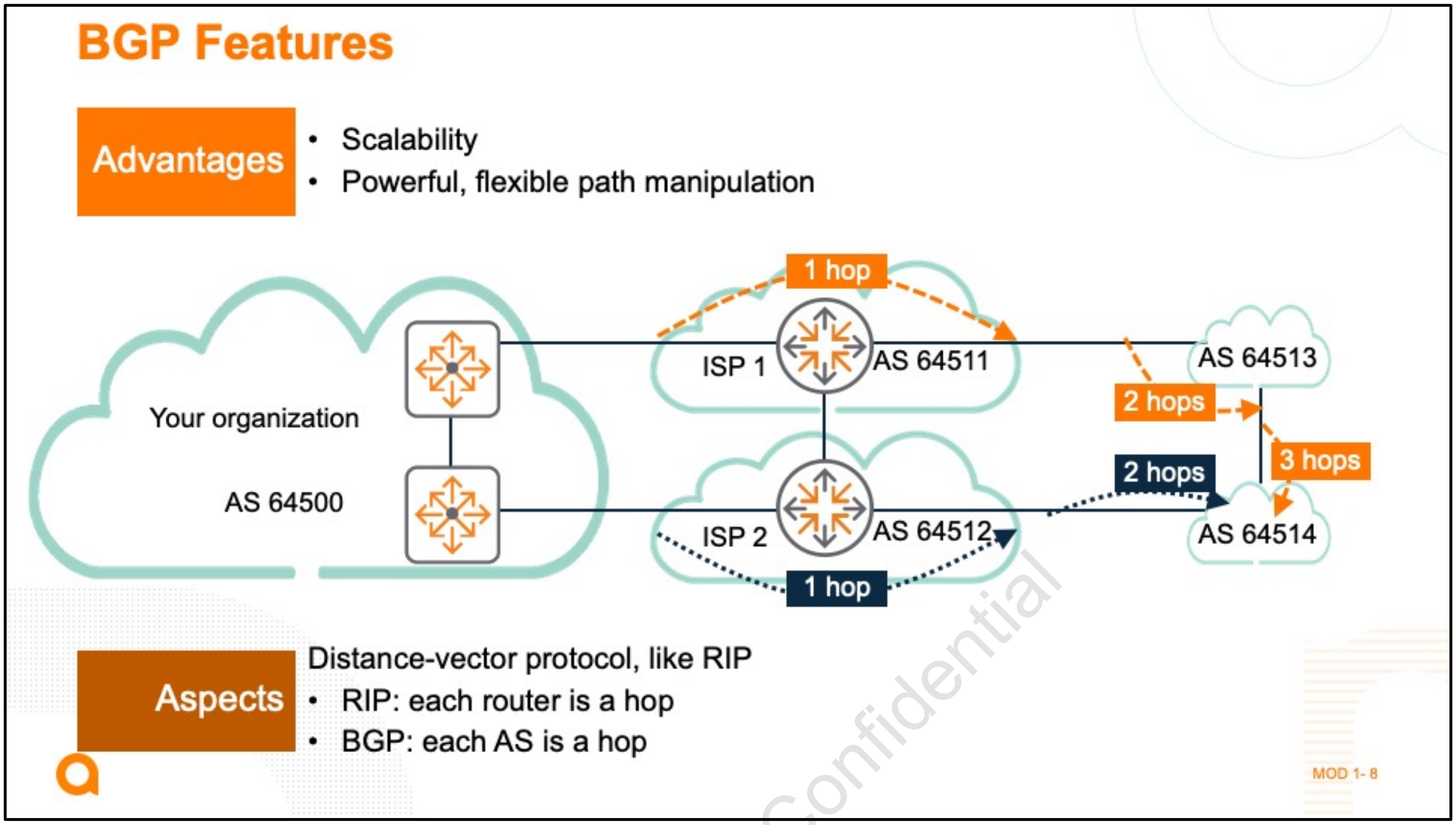
ISPs and large companies use BGP to share routes

- Within their own domains
- With other ISPs
- With customers and partners

ISPs and some bigger companies can have very large, complex networks – too large for OSPF. These organizations may use BGP to share routes within their domains, with other ISPs, and with customers and partners.

Note: Implementing a complete BGP solution for a service provider or ISP or for a large organization that involves confederations is beyond the scope of this course. This course focuses on how smaller organizations can connect to a service provider using BGP when static routes are not sufficient or do not meet the company's accessibility policies.

Aruba Training-Confidential



BGP Features

Advantages

- Scalability
- Powerful, flexible path manipulation

Aspects

Distance-vector protocol, like RIP

- RIP: each router is a hop
- BGP: each AS is a hop

MOD 1- 8

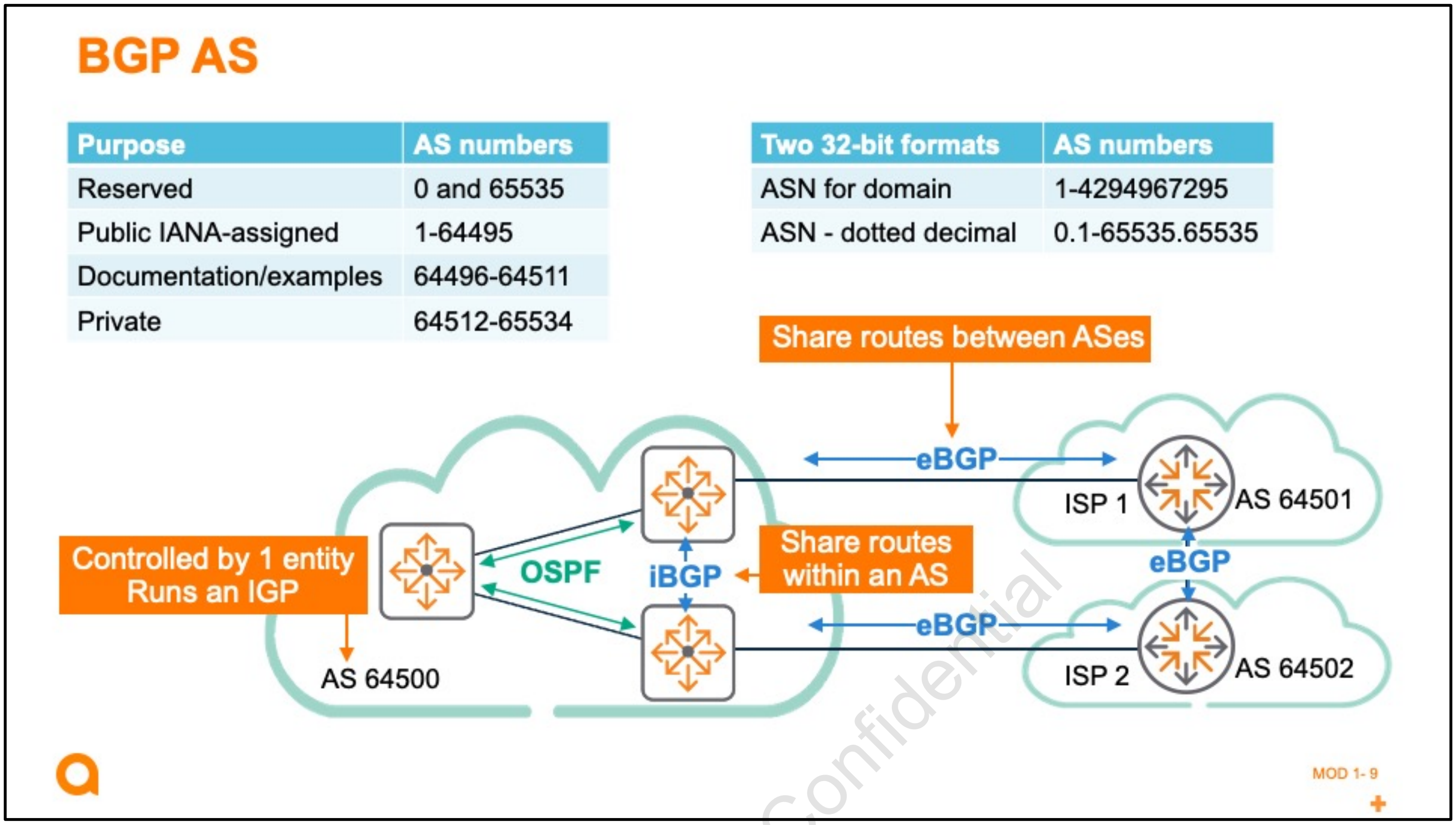
Aruba Training-Confidential

Because BGP is designed for use on the Internet, it can handle millions of routes. The protocol also permits a high level of administrative control for each organization. Organizations can aggregate routes, filter routes, and manipulate attributes in routes.

Like RIP, BGP is a kind of distance-vector routing protocol. However, with RIP the “hop count” is based on the number of routers to traverse or hop over. With BGP, it is based on the number of ASs to traverse. As shown in the figure, using ISP 1 your company is three ASs away from any subnet inside AS 64514 (including the target AS). However, your company must only traverse two ASs to those subnets via ISP 2. All other things being equal, the shortest path is considered the best path. You will soon learn to use other criteria to control BGP path decisions – you will soon be the BGP puppet master!

So, if you don’t care about pathing into or out of your AS, you can simply use static routes. Some paths may be sub-optimal, but in many scenarios, it will not affect the user experience, nor your performance or security requirements. However, if you need this level of control, you need BGP.

You can also use VSF to combine two physical AOS-CX switches that support BGP, such as 6300 switches. You can then connect to the two ISPs on different physical chassis, but still retain the simplicity of a single (logical) BGP router for both connections.



BGP AS

Purpose	AS numbers
Reserved	0 and 65535
Public IANA-assigned	1-64495
Documentation/examples	64496-64511
Private	64512-65534

Two 32-bit formats	AS numbers
ASN for domain	1-4294967295
ASN - dotted decimal	0.1-65535.65535

Share routes between ASes

Controlled by 1 entity Runs an IGP

Share routes within an AS

MOD 1- 9

Aruba Training-Confidential

The AS forms the fundamental building block of BGP, and each organization such as a company or ISP generally has its own AS, identified by a unique number. Originally, BGP supported only 16-bit AS numbers, ranging between 0 and 65535. The figure shows how AS numbers are used.

Because 16-bit numbers allow for only 65535 unique numbers, Request for Comments (RFC) 4893 establishes 32-bit AS numbers, increasing the range of AS numbers to 4 billion. (AOS-CX switches supports 32-bit numbers in one of two formats:

1-4294967295: The autonomous system number (ASN) for the domain.

0.1-65535.65535: The autonomous system number (ASN) for the domain in dotted decimal format.

When a company determines that it needs to run BGP in a multihomed scenario, it must obtain a valid public AS from the proper regional authority.

BGP operates in one of two forms depending on the autonomous systems to which the communicating neighbors belong:

| Neighbors in different autonomous systems communicate with external BGP (eBGP).

| Neighbors in the same AS communicate with internal BGP (iBGP).

The same router might use eBGP with one neighbor and iBGP with another, as shown in the figure.

Please understand that iBGP is not the only routing protocol running within the AS.

| The AS also runs an interior gateway protocol (IGP) such as OSPF or RIP, which is localized to the AS and handles routing the local traffic. Typically, just a few routers in the AS—the routers that connect to other autonomous systems and perhaps some core routers—run BGP in addition to the IGP.

Note: In this course and in the lab material, private AS numbers are used.

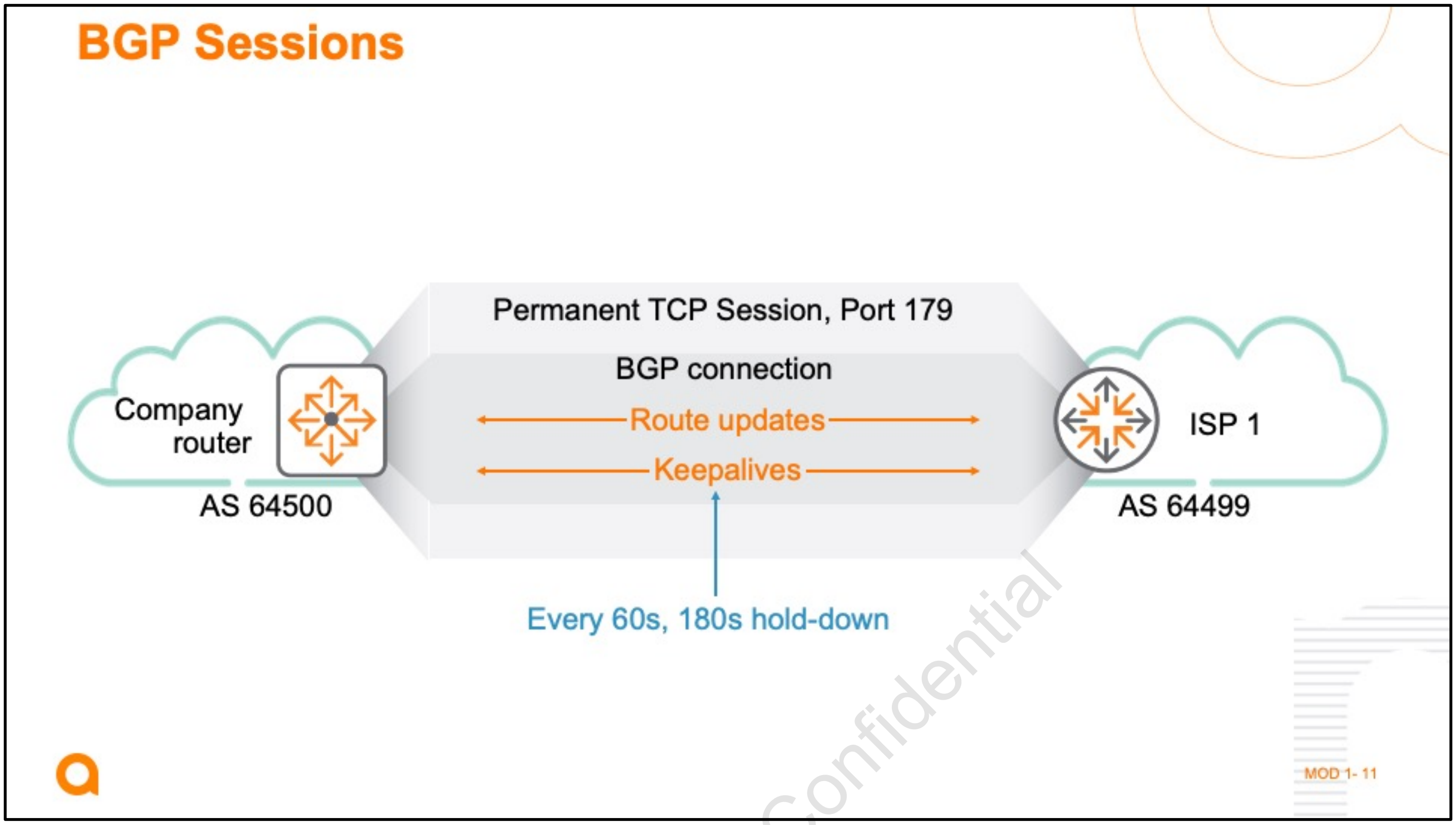
Aruba Training-Confidential

Establishing BGP Connections

MOD 1- 10

You now learn about establishing BGP connections.

Aruba Training-Confidential



BGP Sessions

Permanent TCP Session, Port 179

BGP connection

Route updates

Keepalives

Company router

AS 64500

ISP 1

AS 64499

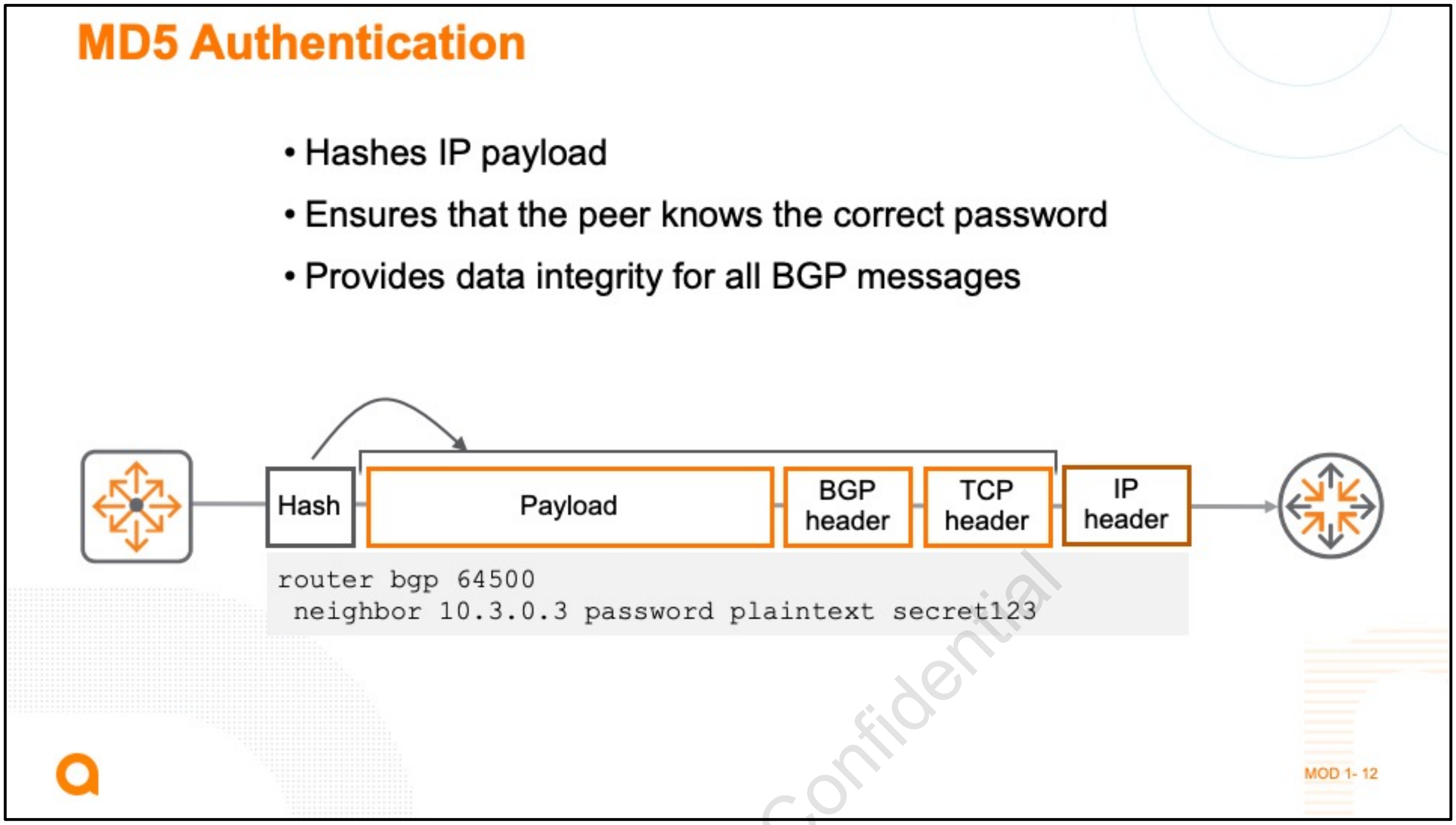
Every 60s, 180s hold-down

MOD 1- 11

When you enable OSPF on an interface, and it automatically discovers neighbors. With BGP you must manually define BGP neighbors to establish connections. BGP neighbors establish a TCP session on port 179. They then open a BGP connection within the session. This connection carries all messages and routing updates between the neighbors, and the TCP session provides a degree of reliability for the communications. The BGP connection also features keepalives, which helps each router determine its neighbor’s reachability. AOS-CX switches send keepalives every 60 second by default, and have a hold-down timer of 180 seconds, three times the keepalive interval.

With this BGP connection in place, routers can exchange their BGP routing tables.

Aruba Training-Confidential



MD5 Authentication

- Hashes IP payload
- Ensures that the peer knows the correct password
- Provides data integrity for all BGP messages

```
router bgp 64500
 neighbor 10.3.0.3 password plaintext secret123
```

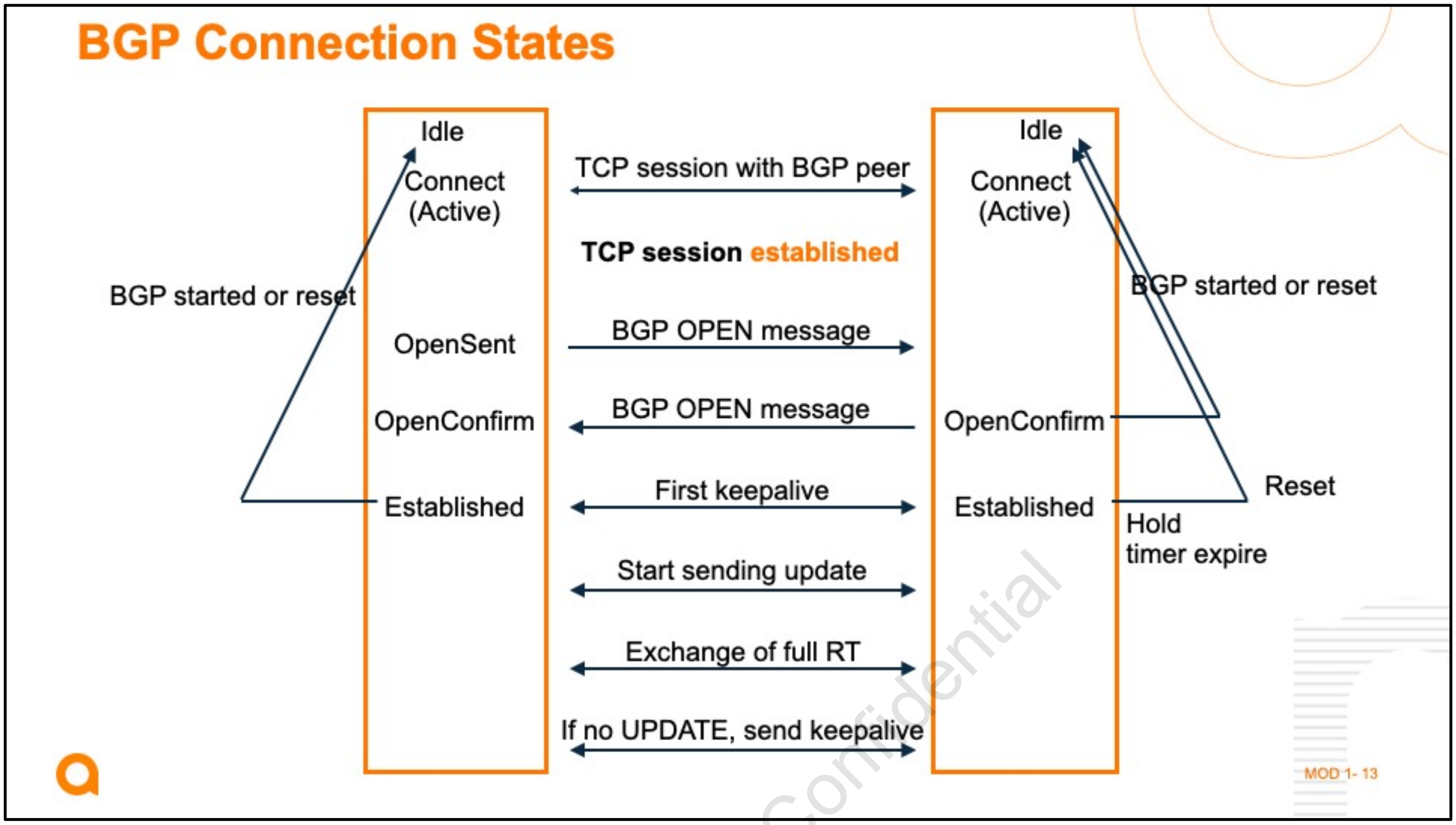
MOD 1- 12

Aruba Training-Confidential

The ISP might require your router to authenticate. You simply add a password when you define the neighbor. The BGP neighbors use this password to hash the TCP payload of packets that carry the BGP messages, appending the hash to the packet. If the neighbors' passwords do not match, the TCP session for the BGP connection fails to establish.

Because the neighbors add the MD5 hash to each packet sent on the TCP session, MD5 authentication provides data integrity for all BGP messages and for underlying TCP session. In other words, each neighbor knows that the proper neighbor sent the message and that no one has altered the message. However, MD5 does not provide protection against eavesdropping.

To assign the password, use the syntax shown in the figure.



BGP Connection States

Idle
Connect (Active)
OpenSent
OpenConfirm
Established

BGP started or reset

TCP session with BGP peer

TCP session established

BGP OPEN message

BGP OPEN message

First keepalive

Start sending update

Exchange of full RT

If no UPDATE, send keepalive

Idle
Connect (Active)
OpenConfirm
Established

BGP started or reset

Reset

Hold timer expire

MOD 1- 13

Aruba Training-Confidential

In the lab, you will see how the BGP neighbors pass into Active state and stayed there when they encountered issues. You will see that a healthy session should have the Established state.

The figure above and the following notes provide more details about the session establishment for your reference.

To establish the BGP connection, the neighbors first open the TCP session on port 179. They then send BGP open messages with a list of parameters for the session and, assuming the parameters match, finally send the first keepalive of the session.

You can trace the states through which a BGP connection might pass:

The BGP connection begins in the Idle state and returns to this state when the session establishment times out without success or an error occurs.

The connection typically transitions immediately to the Connect state, in which the router attempts to contact its neighbor at TCP port 179 and waits for a response.

If the router is set to Passive TCP mode for this neighbor, the session transitions to Active mode, rather than to Connect mode. In this mode, the router waits for the neighbor to initiate the session.

The connection also moves to the Active state if the router does not receive a response to its TCP session initiation. In this state, the router waits for the neighbor's response and also, after a timer expires, might try to initiate the connection again. The router might wait in Active state for a significant period of time if the TCP session fails to establish. If you see a router in Active mode after several minutes, the neighbor might be unreachable.

After the BGP neighbors have established the TCP session, the router sends an open message, either immediately or after a configurable delay. The connection also transitions to the OpenSent state or directly to the OpenConfirm state depending on whether the router has already received the neighbor's open message.

In the OpenSent state, the router waits for the neighbor's open message. Once it receives that message, it verifies that necessary parameters match (for example, the neighbor's AS matches the remote AS configured on your router). Assuming that the parameters match, the router sends a keepalive and transitions to the OpenConfirm state.

If the parameters do not match, the router sends an error notification and returns to the Idle state. Therefore, if you see a router stuck in the Idle state, you might want to check that you configured the local AS and neighbor AS correctly.

In the OpenConfirm state, the router waits for neighbor's keepalive message and also sends its own keepalive. It also listens for error notifications such as one indicating that the open message parameters do not match.

After receiving the neighbor's keepalive, the connection transitions to the Established state.

The connection remains in this state during normal operation. The neighbors continue to send keepalives. If a BGP router fails to receive a number of keepalives, it ends the connection. Note that the keepalive period is quite long (5 minutes, by default).

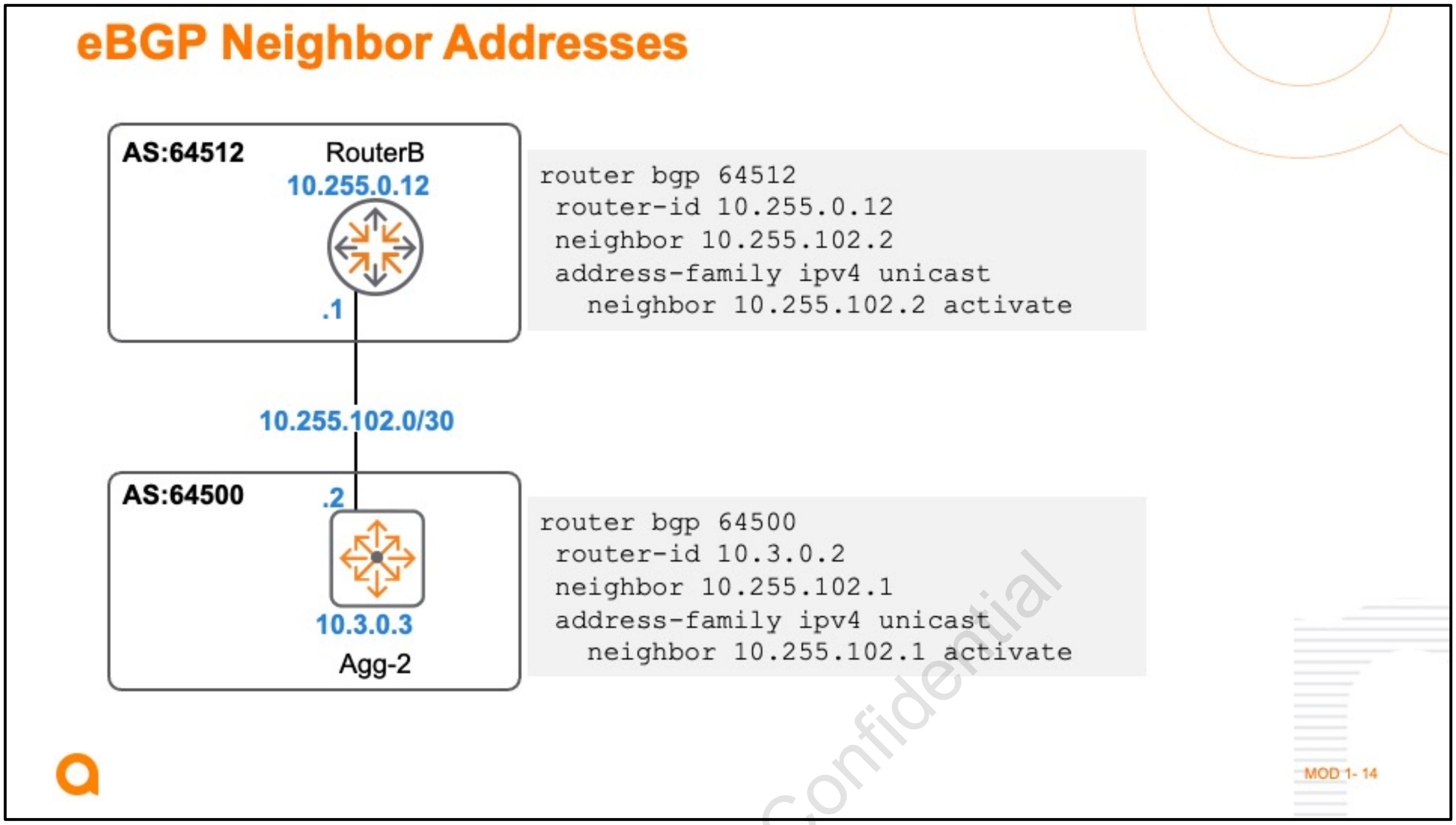
New eBGP neighbors exchange their full BGP routing tables. After the eBGP neighbors have exchanged their BGP routing tables, they only exchange updates. For example, a BGP router sends an update if configuration changes affect routing policies or if a topology change has occurred.

