VMware vSphere: **Design**

Lecture Manual

ESXi 7 and vCenter Server 7

VMware® Education Services
VMware, Inc.
www.vmware.com/education

VMware vSphere: Design [V7]

Lecture Manual

ESXi 7 and vCenter Server 7

Part Number EDU-EN-VSDW7-LECT (05/2020),

The training material is provided “as is,” and all express or implied conditions, representations, and warranties, including any implied warranty of merchantability, fitness for a particular purpose or noninfringement, are disclaimed, even if VMware, Inc., has been advised of the possibility of such claims. This training material is designed to support an instructor-led training course and is intended to be used for reference purposes in conjunction with the instructor-led training course.

The training material is not a standalone training tool. Use of the training material for self-study without class attendance is not recommended. These materials and the computer programs to which it relates are the property of, and embody trade secrets and confidential information proprietary to, VMware, Inc., and may not be reproduced, copied, disclosed, transferred, adapted or modified without the express written approval of VMware, Inc.

Contents

Module 1
Course Introduction

1-2 Course Introduction

1-3 Importance

A successful vSphere virtualization solution is available, manageable, scalable, recoverable, and secure, and it follows VMware best practices. To design a virtualization solution, you must understand and develop skills in analyzing the requirements of your environment and applying VMware design guidelines. In this way, you can implement your solution quickly and efficiently.

1-4 Learner Objectives

After completing this course, you should be able to meet the following objectives:

- Identify and assess the business objectives of the vSphere environment
- Identify business requirements, constraints, assumptions, and risks for all layers in the vSphere environment
- Apply a framework to a design
- Analyze design choices and best-practice recommendations
- Create a design that ensures availability, manageability, performance, recoverability, and security
- Design the core management infrastructure for an enterprise
- Design the virtual data center for an enterprise
- Design the compute infrastructure for an enterprise
- Design the storage and network infrastructures for an enterprise
- Design virtual machines to run applications in a vSphere infrastructure
- Design security, manageability, and recoverability features for an enterprise

1-5 Course Outline

1. Course Introduction
2. Infrastructure Assessment
3. Core Management Infrastructure
4. Virtual Data Center Infrastructure
5. Compute Infrastructure
6. Storage Infrastructure
7. Network Infrastructure
8. Virtual Machine Design
9. Infrastructure Security
10. Infrastructure Manageability
11. Infrastructure Recoverability

1-6 Typographical Conventions

The following typographical conventions are used in this course.

Conventions	Usage and Examples
`Monospace`	Identifies command names, command options, parameters, code fragments, error messages, filenames, folder names, directory names, and path names: • Run the `esxtop` command. • ... found in the `var/log/messages` file.
`Monospace Bold`	Identifies user inputs: • Enter **`ipconfig /release`**.
Boldface	Identifies user interface controls: • Click the **Configuration** tab.
Italic	Identifies book titles: • *vSphere Virtual Machine Administration*
< >	Indicates placeholder variables: • <ESXi_host_name> • ... the `Settings/<Your_Name>.txt` file

1-7 References

Title	Location
vSphere Installation and Setup	https://docs.vmware.com/en/VMware-vSphere/index.html
vCenter Server and Host Management	https://docs.vmware.com/en/VMware-vSphere/index.html
vSphere Virtual Machine Administration	https://docs.vmware.com/en/VMware-vSphere/index.html
v*Sphere Networking*	https://docs.vmware.com/en/VMware-vSphere/index.html
vSphere Security	https://docs.vmware.com/en/VMware-vSphere/index.html
VMware Validated Design	https://docs.vmware.com/en/VMware-Validated-Design/index.html

1-8 VMware Online Resources

Documentation for vSphere: https://docs.vmware.com/

VMware Communities: http://communities.vmware.com

- Start a discussion.
- Access the knowledge base.
- Access documentation, technical papers, and compatibility guides.
- Access communities.
- Access user groups.

VMware Support: http://www.vmware.com/support

VMware Hands-on Labs: http://hol.vmware.com

VMware Education: http://www.vmware.com/education

- Access course catalog and worldwide course schedule.

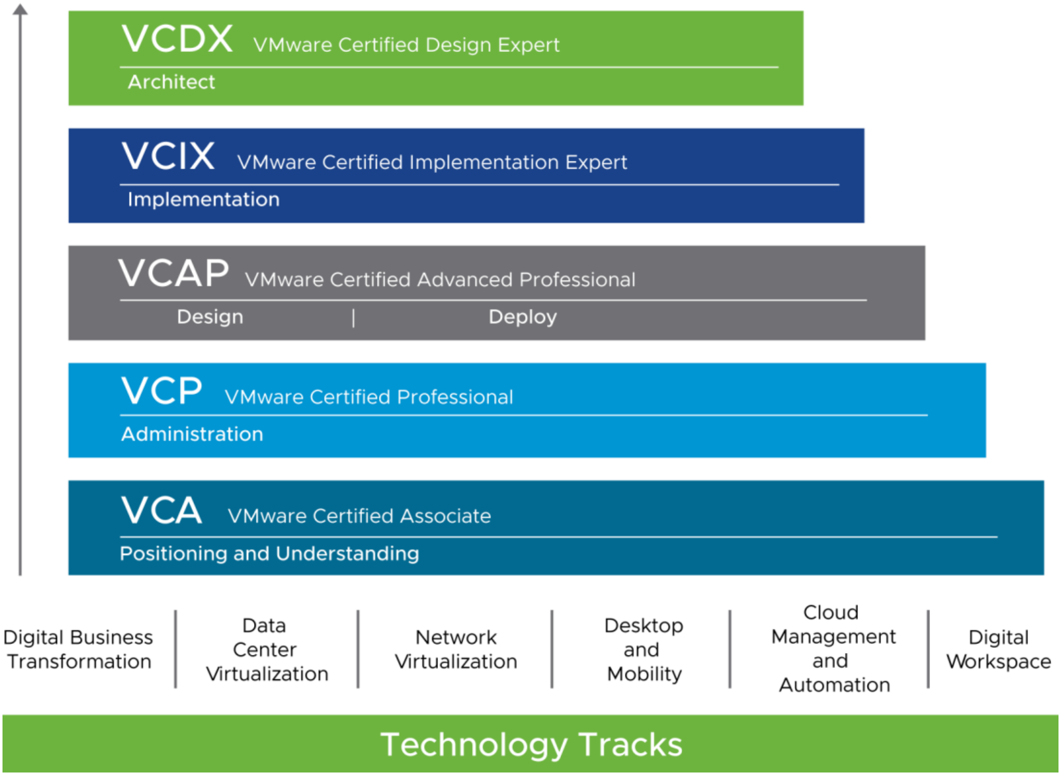
1-9 VMware Education Overview

Your instructor will introduce other Education Services offerings available to you:

- VMware Learning Paths:
  - Help you find the course that you need based on the product, your role, and your level of experience
  - Can be accessed at https://vmware.com/education
- VMware Learning Zone, which is the official source of digital training, includes the following options:
  - On Demand Courses: Self-paced learning that combines lecture modules with hands-on practice labs
  - VMware Lab Connect: Self-paced, technical lab environment that lets you practice skills learned during instructor-led training
  - Certification Exam Prep: Comprehensive video-based reviews of exam topics and objectives to help you take your certification exam
- For more information, see https://vmwarelearningzone.vmware.com.

1-10 VMware Certification Overview

VMware certifications validate your expertise and recognize your technical knowledge and skills with VMware technology.


VCDX VMware Certified Design Expert
Architect
VCIX VMware Certified Implementation Expert
Implementation
VCAP VMware Certified Advanced Professional
Design | Deploy
VCP VMware Certified Professional
Administration
VCA VMware Certified Associate
Positioning and Understanding
Digital Business Transformation
Data Center Virtualization
Network Virtualization
Desktop and Mobility
Cloud Management and Automation
Digital Workspace
Technology Tracks

VMware certification sets the standards for IT professionals who work with VMware technology. Certifications are grouped into technology tracks. Each track offers one or more levels of certification (up to five levels).

For the complete list of certifications and details about how to attain these certifications, see https://vmware.com/certification.

1-11 VMware Badge Overview

VMware badges are digital emblems of skills and achievements.

vmware
PROFESSIONAL
DATA CENTER VIRTUALIZATION 2020
CERTIFICATION

vmware
IMPLEMENTATION EXPERT
DATA CENTER VIRTUALIZATION 2020
MILESTONE

vmware
SPECIALIST
vSAN 2020
SPECIALIST

vmware
IT ACADEMY
NETWORK VIRTUALIZATION CONCEPTS
KNOWLEDGE

vmware
HYPER-CONVERGED INFRASTRUCTURE
HCI TECHNICAL WHITEBOARD 2020
SKILL

vmware
SUBJECT MATTER EXPERT
CLOUD MANAGEMENT & AUTOMATION 2020
CONTRIBUTOR

Digital badges have the following features:

- Easy to share in social media (LinkedIn, Twitter, Facebook, blogs, and so on)
- Tethered to VMware to validate and verify achievement
- Contain metadata with skill tags and accomplishments
- Based on Mozilla's Open Badges standard

For the complete list of digital badges, see http://www.pearsonvue.com/vmware/badging.

Module 2
Infrastructure Assessment

2-2 Importance

To keep the vSphere virtual infrastructure design on track, you must clearly define goals, requirements, assumptions, risks, and constraints. These factors are the guideposts of the design.