Routes remain in the BGP routing table unless an update occurs. A router can

Aruba Training-confidential

also request a route refresh from a neighbor which might be necessary so that the router can apply new policies to routes that it has already received.

Aruba Training-Confidential



eBGP Neighbor Addresses

```
router bgp 64512
 router-id 10.255.0.12
 neighbor 10.255.102.2
 address-family ipv4 unicast
   neighbor 10.255.102.2 activate
```

```
router bgp 64500
 router-id 10.3.0.2
 neighbor 10.255.102.1
 address-family ipv4 unicast
   neighbor 10.255.102.1 activate
```

MOD 1- 14

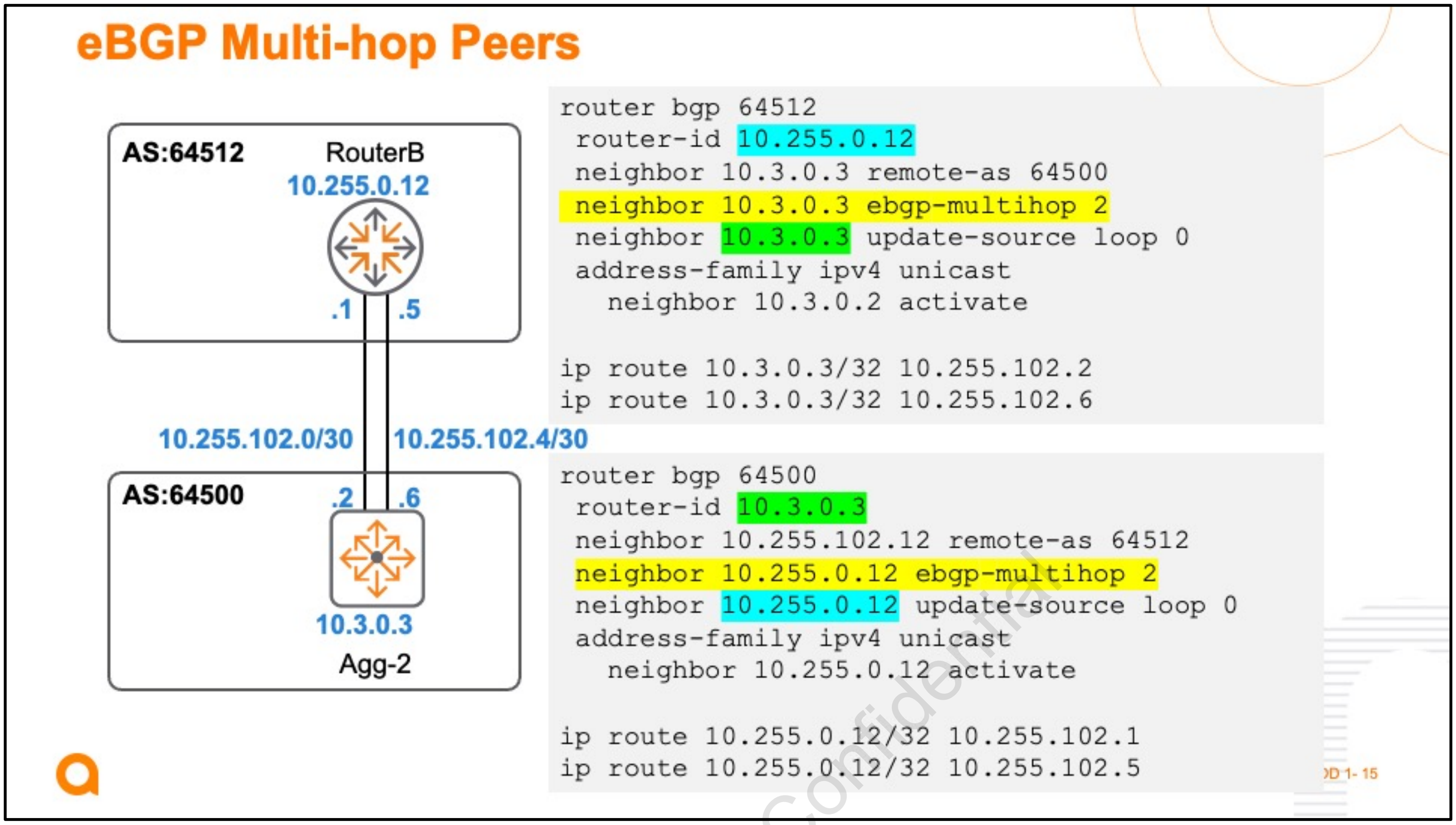
Aruba Training-Confidential

The figure shows connectivity for two eBGP neighbors.

By default, eBGP neighbors must establish the session on a direct Layer 3 connection. In other words, they have IP addresses on the same subnet - a Layer 2 switch between them is no problem. The eBGP protocol enforces this criterion by checking that the neighbor IP address is on a network connected to the local BGP router and by using messages with a TTL of 1.

The configuration is shown in the figure. You configure each router with their respective AS. Since the AS numbers are different, the peers automatically know that this is an eBGP connection. Then you specify a router ID, just like with OSPF – often a loopback address. Now you specify the physical IP address of the directly-connected neighbor. Finally, you must activate the neighbor. This begins the TCP port 179 handshake and establishes a BGP session.

One issue with this directly-connected eBGP requirement is redundancy. If that single link fails, there is no alternate path. In many cases, there is only one link connecting eBGP peers anyway, so this is not an issue. But what if you want better redundancy? What if you want to have multiple links between eBGP peers?



eBGP Multi-hop Peers

```
router bgp 64512
 router-id 10.255.0.12
 neighbor 10.3.0.3 remote-as 64500
 neighbor 10.3.0.3 ebgp-multihop 2
 neighbor 10.3.0.3 update-source loop 0
 address-family ipv4 unicast
   neighbor 10.3.0.2 activate

ip route 10.3.0.3/32 10.255.102.2
ip route 10.3.0.3/32 10.255.102.6
```

```
router bgp 64500
 router-id 10.3.0.3
 neighbor 10.255.102.12 remote-as 64512
 neighbor 10.255.0.12 ebgp-multihop 2
 neighbor 10.255.0.12 update-source loop 0
 address-family ipv4 unicast
   neighbor 10.255.0.12 activate

ip route 10.255.0.12/32 10.255.102.1
ip route 10.255.0.12/32 10.255.102.5
```

Aruba Training-Confidential

The figure shows an exception to the directly-connected eBGP rule. If you have multiple physical links between eBGP peers, they will benefit by peering via their loopback addresses – to maximize resiliency.

As shown in the figure, you configure eBGP peers to use their loopback addresses, just like you typically do with iBGP peers. The difference is that with eBGP you must configure the multihop command highlighted in the figure. The number 2 at the end of the command is a hop count – it sets the TTL field to 2, so the peer can only be 2 router hops away.

The static routes are configured so BGP can actually reach the loopback addresses. Router B has two static routes to Agg-2's loopback address via the directly connected links shown in the figure. Similarly, Agg-2 has two static routes to RouterB's loopback interface.

As configured, both links have the same administrative distance. The router will load-share across these links. If desired, you could adjust administrative distance or some other BGP metric, so that one link was the primary link. The other link is only used if the primary link fails.

Also understand that by default, each router sends BGP packets with its physical interface IP as the source address. Router-B would send a packet to Agg-2 with source address

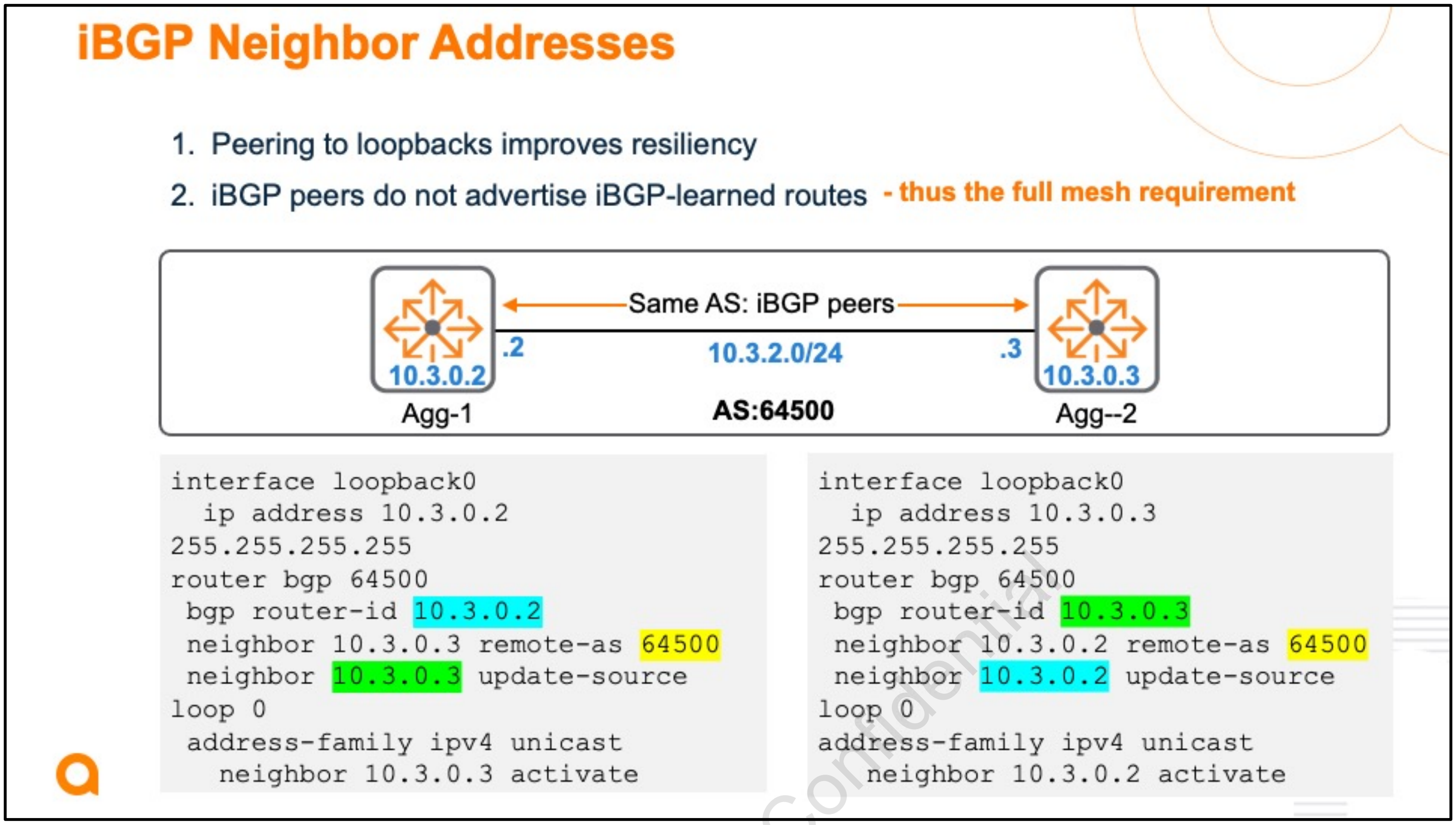
10.255.102.1 or .5. Agg-2 would receive this and think, “Who is this? I’m configured with neighbor 10.255.0.12.” Agg-2 would discard this packet, and peering fails. Thus the neighbor x update-source loopback 0 commands. You instruct Router B to send all BGP packets to Agg-2 with source address 10.255.0.12, which matches Agg-2’s configured neighbor. Packets are accepted, peering succeeds.

Another use case for eBGP multihop is when you must reach the ISP router that runs BGP through another router or routers. To set up this solution, you must set the eBGP multihop to the correct number (2 if one intermediate router intervenes, three if two routers intervene, and so on). Your router also requires a route to the ISP router’s IP address, of course. However, it is not recommended to have intervening routers between eBGP neighbors, and you should avoid this setup whenever possible.

Note: Understand that the multi-hop feature described here is fine to deploy, and quite common. You are only being asked to avoid having routers between eBGP peers.

As of AOS-CX 10.2, routing switches support BGP connections over GRE tunnels.

Aruba Training-Confidential



iBGP Neighbor Addresses

1. Peering to loopbacks improves resiliency
2. iBGP peers do not advertise iBGP-learned routes - thus the full mesh requirement

Same AS: iBGP peers

10.3.0.2 .2 10.3.2.0/24 .3 10.3.0.3

Agg-1 AS:64500 Agg--2

```
interface loopback0
  ip address 10.3.0.2
255.255.255.255
router bgp 64500
 bgp router-id 10.3.0.2
 neighbor 10.3.0.3 remote-as 64500
 neighbor 10.3.0.3 update-source
loop 0
 address-family ipv4 unicast
   neighbor 10.3.0.3 activate
```

```
interface loopback0
  ip address 10.3.0.3
255.255.255.255
router bgp 64500
 bgp router-id 10.3.0.3
 neighbor 10.3.0.2 remote-as 64500
 neighbor 10.3.0.2 update-source
loop 0
address-family ipv4 unicast
   neighbor 10.3.0.2 activate
```

Aruba Training-Confidential

Peering to loopbacks improves resiliency

Unlike eBGP neighbors, iBGP neighbors need not be directly connected. The iBGP peers in the example shown are directly connected, but there could be a large cloud of routers running OSPF between them. If you peer iBGP routers using a physical interface, and that physical interface or link fails, the peer is down, and routing is broken.

Instead, you configure iBGP routers to peer to each other's loopback address, and they rely on a highly redundant OSPF system to route packets between these loopbacks. If one path fails, the peer session can be maintained via another path.

The configuration is shown in the figure – very similar to an eBGP multi-hop configuration. Please take a moment to review. Especially important is to use the update-source command and activate the neighbor

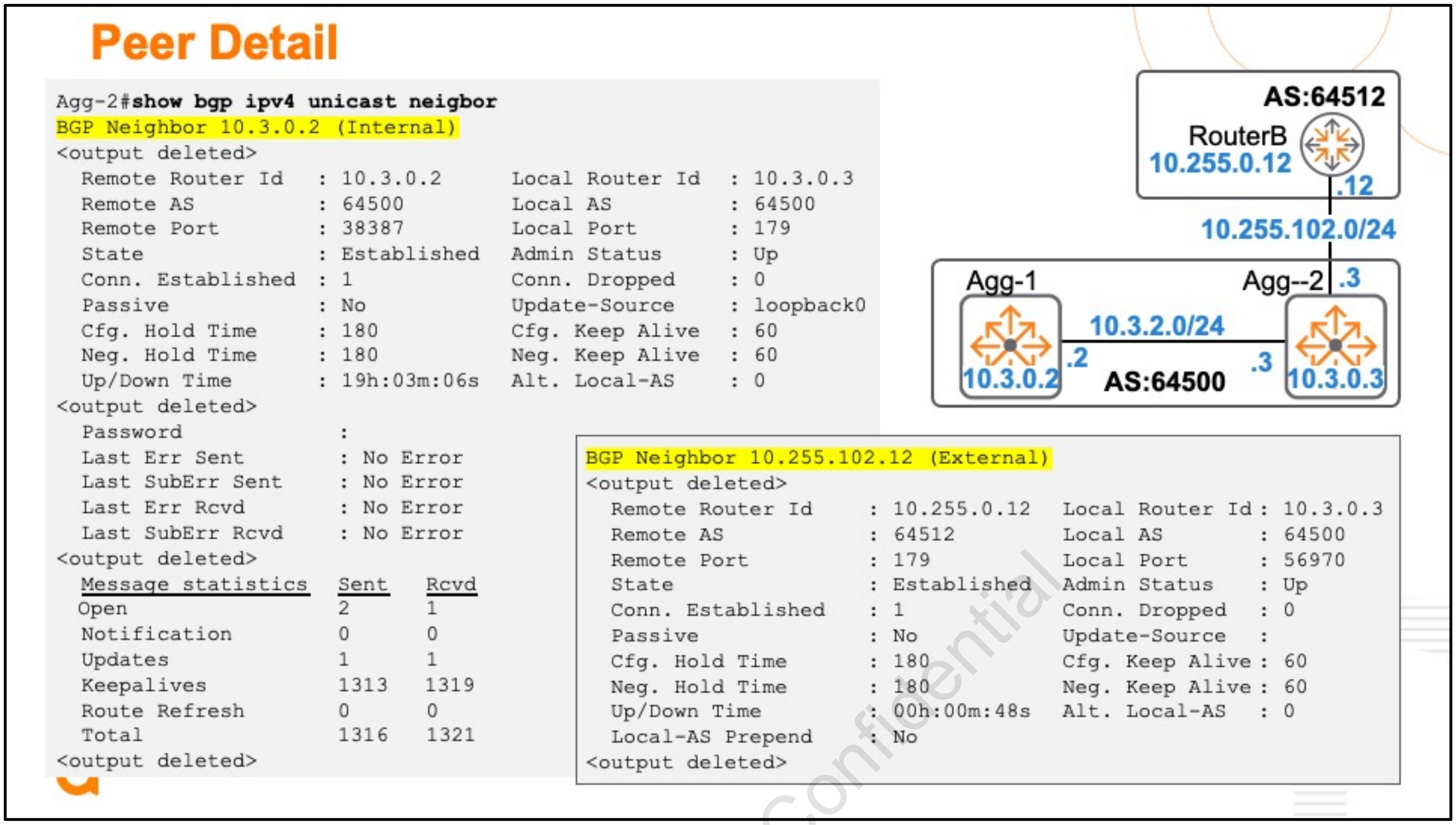
iBGP peers do not advertise iBGP-learned routes

iBGP peers do not advertise routes learned from other iBGP peers. This may seem odd if you are used to OSPF, or any other IGP. OSPF is based on the fundamental idea that OSPF Router A shares what it learns from OSPF Router B with all other routers. Not so with iBGP. This leads to the iBGP full-mesh requirement. Every iBGP router must peer with every other iBGP router in your AS.

Most organizations only have a few BGP routers, so this is not a problem. Two router – two peers, 3 routers – 6 peers (2 per router), 4 routers – 12 peers (3 per router).

But large organizations may have dozens or even hundreds of iBGP routers. Imagine how large and unwieldy the configuration shown in the figure would be if you had to configure 900 peers on each router! Later in this module, you will learn about Route Reflectors and Confederations to mitigate this issue.

Aruba Training-Confidential



Peer Detail

```
Agg-2#show bgp ipv4 unicast neigbor
BGP Neighbor 10.3.0.2 (Internal)
<output deleted>
  Remote Router Id   : 10.3.0.2      Local Router Id   : 10.3.0.3
  Remote AS          : 64500         Local AS          : 64500
  Remote Port        : 38387         Local Port        : 179
  State              : Established   Admin Status      : Up
  Conn. Established  : 1             Conn. Dropped     : 0
  Passive            : No            Update-Source     : loopback0
  Cfg. Hold Time     : 180           Cfg. Keep Alive   : 60
  Neg. Hold Time     : 180           Neg. Keep Alive   : 60
  Up/Down Time       : 19h:03m:06s   Alt. Local-AS     : 0
<output deleted>
  Password              :
  Last Err Sent         : No Error
  Last SubErr Sent      : No Error
  Last Err Rcvd         : No Error
  Last SubErr Rcvd      : No Error
<output deleted>
  Message statistics    Sent     Rcvd
  Open                  2        1
  Notification          0        0
  Updates               1        1
  Keepalives            1313     1319
  Route Refresh         0        0
  Total                 1316     1321
<output deleted>
```

```
BGP Neighbor 10.255.102.12 (External)
<output deleted>
  Remote Router Id     : 10.255.0.12    Local Router Id : 10.3.0.3
  Remote AS            : 64512          Local AS        : 64500
  Remote Port          : 179            Local Port      : 56970
  State                : Established    Admin Status    : Up
  Conn. Established    : 1              Conn. Dropped   : 0
  Passive              : No             Update-Source   :
  Cfg. Hold Time       : 180            Cfg. Keep Alive : 60
  Neg. Hold Time       : 180            Neg. Keep Alive : 60
  Up/Down Time         : 00h:00m:48s    Alt. Local-AS   : 0
  Local-AS Prepend     : No
<output deleted>
```

Aruba Training-Confidential

Now that you can describe and configure BGP peering, you should be able to validate it. The figure shows output for Agg-2's show bgp ipv4 unicast neighbor command, edited to focus our attention. The output is split into two sections – one for each BGP neighbor or peer.

Neighbor 10.3.0.2 is an Internal (iBGP) peer. You see the local and remote router IDs and ASes, and you see that indeed, BGP uses TCP port 179.

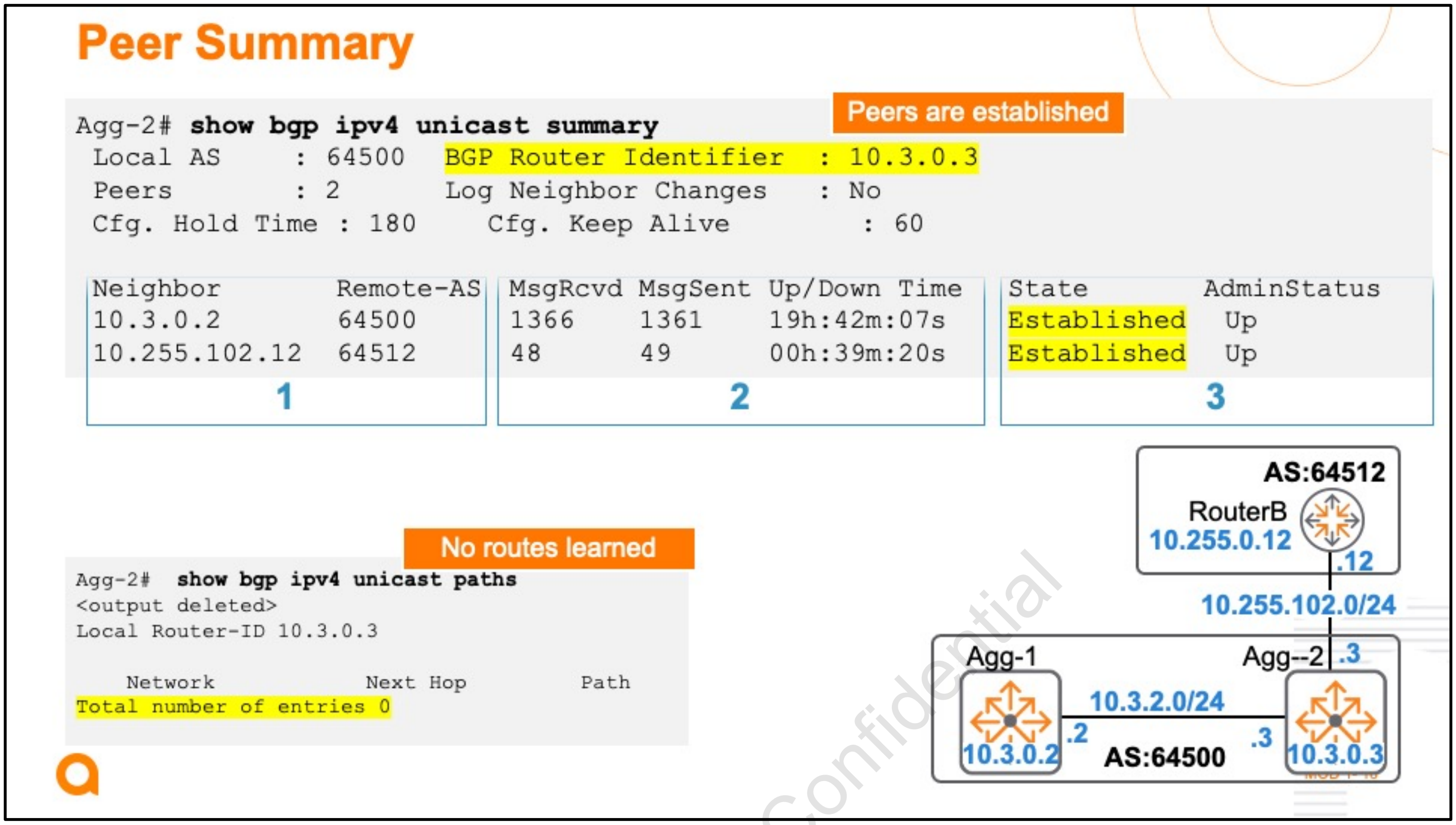
The state/status is Established/Up. You can validate that the update source is loopback 0, as well as the Hold and Keepalive timers of 180 and 60 seconds, respectively.

Pay attention to the Up/Down Time – consistently short up times mean that the peers are consistently losing connectivity – you may have a network stability issue.

You can see that there is no password configured, and no error messages received – nice!

Finally, you see the messages you recently learned about – Open, Notification, updates, keepalives, route refresh.

The next section is for neighbor 10.255.102.12, Agg-2's eBGP peer, which holds the same information as the first section.



Peer Summary

Peers are established

```
Agg-2# show bgp ipv4 unicast summary
 Local AS      : 64500   BGP Router Identifier   : 10.3.0.3
 Peers         : 2       Log Neighbor Changes    : No
 Cfg. Hold Time : 180     Cfg. Keep Alive          : 60

 Neighbor          Remote-AS  MsgRcvd MsgSent Up/Down Time   State        AdminStatus
 10.3.0.2          64500      1366    1361    19h:42m:07s    Established   Up
 10.255.102.12     64512      48      49      00h:39m:20s    Established   Up
```

No routes learned

```
Agg-2#  show bgp ipv4 unicast paths
<output deleted>
Local Router-ID 10.3.0.3

    Network              Next Hop          Path
Total number of entries 0
```

Aruba Training-Confidential

The detail you get with the previous show bgp ipv4 unicast neighbor command is often quite useful. However, you can see a nice summary of this information – often all you really need, with the command show bgp ipv4 unicast summary.

Click on the numbers below each column of output to learn more.

The summary command proves that you have successful BGP peering, but the show bpg ipv4 unicast paths command shows that no routes have been learned. So, you explore BGP advertisements next.

Focus on the sections in the figure's CLI output labeled "1", "2", and "3" while studying the information below.

CLI Output Section 1

The neighbor column shows who this router is peering with. It is very effective for both learning and troubleshooting to compare the output of show commands to a network diagram. Play a game – look at the diagram, and try to predict what you will see in the output. The diagram shows that router Agg-2 has two peers. One is an eBGP peer connected on a single physical link, so the peer is formed with the physical address – 10.255.102.12. the other peer is an iBGP neighbor, and so should peer with the loopback address – 10.3.0.2.

You know how to confirm whether a peer is iBGP or eBGP, yes? The first line of the output shows that Agg-2's Local AS is 64500. Neighbor 10.3.0.2 has the same AS – iBGP. Neighbor 10.255.102.12 is in a different AS – 64512. Thus, you know this is an eBGP neighbor.

CLI Output Section 2

The Messages Received and Sent columns are greater than zero, so you know that the peers are successfully communicating with each other. You will likely see these number slowly increment over time.

The Up/Down Time can help you with forensics. Neighbor 10.3.0.2 has been up for over 19 hours and 42 minutes, while the other peer has only been up for about 39 minutes. You might ask, “What happened 39 minutes ago?” Perhaps this router was just brought up for the first time, so this is normal. Or maybe this peer was established months ago. You might explore log files and other resources to determine why this connection was dropped.

CLI Output Section 3

The State and AdminStatus columns look good – Established and Up, respectively. This means you have fully successful peer relationships with every neighbor shown in your network diagram. During a high-pressure troubleshooting situation, you might pull up your network diagram, and see that Agg-2 should have 3 connected neighbors, so you look at the output and validate that you see the rows with Established and Up. If you only see 2 rows, you know that one router is missing (or your diagram is out of date). Or you see three rows, one for each neighbor, but one peer is stuck in one of the other states you learned about.

Do you have those states memorized yet? Go back and review the figure titled “BGP Connection States” (Idle, Connect, OpenSent, OpenConfirm, Established)

Aruba Training Confidential

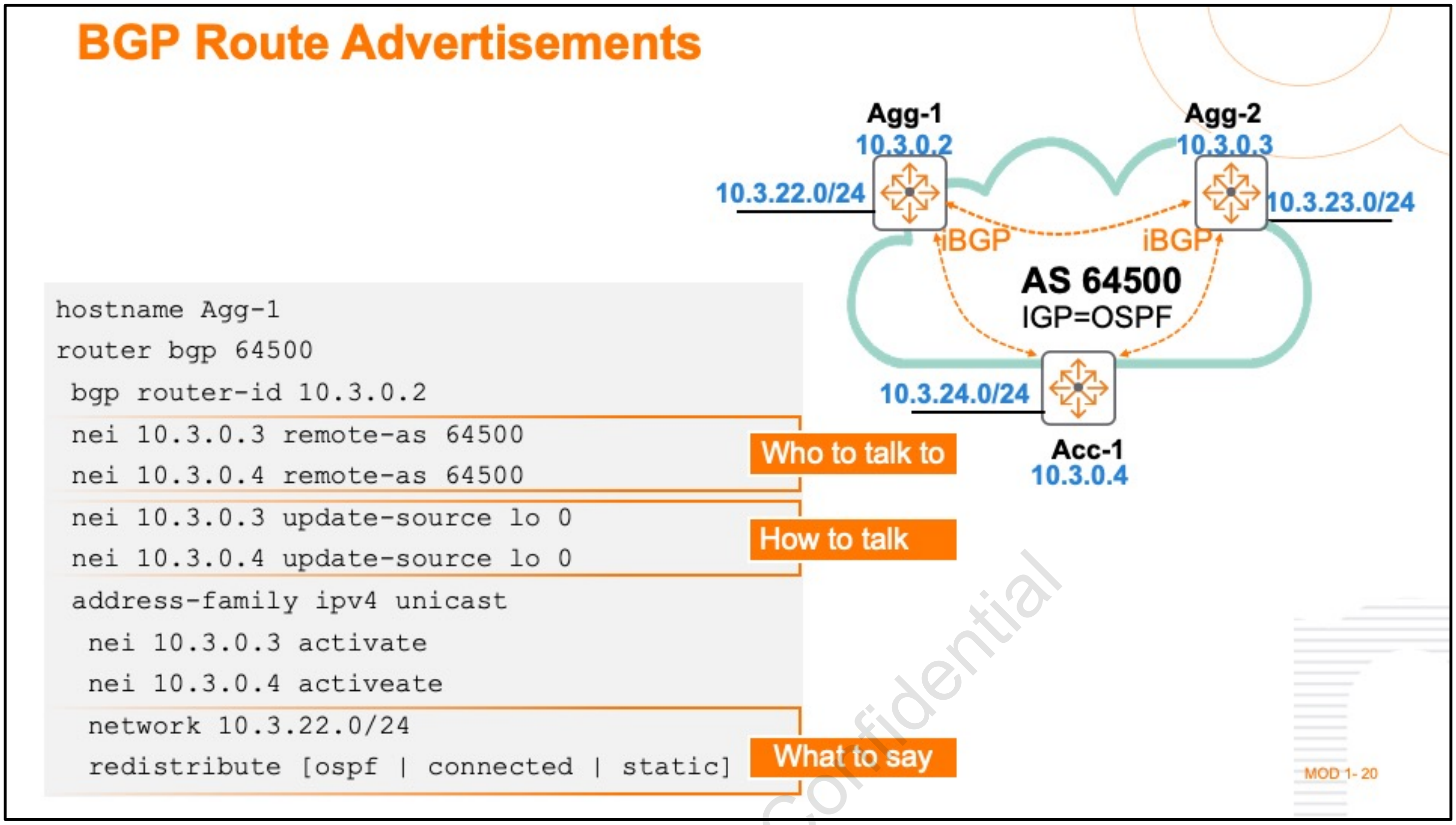
BGP Advertisements

aruba

a Hewlett Packard Enterprise company

Now we explore BGP advertisements.

Aruba Training-Confidential



BGP Route Advertisements

```
hostname Agg-1
router bgp 64500
 bgp router-id 10.3.0.2
 nei 10.3.0.3 remote-as 64500
 nei 10.3.0.4 remote-as 64500
 nei 10.3.0.3 update-source lo 0
 nei 10.3.0.4 update-source lo 0
 address-family ipv4 unicast
  nei 10.3.0.3 activate
  nei 10.3.0.4 activeate
  network 10.3.22.0/24
  redistribute [ospf | connected | static]
```

Who to talk to

How to talk

What to say

Aruba Training-Confidential

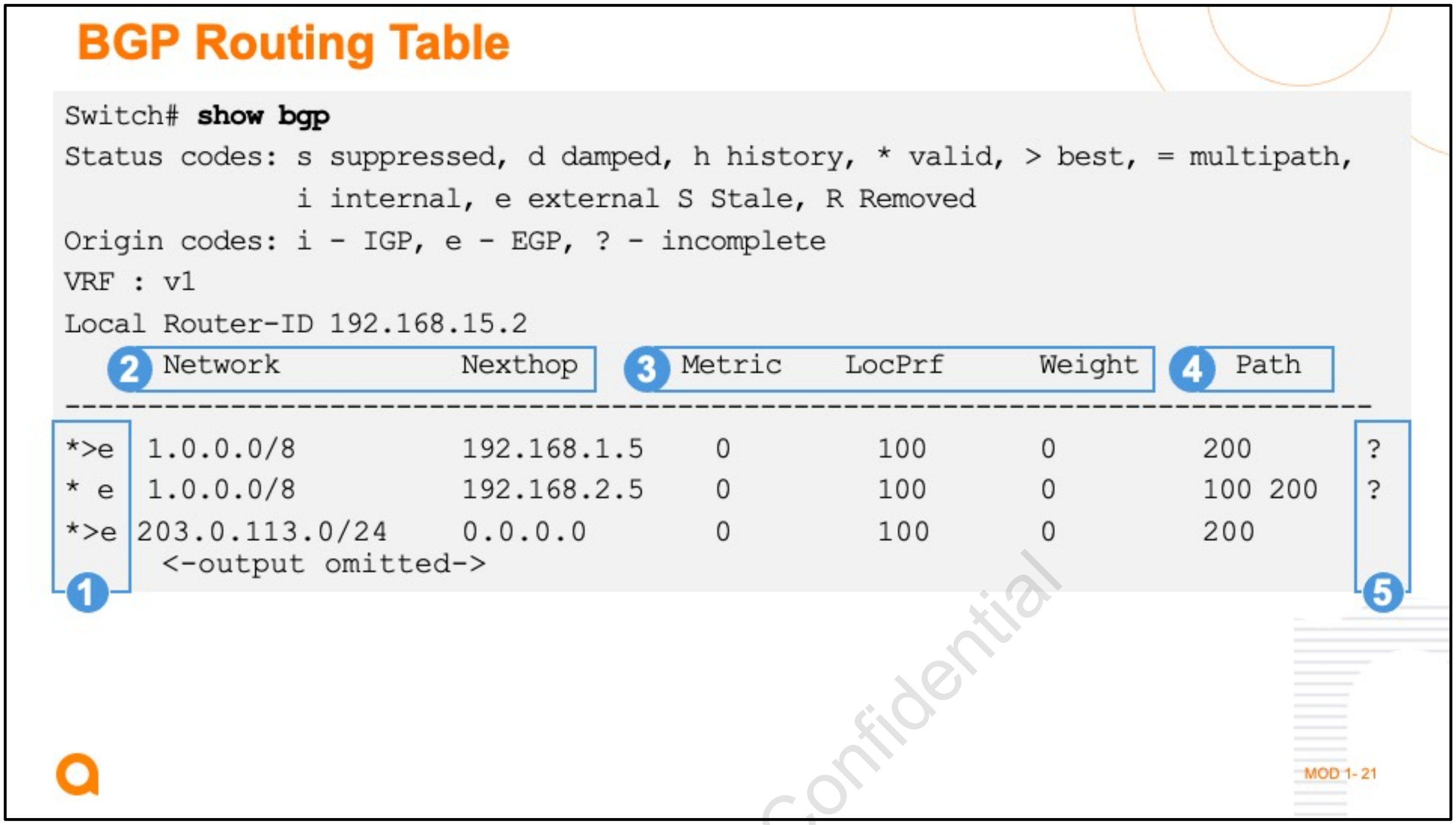
You just learned how to configure BGP peers – you told the routers who to talk to. Now you must ensure they know what to say. What routes will BGP neighbors advertise? How do you get them to advertise routes as desired?

You know that iBGP routers do not advertise iBGP-learned routes – thus the full-mesh iBGP requirement. But they do advertise eBGP learned routes. So, how do you get a router to originate a route advertisement that will propagate through the system?

If you want BGP routers to originate advertise any other routes, such as routes to local networks, you must add those routes to the BGP routing table explicitly. For example, your BGP router needs to advertise the company's IP block to the ISP. You must add a route to that block to the BGP routing table.

The figure shows the BGP configuration for Agg-1, with new syntax – network 10.3.22.0/24. You are telling Agg-1 that it should advertise this route. This makes sense, since this network is directly connected to Agg-1. Similarly, Agg-2 has network 10.3.23.0/24, and Access-1 has network 10.3.24.0/24.

Another way to get routers to originate a network is to redistribute it from another routing protocol – OSPF, directly connected, or static routes.
Let's look at some details about how BGP advertises these routes.



BGP Routing Table

```
Switch# show bgp
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed
Origin codes: i - IGP, e - EGP, ? - incomplete
VRF : v1
Local Router-ID 192.168.15.2
     Network              Nexthop        Metric    LocPrf      Weight     Path
------------------------------------------------------------------------------
*>e  1.0.0.0/8            192.168.1.5    0         100         0          200       ?
* e  1.0.0.0/8            192.168.2.5    0         100         0          100 200   ?
*>e 203.0.113.0/24        0.0.0.0        0         100         0          200
     <-output omitted->
```

Aruba Training-Confidential

Now your BGP routers are establishing a connection, advertising routes, receiving routes, and building a BGP routing table. You should be good at interpreting this table. Click on the numbers to explore.

Focus on the number-labeled section of the CLI output in the figure while reading the information below.

1

Each route has several symbols at the beginning. The status codes at the top of the output summarize the meaning of these symbols. An asterisk (*) indicates that the route is valid. This means that the next hop is reachable – it will be added to the route table for packet forwarding. The “e” means the route is from a foreign AS or protocol, while “i” means the route is a native BGP route from the local AS. The greater than sign (>) indicates that this is the best path to reach a destination. In the example, two paths are listed for destination network 1.0.0.0/8. The path via 192.168.1.5 is the one that is added to the route table and used to forward packets. The other path is not used unless the best path is lost. This “best path” selection is based on the other columns shown in the output – Metric, Local Preference (LocPrf), Weight, path, and origin, which you will explore soon.

2

These fields indicate the destination network address and the next hop. You can use these fields to validate the expected networks exist, and that the next hop is as expected. You compare your network documentation to the output of this command. If you expect 100 networks and only 99 are listed, you determine the missing route and trace it back to its source. If route exists but the next hop incorrect, there would not be an asterisk (*) in the first column – you may have forgotten to add the next-hop-self parameter to one of this router's peers.

3

All these columns help BGP to determine the best path, along with the Path and Origin Code columns. You will explore this in detail soon.

The metric indicates the cost for reaching the next hop; unlike with OSPF, this metric is not the most important criteria for selecting between routes. You might modify this to suggest how other AS's forward traffic into your AS.

You might modify Local Preference (LocPrf) to define preferred paths toward external destinations in other Ass.

You can also use weight to control route selection, but this value has local significance only. It is not advertised to other routers.

4

The AS path attribute is included with every BGP route advertisement. A router in the AS that owns this route includes its own AS number in the path attribute, then sends it out to other ASs. A BGP router in a different AS receives the route, adds its AS to the path, and so on. In this way, the AS path indicates every AS which traffic will traverse to reach the destination. A shorter AS path is preferred.

Now you know why there is a ">" symbol next to the top entry for 1.0.0.0/8 – why it is the best path. This router learned about 1.0.0.0/8 from router 192.168.1.5, which is in AS 200. You know this because only AS 200 is listed in the Path attribute. It also learned it from next-hop router 192.168.2.5 in AS 100, which learned it from 200. Since this is a longer AS path, and all other criteria are equal, the shorter path is considered "best". The AS path attribute is also used as BGP's primary loop elimination technique. If a router receives a route with its own AS in the AS path, it drops the route.

Aruba Training Confidential

5

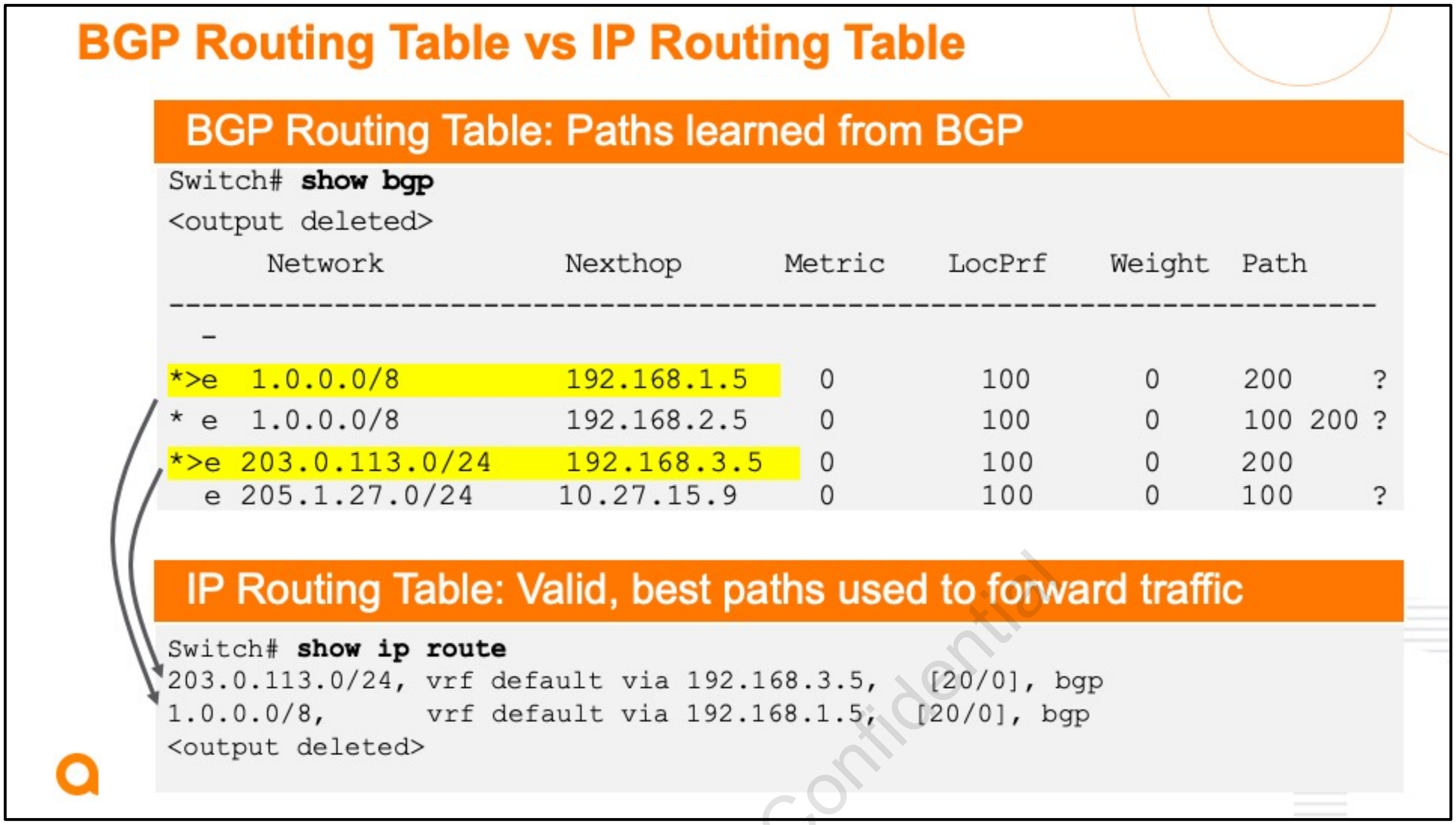
The last column is the “origin code” – where the route came from.

“i” means the route probably came from an IGP (OSPF, RIP, etc) and was injected into BGP with the network command.

, e – from the old Exterior Gateway Protocol (EGP) which is deprecated. If you see this code, you may have time traveled back to the 1980’s!

The “?” symbol means “incomplete” – there is not enough information available to determine how this route originated. This typically means the route was redistributed into BGP from a foreign routing protocol – perhaps OSPF or static.

Aruba Training-Confidential



BGP Routing Table vs IP Routing Table

BGP Routing Table: Paths learned from BGP

```
Switch# show bgp
<output deleted>
      Network              Nexthop       Metric     LocPrf     Weight  Path
---------------------------------------------------------------------------
  -
*>e  1.0.0.0/8            192.168.1.5     0          100         0      200     ?
* e  1.0.0.0/8            192.168.2.5     0          100         0      100 200 ?
*>e 203.0.113.0/24        192.168.3.5     0          100         0      200
  e 205.1.27.0/24        10.27.15.9       0          100         0      100     ?
```

IP Routing Table: Valid, best paths used to forward traffic

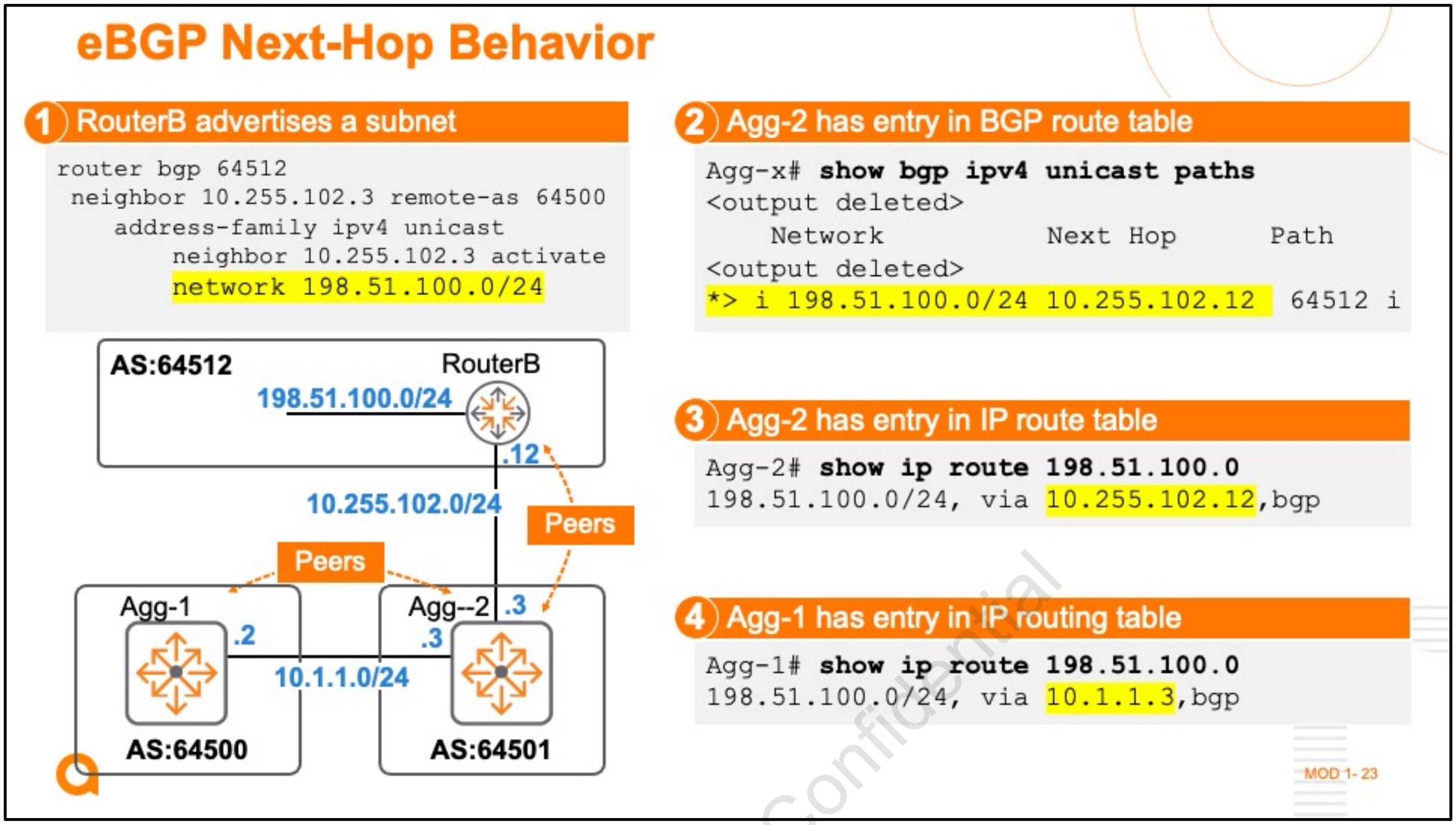
```
Switch# show ip route
203.0.113.0/24, vrf default via 192.168.3.5,  [20/0], bgp
1.0.0.0/8,      vrf default via 192.168.1.5,  [20/0], bgp
<output deleted>
```

Aruba Training-Confidential

You must distinguish between the BGP route table and the IP route table. The BGP table includes all routes learned via BGP – some of these routes are valid paths, some not. Some of these routes are the best paths, some are not. Only the valid, best paths are added to the IP routing table, which is actually used to forward packets.

So, to see what paths BGP has learned, look at the BGP routing table. To see which of those paths are used to forward packets, look at the IP routing table. Two of the four BGP route table entries were added to the IP route table. The other 1.0.0.0/8 route was not added because it is not the best path – only best paths are added to the IP route table. The 205.1.27.0/24 route was not added because it is not a valid route. You can tell this because the route lacks the * symbol.

The route may be valid because although the router knows about this route, it cannot find the next-hop IP address. Let’s explore this issue of invalid next hop addresses.



eBGP Next-Hop Behavior

1 RouterB advertises a subnet

```
router bgp 64512
 neighbor 10.255.102.3 remote-as 64500
    address-family ipv4 unicast
        neighbor 10.255.102.3 activate
        network 198.51.100.0/24
```

2 Agg-2 has entry in BGP route table

```
Agg-x# show bgp ipv4 unicast paths
<output deleted>
     Network            Next Hop       Path
<output deleted>
*> i 198.51.100.0/24 10.255.102.12  64512 i
```

3 Agg-2 has entry in IP route table

```
Agg-2# show ip route 198.51.100.0
198.51.100.0/24, via 10.255.102.12,bgp
```

4 Agg-1 has entry in IP routing table

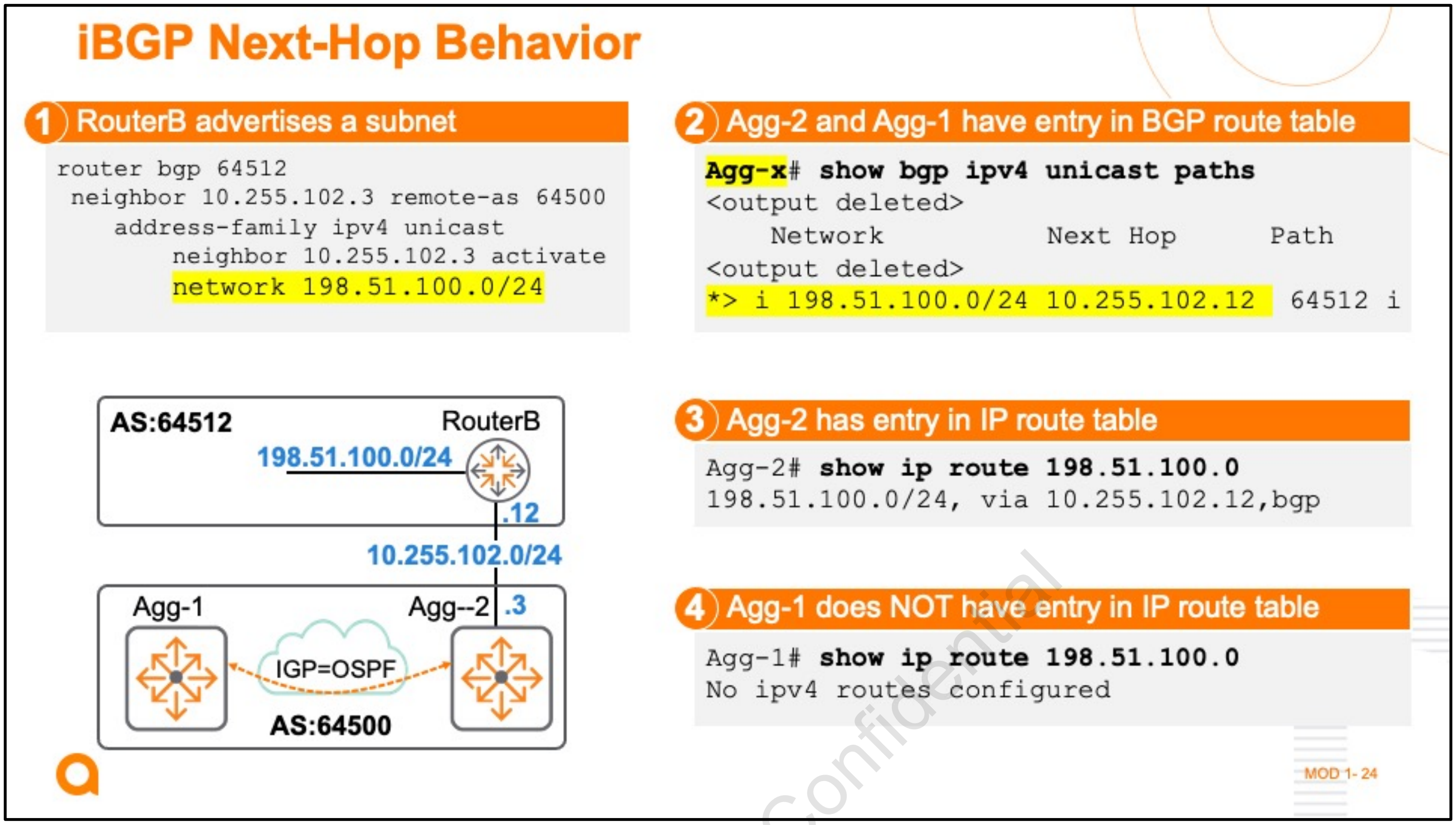
```
Agg-1# show ip route 198.51.100.0
198.51.100.0/24, via 10.1.1.3,bgp
```

In the figure, notice that Agg-2 has two eBGP peers – 10.1.1.2 and 10.255.102.12. Router B is configured to advertise network 198.51.100.0/24. Router B advertises this route out its interface with IP 10.255.102.12.