2-3 Module Lessons

1. Business Objectives and Requirements
2. Conceptual, Logical, and Physical Designs
3. Overview of Architecture Frameworks

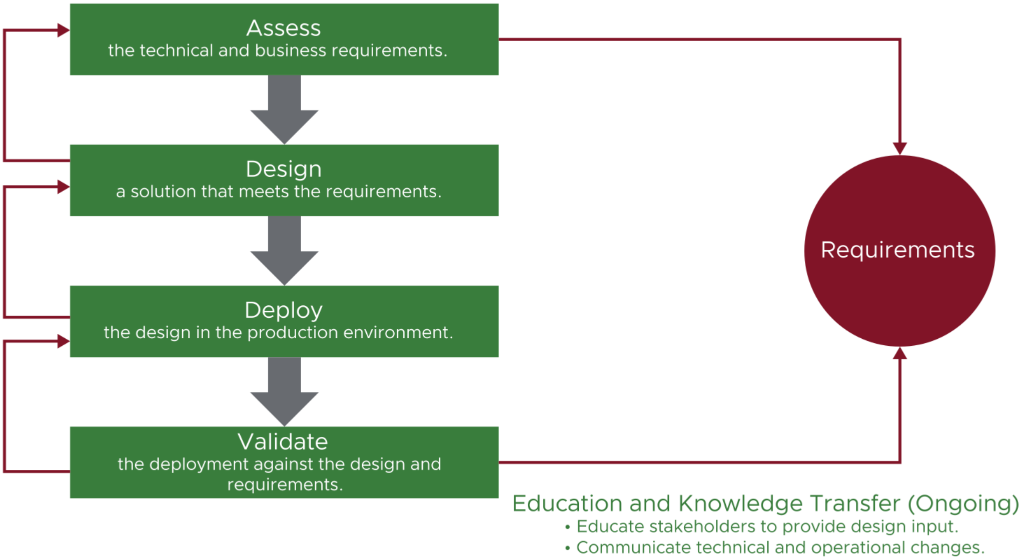
2-4 **Lesson 1: Business Objectives and Requirements**

2-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Follow a proven process to design a virtualization solution
- Define customer business objectives
- Gather and analyze business and application requirements

2-6 Design Process Overview


Assess
the technical and business requirements.
Design
a solution that meets the requirements.
Deploy
the design in the production environment.
Validate
the deployment against the design and requirements.
Requirements
Education and Knowledge Transfer (Ongoing)
• Educate stakeholders to provide design input.
• Communicate technical and operational changes.

Following a proven design methodology is the most efficient way to conduct a design-and-deploy engagement. Using a proven methodology also helps make the design repeatable and predictable.

This course follows a design process based on the V-model of systems engineering and verification. The foundation of the process is a thorough set of business and technical requirements. Subsequent phases build on the work in the previous phase and are validated against the requirements to ensure that the project stays on track. This design-and-deploy process includes the following phases:

- Assess: You define the scope of the project and gather data for the design. Understanding customer requirements and project specifics and identifying the key stakeholders are critical success factors for delivering a design that meets customer needs.
- Design: With the information gathered in the assessment phase, you can begin the design work. The design phase is iterative. This phase involves a series of individual and group interviews with stakeholders, and individual work by the designer. This phase also involves considering the organization's goals, requirements, and constraints, and current best practices.

- Deploy: The objective of the deploy phase is to build and configure the production environment according to the design documents. The build typically goes through changes and iterations.
- Validate: After the solution is deployed, you conduct tests to verify that the solution is built to the design specifications and that the solution behaves as required. During the validation phase, the customer learns the mechanics of the solution. This phase might be iterative and require additional changes to resolve configuration issues and other concerns.

Customer involvement during all phases of the process is key to the success of the project. During the assess and design phases, stakeholders must understand vSphere and virtualization concepts so that they can provide appropriate inputs in the design process. During the deploy and validate phases, customers, operators, and administrators must understand the new technology and operational changes so that they can successfully adopt the design.

2-7 Examples of Deliverables by Phase

The documentation requirements for each phase of a project vary based on the cost and length of the project.

Phase	Documentation
Assessment	Current state analysis report Conceptual design
Design	Design blueprints, which include a logical design and a physical design
Deployment	Installation and configuration documentation
Validation	Validation plan with test results

Each phase of a project has different deliverables.

Documentation requirements vary based on whether the design is product-oriented or solution-oriented:

- In a product-oriented design, the performance of the product is the measure of success. The design is often more generic and less tailored to the organization’s specific needs. Documentation typically includes a larger proportion of references to instructions in the product documentation than a solution-oriented design.
- In a solution-oriented design, how well the design achieves the goals of the organization is the measure of success. This design type is used in this course. In general, a solution-oriented project requires more extensive documentation because the solution is uniquely tailored to the conditions of the organization.

2-8 Project Schedule Example

Project time frames vary by engagement. This example shows the relative time frames to complete the different phases in a project.

Activities by Week	Days to Complete	1	2	3	4	5	6	7	8	9
1. Planning	1/2	🟩								
2. Current State Analysis	10		🟩	🟩						
3. Kickoff	1				🟩					
4. Educate	1/2				🟩					
5. Assessment	2				🟩					
6. Design	8					🟩	🟩			
7. Deploy	9							🟩	🟩	
8. Validate	3									🟩
9. Knowledge Transfer	5									🟩
10. Conclusion	1									🟩

This course focuses on
the Assessment and

The slide shows a typical schedule for a virtualization design and deploy project.

2-9 Working with Stakeholders

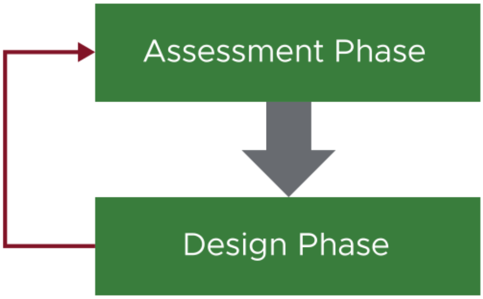
Stakeholders should include representatives from all groups that are affected by the design. Stakeholders can include the following individuals:

- A project sponsor (for example, the CIO, VP of infrastructure, or IT director)
- Virtualization architects
- Business decision makers
- Core technical teams, such as product development, server, storage, networking, security, and backup and recovery

You interview stakeholders and conduct workshops to gather requirements and build consensus.

Gathering requirements is an iterative process, which might require multiple rounds of interviews.

Asking the right questions is vital, and you must gather both functional and nonfunctional requirements.



A good strategy for a successful project is to bring the correct people together and build consensus. A design that is viewed as successful by only a portion of the project participants and sponsors works against a large-scale adoption. You must identify project advocates and project participants who might require some extra attention. Your goal is to be an unbiased project champion. You must show concern for each department's unique challenges and perspectives while having a firm grasp on the vision of the project as a whole.

The list of stakeholders and subject matter expert areas is not exhaustive. You might identify additional people to meet with during your design discussions.

When interviewing stakeholders, requirements can be functional and nonfunctional:

- Functional requirements describe what the solution must accomplish or how the solution must behave, for example, the system must limit access to authorized users.
- Nonfunctional requirements describe the overall characteristics of a solution, for example, the system must be located in the DMZ to satisfy security requirement RS05.

2-10 Identifying Business Factors That Affect Designs

When conducting interviews in the assessment phase, ask questions to learn about business conditions and practices that affect the design. The following common factors affect the design:

- Organizational boundaries: Physical, political, and cultural
- Experience with virtualization and VMware products
- Service-level agreements (SLAs)
- Security and compliance requirements
- Time and project urgency
- Policies and procedures
- Training requirements
- Budget constraints
- Future scaling

Interviews early in the design process can uncover conditions that affect the end design. For example, you might discover that the primary goal of one department in an organization is to reduce its data center footprint and the associated costs. The goal of another department might be to virtualize to meet the business continuity and disaster recovery initiatives of its business. The first department might be seeking to achieve 10:1, 20:1, or higher consolidation ratios. However, the second department might be entirely satisfied with a 5:1 consolidation ratio if the virtual infrastructure meets its disaster recovery requirements.

Each organization has unique conditions that must be identified. Use the list on the slide as a starting point. Other factors include application requirements and support factors, licensing with other platforms, use of the existing hardware, amortization of the investment, and preference for a certain vendor.

Future scaling of the environment is a common factor. Clients might not know the exact scale of their environment when designing the environment. However, having a rough idea of scaling objectives can often prevent redesigning or redeploying components later.

For example, an organization deploys a new virtual infrastructure. If that project is a success, the organization might consider bringing in additional virtual machines from some of their other environments.

2-11 Desired Business Outcomes

Begin the design process when you have a set of clear desired business outcomes.

Business objectives or desired business outcomes help to define the scope of the project and to keep the design on track.

The following examples describe common business objectives:

- Extend the vSphere platform to other areas of the business.
- Build a strong foundation for future projects, including cloud infrastructures.
- Optimize an existing implementation to change and expand over time.
- Consolidate servers across multiple data centers.
- Virtualize the environment for business-critical applications.

You must begin the project by determining the primary business and technical problems that your customer is trying to solve. Typically, these objectives align with common virtualization use cases.

2-12 Analyzing the Current State

Use monitoring tools to do an inventory of the existing infrastructure and report the resource usage.

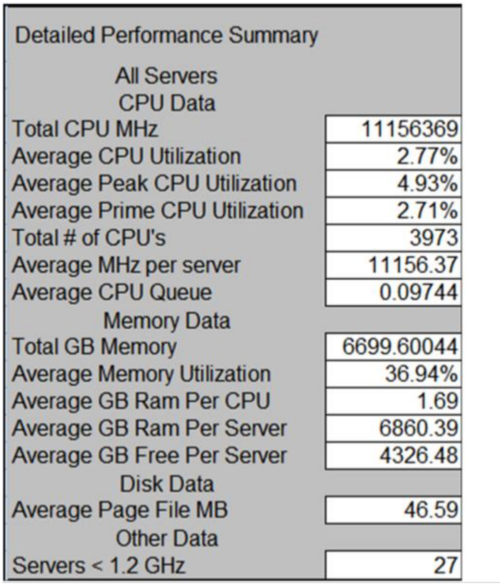
For each system, capture peak and average utilization for the following items:

- CPU
- RAM
- IOPS
- Network utilization

Use the application vendor's requirements for proper sizing of the application.

You can use the following tools to gather inventory and capacity analysis information:

- vRealize Operations Manager
- Operating system-based tools
- Third-party inventory and sizing tools



Detailed Performance Summary

All Servers	
CPU Data	
Total CPU MHz	11156369
Average CPU Utilization	2.77%
Average Peak CPU Utilization	4.93%
Average Prime CPU Utilization	2.71%
Total # of CPU's	3973
Average MHz per server	11156.37
Average CPU Queue	0.09744
Memory Data	
Total GB Memory	6699.60044
Average Memory Utilization	36.94%
Average GB Ram Per CPU	1.69
Average GB Ram Per Server	6860.39
Average GB Free Per Server	4326.48
Disk Data	
Average Page File MB	46.59
Other Data	
Servers < 1.2 GHz	27

Sample Section from a Capacity Planner Report

Several methods are available to collect workload information from an existing infrastructure, although some methods are easier to use and are more comprehensive than other methods.

You can use vRealize Operations Manager to analyze the current state of your environment and help you make resource sizing decisions.

You can also use third-party workload-sizing tools. Prices and capabilities of these tools vary. See the vendor documentation for information about these tools. You can also use operating system-based tools, but the task is time-consuming even in a relatively small data center with only a dozen or so machines. Acquiring meaningful and useful inventory information requires a manual aggregation of the results and the approach is impractical in medium-to-large data centers.

2-13 Gathering Application Requirements

Application requirements determine how to optimize the vSphere design for the applications that run on it.

Gather application requirements from the following sources:

- Interviews with application owners
- Application SLAs
- Vendor documentation and industry averages
- VMware best practice guides

The following examples describe the type of information to gather:

- Workload characteristics (business-critical, VDI, ITaaS, and so on)
- Hardware and software requirements
- Service dependencies
- Communication requirements between applications and with the outside world
- Security zoning requirements
- Specifications for performance, availability, and so on

Through interviews with application owners and the current state analysis, you discover a list of applications that the vSphere infrastructure must support. You must gather the requirements for those applications.

Application requirements are often stated in service-level agreements (SLAs). Sometimes, industry averages or vendor information might be available to determine the current workload and required virtualization resources.

For VMware best practice guides, go to http://www.vmware.com/techpapers.html and search for best practices.

For VMware best practices on virtualizing business-critical applications, see Virtualizing Business Critical Applications at https://www.vmware.com/solutions/business-critical-apps.html.

2-14 Activity: Calculating the SLA (1)

In this activity, you calculate the SLA given the amount of targeted downtime.

- Calculate the constants using the following parameters:
  - 24 hours x 60 minutes = 1,440 minutes per day
  - 1,440 minutes x 365 days = 525,600 minutes in a given calendar year
- Determine the % uptime using the following calculations:
  - If you can only sustain 5 minutes of downtime, subtract 5 minutes from 525,600.
  - Divide the result by the number of minutes in the year to determine the % uptime.
  - In this example, 525,595/525,600 = .99999048… = 99.999% uptime

What is the SLA for 45 minutes of downtime per month?

2-15 Activity: Calculating the SLA (2)

What is the SLA for 45 minutes of downtime per month?

- The example parameter is 1,440 minutes per day (24 hours x 60 minutes).
- Assuming 30 days per month, you calculate 30 x 1,440 minutes in the month = 43,200 minutes.
- SLA = Total Available Time - Acceptable Downtime/Total Available Time:
  - 43,200 - 45 / 43,200 = 43,155 mins
  - 43,155 / 43,200 = 0.99
  - SLA of 99.9%

2-16 Lab 1: Determining Business Objectives

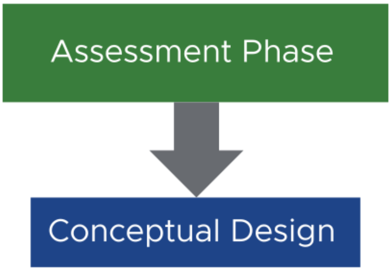
Identify and document the business objectives:

1. Read the Case Study
2. Determine the Business Objectives

2-17 Review of Lab

The instructor facilitates a class discussion to answer the following questions:

- Are you thinking of a solution?
- What are your initial thoughts?
- Do you have any concerns?
- What are the business objectives?



Assessment Phase

Conceptual Design

2-18 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Follow a proven process to design a virtualization solution
- Define customer business objectives
- Gather and analyze business and application requirements

2-19 Lesson 2: Conceptual, Logical, and Physical Designs

2-20 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

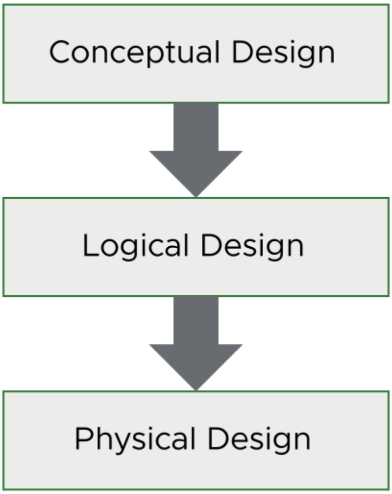
- Document design requirements, constraints, assumptions, and risks
- Use a systematic method to evaluate and document design decisions
- Create a conceptual design

2-21 The Design Model

After you collect the assessment data, you must document the findings so that they can be referenced throughout the design process.

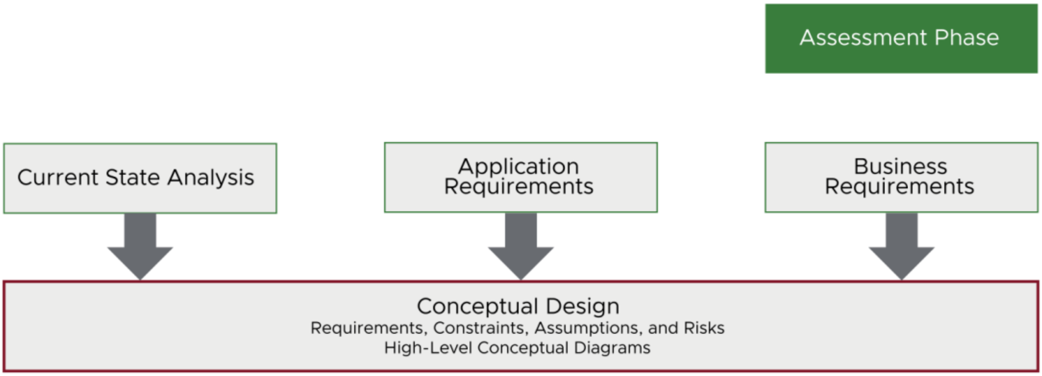
Enterprise infrastructure designs follow a three-step design model:

1. Create a conceptual design.
2. Create a logical design.
3. Create a physical design.



2-22 Creating the Conceptual Design

The conceptual design captures the assessment findings to ensure that the solution meets goals and requirements while staying within the constraints.



The conceptual design is a high-level design based on the information gathered during the assessment phase.

2-23 Conceptual Design: Requirements

Business requirements are the requirements that the designed solution must meet.

The requirements describe what should be achieved in the project and what the solution looks like.

Requirement examples:

- The business must be able to move applications between organization locations in real time.
- The organization must comply with Sarbanes-Oxley regulations.
- The application availability must be 99.99%.

Business requirements are provided by key stakeholders, and the goal of every solution is to achieve each of these requirements.

The requirements include functional and nonfunctional requirements.

2-24 Functional and Nonfunctional Requirements

A functional requirement describes what a system must do, whereas nonfunctional requirements place constraints on how the system performs.

Requirement	Definition	Example
Functional	The functional requirement describes the behavior of the system as it relates to the system's functionality.	A system must send an email whenever an order is placed.
Nonfunctional	The nonfunctional requirement describes a performance characteristic of the system.	Emails should be sent with a latency of no greater than 12 hours from such an activity.

Typically, nonfunctional requirements fall into areas such as availability, manageability, performance, recoverability, and security.

Consider functional requirements as must-do requirements and nonfunctional requirements as must-be requirements, for example:

- The system must process customer records to identify any potential fraudulent activity.
- The system must provide notifications when fraud is detected.
- The system must be scalable to 10 million customer records.
- The system must be recoverable within 15 minutes after failure.

2-25 Examples of Functional Requirements

Functional requirements describe what a system or solution must do. The requirements include the following categories:

- Business rules: For example, the architecture must support both the primary and secondary data centers.
- Administrative functions: For example, network and security administrators must monitor the network traffic of the desired systems.
- Authentication: For example, the system limits access to authorized users.
- Audit tracking
- Certification requirements
- Reporting requirements
- Legal or regulatory requirements

Functional requirements specify functions that the system must perform. Other examples include the following requirements:

- The solution must support 200 virtual machines running concurrently.
- The solution must scale by 20 percent over the next three years.

2-26 Examples of Nonfunctional Requirements

Quality	Description	Measurements (Often in SLAs)
Availability	Available when needed No single point of failure Redundancy	Class of nines Redundancy levels (N+1, 2N, and so on)
Manageability	Managing the infrastructure in terms of lifecycle management, scalability, and capacity planning	Cluster nodes must scale when compute resources are sustained at 70%+ for 5 business days. ESXi host updates must be installed within 1 week of release.
Performance	Provides the required amount of work using the least amount of time and resources	Throughput, latency, transactions per second
Recoverability	Easy to restore after an outage	MTD, RTO, RPO
Security	Minimizes risk and unnecessary complexity Defense-in-depth architecture	Government regulations. Industry standards. Third-party penetration testing and auditing must be executed against environment quarterly with pass rate of 80% or higher.

Nonfunctional requirements describe the quality of a system or solution. The table shows some of the most common nonfunctional qualities and industry standards for measuring the qualities and defining success.

An available design is reliable, eliminates single points of failure, and implements mechanisms to quickly restore services if a failure occurs. Service-level agreements (SLAs) are often used to determine the required level of availability. One common measurement of availability is the class of nines, which refers to the number of nines in the percentage of time a system is available. For example, if a system is available 90 percent of the time, it is called one nine or class one. If a system is available 99.99 percent of the time, it is called four nines or class four. Redundancy requirements are often defined by the number of additional components that are required to ensure availability when a failure occurs. For example, an N+1 redundancy requirement means that one additional component is required. 2N means that twice as many components are required.

A manageable design can be consistently deployed, administered, operated, and supported. This design evolves with organizational needs. When you think of manageability, think of simplification. When a design is as simple as possible, it is easier to understand, easier to explain to others, and easier to expand. Identifying problems and recovering from disasters are also easier. Manageability requirements are often defined by IT service management frameworks, such as ITIL.

A design performs well if it provides the required amount of work using a minimal amount of time and resources. SLAs often define the acceptable amounts of work, time, and resources, and how these factors are measured. If no SLAs are in place, use stakeholder interviews, industry standards, and vendor recommendations.

An infrastructure should be recoverable from any kind of outage. The method of recovery can be as simple as applying backups to implementing disaster recovery procedures. The following metrics dictate the recovery method that is used:

- Maximum tolerable downtime (MTD) defines the duration a system can be down before it causes irreversible consequences.
- Recovery time objective (RTO) defines the amount of time taken to restore the services of the architecture after a disaster or disruption.
- Recovery point objective (RPO) defines the point in time to which the data must be recovered after a disaster or disruption.