Aruba Training-Confidential

To validate this, you go to Agg-2 and use show bgp ipv4 unicast paths. Agg-2 has this route in their BGP routing table. Agg-2 learned this route, with a next hop address of 10.255.102.12. Agg-2 is directly connected to this route, and so can reach the next-hop address of 10.255.102.12. so, Agg-2 adds the route to its IP routing table – it knows it can actually reach this route.

eBGP peers share eBGP-learned routes, so Agg-2 advertises this route to its other peer – Agg-1. when Agg-1 advertises this route to an eBGP peer it changes the next-hop address to its own – 10.1.1.3. Agg-1 is directly connected to this next hop address, so it can definitely reach it. As shown in the figure, the path is added to Agg-1's IP route table.



iBGP Next-Hop Behavior

1) RouterB advertises a subnet

```
router bgp 64512
 neighbor 10.255.102.3 remote-as 64500
    address-family ipv4 unicast
        neighbor 10.255.102.3 activate
        network 198.51.100.0/24
```

2) Agg-2 and Agg-1 have entry in BGP route table

```
Agg-x# show bgp ipv4 unicast paths
<output deleted>
    Network           Next Hop       Path
<output deleted>
*> i 198.51.100.0/24 10.255.102.12   64512 i
```

3) Agg-2 has entry in IP route table

```
Agg-2# show ip route 198.51.100.0
198.51.100.0/24, via 10.255.102.12,bgp
```

4) Agg-1 does NOT have entry in IP route table

```
Agg-1# show ip route 198.51.100.0
No ipv4 routes configured
```

MOD 1- 24

Aruba Training-Confidential

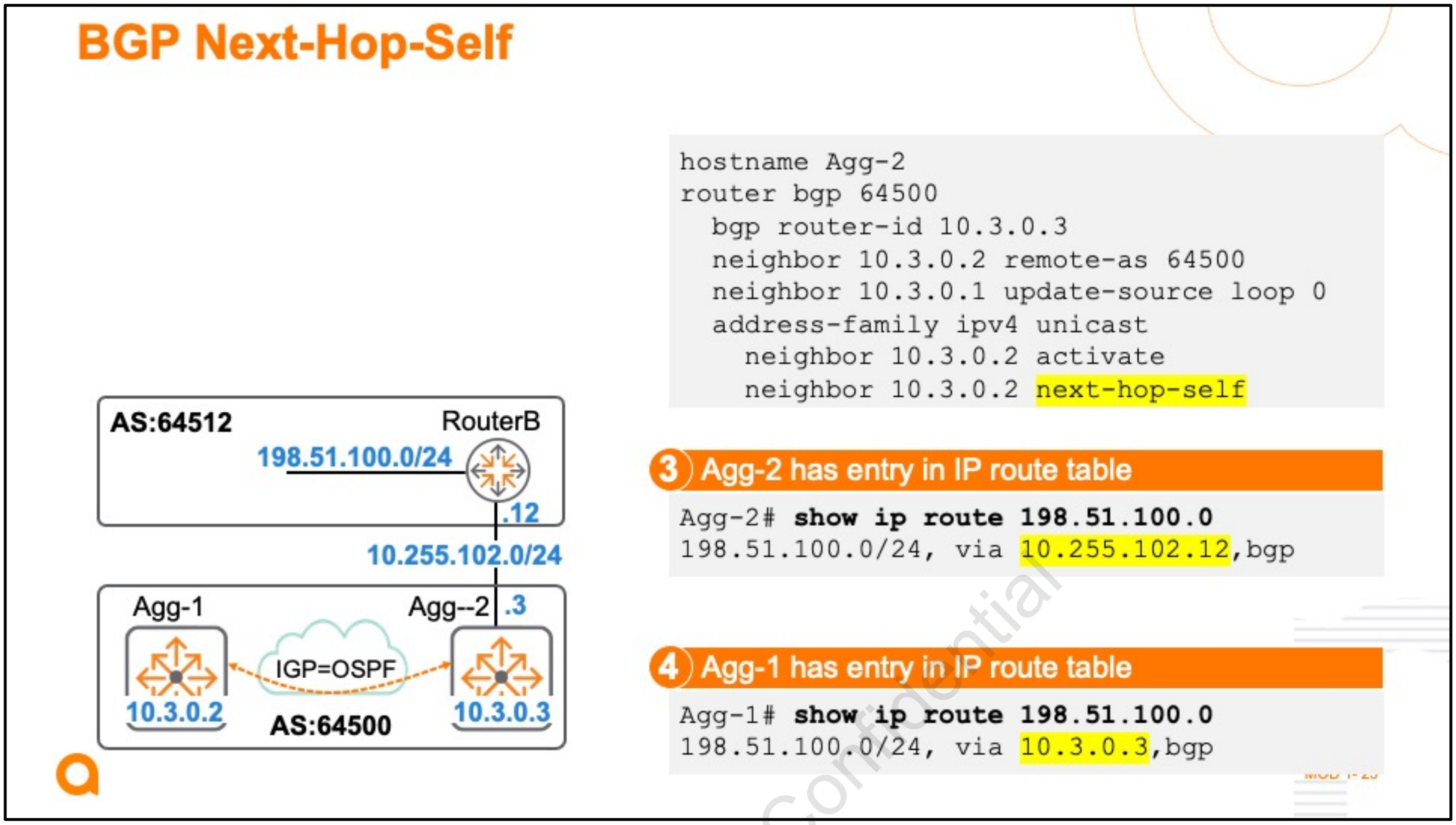
As before, Router B advertises 198.51.100.0/24, out its interface with IP 10.255.102.12.

To validate this, you go to Agg-2 and use show gp ipv4 unicast paths. Both routers show this route in their BGP route table – BGP has learned this route, with a next hop address of 10.255.102.12. Agg-2 is directly connected to that next hop, so it can reach it – the entry is added to its IP route table, and it routes packets to 198.51.100.0/24 – all good. To verify this, you look at the IP routing table and see that indeed the route was added to Agg-2's route table.

However, this route is NOT in the IP route table of Agg-1. In this scenario, Agg-1 and Agg-2 are iBGP peers, and iBGP peers do not change the next-hop IP address of eBGP-learned routes. Router B advertises the target subnet with next-hop 10.255.102.12. Agg-2 receives this and forwards it to Agg-1 with the same next hop - 10.255.102.12. Agg-1 is not directly connected to this next-hop, and OSPF has not learned about this route.

So, although Agg-1 has learned about the route via BGP, it cannot use the route – the next-hop is not valid, the entry is not added to the IP route table, and routing is broken for Agg-1

How do you solve this issue?



BGP Next-Hop-Self

```
hostname Agg-2
router bgp 64500
  bgp router-id 10.3.0.3
  neighbor 10.3.0.2 remote-as 64500
  neighbor 10.3.0.1 update-source loop 0
  address-family ipv4 unicast
    neighbor 10.3.0.2 activate
    neighbor 10.3.0.2 next-hop-self
```

3 Agg-2 has entry in IP route table

```
Agg-2# show ip route 198.51.100.0
198.51.100.0/24, via 10.255.102.12,bgp
```

4 Agg-1 has entry in IP route table

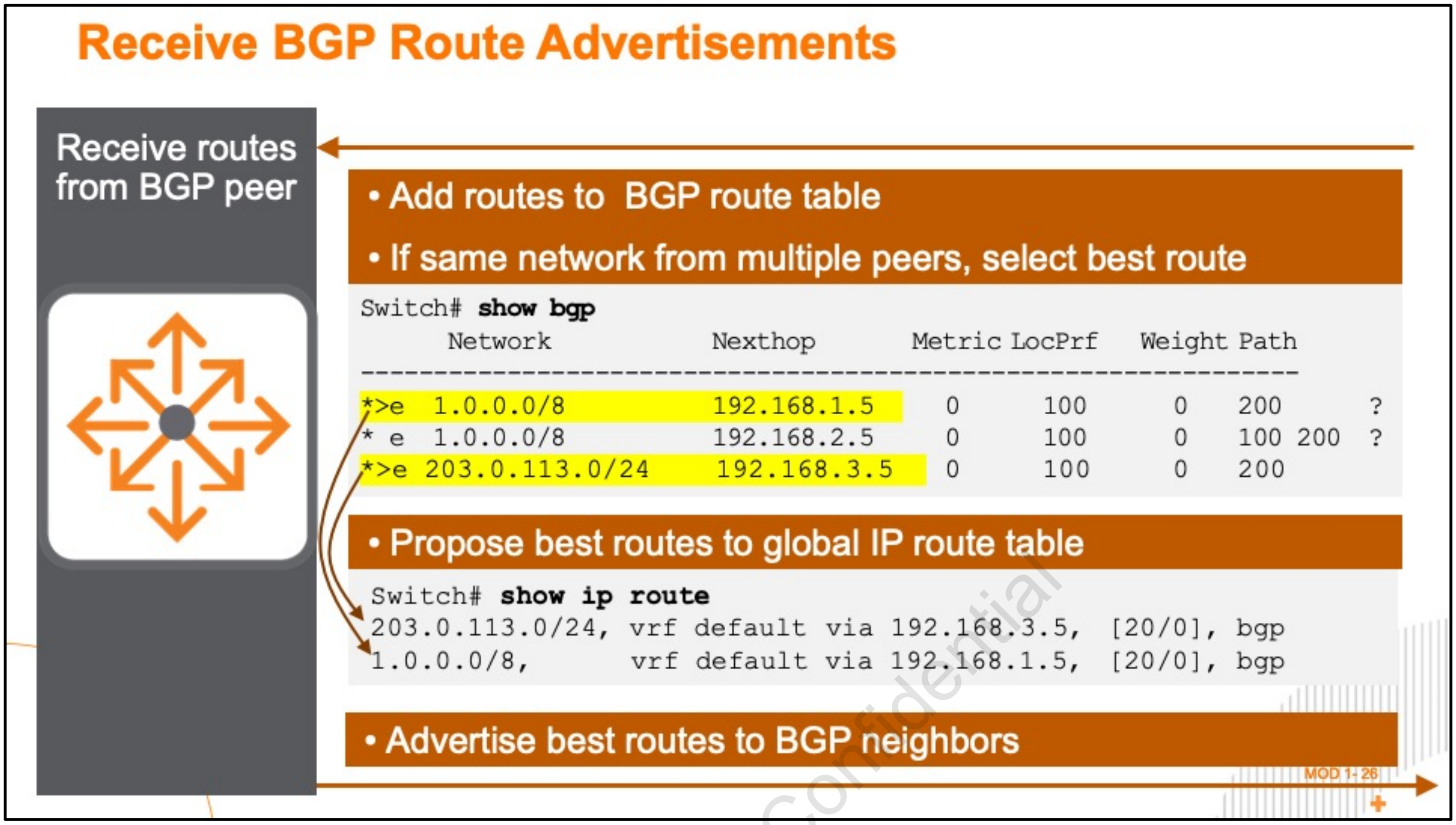
```
Agg-1# show ip route 198.51.100.0
198.51.100.0/24, via 10.3.0.3,bgp
```

Aruba Training-Confidential

The typical method to resolve the next-hop issue is by configuring your iBGP peers with the next-hop-self parameter, as shown in the figure. This makes BGP act a bit more like OSPF. Agg-2 learned about the target subnet from 10.255.102.12, and forwards it on to Agg-1 with itself as the new next hop – 10.3.0.3. This is proven by looking at the route on Agg-1 – it has the expected next hop address, as shown in the figure.

Another option is to tell Agg-2 to redistribute its directly connected network (10.255.102.0/24) into OSPF. All OSPF routers would thus know about this route, eliminating the need for the next-hop-self parameter. This works, but now all OSPF routers now have this extra route in their route table from outside the network. This is not so bad when its only one route, but as the network grows and changes, you must constantly keep up with all this redistribution and cluttering the route tables with external routes. In certain situations, it might also create security weaknesses – adding additional attack vectors for bad actors.

For these reasons, it is almost second nature for many BGP network engineers – for iBGP peers, configure next-hop-self.



Receive BGP Route Advertisements

Receive routes from BGP peer

- Add routes to BGP route table
- If same network from multiple peers, select best route

```
Switch# show bgp
       Network              Nexthop        Metric LocPrf   Weight Path
--------------------------------------------------------------------
*>e  1.0.0.0/8             192.168.1.5      0      100       0    200       ?
* e  1.0.0.0/8             192.168.2.5      0      100       0    100 200   ?
*>e 203.0.113.0/24          192.168.3.5     0      100       0    200
```

- Propose best routes to global IP route table

```
Switch# show ip route
203.0.113.0/24, vrf default via 192.168.3.5,  [20/0], bgp
1.0.0.0/8,      vrf default via 192.168.1.5,  [20/0], bgp
```

- Advertise best routes to BGP neighbors

MOD 1-26

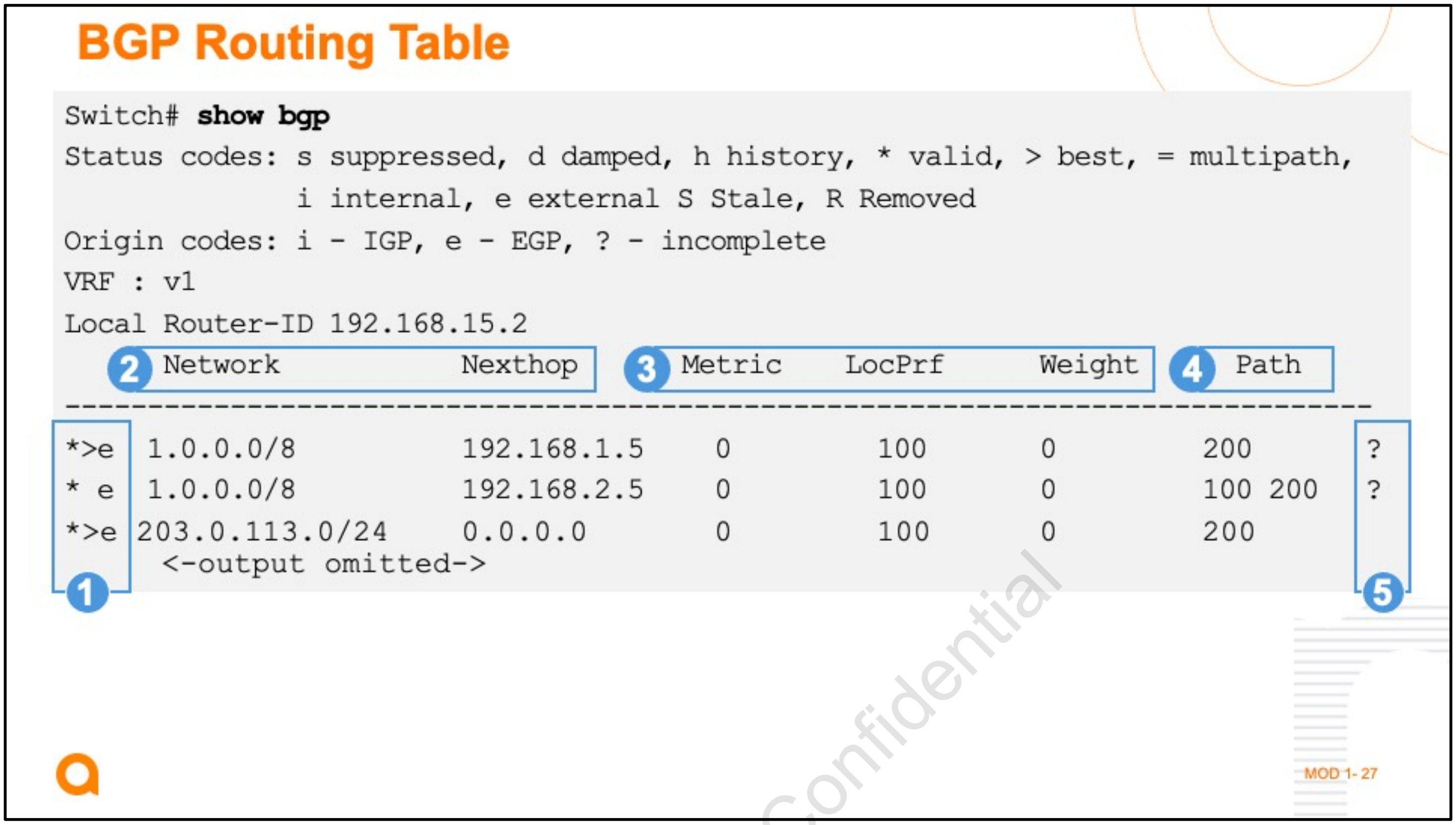
You must distinguish between the BGP route table and the IP route table. The BGP table includes all routes learned via BGP – some of these routes are valid paths, some not. Some of these routes are the best paths, some are not. Only the valid, best paths are added to the IP routing table, which is actually used to forward packets.

So, to see what paths BGP has learned, look at the BGP routing table. To see which of those paths are used to forward packets, look at the IP routing table. Two of the four BGP route table entries were added to the IP route table. The other 1.0.0.0/8 route was not added because it is not the best path | – only best paths are added to the IP route table – those with an * symbol on the far right.

Sometimes the router may only have one path in this table, but has no * symbol – it is an invalid route. The route may be valid because although the router knows about this route, it cannot find the next-hop IP address.

| All valid BGP best path routes are then advertised to BGP neighbors

Aruba Training-Confidential



BGP Routing Table

```
Switch# show bgp
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed
Origin codes: i - IGP, e - EGP, ? - incomplete
VRF : v1
Local Router-ID 192.168.15.2
    Network            Nexthop        Metric     LocPrf      Weight      Path
-----------------------------------------------------------------------------
*>e  1.0.0.0/8         192.168.1.5     0          100         0          200       ?
* e  1.0.0.0/8         192.168.2.5     0          100         0          100 200   ?
*>e 203.0.113.0/24     0.0.0.0         0          100         0          200
     <-output omitted->
```

Aruba Training-Confidential

Now your BGP routers are establishing a connection, advertising routes, receiving routes, and building a BGP routing table. You should be good at interpreting this table. Click on the numbers to explore.

Focus on the number-labeled section of the CLI output in the figure while reading the information below.

1

Each route has several symbols at the beginning. The status codes at the top of the output summarize the meaning of these symbols. An asterisk (*) indicates that the route is valid. This means that the next hop is reachable – it will be added to the route table for packet forwarding. The “e” means the route is from a foreign AS or protocol, while “i” means the route is a native BGP route from the local AS. The greater than sign (>) indicates that this is the best path to reach a destination. In the example, two paths are listed for destination network 1.0.0.0/8. The path via 192.168.1.5 is the one that is added to the route table and used to forward packets. The other path is not used unless the best path is lost. This “best path” selection is based on the other columns shown in the output – Metric, Local Preference (LocPrf), Weight, path, and origin, which you will explore soon.

2

These fields indicate the destination network address and the next hop. You can use these fields to validate the expected networks exist, and that the next hop is as expected. You compare your network documentation to the output of this command. If you expect 100 networks and only 99 are listed, you determine the missing route and trace it back to its source. If route exists but the next hop incorrect, there would not be an asterisk (*) in the first column – you may have forgotten to add the next-hop-self parameter to one of this router's peers.

3

All these columns help BGP to determine the best path, along with the Path and Origin Code columns. You will explore this in detail soon.

The metric indicates the cost for reaching the next hop; unlike with OSPF, this metric is not the most important criteria for selecting between routes. You might modify this to suggest how other AS's forward traffic into your AS.

You might modify Local Preference (LocPrf) to define preferred paths toward external destinations in other Ass.

You can also use weight to control route selection, but this value has local significance only. It is not advertised to other routers.

4

The AS path attribute is included with every BGP route advertisement. A router in the AS that owns this route includes its own AS number in the path attribute, then sends it out to other ASs. A BGP router in a different AS receives the route, adds its AS to the path, and so on. In this way, the AS path indicates every AS which traffic will traverse to reach the destination. A shorter AS path is preferred.

Now you know why there is a ">" symbol next to the top entry for 1.0.0.0/8 – why it is the best path. This router learned about 1.0.0.0/8 from router 192.168.1.5, which is in AS 200. You know this because only AS 200 is listed in the Path attribute. It also learned it from next-hop router 192.168.2.5 in AS 100, which learned it from 200. Since this is a longer AS path, and all other criteria are equal, the shorter path is considered "best". The AS path attribute is also used as BGP's primary loop elimination technique. If a router receives a route with its own AS in the AS path, it drops the route.

Aruba Training Confidential

5

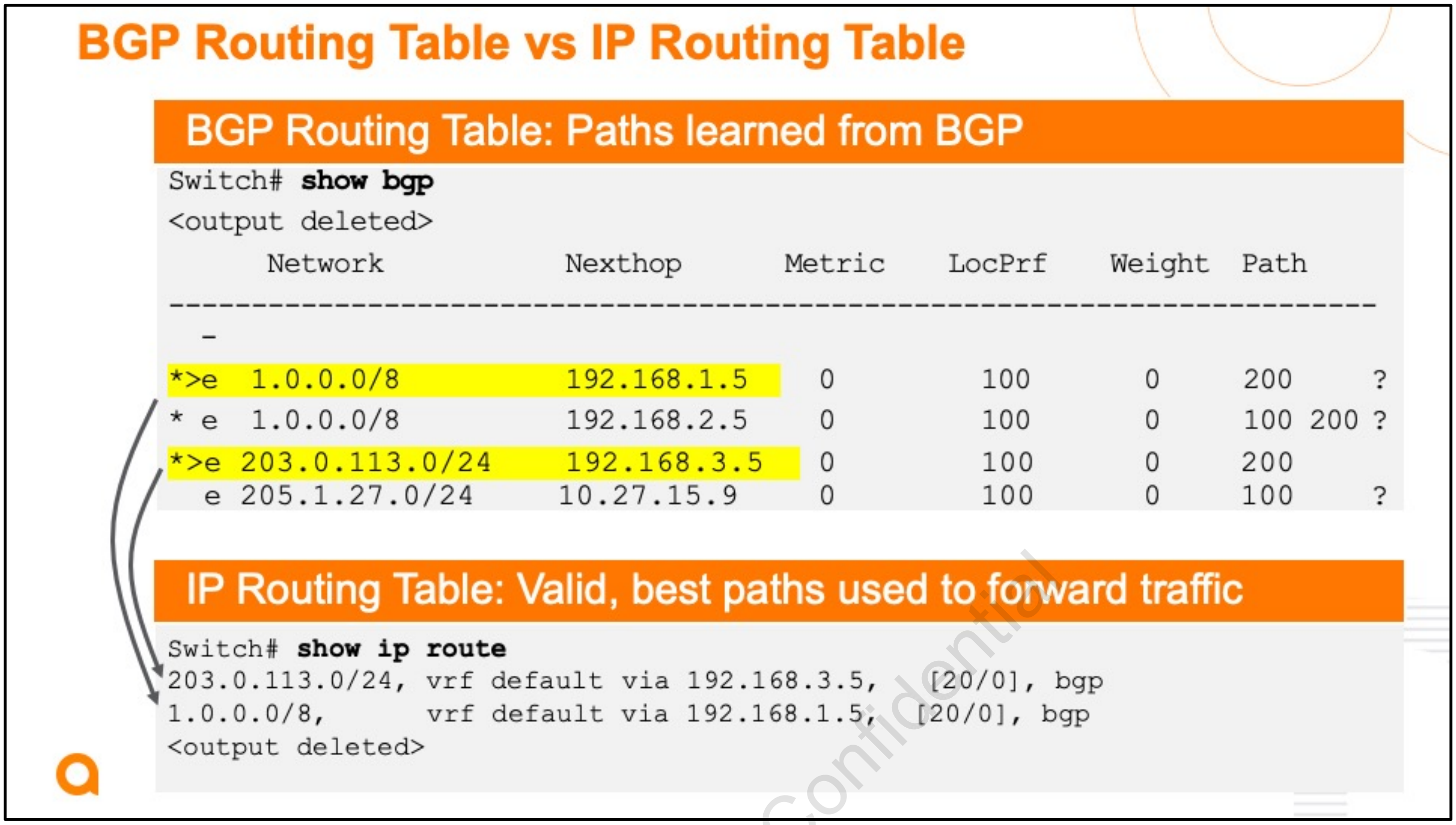
The last column is the “origin code” – where the route came from.

“i” means the route probably came from an IGP (OSPF, RIP, etc) and was injected into BGP with the network command.

, e – from the old Exterior Gateway Protocol (EGP) which is deprecated. If you see this code, you may have time traveled back to the 1980’s!

The “?” symbol means “incomplete” – there is not enough information available to determine how this route originated. This typically means the route was redistributed into BGP from a foreign routing protocol – perhaps OSPF or static.

Aruba Training-Confidential



BGP Routing Table vs IP Routing Table

BGP Routing Table: Paths learned from BGP

```
Switch# show bgp
<output deleted>
       Network             Nexthop      Metric    LocPrf    Weight  Path
---------------------------------------------------------------------------
  -
*>e  1.0.0.0/8           192.168.1.5     0          100        0      200     ?
* e  1.0.0.0/8           192.168.2.5     0          100        0      100  200 ?
*>e 203.0.113.0/24       192.168.3.5     0          100        0      200
  e 205.1.27.0/24       10.27.15.9       0          100        0      100     ?
```

IP Routing Table: Valid, best paths used to forward traffic

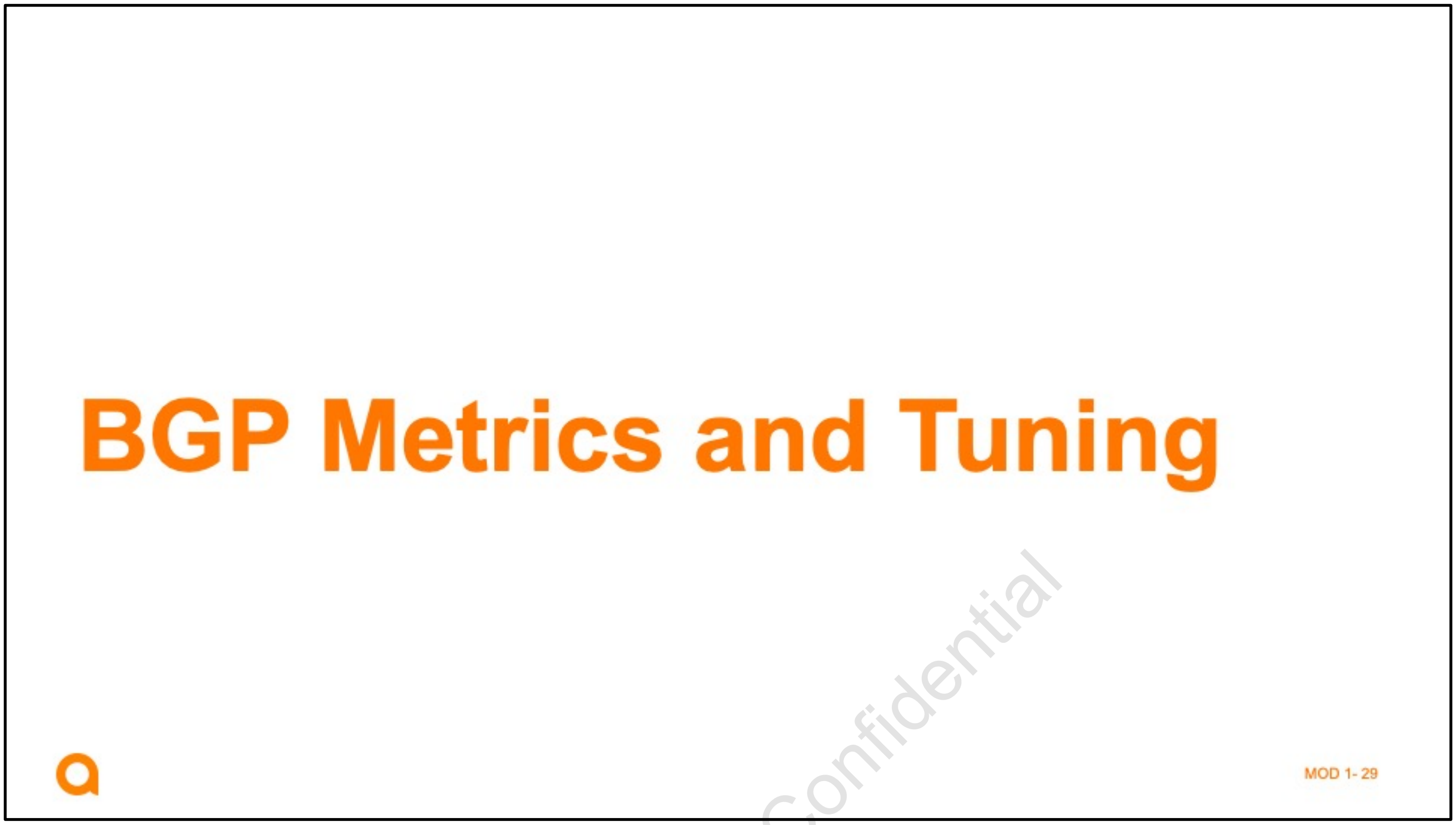
```
Switch# show ip route
203.0.113.0/24, vrf default via 192.168.3.5,  [20/0], bgp
1.0.0.0/8,      vrf default via 192.168.1.5,  [20/0], bgp
<output deleted>
```

Aruba Training-Confidential

You must distinguish between the BGP route table and the IP route table. The BGP table includes all routes learned via BGP – some of these routes are valid paths, some not. Some of these routes are the best paths, some are not. Only the valid, best paths are added to the IP routing table, which is actually used to forward packets.

So, to see what paths BGP has learned, look at the BGP routing table. To see which of those paths are used to forward packets, look at the IP routing table. Two of the four BGP route table entries were added to the IP route table. The other 1.0.0.0/8 route was not added because it is not the best path – only best paths are added to the IP route table. The 205.1.27.0/24 route was not added because it is not a valid route. You can tell this because the route lacks the * symbol.

The route may be valid because although the router knows about this route, it cannot find the next-hop IP address. Let's explore this issue of invalid next hop addresses.