A good security design minimizes risks and unnecessary complexity. Typically, security designs apply the concept of defense in depth. Defense in depth is a strategy of multiple layers of protection, where each layer must be individually circumvented to gain access. In theory, the more layers of security, the greater the level of overall protection. The number and type of measures must be weighed against the complexity of implementing and administering the various layers.

Many governments and industry associations worldwide have set up regulations and standards that mandate the protection of various types of information. In the United States, industry-specific privacy regulations and state regulations dictate requirements for handling personal information. Industry-specific privacy regulations include the Health Insurance Portability and Accountability Act (HIPPA) for health care services. For financial services, regulations include the Sarbanes-Oxley Act of 2002 (SOX). Other regulations include ISO 27001, FIPS, and NIST. Companies operating in Europe must abide by the European Union’s Data Protection Directive and its specific implementations by individual countries. Globally, credit card data must be protected as per the Payment Card Industry (PCI) Data Security Standard.

2-27 Conceptual Design: Constraints

Constraints are conditions that provide boundaries to the design and often get confused with requirements.

With a requirement, the architect can evaluate multiple options and make a design decision, whereas a constraint dictates the answers and removes the ability of the architect to decide.

For example, a constraint is where the design must use the existing shared storage array, whereas a requirement is where the design must use shared storage.

A straightforward approach to identifying constraints is to consider anything that limits the design choice made by the architect. If multiple options are not available to make a design decision, it is a constraint. For example, because of an existing vendor relationship, host hardware is already selected.

2-28 Conceptual Design: Assumptions

Assumptions list the conditions that are believed to be true but are not confirmed.

All assumptions should be validated before the deployment.

An example of an assumption is that enough unused capacity is available on the storage array for the new workloads.

Whenever you have an assumption, you must have an accompanying risk. Every risk requires a mitigation strategy.

Assumptions are design components that are presumed to be valid without proof. Documented assumptions should be validated during the design process so that, by the time the design is implemented, no assumptions remain.

The organization has sufficient network bandwidth between sites to facilitate replication is an example of an assumption.

If something is documented or stated without empirical proof, it is an assumption and, therefore, must be validated.

2-29 Conceptual Design: Risk and Risk Mitigation

Risks are factors that might have a negative effect on the design. Risk mitigation includes the proactive steps that you can take to accept or prevent these potential negative effects.

Mitigation Action	Description
Accept	Acknowledge that a risk impacts the project. Make an explicit decision to accept the risk without any changes to the project. Project management approval is mandatory here.
Avoid	Adjust the project scope, schedule, or constraints to minimize the effects of the risk.
Control	Take action to minimize the impact or reduce the intensification of the risk.
Transfer	Implement an organizational shift in accountability, responsibility, or authority to other stakeholders that accept the risk.
Continue monitoring	Often suitable for low-impact risks. Monitor the project environment for potentially increasing impact of the risk.

A risk is anything that might prevent the achievement of project goals. All risks should be clearly mitigated, ideally with a standard operating procedure (SOP).

No design is perfect, and it is important to document as many risks as you can identify so that you are prepared and can craft mitigation plans. Not paying close attention to risks might leave the design in a vulnerable state.

2-30 Activity: Identifying Design Categories (1)

Identify the category for each description. Is the category an assumption, constraint, requirement, or risk?

Description	Category
Having vSphere vMotion traffic and data traffic on the same physical network can lead to network disruptions if not designed carefully.	?
The design must provide a centralized management console to manage both data centers.	?
The customer provides sufficient storage for building the environment.	?
No funding exists for a new storage array and, therefore, existing storage hardware must be used.	?
The design must address security zone requirements for management, production, dev/test, and QA workloads.	?

2-31 Activity: Identifying Design Categories (2)

Identify the category for each description. Is the category an assumption, constraint, requirement, or risk?

Description	Category
Having vSphere vMotion traffic and data traffic on the same physical network can lead to network disruptions if not designed carefully.	Risk
The design must provide a centralized management console to manage both data centers.	Requirement
The customer provides sufficient storage for building the environment.	Assumption
No funding exists for a new storage array and therefore, existing storage hardware must be used.	Constraint
The design must address security zone requirements for management, production, dev/test, and QA workloads.	Requirement

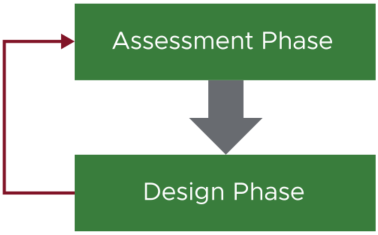
Understanding the differences between requirements, constraints, assumptions, and risks helps you to create a conceptual design that meets the organization’s business goals and requirements.

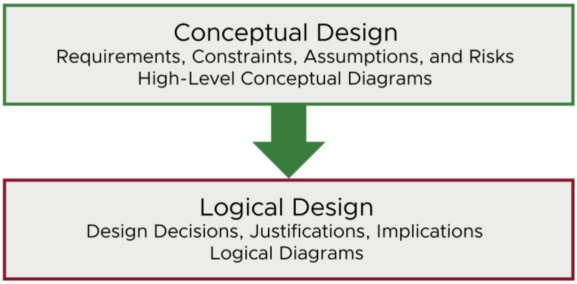
2-32 Creating the Logical Design

During the logical design, you decide how to arrange all major infrastructure components to satisfy service dependencies and requirements form the conceptual design.

The design phase is an iterative process. vSphere design decisions are made for several components:

- Management
- Clusters
- Storage
- Networking
- Virtual machines
- Security




Conceptual Design
Requirements, Constraints, Assumptions, and Risks
High-Level Conceptual Diagrams
Logical Design
Design Decisions, Justifications, Implications
Logical Diagrams

The logical design follows the conceptual design. The logical design defines how to arrange and use components and features of the infrastructure to satisfy service dependencies and other relationships specified in the conceptual design. The logical design does not provide specific details about, for example, server hardware, port connections, or Fibre Channel zones.

2-33 Design Decision Justifications

Designs are a series of compromises. When a design decision does not directly relate to a requirement, use nonfunctional requirements to evaluate and justify the decision.

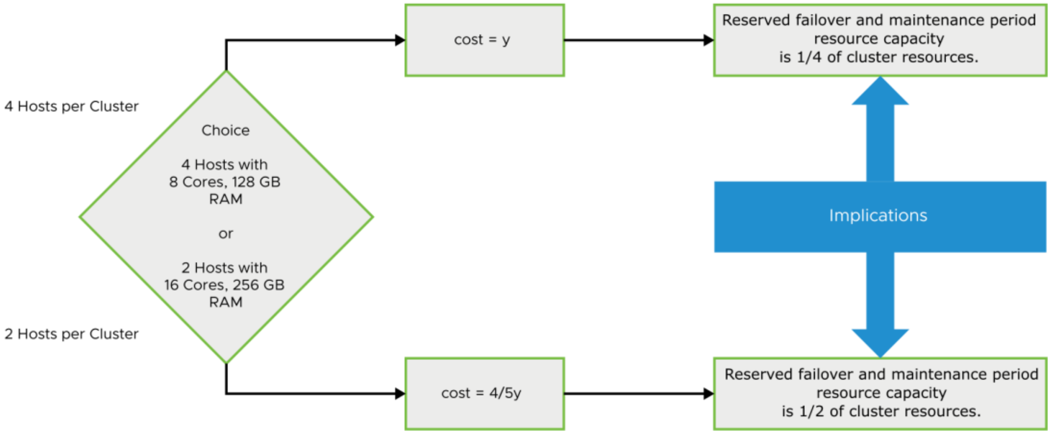
Principle	Description
Availability	How well does the solution ensure that services are available to meet business goals and requirements?
Manageability	How easily can the solution expand for future growth and ensure that enough resources exist to meet performance SLAs? How does this solution manage the life cycle of the components in the environment? Does the solution make operations simpler or more complex?
Performance	How does this solution affect infrastructure performance?
Security	Does this solution make the infrastructure more or less secure?
Recoverability	How well does this solution meet RTO and RPO requirements?

Designing is a process of balancing customer and application requirements with general VMware best practices. Often, a design choice involves a compromise. For example, an organization’s requirement to maintain hardware isolation between departments might force a reduction in the achievable consolidation ratio. The reduction in the consolidation ratio might affect both the capital and the operational costs. Likewise, an organization’s policy of doing business with a preferred hardware vendor might affect the achievable price/performance ratio. You must weigh the implications of each design choice to determine the best solution. A good method of evaluating design decisions is to weigh implications against nonfunctional design qualities.

Cost indicates the one-time and ongoing expense to implement the solution. Even with a nearly unlimited budget, a perfect design might not exist. Every organization has a budget. The design should be good enough to meet the organization’s business goals while staying within the budget. If you cannot meet the goals and stay within the budget, you must carefully explain the options so that the organization’s stakeholders can decide the course of action.

2-34 Design Decision Implications

You must understand the best practice and decide whether it can be included in the design. If a best practice is not optimal, document implications and communicate risks.



4 Hosts per Cluster
Choice
4 Hosts with
8 Cores, 128 GB
RAM
or
2 Hosts with
16 Cores, 256 GB
RAM
2 Hosts per Cluster
cost = y
cost = 4/5y
Reserved failover and maintenance period resource capacity is 1/4 of cluster resources.
Implications
Reserved failover and maintenance period resource capacity is 1/2 of cluster resources.

In the design process, multiple inputs must be simultaneously evaluated. For example, information gathered during interviews must be evaluated against current best practice information and against the current infrastructure inventory results. The evaluation might result in multiple design choices. Most choices have design implications.

The example shows the implications of choosing between two hardware platforms. Both hardware platforms support a cluster that is enabled for vSphere DRS and vSphere HA. Both hardware platform options meet the necessary CPU and memory requirements.

Purchasing the 16-core hosts is slightly less expensive. If you reserve one host in the cluster for failover and maintenance periods, you must reserve 50 percent of the available compute capacity. Purchasing the eight-core hosts is slightly more expensive. Reserving one host in the cluster for failover and maintenance results in reserving only 25 percent of the available compute capacity.

As you work with stakeholders to analyze the design alternatives, always consider and document the effects of a design choice. Consider the risks involved with the design choice and how you can mitigate those risks.

2-35 Example of a Logical Design

The logical design includes design decisions, justifications, and implications for the major infrastructure components.

Design decisions must support the requirements, constraints, assumptions, and risks that are outlined in the conceptual design.

Assign an identification number to each design decision so that the architect and stakeholders easily reference decisions using the ID no.

ID	Design Decision	Design Justification	Design Implication
DD01	Two vCenter Server instances are used.	The client has a policy to separate the development and test environment from the production environment. So, each environment has its own vCenter Server instance.	The client must buy a license for each site. However, manageability is improved.
DD02	The vCenter Server instances are in linked mode.	The client wants to manage both sites from the same interface. The sites use roles and permissions to limit access.	The vCenter Server Appliances must be configured for Enhanced Linked Mode.

The logical design is useful for understanding and evaluating the design of the infrastructure without becoming lost in the connection and configuration details of the physical design.

In addition, several physical designs can be built based on a good quality logical design.

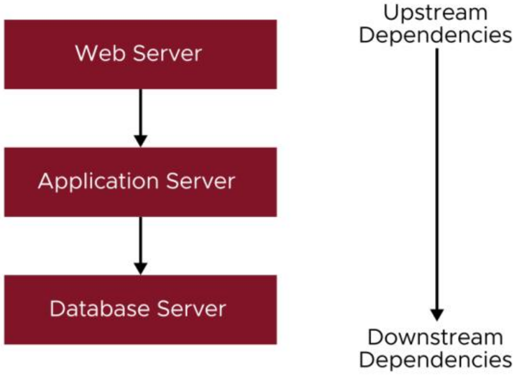
By assigning an identification number to each design decision, the architect can easily reference and understand all the requirements that are met. Peer reviews and discussions of these decisions can be easily identified in the workshops.

2-36 Service Dependencies

The logical design must include information about the dependencies for infrastructure and application services.

Service dependencies can be described with an entity relationship diagram:

- Each service in the diagram can have upstream and downstream dependencies:
  - — An upstream dependency is an entity that depends on the service.
  - — A downstream dependency is an entity on which the service depends.
- The arrow is a dependency connection that points to what the service depends on:
  - — For example, the database server depends on the vSphere layer.



An entity relationship diagram is a simple and effective way of documenting and communicating to stakeholders the various aspects of the vSphere design. The diagram shows the relationships between components, which can be useful from a design perspective and an operational perspective.

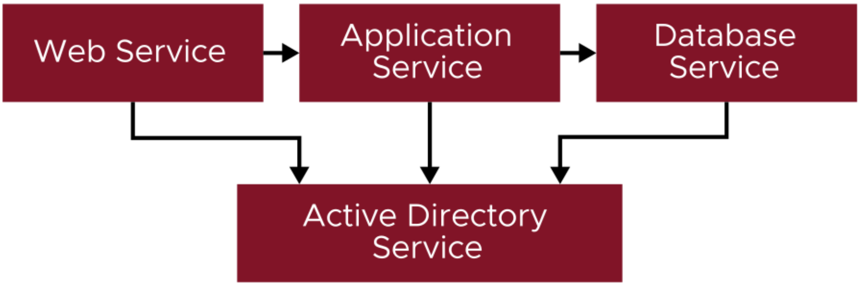
Relationships between components can be defined in the following terms:

- Runs on versus runs
- Depends on versus used by
- Contains versus contained by
- Hosts versus hosted by

2-37 Example of Service Dependencies (1)

Events occurring downstream of the service affect all the components that are upstream.

The example shows application service dependencies. If the Active Directory service fails, the web, application, and database services are impacted.

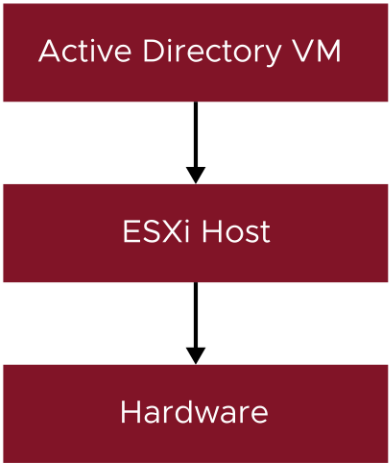


Entity relationship diagrams are used to map service dependencies. These dependency mappings are useful in several ways. You can make design decisions based on these mappings. For example, you can identify which VMs can be placed into vSphere vApps. A web server, application server, and database server might be grouped into a single, logical unit and therefore be placed into a vApp.

These diagrams can also help to identify the root causes of issues. For example, if the application service is not responding, you can identify the services or components that the application service depends on and determine whether one or more of these dependencies is the root cause.

2-38 Example of Service Dependencies (2)

For infrastructure service dependencies, when hardware fails the ESXi host is affected, which in turn affects the Active Directory VM.

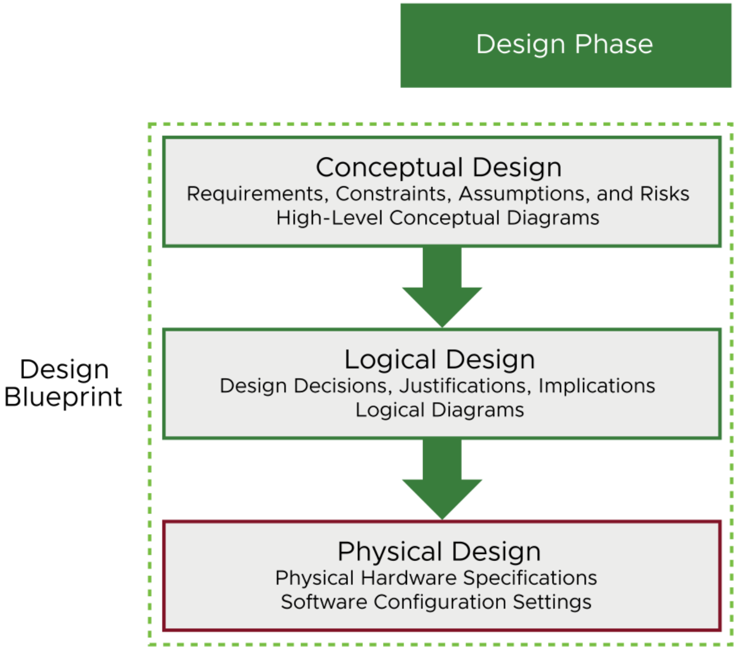


An ESXi host can also have a dependency on the Active Directory service.

You can use entity relationship diagrams to map entities in your design. The entities can be application services, infrastructure components, hardware components, and so on.

2-39 Creating the Physical Design

The physical design provides the detailed specifications for purchasing hardware and ultimately deploying the solution.


Design Phase
Design Blueprint
Conceptual Design
Requirements, Constraints, Assumptions, and Risks
High-Level Conceptual Diagrams
Logical Design
Design Decisions, Justifications, Implications
Logical Diagrams
Physical Design
Physical Hardware Specifications
Software Configuration Settings

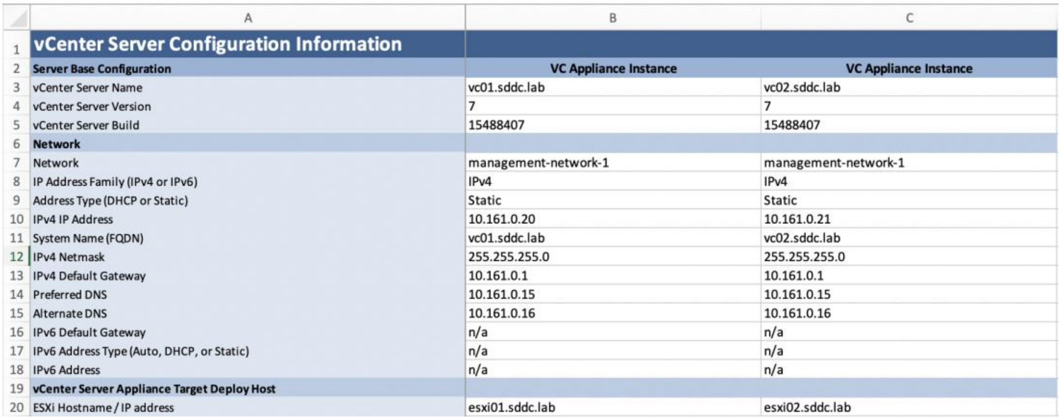
The physical design is the last design. The physical design is based on the logical design. The physical design includes specific hardware from specific vendors. This design also lists specific configurations for each of the components that are deployed.

2-40 Example of a Physical Design

The physical design describes implementation details based on the logical design.

More than one physical design can be built based on a good quality logical design.

Physical specifications can be recorded in a spreadsheet.



	A	B	C
1	**vCenter Server Configuration Information**		
2	**Server Base Configuration**	**VC Appliance Instance**	**VC Appliance Instance**
3	vCenter Server Name	vc01.sddc.lab	vc02.sddc.lab
4	vCenter Server Version	7	7
5	vCenter Server Build	15488407	15488407
6	**Network**		
7	Network	management-network-1	management-network-1
8	IP Address Family (IPv4 or IPv6)	IPv4	IPv4
9	Address Type (DHCP or Static)	Static	Static
10	IPv4 IP Address	10.161.0.20	10.161.0.21
11	System Name (FQDN)	vc01.sddc.lab	vc02.sddc.lab
12	IPv4 Netmask	255.255.255.0	255.255.255.0
13	IPv4 Default Gateway	10.161.0.1	10.161.0.1
14	Preferred DNS	10.161.0.15	10.161.0.15
15	Alternate DNS	10.161.0.16	10.161.0.16
16	IPv6 Default Gateway	n/a	n/a
17	IPv6 Address Type (Auto, DHCP, or Static)	n/a	n/a
18	IPv6 Address	n/a	n/a
19	**vCenter Server Appliance Target Deploy Host**		
20	ESXi Hostname / IP address	esxi01.sddc.lab	esxi02.sddc.lab

Although the logical design is useful for understanding and evaluating the design of the infrastructure, the physical design is more useful for purchasing hardware and building the infrastructure. For example, you can name specific host hardware, define PCI slot assignments, specify specific port connections, provide host name and IP address information, and choose operating systems.