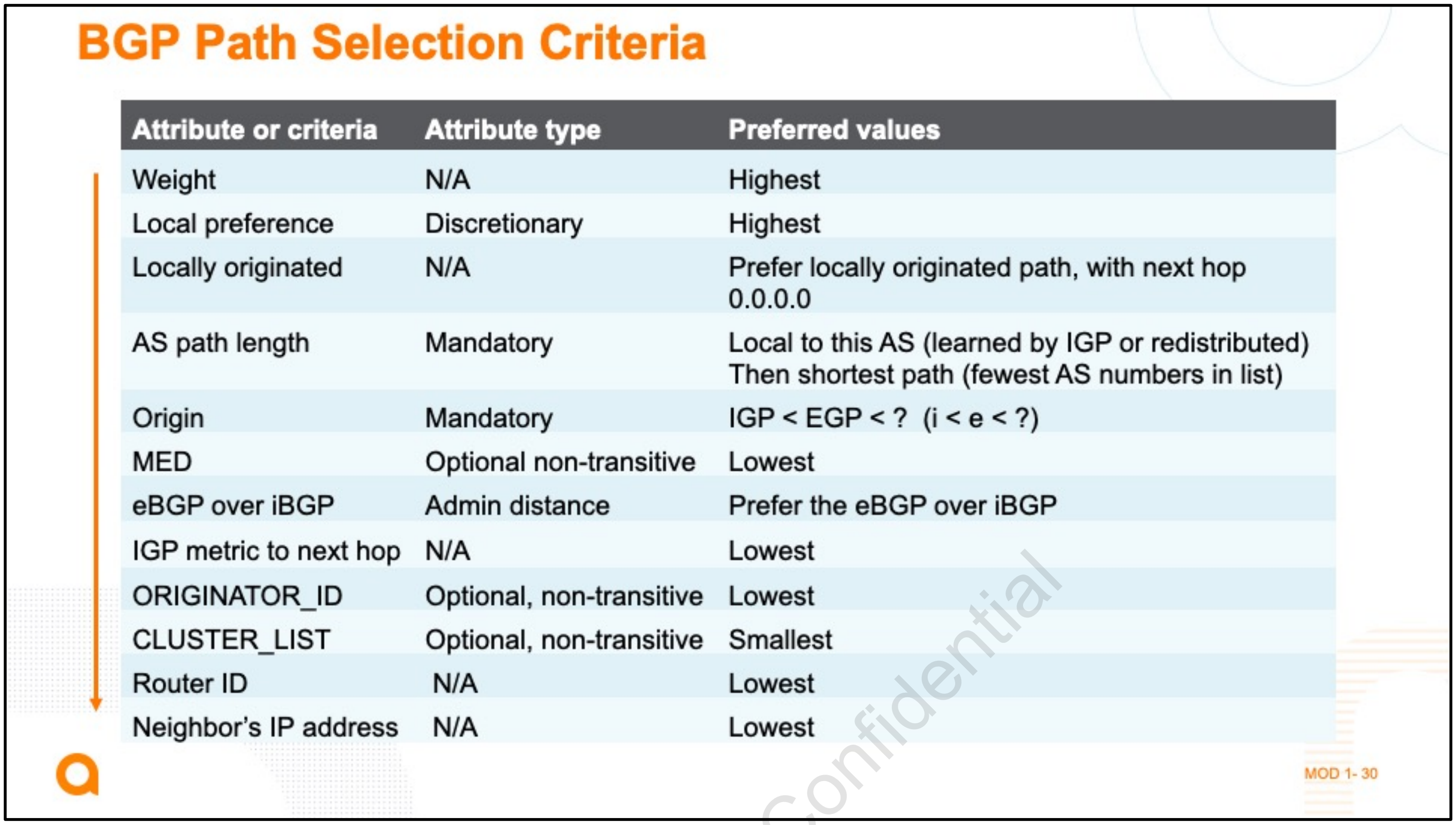


BGP Metrics and Tuning

MOD 1- 29

Aruba Training-Confidential

You know how to form BGP neighbor relationships, and how to describe and configure BGP route advertisements. Now you learn how to control the BGP path selection process.



BGP Path Selection Criteria

Attribute or criteria	Attribute type	Preferred values
Weight	N/A	Highest
Local preference	Discretionary	Highest
Locally originated	N/A	Prefer locally originated path, with next hop 0.0.0.0
AS path length	Mandatory	Local to this AS (learned by IGP or redistributed) Then shortest path (fewest AS numbers in list)
Origin	Mandatory	IGP < EGP < ? (i < e < ?)
MED	Optional non-transitive	Lowest
eBGP over iBGP	Admin distance	Prefer the eBGP over iBGP
IGP metric to next hop	N/A	Lowest
ORIGINATOR_ID	Optional, non-transitive	Lowest
CLUSTER_LIST	Optional, non-transitive	Smallest
Router ID	N/A	Lowest
Neighbor's IP address	N/A	Lowest

MOD 1- 30

Aruba Training-Confidential

When multiple identical BGP routes are valid and have the same administrative distance, BGP uses the criteria shown in the figure to select the best route.

You see the BGP attributes and the preferred values for each one. Attributes are listed in the order by which the router examines them when selecting the best route. That is, the router selects the route with the better weight first. Only if those values are equal does the router move on to examine the next attribute, local preference. You can set a router's weight for specific routes. This value has meaning only on the router itself – it is not advertised. Local preference is not sent to eBGP peers, but it is one of the attributes sent in an update message to iBGP neighbors – those local to this router's AS.

Without special manipulation of weight and local preference on your part, your router will tend to prefer routes based on the AS path. It will accept the route from the ISP that reaches the networks in question through fewer ISPs. Presumably, the shorter AS path indicates a shorter latency; however, this is not necessarily the case.

For routes with the same number of AS IDs in the AS path attribute, the origin might come into play. Routes that were originally injected into BGP are preferred to routes learned from EGP (deprecated) and redistributed routes. However, often the ISPs will have learned the route from a similar ultimate source, so the origin is identical.

Note: Redistributed routes have an origin code of “?” (unknown) in the BGP table, as you will soon see.

The Multi-Exit Discriminator (MED) would take effect only if the ISP routers have explicitly set this attribute. They might do this if they needed to distinguish between multiple paths into your AS. This is only a consideration when you have multiple connections to the same ISP. Generally, a router only compares MED values between routes advertised from the same AS. However, in some cases, you can request different ISPs to send routes with different MED values. You must enter this BGP command to permit the AOS-CX switch to compare the MED values between the two ISPs:

```
switch(config-bgp)# bgp always-compare-med
```

Aruba Training Confidential

However, without such special manipulation, if the AS path length is the same, your router will probably accept the route from the router to which it has a lower cost connection. If the metric to both ISPs is the same, the choice becomes nearly arbitrary. If ECMP is enabled, in fact, the AOS-CX routing switch adds multiple routes with the same next hop cost and load balances over them. If ECMP is disabled, the switch selects the route based on age, then router ID, and then the cluster list—but only if these checks are enabled. The final tie breaker is the advertising router ID.

Similarly, assuming no manipulation of attributes, routers throughout the Internet will tend to route traffic to your company through the ISP to which they are closer.

BGP allows you to control several of these attributes, to fine-tune the route selection. A neighbor router can also manipulate attributes and has ultimate authority over its route selection. Therefore, you can dictate which routes your devices choose, but you might not be able to determine the flow for inbound traffic.

The figure also indicates the type of attribute. Attributes without a type (N/A) are not included in the route advertisement – they only affect the local router’s path decisions. A good grasp of attribute types will help you do more advanced manipulation. Mandatory and discretionary attributes are both well-known attributes that all BGP routers understand; mandatory attributes must be included in every route advertisement, while discretionary attributes are optional. An optional attribute might not be supported by all BGP routers. The “non-transitive” description means that if a BGP router receives this optional attribute and does not support it, it should drop the attribute.

In summary, here are the steps that could be involved in determining which BGP route is preferred:

Do not consider route if the next hop is unreachable

Select the route with the highest weight

Select the route with the highest LOCAL_PREF

Select the route originated by the local router

Select the route with the shortest AS-PATH

Select the IGP, EGP, or Incomplete route in turn , referred to as the “origin” attribute

Select the route with the lowest MED value

Select the route learned from eBGP, confederation, or iBGP in turn

Select the route with the smallest next hop metric

Select the route with the smallest ORIGINATOR_ID

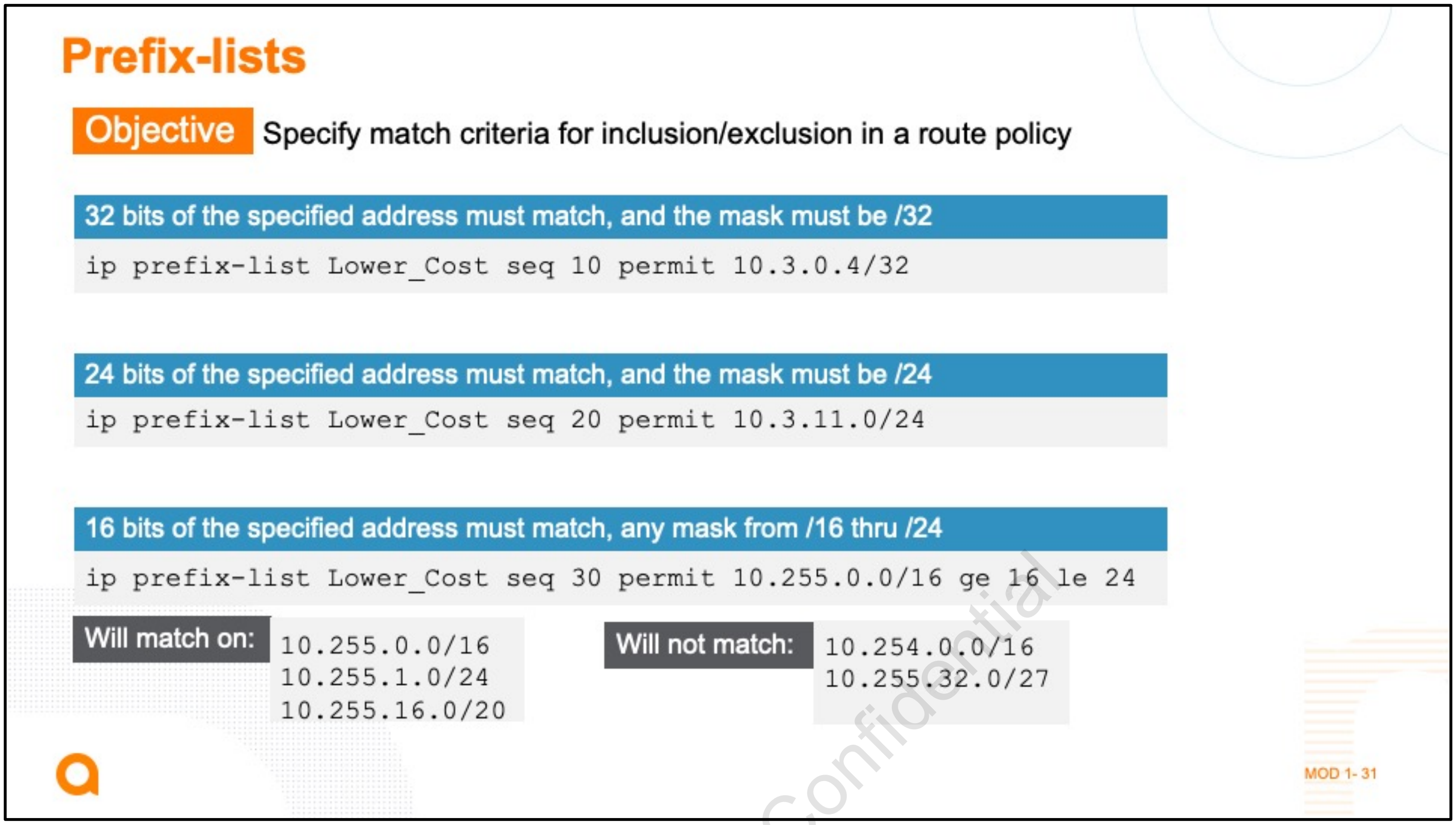
Select the route with the shortest CLUSTER_LIST

Select the route advertised by the router with the smallest Router ID

Select the route advertised by the peer with the lowest IP address

You will often use route maps to manipulate BGP path decisions, sometimes in conjunction with prefix lists.

Aruba Training-Confidential



Prefix-lists

Objective Specify match criteria for inclusion/exclusion in a route policy

32 bits of the specified address must match, and the mask must be /32

```
ip prefix-list Lower_Cost seq 10 permit 10.3.0.4/32
```

24 bits of the specified address must match, and the mask must be /24

```
ip prefix-list Lower_Cost seq 20 permit 10.3.11.0/24
```

16 bits of the specified address must match, any mask from /16 thru /24

```
ip prefix-list Lower_Cost seq 30 permit 10.255.0.0/16 ge 16 le 24
```

Will match on:

```
10.255.0.0/16
10.255.1.0/24
10.255.16.0/20
```

Will not match:

```
10.254.0.0/16
10.255.32.0/27
```

MOD 1- 31

Aruba Training-Confidential

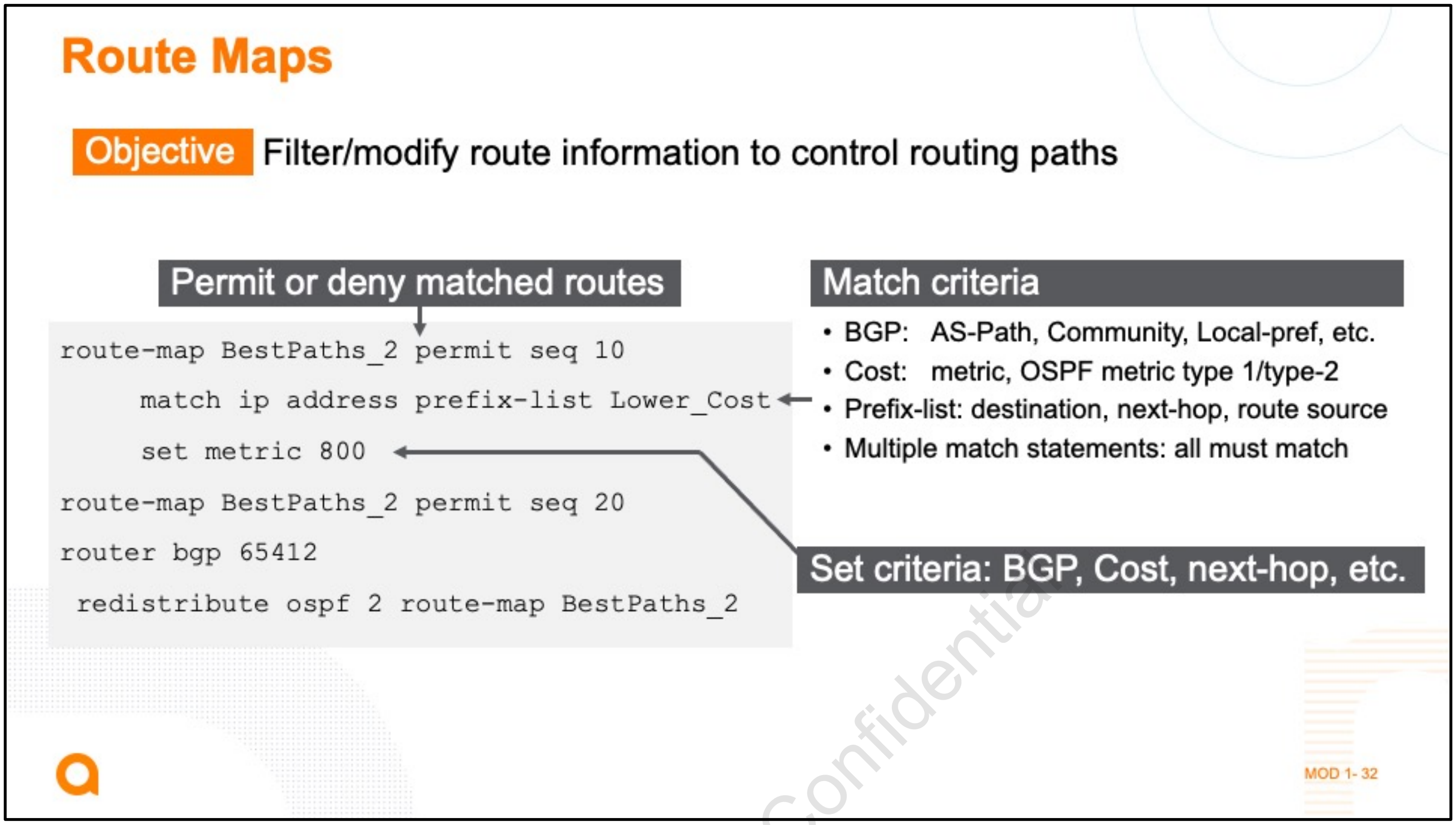
Prefix lists are named lists of route prefixes. They are used to match routes for inclusion in or exclusion from route policies. The sequence number determines the order of matching. The matches are performed starting from the lowest sequence number to the highest sequence number until there is a match. The sequence number is however optional and is autogenerated whenever it is not explicitly mentioned. All prefixes with the same prefix list name are grouped.

The autogenerated sequence number is derived by adding 10 to the highest sequence number available. This technique makes it possible to insert new prefix list sequence number in between.

Use the ge and le parameters to combine prefixes with a range of network mask. For example, 10.255.0.0/16 ge 16 le 24 will match all prefixes within the 10.255.0.0/16 network that have a mask greater than or equal to 16 bits and less than or equal to 24 bits in length. For instance, subnets such as those shown at the left side of the figure would match, because the first 16 bits of each address matches 10.255, and because the mask is between 16 and 24 bits long, inclusive.

The examples on the right would not match, either because the first 16 bits of the address are different, or because the mask is outside of the specified range.

A prefix list is one of several ways to specify match criteria for a route map.



Route Maps

Objective Filter/modify route information to control routing paths

Permit or deny matched routes

```
route-map BestPaths_2 permit seq 10
     match ip address prefix-list Lower_Cost
     set metric 800
route-map BestPaths_2 permit seq 20
router bgp 65412
 redistribute ospf 2 route-map BestPaths_2
```

Match criteria

- BGP: AS-Path, Community, Local-pref, etc.
- Cost: metric, OSPF metric type 1/type-2
- Prefix-list: destination, next-hop, route source
- Multiple match statements: all must match

Set criteria: BGP, Cost, next-hop, etc.

MOD 1- 32

Aruba Training-Confidential

You use route maps to route path selection and availability, by filtering or modifying route information. Routing policies can filter advertised, received, and redistributed routes. They can also modify attributes for specific routes. You use a three-step process to deploy route maps:

Configure filters based on route attributes. You can use prefix-lists to select a destination subnet, next-hop, or route source. The figure lists other criteria you can match on, as you are about to learn.

Create route maps and apply those filters, along with appropriate set criteria as needed

Apply the route map to the redistribution statement

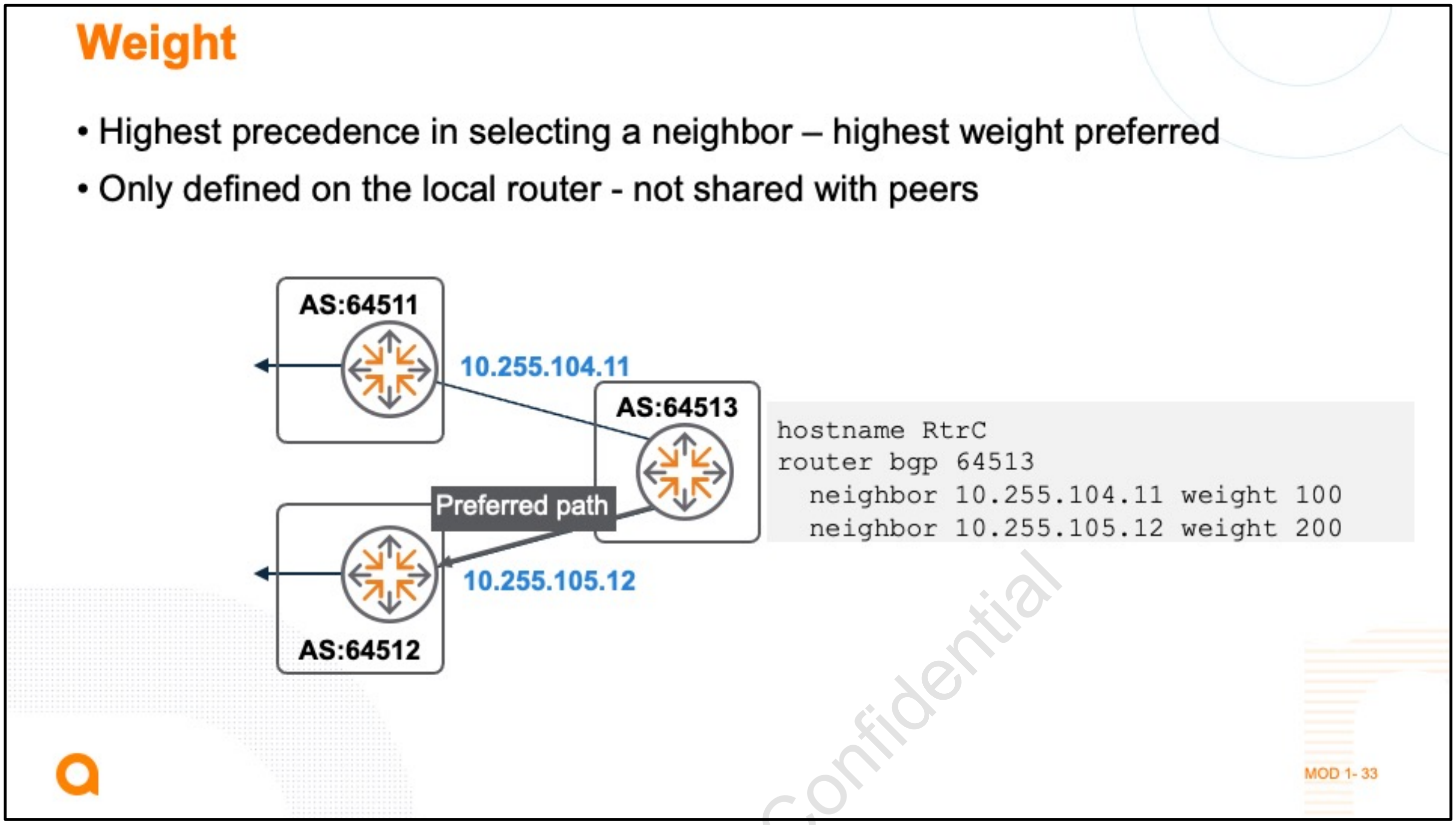
The sequence numbers control top-down processing order, with a permit statement to allow matched routes to be advertised, or a deny statement to prevent their advertisement. Do you see the difference between route maps and ACLs? You often use ACLs to permit or deny user or server data traffic . You use route maps to control the advertisement of routes.

If you specify multiple match criteria within a sequence numbered section, then all criteria must match – it is a logical AND operation. Suppose you added a second match statement under sequence 10 of match ip route-source prefix-list Permitted_Sources. Route advertisements are only permitted and set if they match two criteria – they must be destination subnets that match the prefix-list Lower_Cost, AND they must be sourced (advertised from) ip addresses matched in prefix-list Permitted_Sources.

If both criteria are not met, then there is no match for sequence 10, and sequence 20 is processed. Like an ACL, once there is a match, no more sequences are processed. Also like an ACL, there is an implicit “deny all” at the bottom. If you want unmatched routes to be advertised, simply add a line at the bottom route-map BestPaths_2 permit sequence xx. No need to specify any criteria or set – the statement serves to allow all previously unmatched route advertisements.

Now that you have a better understanding of prefix lists and route maps, you should learn how to used these valuable tools to manipulate BGP path decisions, starting with AS Path Length.

Aruba Training-Confidential



Weight

- Highest precedence in selecting a neighbor – highest weight preferred
- Only defined on the local router - not shared with peers

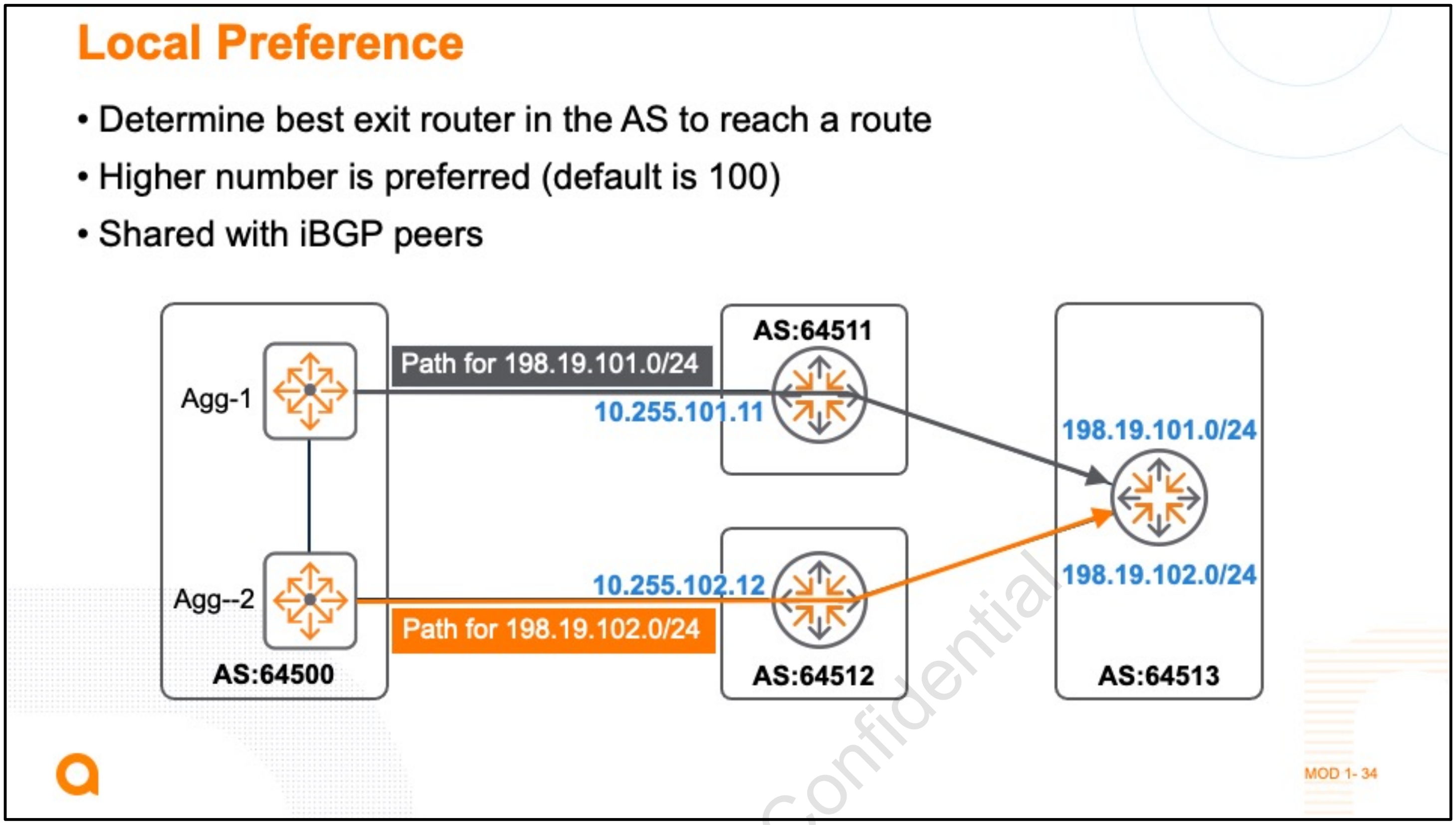
```
hostname RtrC
router bgp 64513
  neighbor 10.255.104.11 weight 100
  neighbor 10.255.105.12 weight 200
```

Aruba Training-Confidential

Weight is a metric that is not really part of the BGP standard, since it's not an attribute shared between BGP peers. Instead, the metric allows you to implement a policy on the local router as to which eBGP peer it prefers. The higher the weight, the more preferred the routes from the respective peer.

The figure shows an example of using weight to determine path selection, by simply adding a weight value to the neighbor statements.

You can also use route maps to adjust weight values for only specific routes from a neighbor, instead of all the routes of a neighbor, like shown in the above example. If using a route map, you would use the set weight command to adjust the weight for matching routes.