2-41 Lab 2: Creating a Conceptual Design

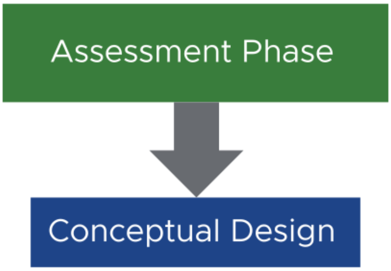
Create a conceptual design using the business objectives defined in the case study:

1. Create the Conceptual Design

2-42 Review of Lab

The instructor facilitates a class discussion to answer the following questions:

- What are the functional business requirements?
- What are the nonfunctional business requirements?
  - Did you consider availability, manageability, performance, recoverability, and security (AMPRS)?
- What are the assumptions?
- What are the risks?
- What are the constraints?



2-43 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Document design requirements, constraints, assumptions, and risks
- Use a systematic method to evaluate and document design decisions
- Create a conceptual design

2-44 **Lesson 3: Overview of Architecture Frameworks**

2-45 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the Open Group Architecture Framework principles
- Describe VMware Validated Design architecture principles
- Describe the VMware Cloud Foundation architecture principles

2-46 About TOGAF

The Open Group Architecture Framework (TOGAF) standard is a proven enterprise architecture methodology and framework that is used by the world's leading organizations to improve business efficiency. TOGAF has the following key features:

- It is a standard approach for helping with the acceptance, production, use, and maintenance of enterprise architectures.
- It is based on an iterative process model that is supported by best practices and a reusable set of existing architectural assets.
- The TOGAF standard can be used for developing a broad range of different enterprise architectures.
- The TOGAF Architecture Development Method (ADM) is used for developing an enterprise architecture that addresses business needs.

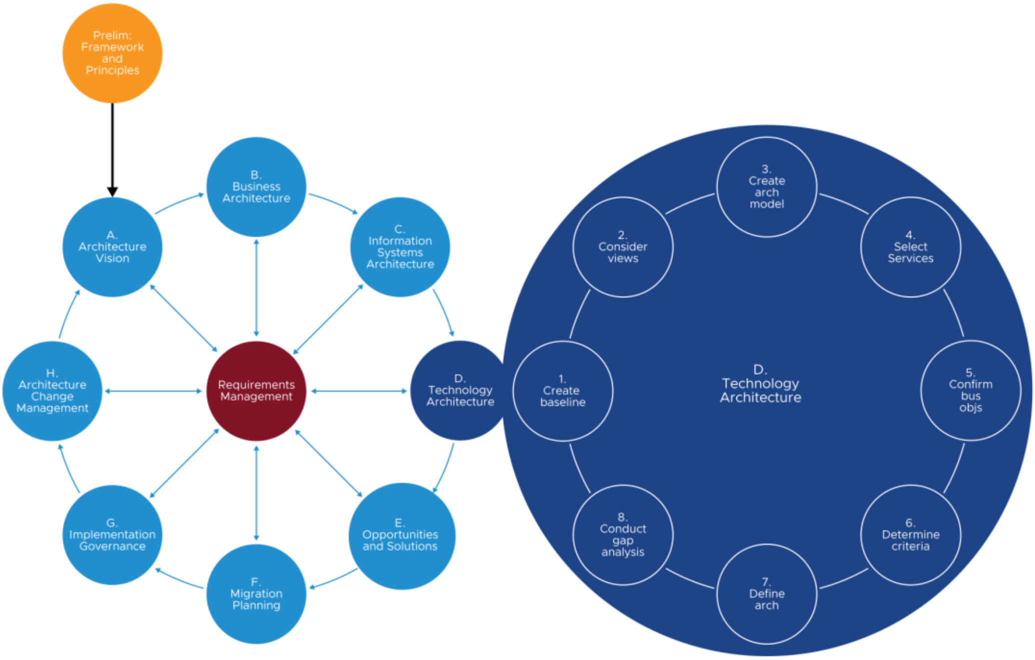
The Open Group Architecture Framework (TOGAF) standard is developed and maintained by members of the Open Group, working in the Architecture Forum. The original development of TOGAF version 1 in 1995 was based on the US Department of Defense Technical Architecture Framework for Information Management (TAFIM).

Accompanying the standard is the TOGAF Library. The TOGAF Library is a reference library containing guidelines, templates, patterns, and other forms of reference material to accelerate the creation of new architectures for the enterprise.

TOGAF complements and can be used with other frameworks that are more focused on specific deliverables for particular vertical sectors such as government, telecommunications, manufacturing, defense, and finance.

2-47 Architecture Development Method

ADM describes a method for developing an enterprise architecture and forms the core of TOGAF.



The TOGAF ADM is the result of continuous contributions from many architecture practitioners. The phases of the ADM cycle shown in Architecture Development Cycle are further divided into steps. It integrates elements of TOGAF and other available architectural assets, to meet the business and IT needs of an organization.

The ADM works in the following ways:

- The ADM is iterative, over the whole process, between phases, and within phases. For each iteration of the ADM, a fresh decision must be taken on several types of issues:
  - The breadth of coverage of the enterprise to be defined
  - The level of detail to be defined

  - The extent of the time horizon aimed at, including the number and extent of any intermediate time horizons

  - The architectural assets to be used in the organization's Enterprise Continuum, including:

    - Assets created in previous iterations of the ADM cycle within the enterprise

    - Assets available elsewhere in the industry (other frameworks, systems models, vertical industry models, and so on)

- The decisions must be made based on a practical assessment of resource and competence availability, and the value that can realistically be expected to accrue to the enterprise from the scope of the architecture work.

- As a generic method, the ADM is used by enterprises in a wide variety of different geographies and applied in different vertical sectors or industry types. As such, it can be, but does not necessarily have to be, tailored to specific needs. For example:

  - It can be used with the set of deliverables of another framework, where these frameworks are deemed to be more appropriate for a specific organization. (For example, many US federal agencies have developed individual frameworks that define the deliverables specific to their particular departmental needs.)

  - It can be used with the well-known Zachman Framework, which is an excellent classification scheme, but lacks an openly available, well-defined methodology.

2-48 About VMware Validated Design Principles (1)

VMware Validated Design includes prescriptive blueprints with comprehensive deployment and operational practices and has the following advantages:

- Complete data center level design
- Standardized
- Proven and robust
- Applicable to broad use cases

VMware Validated Design is a family of solutions for data center designs that span compute, storage, networking, and management, serving as a blueprint for your software-defined data center (SDDC) implementation. The documentation of VMware Validated Design consists of succeeding deliverables for all stages of the SDDC life cycle.

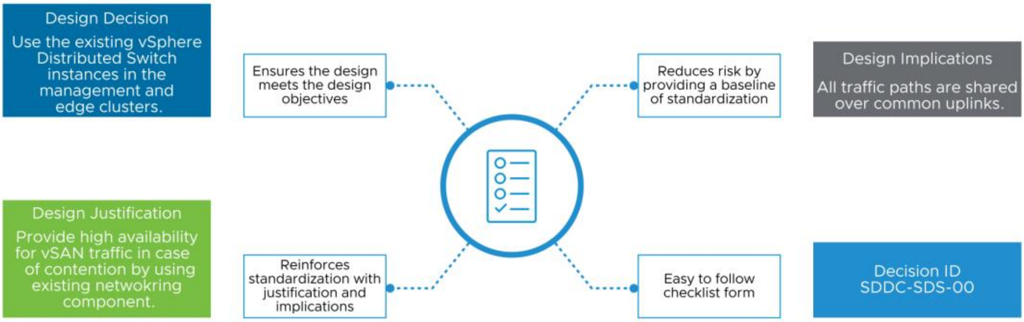
VMware Validated Design supports three SDDC implementations: Standard SDDC, remote office/branch office (ROBO) SDDC, and consolidated SDDC.

After you deploy the SDDC, VMware Validated Design provides several types of postdeployment guidance:

- Operational guidance
- Security and compliance guidance
- Scenarios for using the system according to an IT use case, such as workload provisioning and intelligent monitoring
- Guidance on adding workload domains with an architecture for VMware software that is additional to the list of required products.

You can also use technical notes to adjust your system to a certain network and data center setup.

2-49 About VMware Validated Design Principles (2)



Design Decision

Use the existing vSphere Distributed Switch instances in the management and edge clusters.

Ensures the design meets the design objectives

Reduces risk by providing a baseline of standardization

Design Implications

All traffic paths are shared over common uplinks.

Design Justification

Provide high availability for vSAN traffic in case of contention by using existing netwokring component.

Reinforces standardization with justification and implications

Easy to follow checklist form

Decision ID SDDC-SDS-00

VMware Validated Design for SDDC is tested by VMware to ensure that all components and their individual versions work together, scale, and perform as expected. Unlike reference architectures, which focus on an individual product or purpose, VMware Validated Design takes a holistic approach to design, encompassing many products in a full stack for a broad set of use case scenarios in an SDDC.

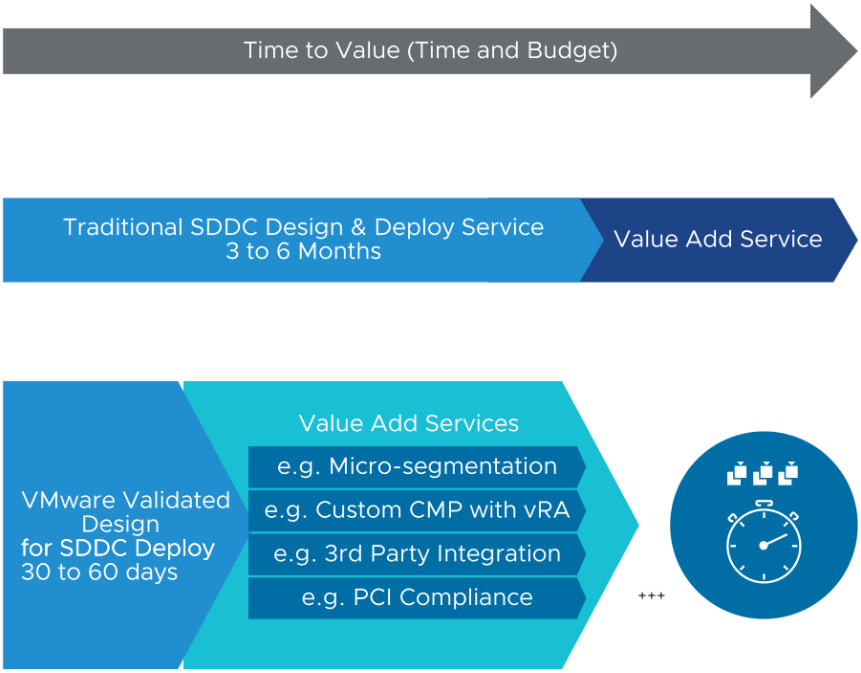
The VMware Validated Design Architecture and Design document contains a validated model of the SDDC and provides a detailed design of each management component of the SDDC stack.

VMware Validated Design is based on a set of individual VMware products with different versions, which are included in the VMware Customer Experience Improvement Program (CEIP). This program provides information that can be used to improve VMware products and services, fix problems, and advise on how best to deploy and use VMware products.

For more information on VMware Validated Design, see https://www.vmware.com/support/pubs/vmware-validated-design-pubs.html.

2-50 VMware Validated Design Benefits

VMware Validated Design accelerates the time to value by reducing the deployment time and budget.



Typically, the two factors that determine what services can be delivered are available time and budget. In a VMware-based SDDC, engagements for simply standing up the SDDC impact heavily on both factors.

Usually an SDDC design and deploy service, which can take 4 to 6 months to complete, might not leave much time or money for additional services that the customer considers of high value. For example, you might only be able to add a micro-segmentation service or something similar. Even if an extra service is added within the time and budget constraints, parallelizing other services is difficult before the SDDC itself is deployed.

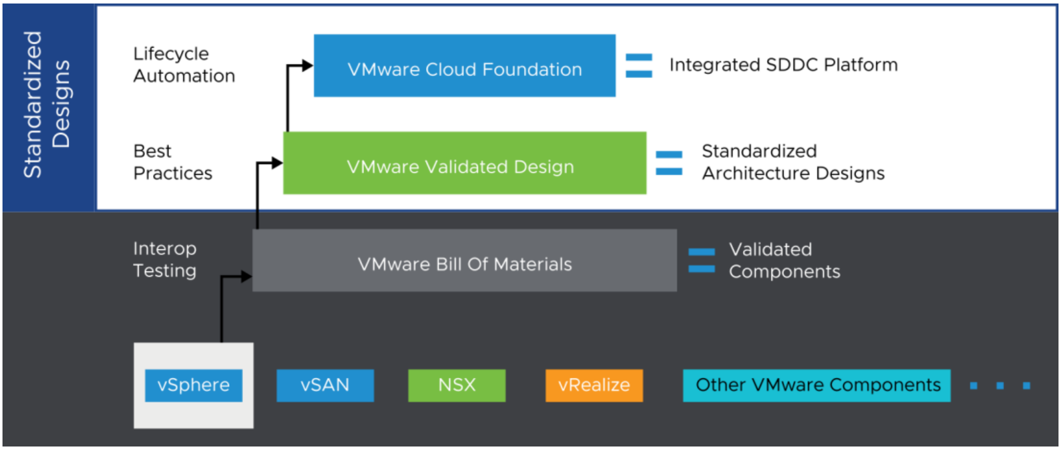
Delivery of any additional services are a great outcome for the customer. However, they must acquire software licenses for products that provide more capabilities, which might be delivered by the use of VMware Validated Design to achieve a huge increase on project acceleration.

VMware Validated Design 5.0 (based on vSphere 6.7) introduced a level of deployment automation in the form of VMware Cloud Builder 1.0. The VMware Cloud Builder appliance provides a streamlined automated deployment for the SDDC. It can be used to deploy a single region, a dual region, or a consolidated SDDC. In earlier releases, VMware Cloud Builder and its predecessor the Deployment Tool Kit (DTK) were only available to the VMware Professional Services Organization (PSO) and select partners.

The release of VMware Validated Design 5.0 with VMware Cloud Builder to all VMware SDDC customers further simplifies and standardizes the implementation of VMware deployments. The VMware Validated Design deployment guides are updated to incorporate the use of VMware Cloud Builder 1.0. The time to completion can be reduced significantly if predeployment requirements are met completely and efficiently. This compliance accelerates time to value and time to market.

2-51 VMware Approach to SDDC

VMware Validated Design can be used independently or as part of VMware Cloud Foundation.

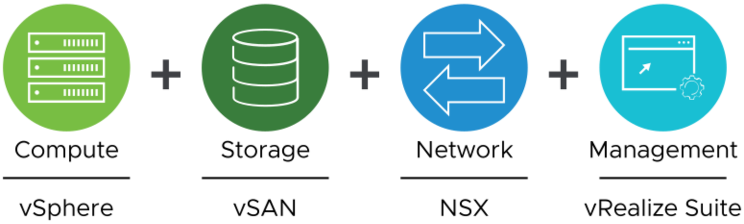


Extensive interoperability testing ensures that software components are validated as working together as a stack. This testing eliminates the need for customers to continually check the version interoperability of software components. They can focus on getting value from their features. Further, with VMware Cloud Foundation, day 2 operations are automated, so the update of the stack is performed following orchestrated automation that ensures the sequencing of the upgrade while maintaining interoperability.

2-52 VMware Cloud Foundation Platform

VMware Cloud Foundation is the unified SDDC platform that brings together VMware's vSphere, vSAN, and NSX into a natively integrated stack to deliver an enterprise-ready cloud infrastructure for the private and public cloud:

- VMware Cloud Foundation is built on VMware Validated Design for the SDDC.
- VMware Cloud Foundation provides automated deployment and life cycle management of the full SDDC.



VMware Cloud Foundation does not change or limit the capabilities of the underlying software building blocks of the SDDC. Customers can take advantage of all their features and capabilities. VMware Cloud Foundation provides an easy way to deploy and maintain them, freeing up the customer for more valuable work in their day jobs.

VMware Cloud Foundation also supports automated life cycle management, which includes nondisruptive patching and upgrading of its core software building blocks vSphere, vCenter, vSAN, and NSX (both NSX-V and NSX-T Data Center) at a per cluster level.

VMware Cloud Foundation can also be used to automate the deployment, expansion, patching, and upgrades for vRealize Lifecycle Manager, vRealize Log Insight, vRealize Operations, and vRealize Automation.

VMware Cloud Foundation can also automate the (optional) deployment of VMware Horizon virtual desktop infrastructure (VDI) in accordance with the VMware Horizon pod and block best practice. This deployment of VMware Horizon can be deployed to an NSX-V supported workload domain.

The automated deployment of VMware Enterprise PKS, the VMware SDDC-supported platform for cloud native Kubernetes container management, is also an option that can be deployed to an NSX-T Data Center supported workload domain.

2-53 VMware Cloud Foundation OEM-Integrated Systems

VMware Cloud Foundation can be deployed as an OEM-integrated system.
OEM-integrated solutions are of the following types:

- Jointly engineered solutions
  - VMware Cloud Foundation/Dell EMC VxRail
- Composable
  - VMware Cloud Foundation/HPE Synergy
  - VMware Cloud Foundation/Dell MX
- Integrated systems
  - Fujitsu PRIMEFLEX for VMware Cloud Foundation
  - Hitachi Unified Compute Platform (UCP) RS
  - QCT QxStack

All integrated systems are delivered ready to install at the customer site.

Jointly engineered systems, such as VMware Cloud Foundation on VxRail, provide integration with VMware Cloud Foundation components. The integration uses and extends existing VxRail HCI integrated system features and operations processes to VMware Cloud Foundation, including, but not limited to, life cycle management of the hardware and software subsystems. The integration is achieved using native SDDC Manager orchestrated workflows that are seamlessly integrated with VxRail Manager.

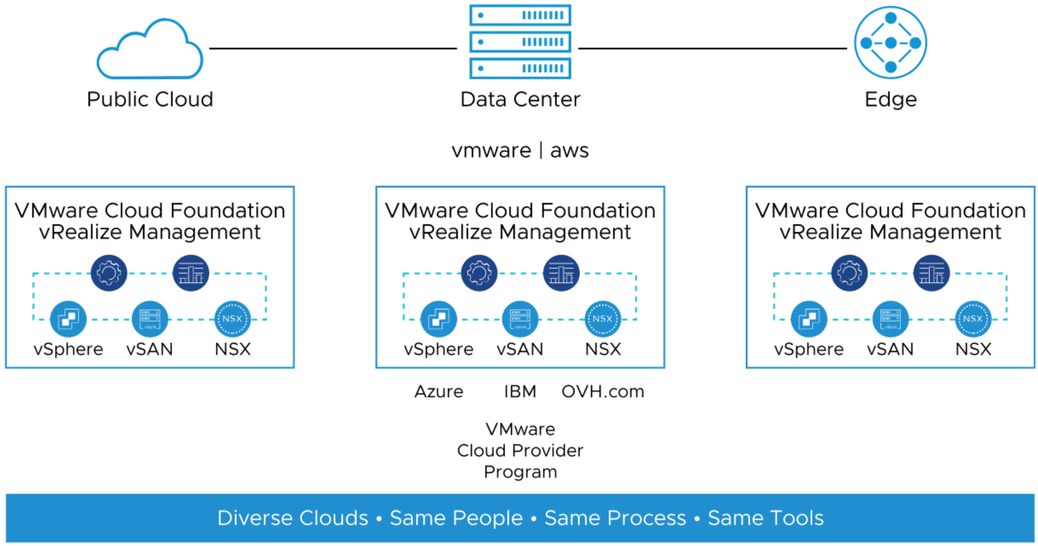
Composable systems, such as Dell MX and HPE Synergy, integrate with VMware Cloud Foundation through the Redfish API. In this way, hardware resources under the control of VMware Cloud Foundation can be composed and decomposed.

Similar to jointly engineered solutions, integrated systems are preassembled and imaged at the factory and arrive ready to run.

2-54 VMware Cloud Foundation Deployment

VMware Cloud Foundation is not limited to on-premises, owned, hardware:

- VMware Cloud on Dell EMC is a subscription-based hardware IaaS built on VMware Cloud Foundation.
- VMware Cloud on AWS is a subscription-based VMware Cloud Foundation deployment inside AWS regions.
- VMware works with several third-party cloud providers to deliver infrastructure services, based on VMware Cloud Foundation, through a subscription model.



VMware Cloud Foundation drives simplification so that you can support applications and data across all your different environments with the followings benefits:

- Offers a standardized, repeatable, and consistent meaning
- Contains the same common heterogenous components that run from commercial hardware across an on-premises, edge, and broad hybrid cloud eco-system
- Includes many leading public clouds and cloud service providers

The SDDC provides intrinsic and intelligent security in every component from the hypervisor to the storage, networking, and management layers.