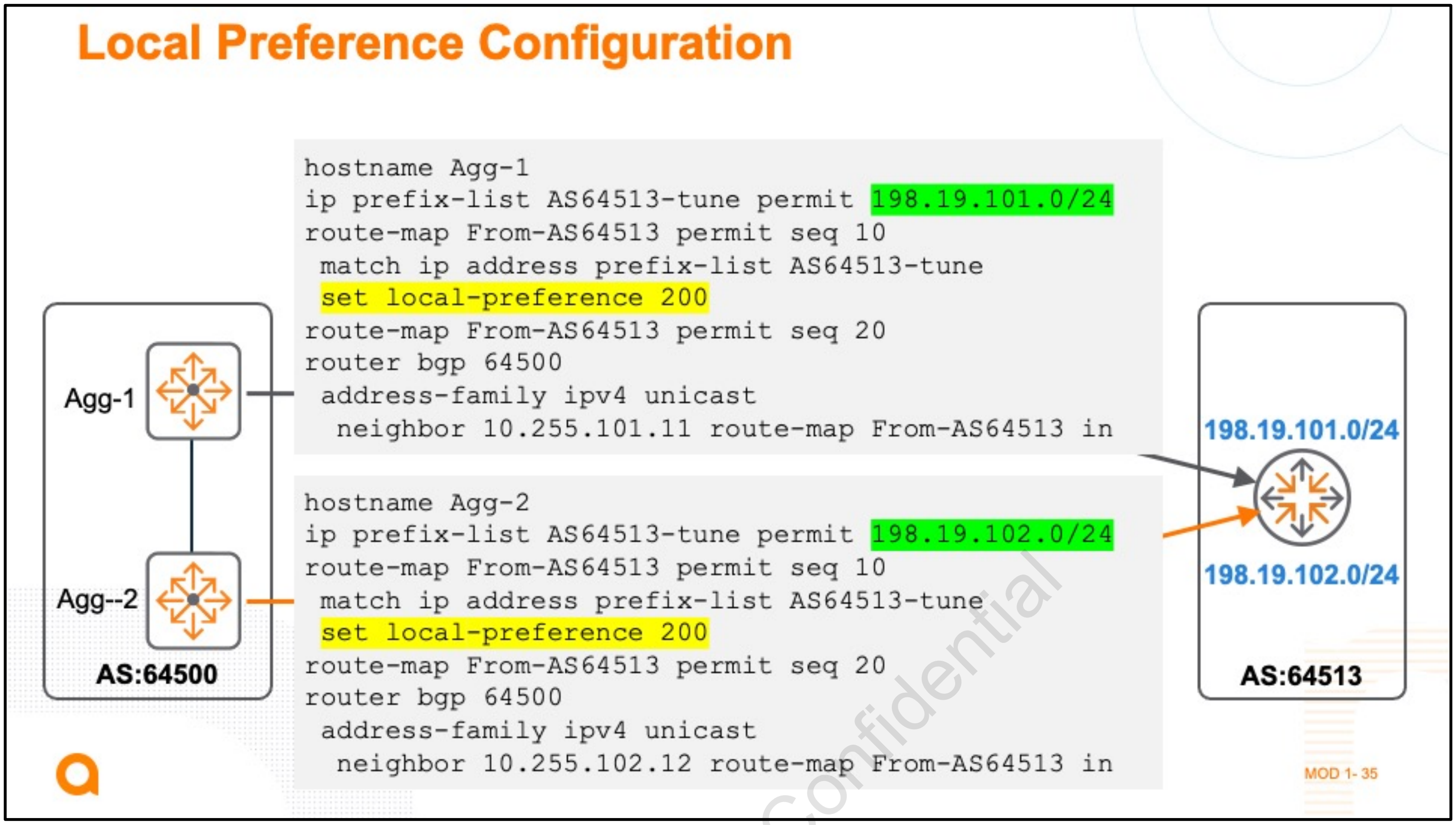
Local Preference

- Determine best exit router in the AS to reach a route
- Higher number is preferred (default is 100)
- Shared with iBGP peers

Local preference is used with iBGP to determine the best exit to reach a particular route. Higher values are preferred - the default is 100, and its only shared with iBGP peers.

This scenario uses local preference, such that AS 64500 prefers some paths to AS 64513 via Agg-1, and others via Agg-2. Let's make it happen.

Aruba Training-Confidential



Local Preference Configuration

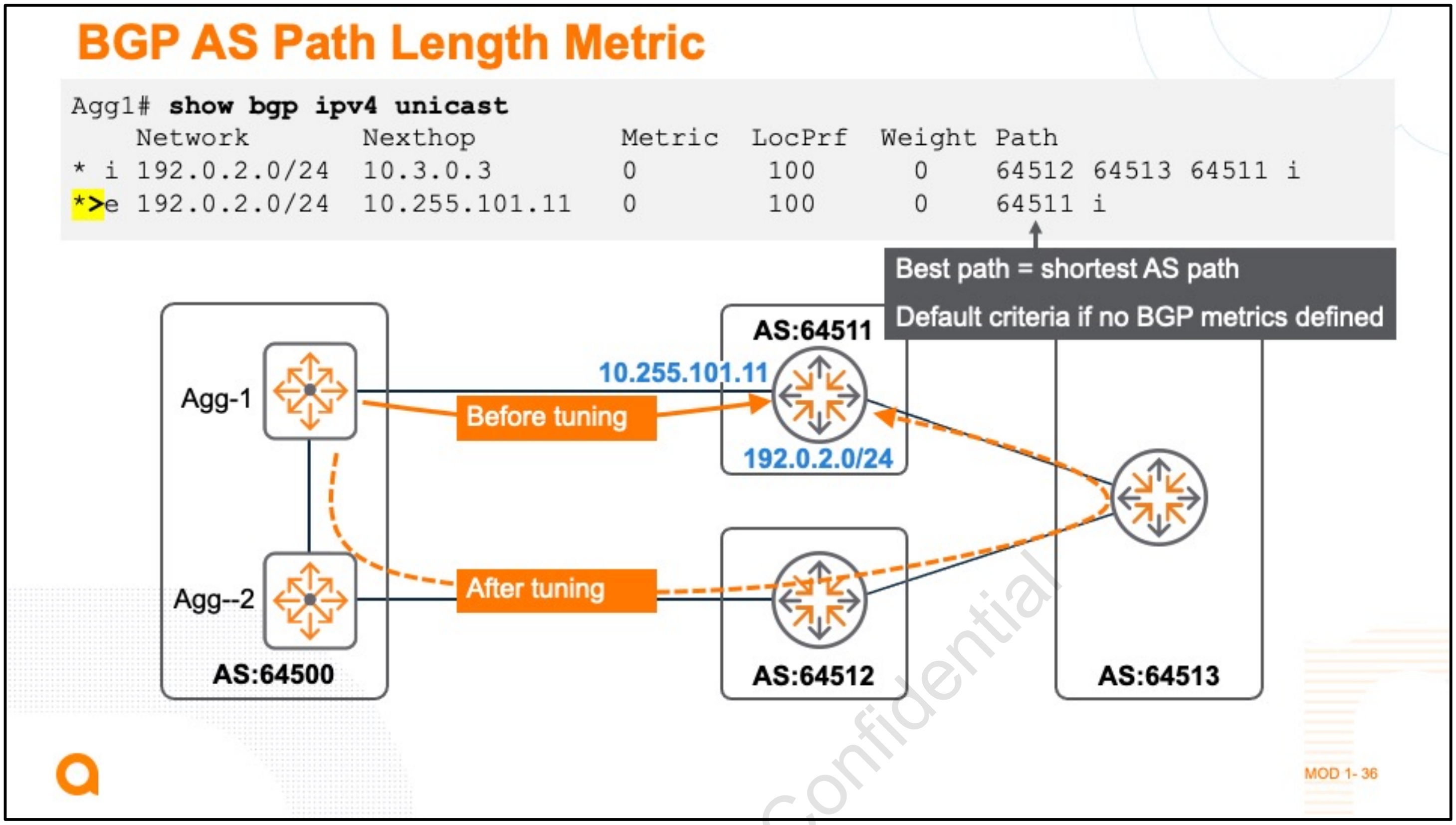
```
hostname Agg-1
ip prefix-list AS64513-tune permit 198.19.101.0/24
route-map From-AS64513 permit seq 10
 match ip address prefix-list AS64513-tune
 set local-preference 200
route-map From-AS64513 permit seq 20
router bgp 64500
 address-family ipv4 unicast
  neighbor 10.255.101.11 route-map From-AS64513 in
```

```
hostname Agg-2
ip prefix-list AS64513-tune permit 198.19.102.0/24
route-map From-AS64513 permit seq 10
 match ip address prefix-list AS64513-tune
 set local-preference 200
route-map From-AS64513 permit seq 20
router bgp 64500
 address-family ipv4 unicast
  neighbor 10.255.102.12 route-map From-AS64513 in
```

Aruba Training-Confidential

The figure shows the configuration used to achieve the Local Preference scenario. The route maps for Agg-1 and Agg-2 are the same, except the prefix-list matches on a different route.

Pause the presentation to study this syntax.



BGP AS Path Length Metric

```
Agg1# show bgp ipv4 unicast
    Network         Nexthop          Metric   LocPrf   Weight  Path
* i 192.0.2.0/24    10.3.0.3         0         100       0     64512 64513 64511 i
*>e 192.0.2.0/24    10.255.101.11    0         100       0     64511 i
```

Best path = shortest AS path

Default criteria if no BGP metrics defined

Aruba Training-Confidential

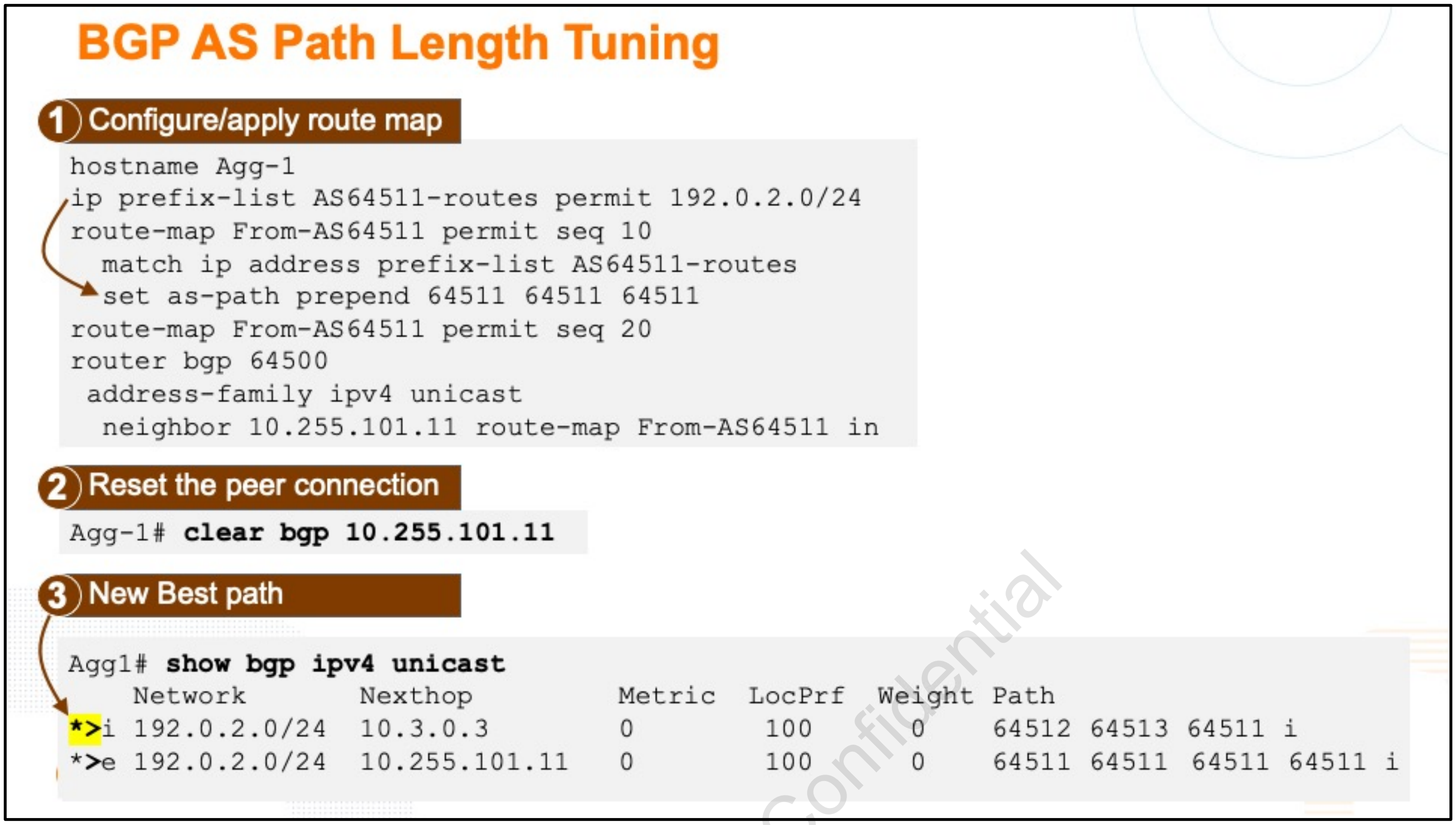
You know that BGP has many metrics to manipulate routing. If weight or local preference is not defined, then AS path length determines the best path, and also helps BGP detect route loops.

In this example, all other things being equal, the best path to 192.0.2.0/24 is out Agg-1, directly to AS 64511 – 1 hop away. The path out Agg-2 via AS 64512 is 3 hops or ASes away.

You can see this in the show bgp ipv4 unicast output, which shows all discovered BGP paths, and what BGP believes to be the best paths. The greater than (>) sign indicates that, of the two possible paths, this one is the best path. Why? The weight and local preference are the same, and so “best” is determined by AS path length.

Note: By default, BGP does not do load sharing: the first “best” path used is what is placed in the routing table unless you enable equal-cost multipath (ECMP) for BGP.

A shorter AS path is not necessarily the best path. Perhaps the link between Agg-1 and AS 64511 is very slow, and you want to tune the AS Path metric to use the other path. Let’s do it.



BGP AS Path Length Tuning

1 Configure/apply route map

```
hostname Agg-1
ip prefix-list AS64511-routes permit 192.0.2.0/24
route-map From-AS64511 permit seq 10
  match ip address prefix-list AS64511-routes
  set as-path prepend 64511 64511 64511
route-map From-AS64511 permit seq 20
router bgp 64500
 address-family ipv4 unicast
  neighbor 10.255.101.11 route-map From-AS64511 in
```

2 Reset the peer connection

```
Agg-1# clear bgp 10.255.101.11
```

3 New Best path

```
Agg1# show bgp ipv4 unicast
    Network         Nexthop           Metric   LocPrf  Weight Path
*>i 192.0.2.0/24    10.3.0.3          0         100      0    64512 64513 64511 i
*>e 192.0.2.0/24    10.255.101.11     0         100      0    64511 64511  64511  64511 i
```

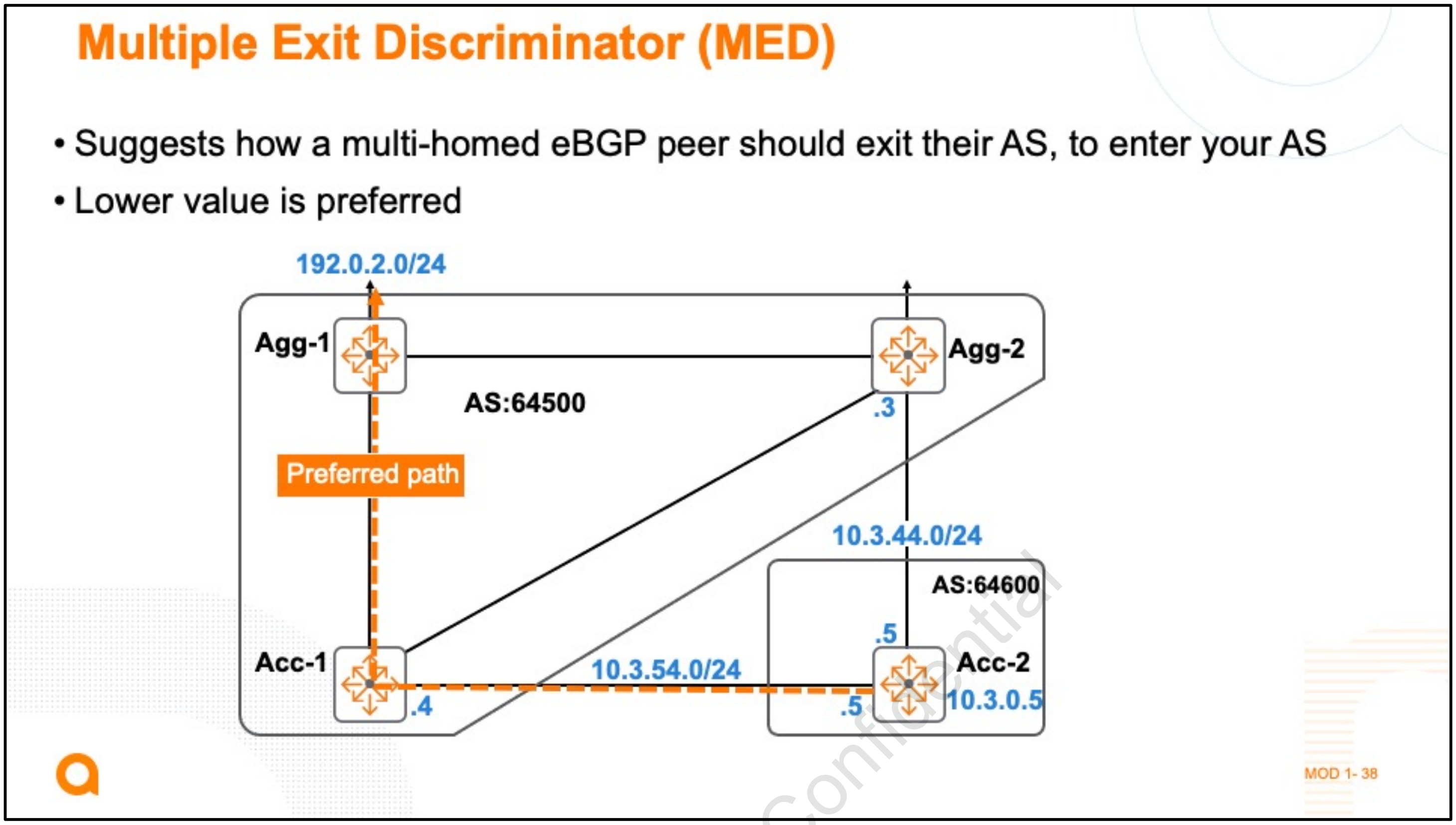
You want traffic to 192.0.2.0/24 to go out Agg-2, via AS 64512 > 64513 > 64511. One solution is to implement AS pre-pending, as shown in the figure.

A prefix list is applied to a route map as match criteria, and the AS 64511 is prepended to this route. This route map is applied to the 10.255.101.11 peer inbound. If you were to check the BGP table at this point, you would probably not see a change. New BGP policies are typically applied when the peer is established. So, you must clear the peer connection, which forces it to reform. Now that you know about the need to clear connections, this will not be mentioned in the following tuning discussions. It is now assumed that you will know to do this.

Now when you check the BGP table, you can see that the AS path prepend worked, and the desired best path is now selected.

Do not forget to add the “permit any” statement at the bottom of your route map, as shown in the figure - route-map From-AS64511 permit seq 20. Otherwise, the router will only advertise the routes that are tuned from sequence number 10. The implicit “deny all” at the bottom will filter all other routes from being advertised.

Aruba Training-Confidential



Multiple Exit Discriminator (MED)

- Suggests how a multi-homed eBGP peer should exit their AS, to enter your AS
- Lower value is preferred

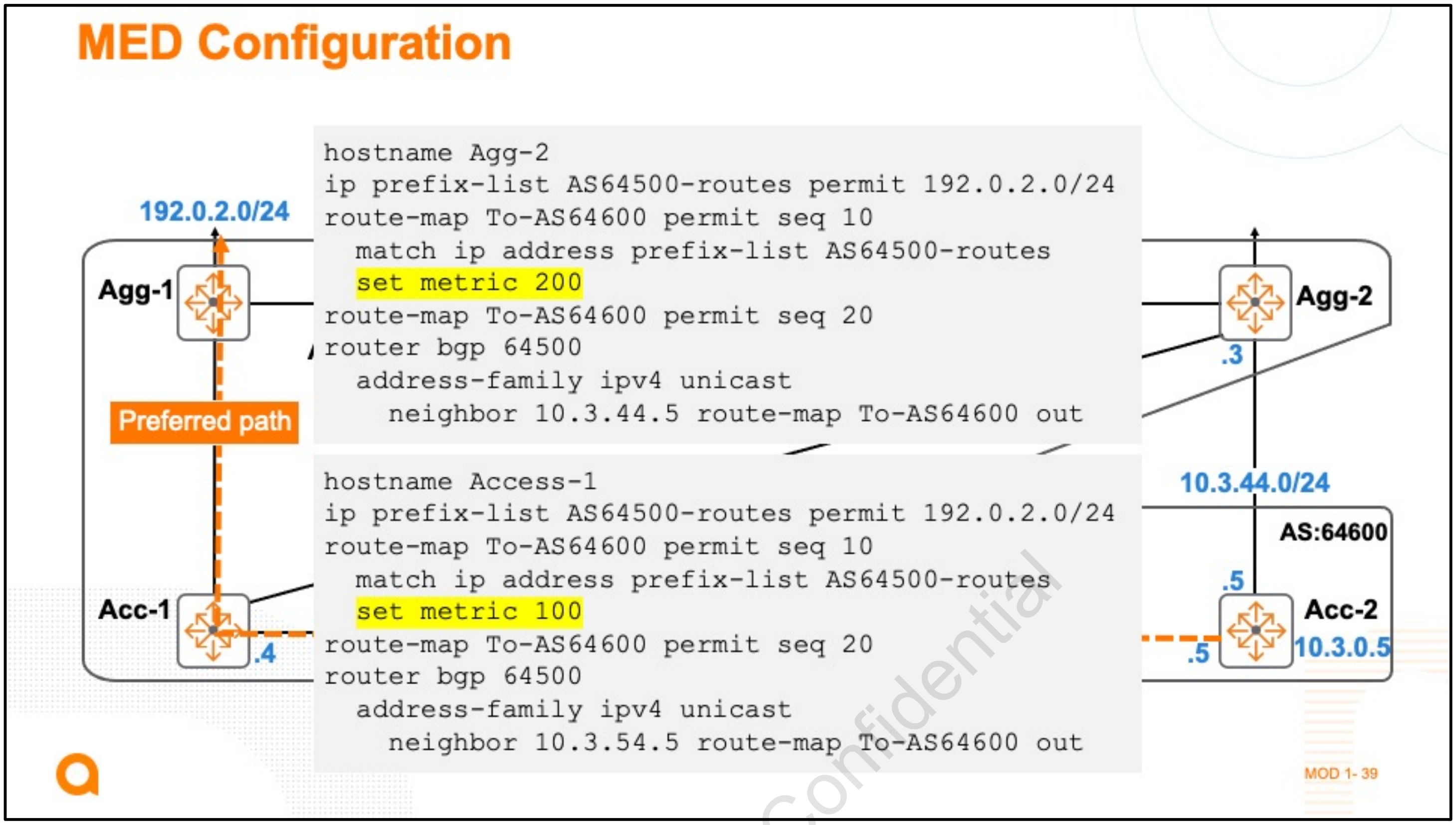
MOD 1- 38

Aruba Training-Confidential

The Multiple Exit Discriminator (MED) attribute is shared with a multi-homed eBGP peer to influence how they will enter your network. The figure shows how AS64500 and 64600 are multi-homed – one connection via the 10.3.54.0 link, and another via 10.3.44.0. This is the use case for MED. You might configure your routers in AS 64500 to send different MED values to the AS 64600 peer. This is a suggestion to these eBGP peers, "When packets exit your network, en route to my network, please use Path A for some routes and path B for other routes".

AS 64500 might be a large cloud with thousands of subnets. You know that it is more efficient reach some of your networks via the 10.3.44.0 link, while others should use the 10.3.54.0 link. So you make this suggestion to the external AS to forward traffic accordingly. Remember it is called an Autonomous System because someone has autonomy (control) over it. You cannot force the AS to pay attention to your MED values. You would need to work with them as a partner, tell them of your intentions, and ask that they configure their routers to heed your request.

Perhaps it makes more sense for traffic destined to 192.0.2.0 to enter AS 64500 via Access-1 – maybe its a faster path, or a more direct path….



MED Configuration

```
hostname Agg-2
ip prefix-list AS64500-routes permit 192.0.2.0/24
route-map To-AS64600 permit seq 10
  match ip address prefix-list AS64500-routes
  set metric 200
route-map To-AS64600 permit seq 20
router bgp 64500
  address-family ipv4 unicast
    neighbor 10.3.44.5 route-map To-AS64600 out
```

```
hostname Access-1
ip prefix-list AS64500-routes permit 192.0.2.0/24
route-map To-AS64600 permit seq 10
  match ip address prefix-list AS64500-routes
  set metric 100
route-map To-AS64600 permit seq 20
router bgp 64500
  address-family ipv4 unicast
    neighbor 10.3.54.5 route-map To-AS64600 out
```

… so you create the configuration shown here, in which Access-1 advertises a lower MED for that route than Agg-2. Pause the presentation to review this configuration.

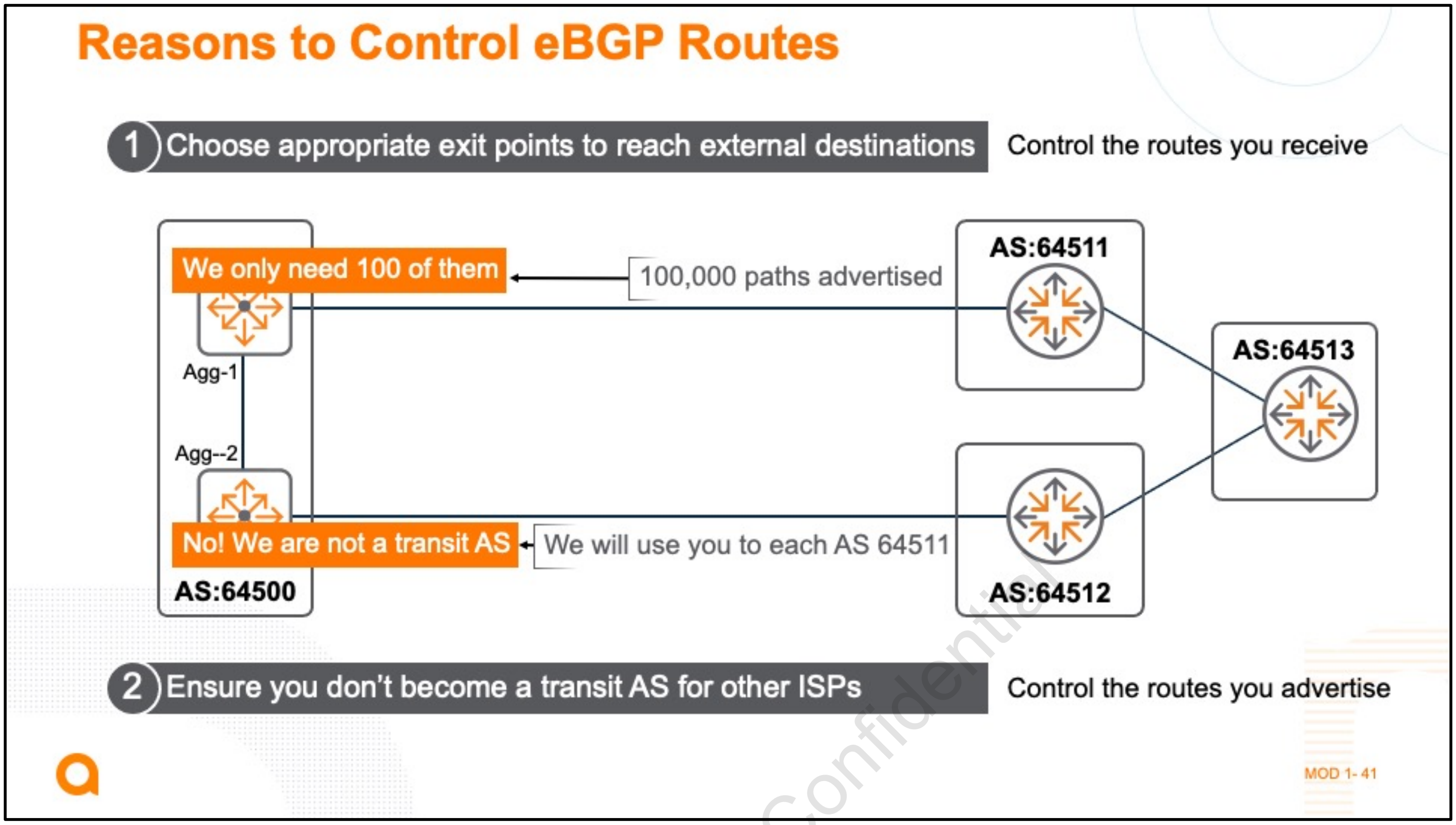
Aruba Training-Confidential

Route Control

MOD 1- 40

This section introduces the basics of controlling routes with BGP.

Aruba Training-Confidential



Reasons to Control eBGP Routes

1 Choose appropriate exit points to reach external destinations — Control the routes you receive

2 Ensure you don't become a transit AS for other ISPs — Control the routes you advertise

MOD 1- 41

There are many reasons to control external BGP (eBGP) routes. The two most common are:

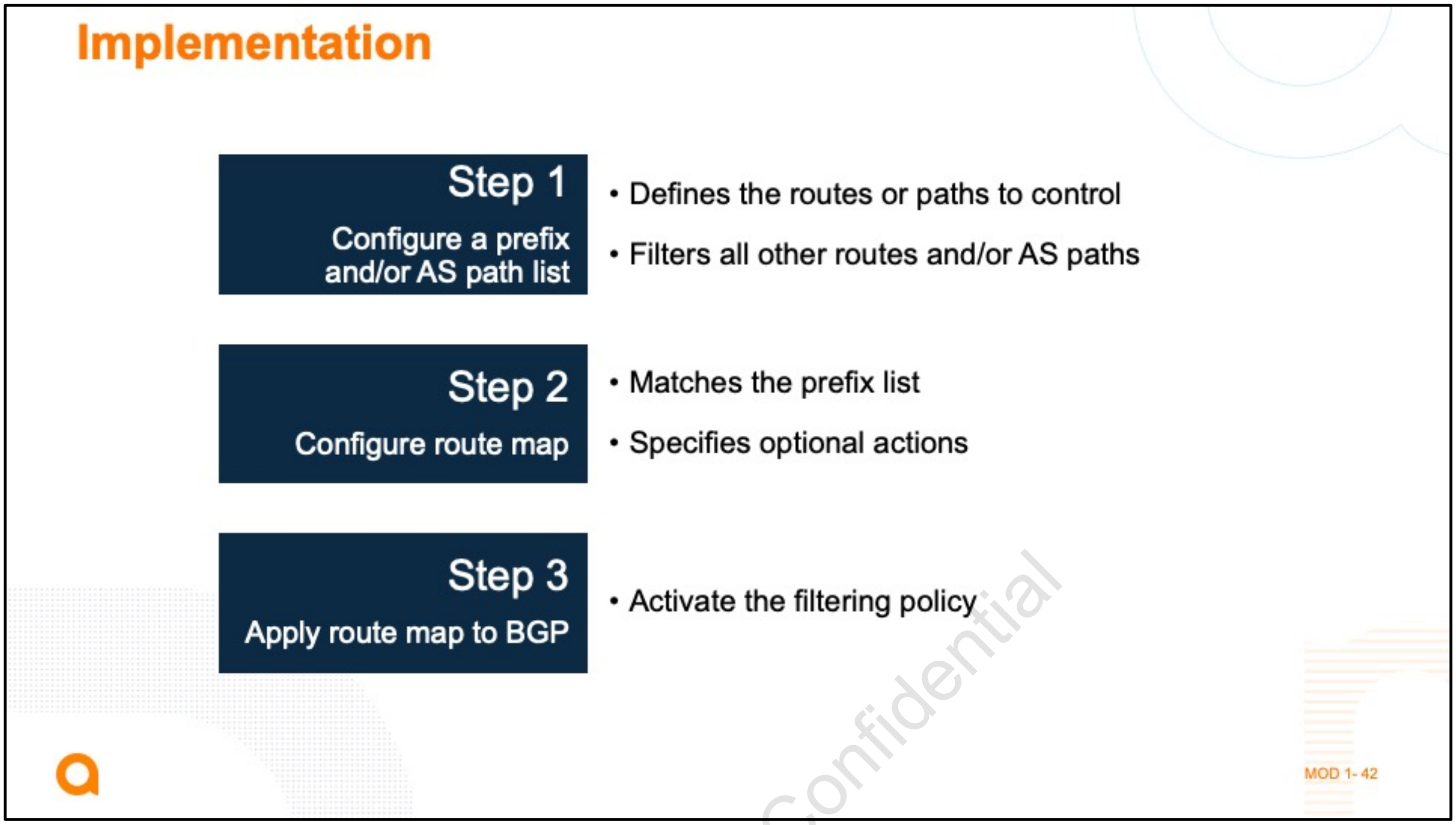
Choose appropriate exit points to reach external destinations

Ensure you don't become a transit AS for other ISPs

For example, if you have resources in ISP1 that you commonly access, it makes sense to use the connection you have to ISP1 to reach them, versus go through an other ISP to reach those resources. Second, you don't want ISPs to use your company's bandwidth in order to connect to other ISPs, which is referred to as a transit AS. Based on these two reasons, you'll need to implement eBGP route control through prefix or AS path lists to control the external BGP routes you receive and/or advertise.

Note: Use Looking Glass to verify the results of your configuration - https://www.gin.ntt.net/looking-glass-landing/

Aruba Training-Confidential



Implementation

Step 1
Configure a prefix and/or AS path list

- Defines the routes or paths to control
- Filters all other routes and/or AS paths

Step 2
Configure route map

- Matches the prefix list
- Specifies optional actions

Step 3
Apply route map to BGP

- Activate the filtering policy

MOD 1- 42

Aruba Training-Confidential

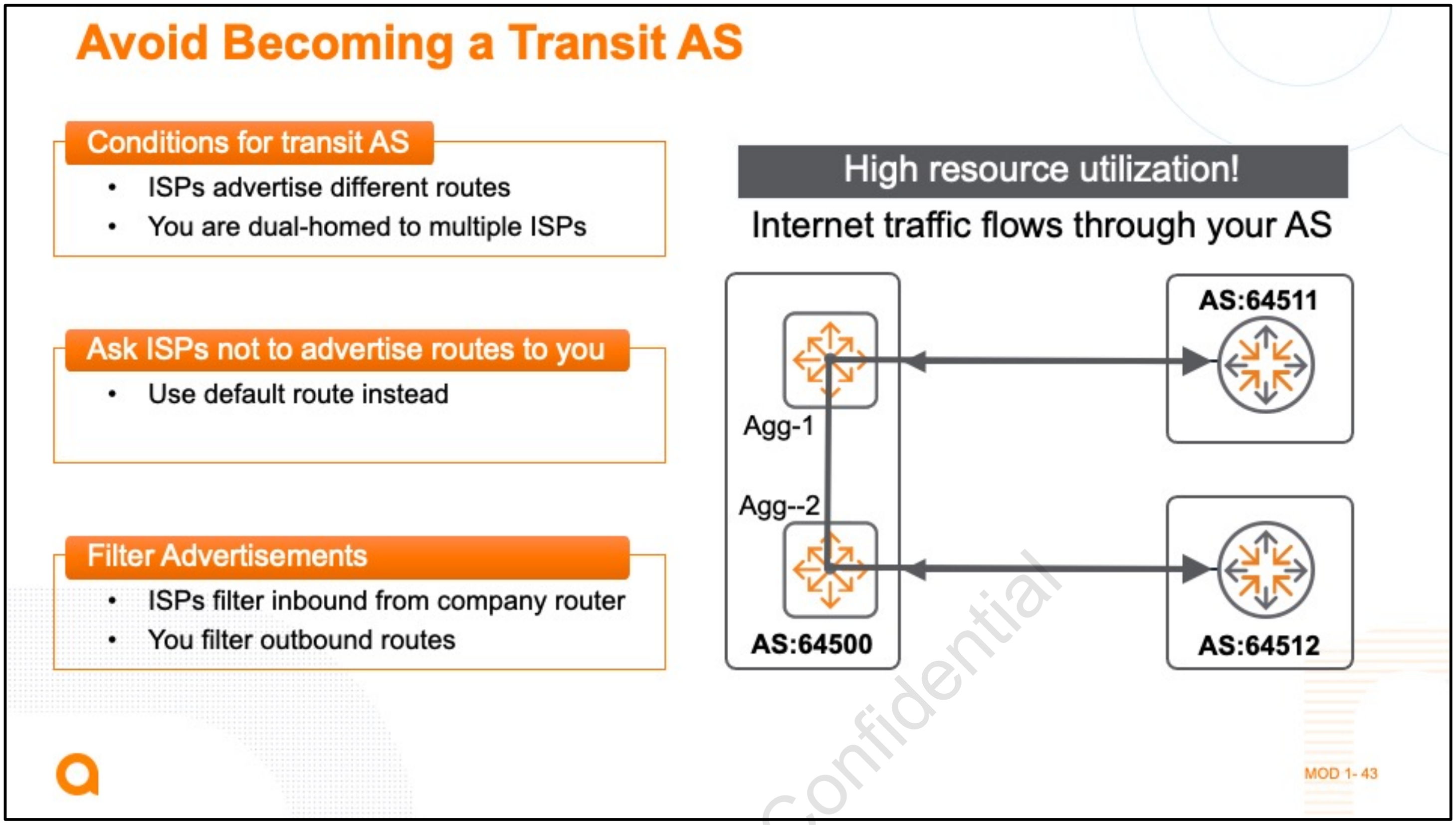
You can manipulate route preference using route maps, which you have already used extensively in this training. So far, you have used route maps to tune routes that already exist in the BGP table. Earlier you used them to control OSPF route redistribution. Now you will use them to control prefixes advertised out to other ASes or accepted inbound from other ASes.

Recall that deploying a route policy is three-step process:

Configure filters based on route attributes, such as destination address and the advertising router address.

Create a route map and apply filters to it.

Apply the route map to BGP



Avoid Becoming a Transit AS

Conditions for transit AS

- ISPs advertise different routes
- You are dual-homed to multiple ISPs

Ask ISPs not to advertise routes to you

- Use default route instead

Filter Advertisements

- ISPs filter inbound from company router
- You filter outbound routes

High resource utilization!

Internet traffic flows through your AS

MOD 1- 43

Aruba Training-Confidential

In the multi-homed ISP scenario, you should only advertise your own public IP addresses to the ISPs. Do not announce routes that your AS learns from ISP 1 to ISP 2, nor vice versa. This can cause ISP routers to use your AS for internet traffic. Your router will quickly become overburdened by external Internet traffic.

Several scenarios could cause your AS to become a transit AS:

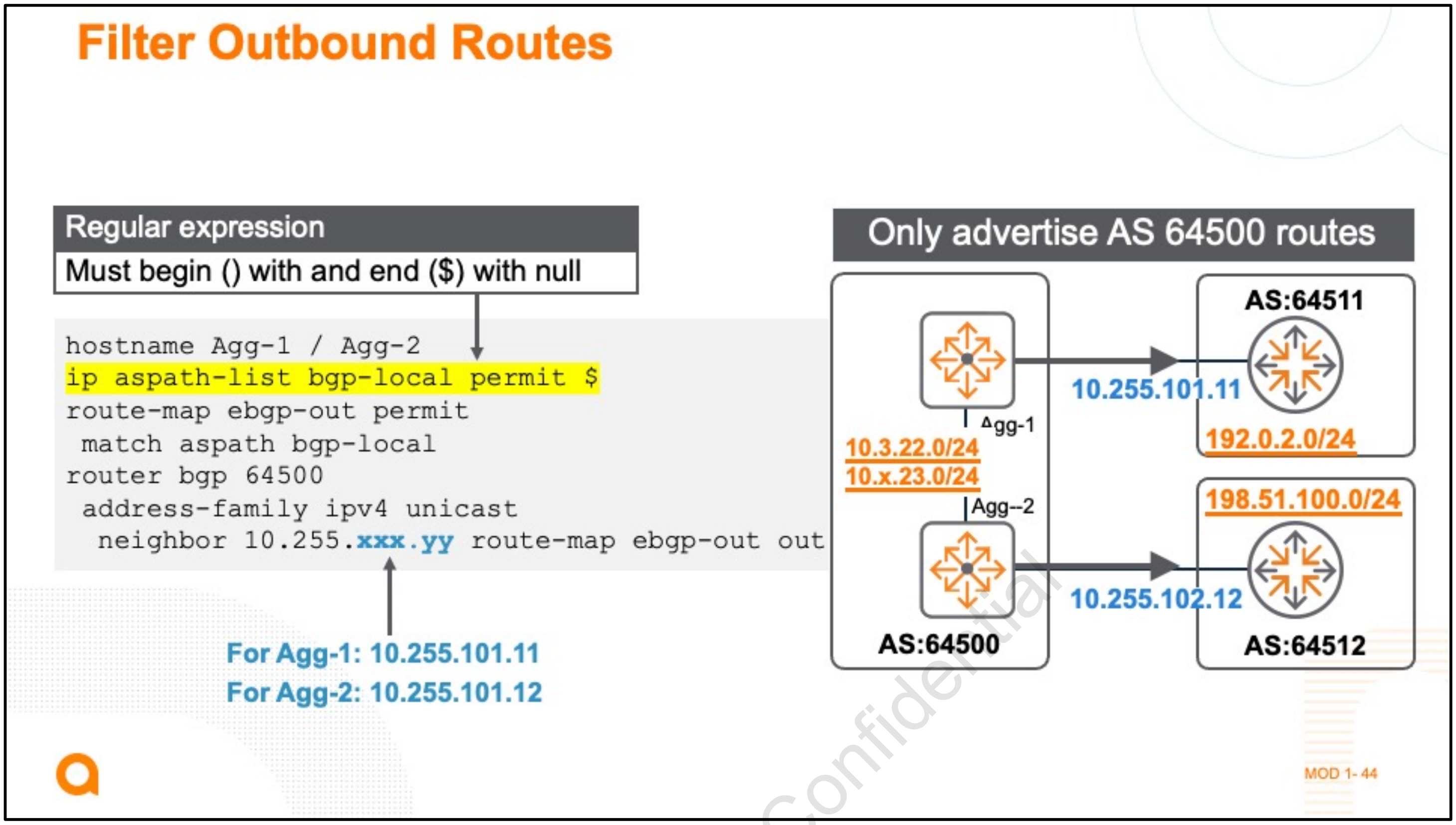
The ISPs advertise Internet routes differently with different aggregation. Any route that your BGP routers receive from only one ISP, they could begin advertising to the other ISP.

You connect a single AOS-CX switch or VSF fabric to both ISPs. In this case, the switch has two eBGP neighbors, and it will advertise best routes received from one to the other.

You should communicate with your ISPs to prevent issues in any of these scenarios.

Sometimes you are using BGP primarily to announce your public IP block. For traffic outbound to the Internet, you do not care about finding the best path. You might want to use the connection to one ISP as the primary connection and the other as a standby only if the first fails, or you might want to load balance across both. In cases such as these, you might ask your ISPs not to send any route advertisements to you. You can then create static default routes on the BGP routers. Or in some cases, you might want the ISP to send only a default route in BGP.

If you want to receive routing advertisements from the ISPs, you must ensure that your router does not pass them on to the other ISP. You can ask the ISP router to filter inbound advertisements from you. But it is best practice to filter outbound advertisements yourself.



Filter Outbound Routes

Regular expression

Must begin () with and end ($) with null

```
hostname Agg-1 / Agg-2
ip aspath-list bgp-local permit $
route-map ebgp-out permit
 match aspath bgp-local
router bgp 64500
 address-family ipv4 unicast
  neighbor 10.255.xxx.yy route-map ebgp-out out
```

For Agg-1: 10.255.101.11

For Agg-2: 10.255.101.12

Only advertise AS 64500 routes

MOD 1- 44

Aruba Training-Confidential

In this example, you are setting up route maps to restrict outbound advertisements to eBGP neighbors. There are a couple of ways that you could solve this issue:

Use an AS Path list where the path list is empty, indicating that the route is internal to your network

Use a prefix list to only allow advertisement of your routes to other ISPs

The ip aspath-list command provides an elegant solution for this use case, because it specifies a specific AS path. AS Path lists are named lists of regular expression rules. They are used to match AS Path attributes in the routes for inclusion in or exclusion from route policies. The sequence number is optional and is autogenerated whenever it is not explicitly mentioned. All AS path list rules with the same name are grouped together. The configuration of IP AS path lists is beyond the scope of this course.

Note: See https://www.regex101.com for regular expression examples and to test your regular expressions.

These regex values are used:

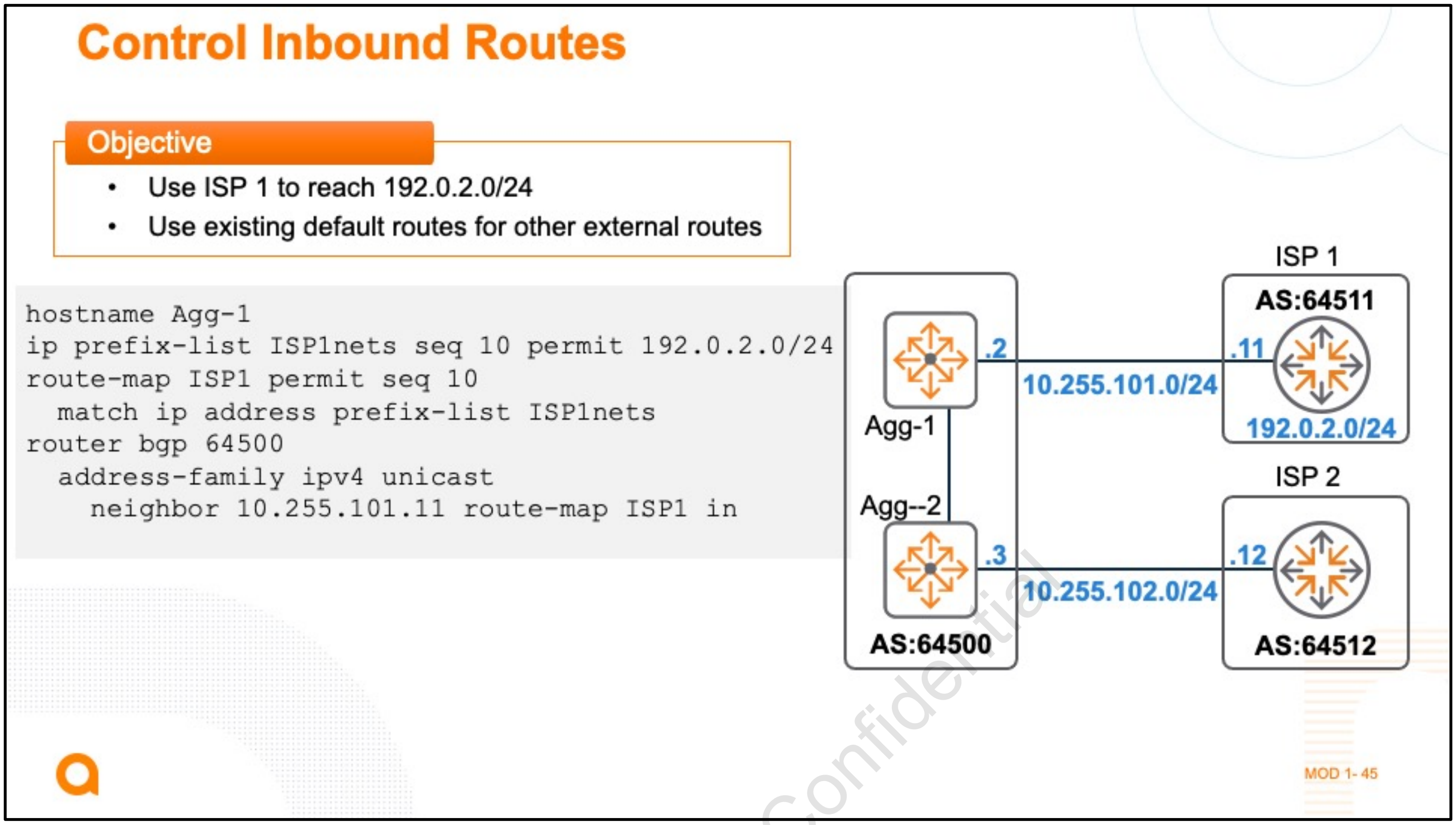
" indicates 'must begin with'

'$' indicates 'must end with’

Combined, they result in '$', which represents an empty string.

All the routes that are originated inside this AS, will have an empty AS-Path attribute, while any route that was received via an external AS (eBGP), will have at least the value of that external AS in the AS-Path. To check the AS-Path of a route, a regular expression can be used. Also notice that the route map is applied to both Agg-1 and Agg-2, when communicating outbound to their respective eBGP peers. You would need to deploy a similar configuration to any other eBGP routers in your network.

Aruba Training-Confidential



Control Inbound Routes

Objective

- Use ISP 1 to reach 192.0.2.0/24
- Use existing default routes for other external routes

```
hostname Agg-1
ip prefix-list ISP1nets seq 10 permit 192.0.2.0/24
route-map ISP1 permit seq 10
  match ip address prefix-list ISP1nets
router bgp 64500
  address-family ipv4 unicast
    neighbor 10.255.101.11 route-map ISP1 in
```

MOD 1- 45

Aruba Training-Confidential

Here is an example route map that filters routes the BGP routers receives from its eBGP neighbor. In this example, you run BGP primarily to announce your public block of IP addresses, and care less about receiving all the BGP routes in the Internet routing tables. You can filter those down in a few different ways. You may choose to just accept routes for larger ranges less than a specific prefix length. Or you might choose to accept just routes that originate in the eBGP neighbor AS because those are the ones where the company will benefit the most by sending the traffic out this specific router. You could also take a combination of these approaches.

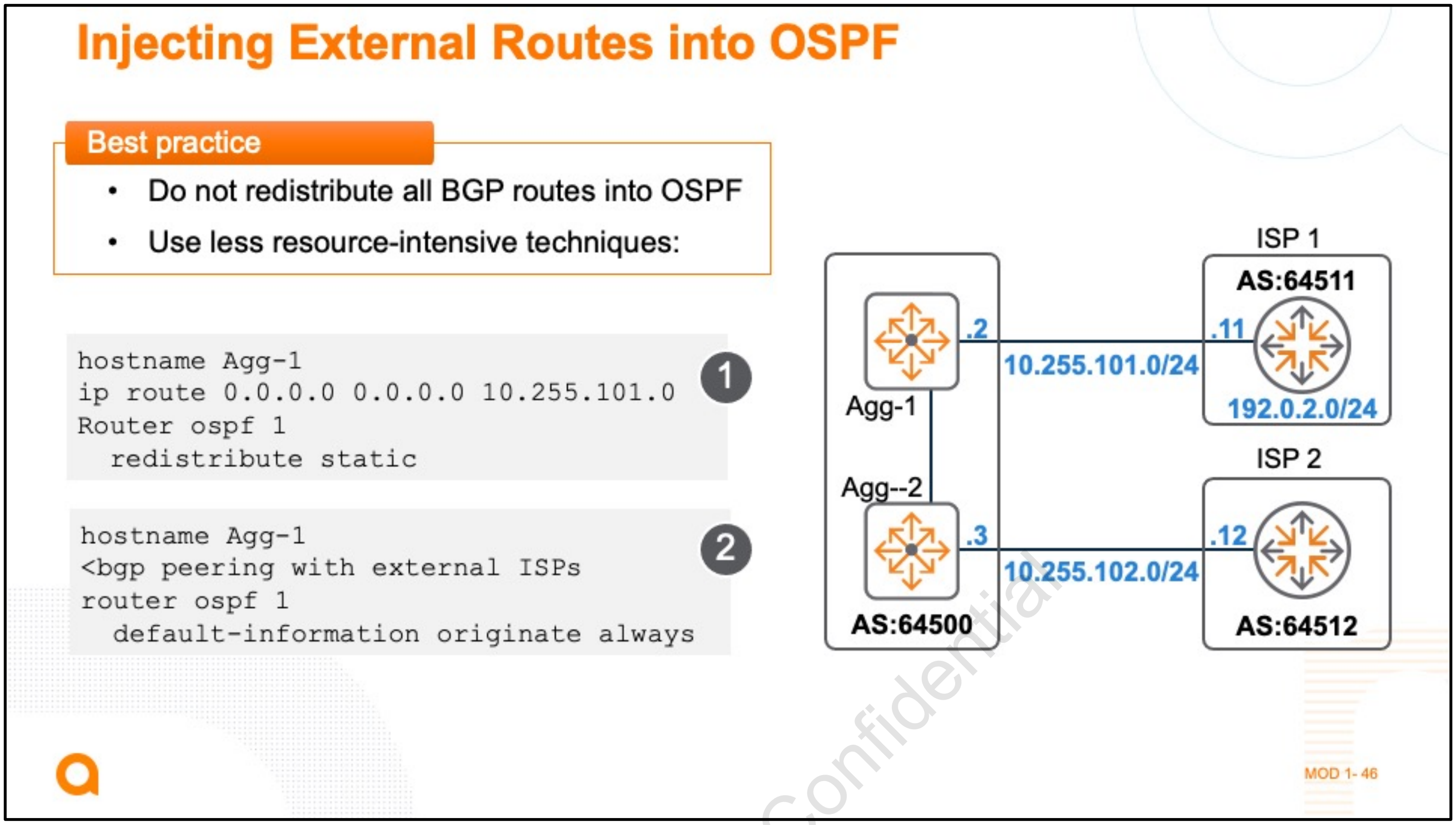
In any case, you must ensure that your BGP edge routers have a default route. You could achieve this with the typical “ip route 0.0.0.0/0” entry, or by using OSPF default-information orginate, or by ensuring that your ISP advertises a default route, which you would then also need to permit with your route map.

In this example, you want to use ISP1 to access resources on the 192.0.2.0/24 network. For other networks, a default route is typically sufficient. The example configuration uses an IP prefix list to select a specific route to learn from ISP1: 192.0.2.0/24. Please note that you might need to perform similar inbound route filtering to make intelligent choices to reaching resources in ISP2 as well.

Important: Make sure your AS is receiving a default route if you filter inbound routes.

What would happen in your link to ISP1 would go down? You would still be able to access 207.1.1.0/24 via a default route to leave your network via ISP2.

Aruba Training-Confidential



Injecting External Routes into OSPF

Best practice

- Do not redistribute all BGP routes into OSPF
- Use less resource-intensive techniques:

1

```
hostname Agg-1
ip route 0.0.0.0 0.0.0.0 10.255.101.0
Router ospf 1
  redistribute static
```

2

```
hostname Agg-1
<bgp peering with external ISPs
router ospf 1
  default-information originate always
```

Aruba Training-Confidential

You should consider the amount of routing information that your BGP routers receive from ISP routers, and how much of that should be available to the company IGP, typically OSPF. Although you can redistribute BGP routes into OSPF, it is not typically recommended. The large BGP routing table might overwhelm the OSPF processes on the local router and on other routers throughout the AS.

Often, you can simply have your BGP routers advertise default routes in OSPF. If they are already acting as ABRs for stub areas, they already advertise default routes in those areas. You might need to have the BGP routers advertise default routes in the OSPF backbone, though. A simple approach is to configure a static route through its ISP connection on the BGP router. Then redistribute static routes into OSPF. This is shown in the figure as option 1.

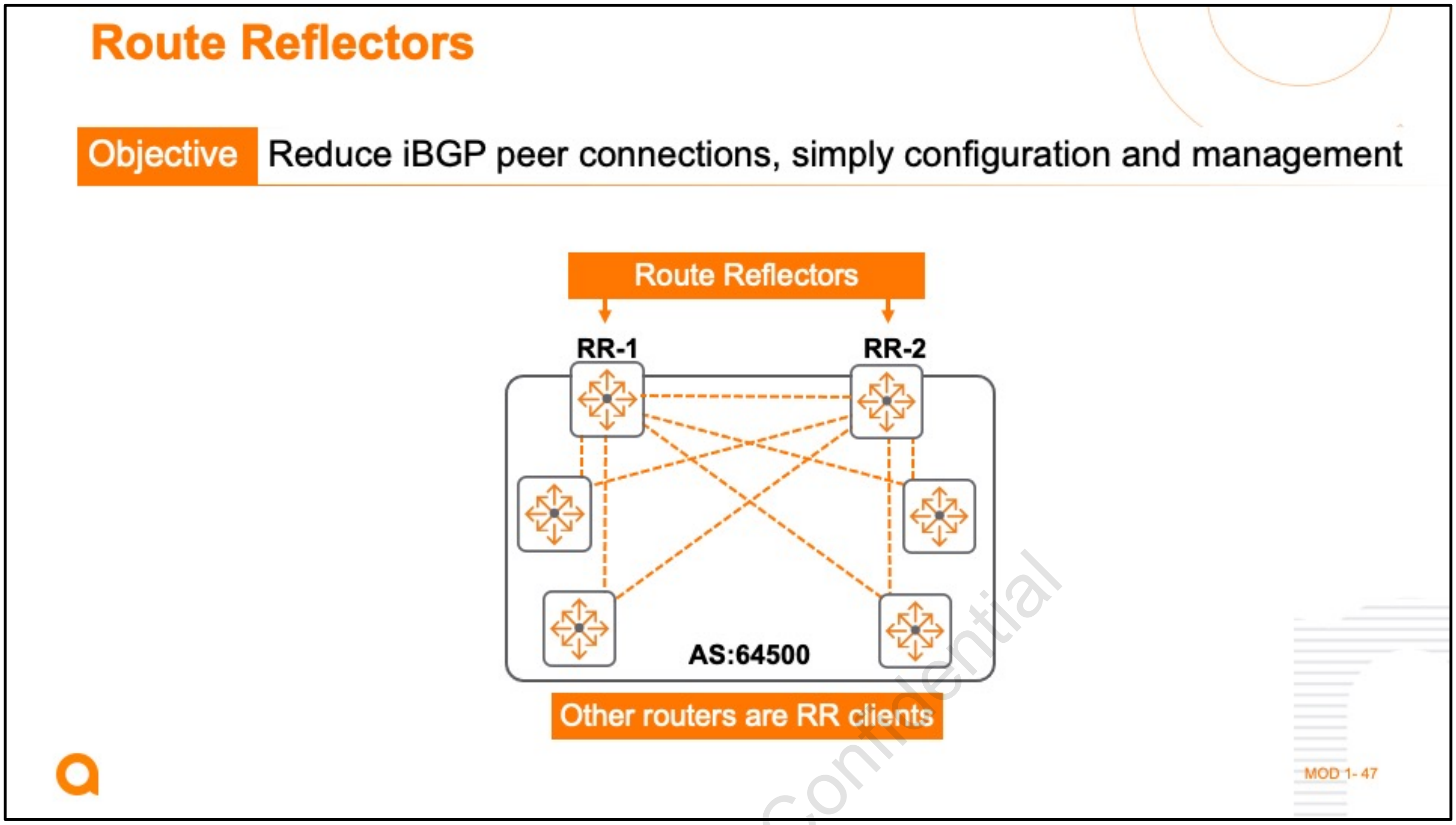
If you are going to redistribute BGP into OSPF, it is highly recommended that you use a route-map to qualify exactly which route(s) you want to redistribute into OSPF. Use the techniques you learned earlier in the redistribution module.

The figure also shows a second option, using the OSPF default-information originate command, with the always option. Without this option, OSPF only advertises default information (gateway of last resort – 0.0.0.0/0) if it actually has a valid path to 0.0.0.0. So, if the interface to 10.255.101.0 goes does down, the static route goes down, so it will not be in OSPF.

If OSPF should have default route, independent of the default route existence in the routing table, use the always.

Note: Your ISP might advertise a default route to your BGP router. In that case, instead of redistributing static routes, you could redistribute just the BGP default route. When you redistribute BGP into OSPF, you apply a route map that permits only that route.

Aruba Training-Confidential



Route Reflectors

Objective Reduce iBGP peer connections, simply configuration and management

Route Reflectors

RR-1

RR-2

AS:64500

Other routers are RR clients

MOD 1- 47

Aruba Training-Confidential

Route Reflectors

A route reflector provides an alternative method of connectivity where the iBGP routers do not need to be fully meshed. In an AS, a single iBGP router (preferably two for redundancy) can be configured as a route reflector. Other iBGP routers in the AS need only be configured as peers to the route reflector(s), as shown in the figure. Route reflectors reflect or “mirror” iBGP routes to other iBGP peers. In the figure Agg-1 and Agg-2 are configured as route reflectors. All other IBGP routers act as route reflector clients.

Understand that a reflector client has no special configuration and is not aware that route reflectors are being used. The big difference is that instead of configuring each client with N-1 peers, you only configure them with 2 peers – Agg-1 and Agg-2 in this example. This greatly improves scalability.

Here’s how it works:

If a route is received from a non-client peer, the route reflector server reflect the route to EBGP peers (as normal) and to reflector clients..

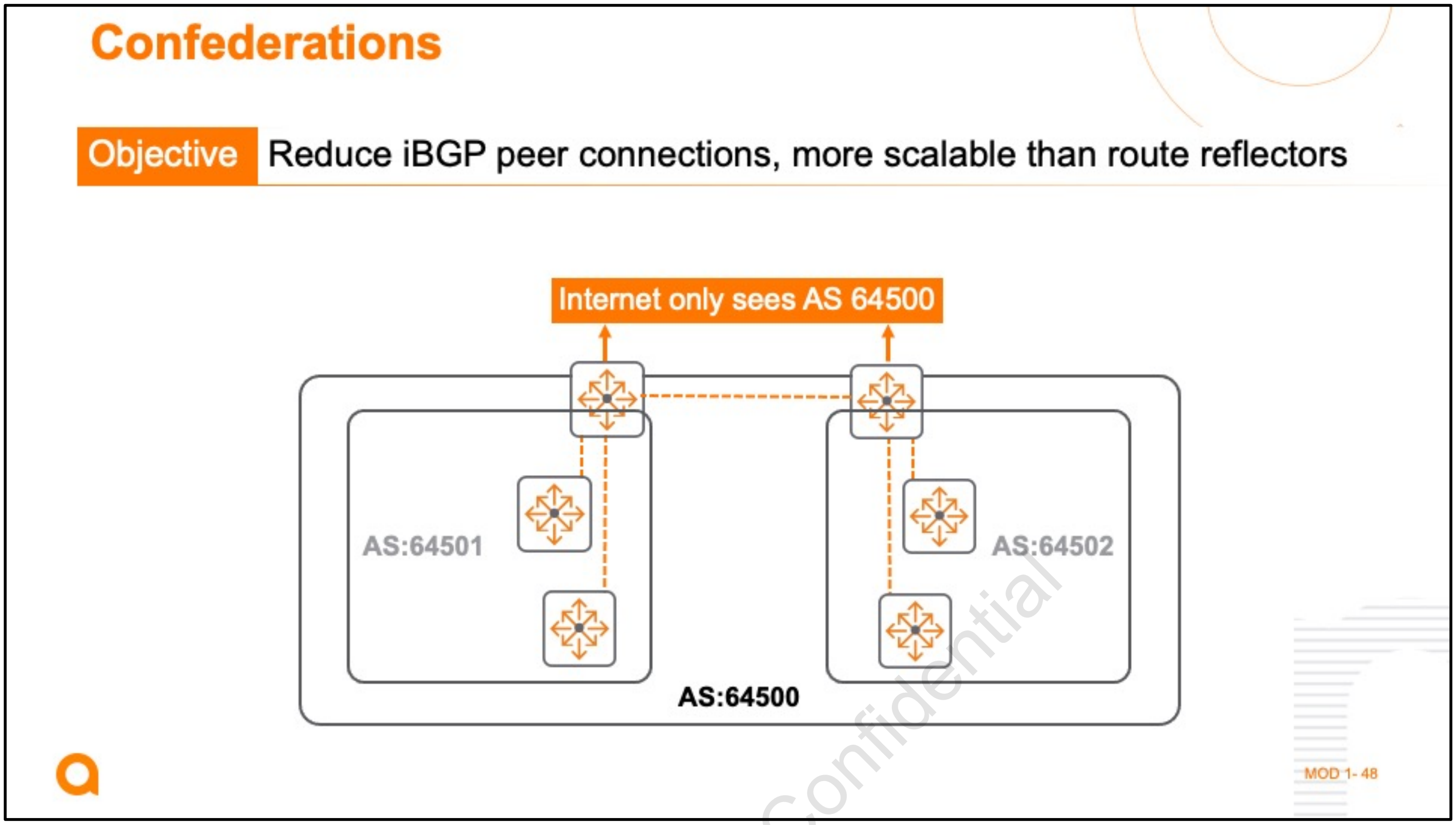
If a route is received from a client peer, the router reflector server reflects the route to all non-client peers and also to client peers, except the originator of the route and reflects it to EBGP peers.

Clusters

A router reflector(s) and its iBGP clients form a cluster. A cluster-ID is associated with every route advertised by the route reflector to its client and/or non-client peers. The cluster-ID is a cumulative, non-transitive BGP attribute and every route reflector must prepend the its local cluster-ID to the cluster-list in order to avoid routing loops. By default, the cluster-ID is the router-ID. The command is shown in the example so you know how to configure it. Route reflectors reduce the number of iBGP peerings to each router and thus reduce processing overhead. To provide scalability, clusters can be interconnected.

Note: Route reflector configuration is beyond the scope of this course.

Aruba Training-Confidential



Confederations

Objective Reduce iBGP peer connections, more scalable than route reflectors

Internet only sees AS 64500

AS:64501

AS:64502

AS:64500

MOD 1- 48

Aruba Training-Confidential

At some point, the use of clusters runs into scalability issues. This is where you can then implement confederations. BGP confederations are sets of “sub-ASs“ within your “real AS”. Typically, the internet only sees this one real AS number. Confederations are used in very large networks where a large AS can be configured to encompass smaller more manageable internal autonomous systems.

Each confederated AS within the company has iBGP fully meshed and has connections to other autonomous systems inside the confederation. Even though these internal autonomous systems have eBGP peers to autonomous systems within the confederation, the autonomous systems exchange routing as if they used iBGP. By doing this, the confederation preserves next hop, metric, and local preference information. To the outside world, the confederation appears to be a single AS. Within this solution, iBGP transit AS problems can be resolved since iBGP requires a full mesh between all BGP routers.

Confederations can be used in conjunction with route reflectors. Both confederations and route reflectors can be subject to issues unless specific design rules, affecting both BGP and the interior routing protocol, are adhered to. Problems that can occur include:

Route oscillation

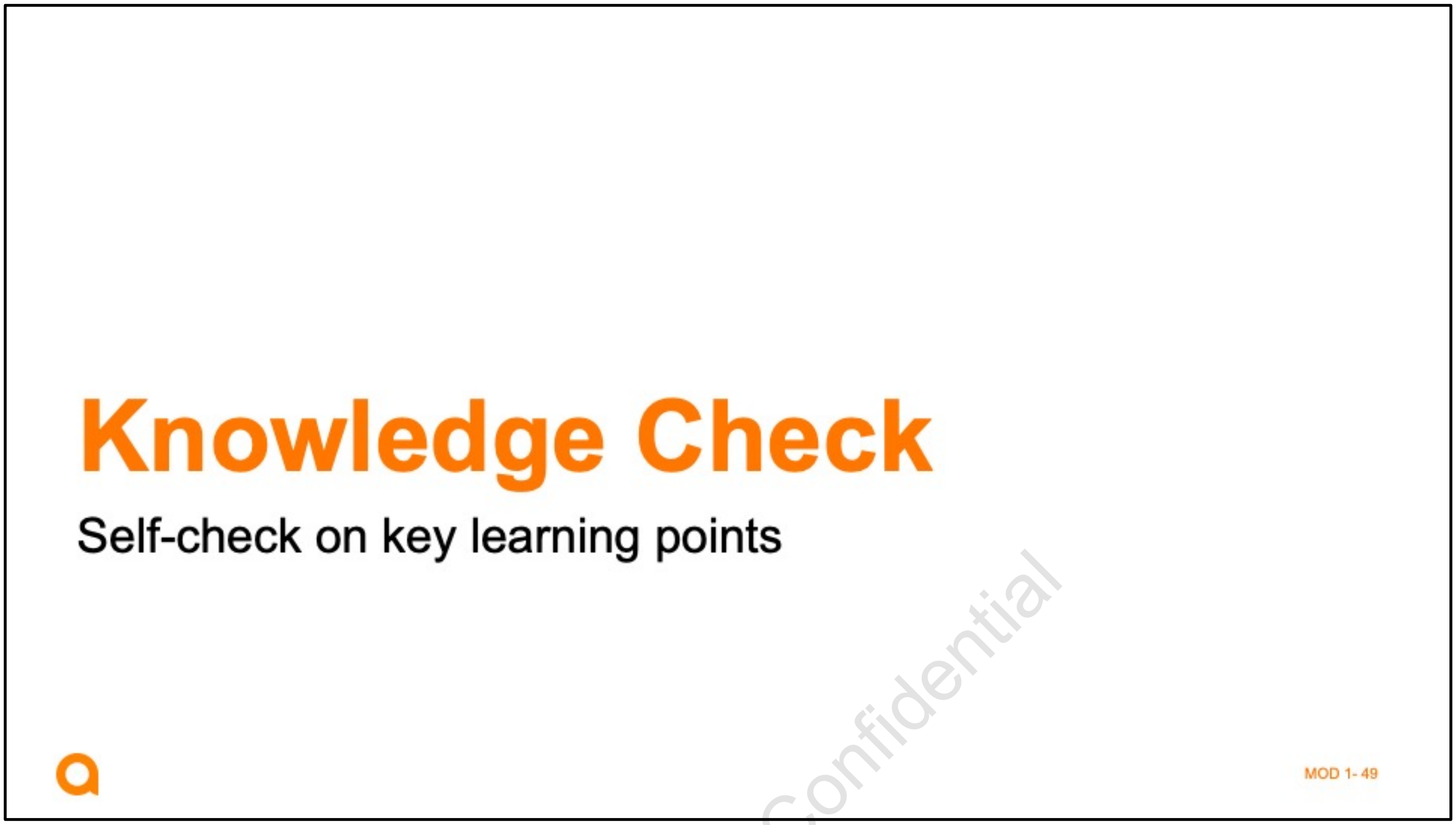
Sub-optimal routing

Increase of BGP convergence time

Additionally, route reflectors, clustering, and BGP confederations were not designed to simplify BGP router configuration. However, these are common tools for experienced BGP network designers and engineers. These tools may be combined, for example, as a hierarchy of route reflectors.

Note: The configuration of confederations is beyond the scope of this course.

Aruba Training-Confidential

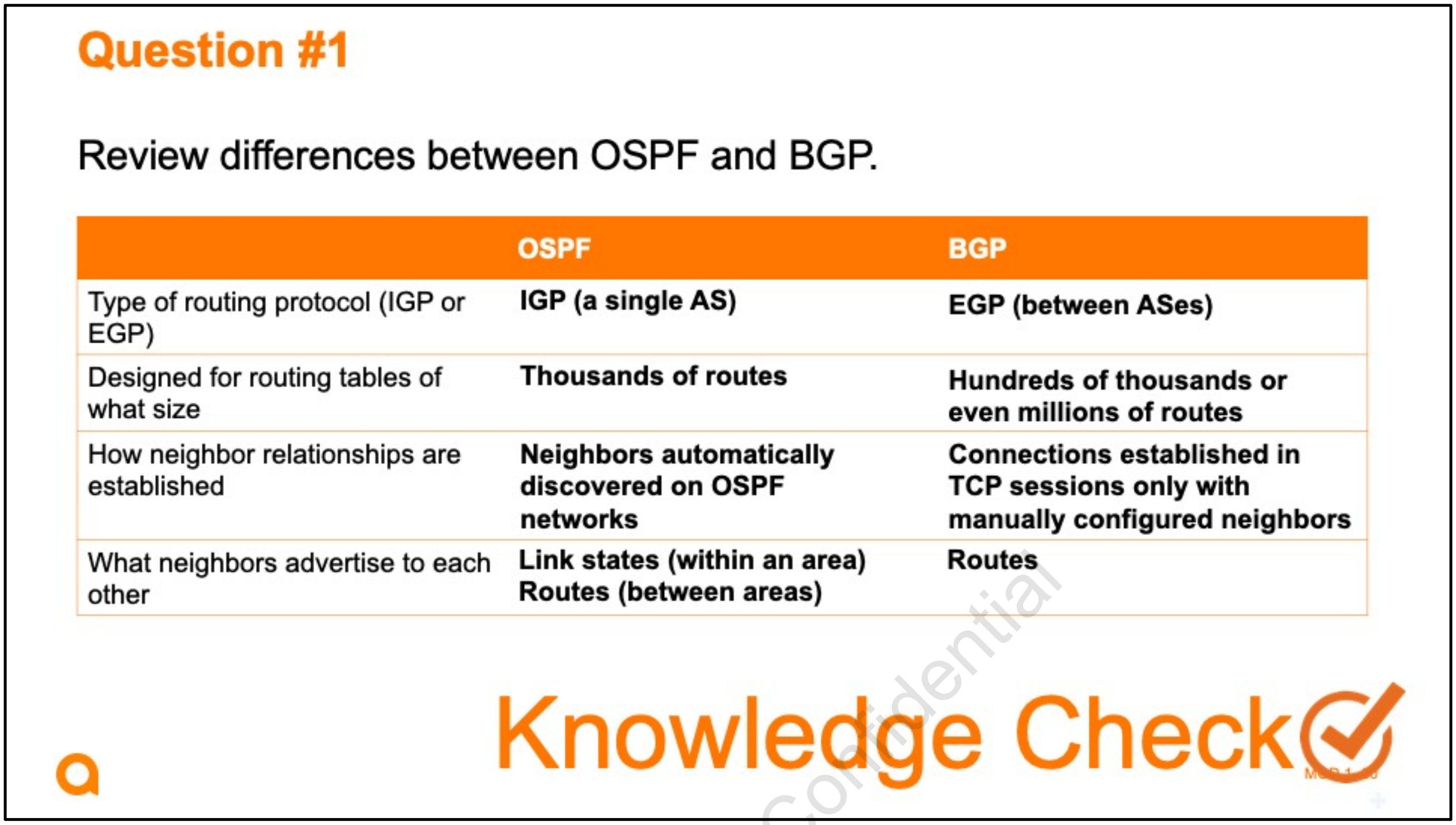


Knowledge Check

Self-check on key learning points

MOD 1- 49

Aruba Training-Confidential



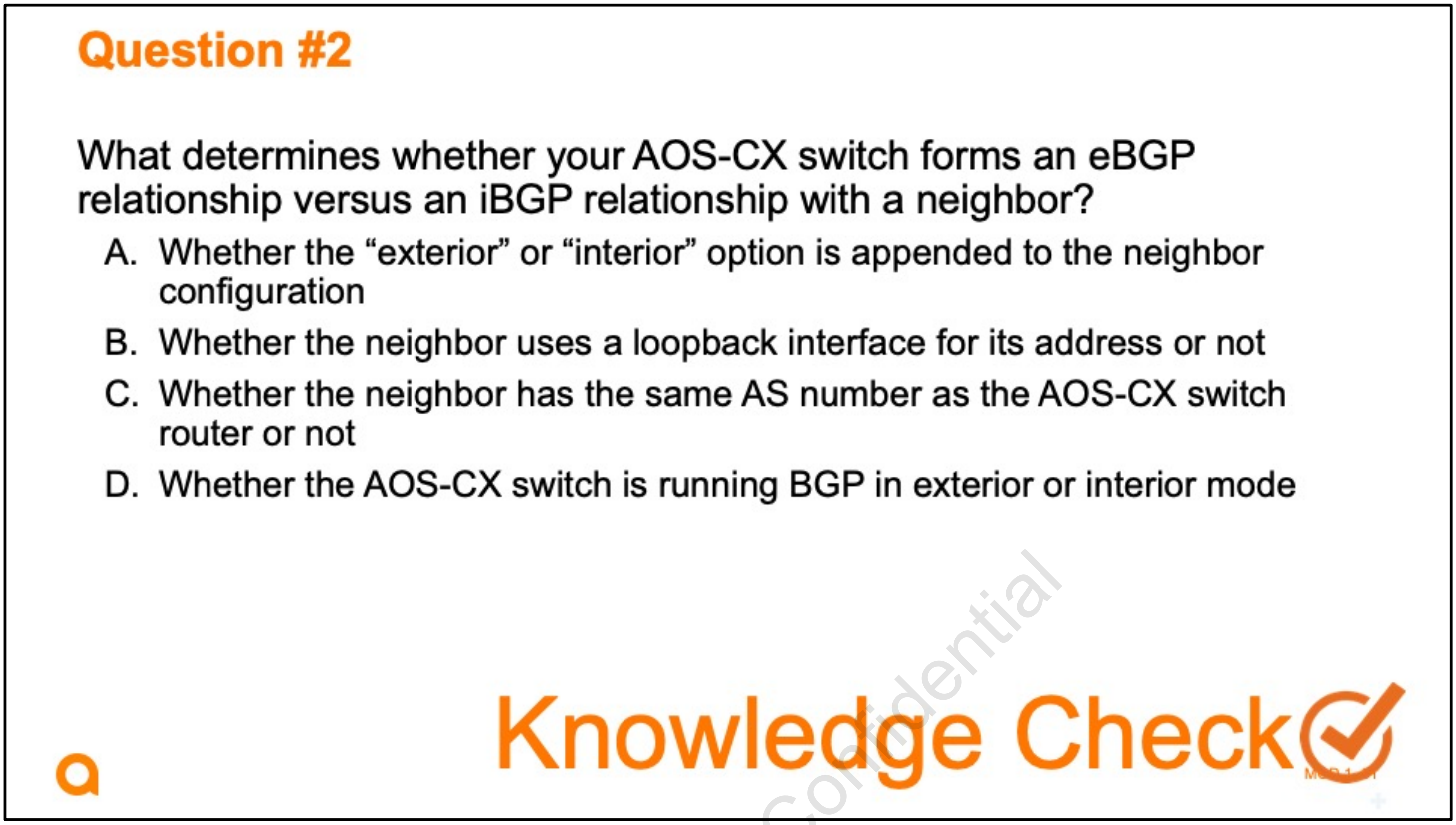
Question #1

Review differences between OSPF and BGP.

	OSPF	BGP
Type of routing protocol (IGP or EGP)	**IGP (a single AS)**	**EGP (between ASes)**
Designed for routing tables of what size	**Thousands of routes**	**Hundreds of thousands or even millions of routes**
How neighbor relationships are established	**Neighbors automatically discovered on OSPF networks**	**Connections established in TCP sessions only with manually configured neighbors**
What neighbors advertise to each other	**Link states (within an area)** **Routes (between areas)**	**Routes**

Knowledge Check

Aruba Training-Confidential



Question #2

What determines whether your AOS-CX switch forms an eBGP relationship versus an iBGP relationship with a neighbor?

A. Whether the “exterior” or “interior” option is appended to the neighbor configuration

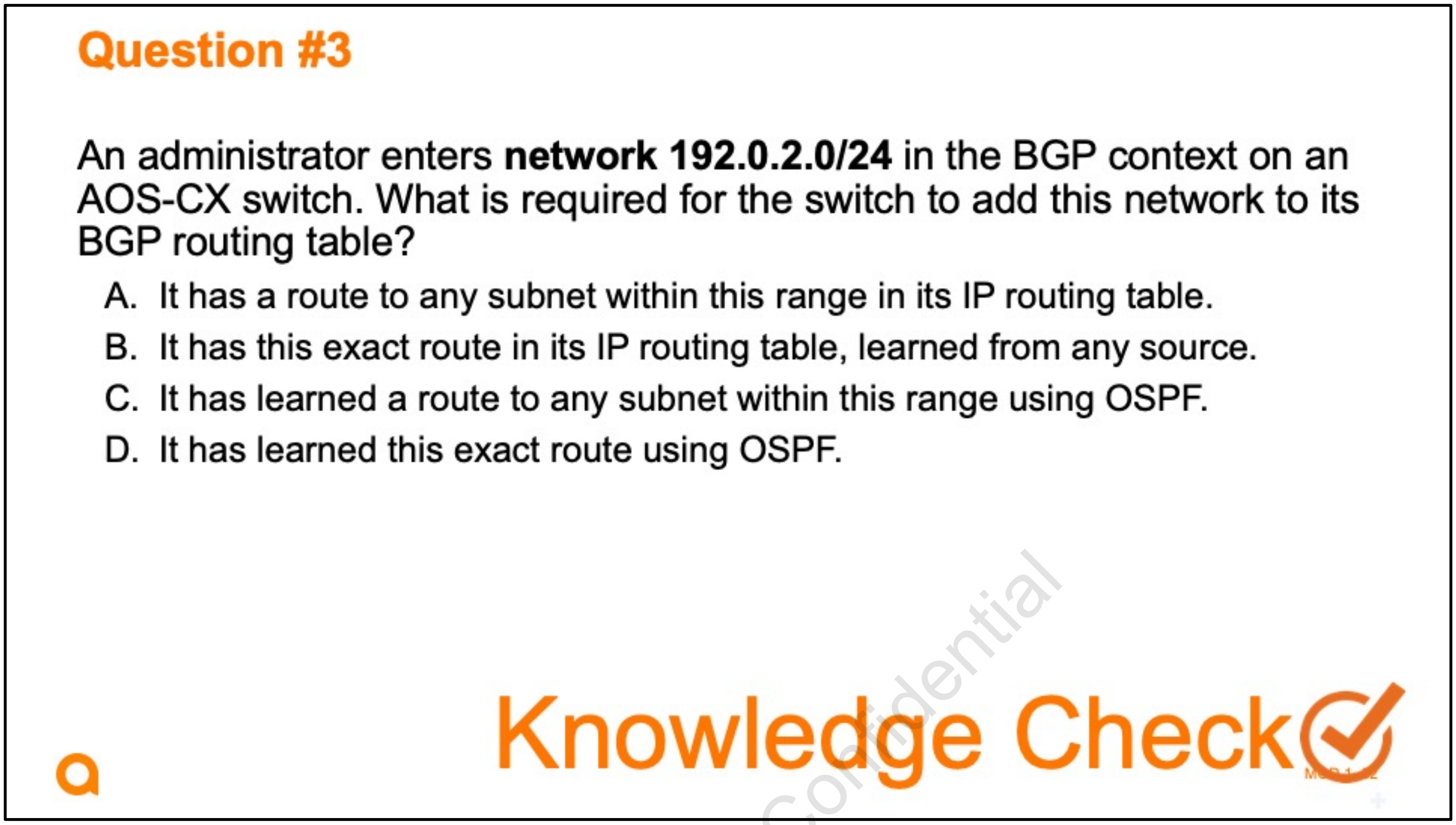
B. Whether the neighbor uses a loopback interface for its address or not

C. Whether the neighbor has the same AS number as the AOS-CX switch router or not

D. Whether the AOS-CX switch is running BGP in exterior or interior mode

Knowledge Check

Aruba Training-Confidential



Question #3

An administrator enters **network 192.0.2.0/24** in the BGP context on an AOS-CX switch. What is required for the switch to add this network to its BGP routing table?

A. It has a route to any subnet within this range in its IP routing table.

B. It has this exact route in its IP routing table, learned from any source.

C. It has learned a route to any subnet within this range using OSPF.

D. It has learned this exact route using OSPF.

Knowledge Check

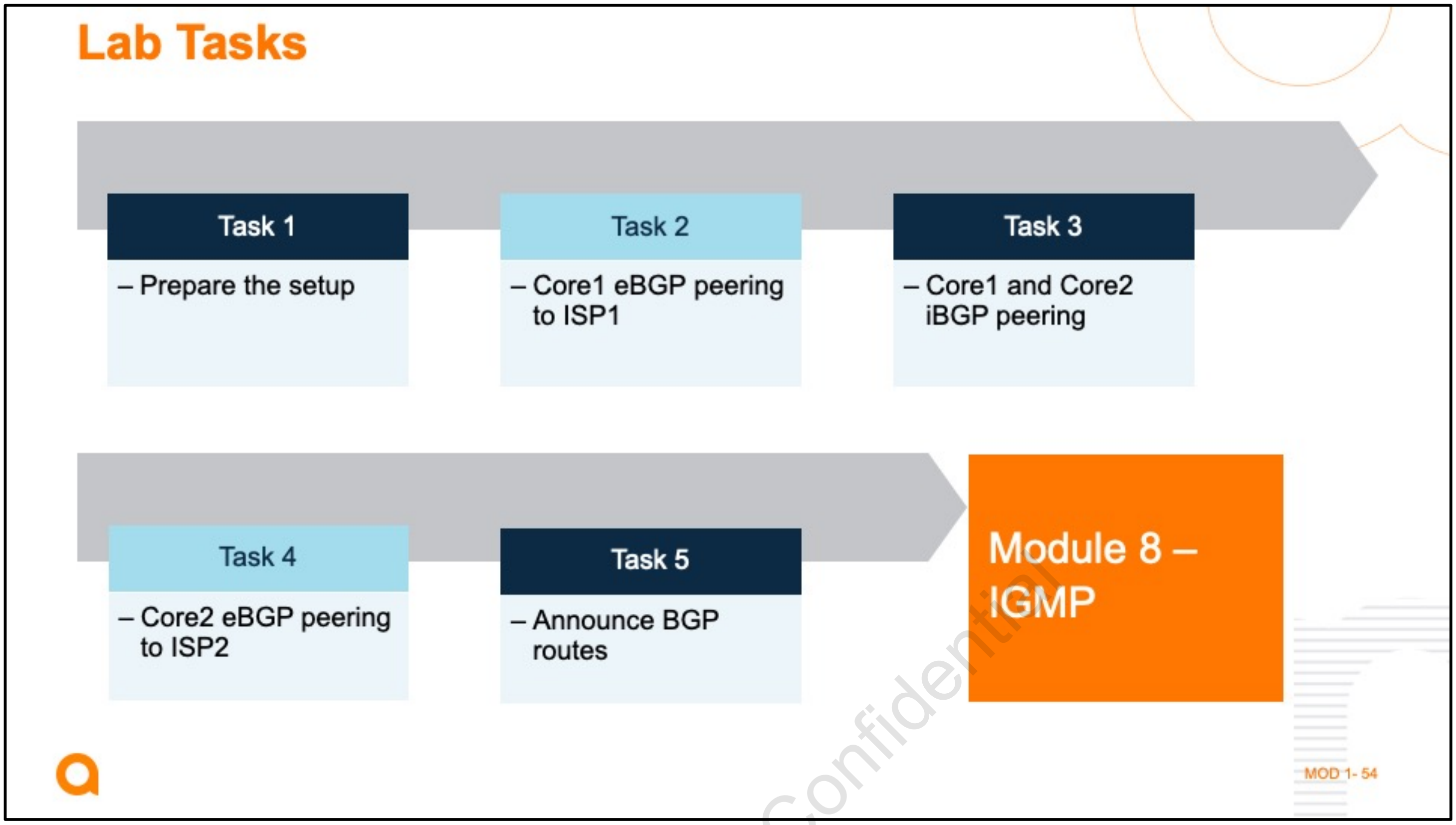
Aruba Training-Confidential

Lab Activity

Lab 6.3

Let’s do the lab.

Aruba Training-C



Lab Tasks

Task 1
- Prepare the setup

Task 2
- Core1 eBGP peering to ISP1

Task 3
- Core1 and Core2 iBGP peering

Task 4
- Core2 eBGP peering to ISP2

Task 5
- Announce BGP routes

Module 8 – IGMP

MOD 1- 54

The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with Module 8 – IGMP.

Aruba Training-Confidential

Aruba Training-Confidential

3333 Scott Blvd, Santa Clara, CA 95054
TEL: 408.227.4500 | FAX: 408.227.4550
www.ARUBANETWORKS.com

EDU-ICX-ILT-v20.211