With these capabilities, organizations can concentrate on building higher value applications and services for consumption with their organizations and by their customers, driving innovation.

2-55 Activity: Framework Comparison

As a group, consider the different frameworks and complete the table.

	Bespoke Design	VMware Validated Design	VCF On-Premises	VCF Jointly Engineered	VCF VMware Cloud Provider Partner	VMware Cloud on AWS
Appropriate Use Case?						
Level of Effort						
Cost Model						
Advantages						
Disadvantages						

The activity does not have one correct answer, but you can align requirements with the frameworks.

Consider the requirements for the following organizations. Are any of the frameworks suited to these requirements?

- Org A: Requires discrete self-contained units of compute and storage at their edge locations.
- Org B: Requires a design for a high-performance compute infrastructure that ingests terabytes of data from an existing on-premises data warehouse and processes it using Field Programmable Gate Array (FPGA) hardware.
- Org C: Requires modern infrastructure so that the organization can evacuate one of its current data centers to avoid being locked into a new building lease.
- Org D: Requires a temporary expansion of its VM estate to accommodate 2,000 additional remote workers who need remote virtual desktops.
- Org E: Requires a new infrastructure on premises, but in a short timescale, with the constraint that it must use the existing shared storage platform.

- Org F: Requires a ready-made solution to deploy on premises and be productive as quickly as possible. The organization has limited resources to operate the platform for routine maintenance tasks.
- Org G: Requires a modern infrastructure platform that is externally hosted but does not lock the organization into any one particular cloud vendor.

In this activity, the level of effort is intended as a simple indicator of approximate design and implementation resource effort.

Categorize the frameworks using the following descriptions:

- High (6 months or more)
- Medium (3 months)
- Low (<1 month)

For the cost model, categorize the architecture frameworks as capital or operational expenditure, or both.

Record advantages or disadvantages that are identified.

2-56 Activity: Framework Comparison Implementation

The activity does not have one correct answer, but you can align requirements with the frameworks.

	Bespoke Design	**VMware Validated Design**	**VCF On-Premises**	**VCF Jointly Engineered**	**VCF VMware Cloud Provider Partner**	**VMware Cloud on AWS**
Appropriate Use Case?	Org B	Org E	Org A	Org F	Org C and Org G	Org D
Level of Effort	High	Medium	Medium/Low	Low	Low	Low
Cost Model	Capex	Capex	Capex or Opex	Capex or Opex	Opex	Opex

2-57 Activity: Framework Comparison of Advantages and Disadvantages

	Bespoke Design	VMware Validated Design	VCF On-Premises	VCF Jointly Engineered	VCF Cloud Provider Partner	VMware Cloud on AWS
Advantage	Meets specific app/ processing	Accelerates time to value without losing flexibility	Short time to value of full SDDC stack	Short time to value of full SDDC stack with lifecycle management built in	Simplest VM migration (hot or cold) and enables network extensions	Simplest VM migration (hot or cold) and enables network extensions while locating near AWS native services
Disadvantage	Resource-intensive and long time to value	Generic template does not cover all use cases	Specific to vSAN vendor offerings	Each joint engineered solution specific to one hardware vendor	Best pricing through term commitments	

2-58 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Describe the Open Group Architecture Framework principles
- Describe VMware Validated Design architecture principles
- Describe the VMware Cloud Foundation architecture principles

2-59 Key Points

- At the beginning of a design project, you must develop a high-level vision for the project, which includes the project scope, goals, requirements, assumptions, and constraints.
- The conceptual design focuses on achieving the organization's business objectives and requirements.
- The logical design includes the relationships between all major infrastructure components. It is useful for understanding and evaluating the design of the infrastructure.
- A physical design includes specific vendor and implementation details.
- Design criteria includes scalability, availability, manageability, performance, recoverability, security, and cost.
- Educate key stakeholders and SMEs about vSphere and virtualization so that they can provide valuable input during the design project.
- Designing the vSphere infrastructure is a balancing act between technical best practices and the organization's goals, requirements, and constraints.
- VMware Validated Design and VMware Cloud Foundation frameworks accelerate the time to value on implementation projects.

Questions?

Module 3

Core Management Infrastructure

3-2 Importance

vCenter Server provides essential data center services such as resource allocation, access control, and performance monitoring.

Designing the correct core management architecture is critical because the design has a significant effect on the manageability and scalability of the virtual infrastructure.

3-3 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Determine the number of vCenter Server instances to include in a design
- Design a vCenter Server deployment topology that is appropriate for the size and requirements of a data center

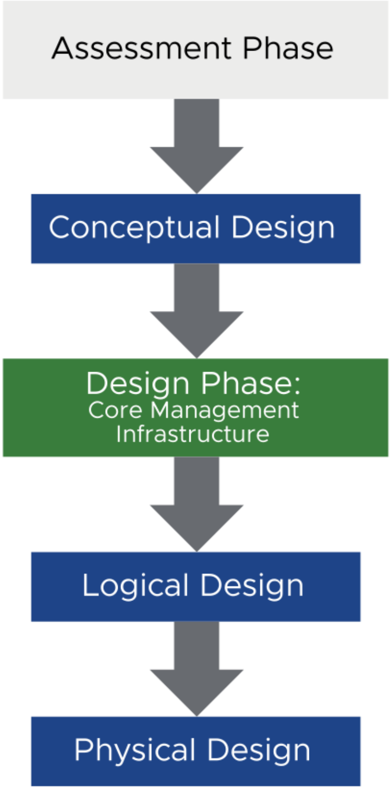
3-4 Designing the Core Management Infrastructure

After you complete the assessment phase and create a conceptual design for your vSphere infrastructure, you can begin the design phase.

Design decisions for the core management infrastructure include determining specifications such as:

- Number of vCenter appliance instances
- High availability

After the logical design is complete, you create the physical design from the logical design specifications.



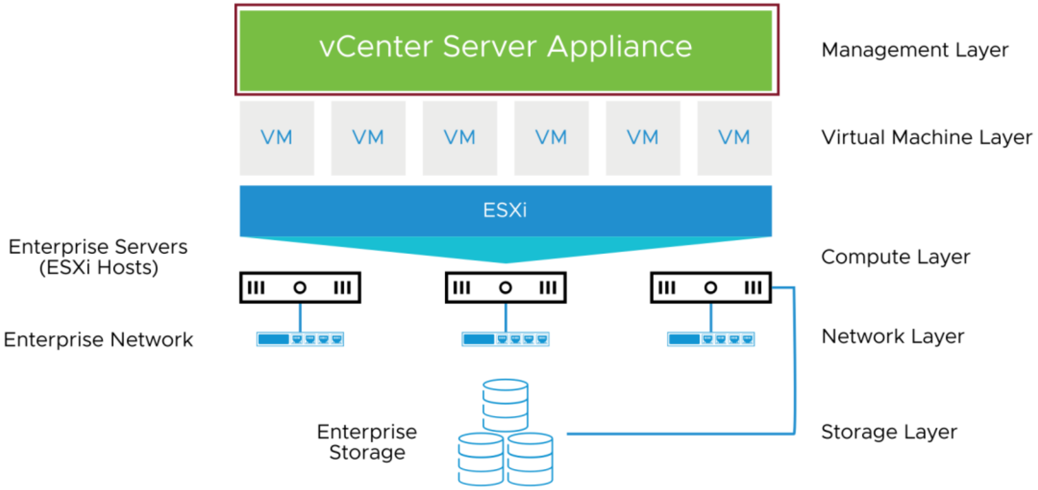
Using the information gathered in the assessment phase, you can begin the design work, starting with the logical design. When creating the logical design, you consider each of the major vSphere infrastructure components and make the appropriate design decisions. You always consider the organization's goals, requirements, constraints, and current best practices.

The design phase is an iterative process. It involves a series of individual and group interviews with stakeholders and some individual work by the designer.

Customer involvement during all phases of the process is key to the success of the project. During the assess and design phases, all the stakeholders must understand vSphere and virtualization concepts so that they can provide appropriate inputs during the design process.

3-5 Management Layer in a vSphere Infrastructure

vCenter Server provides the core services and interfaces for managing and monitoring the virtual infrastructure.

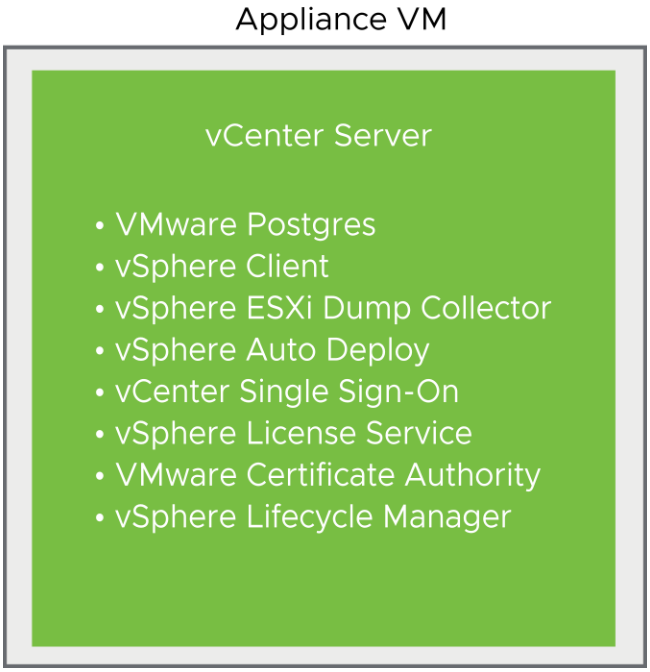


The management layer centers on vCenter Server and consists of several services and components. vCenter Server components provide the services and interfaces for the management and monitoring of all components in the virtual infrastructure.

vCenter Server also unifies the resources from the individual computing servers to be shared among VMs in the entire data center. vCenter Server manages the assignment of VMs to the computing servers and the assignment of resources to the VMs in a given computing server based on the policies that the system administrator sets.

3-6 Review of vCenter Server Components

vCenter Server Appliance provides all the services for a virtual data center.



Appliance VM

vCenter Server

- VMware Postgres
- vSphere Client
- vSphere ESXi Dump Collector
- vSphere Auto Deploy
- vCenter Single Sign-On
- vSphere License Service
- VMware Certificate Authority
- vSphere Lifecycle Manager

When deploying vCenter Server 7.0, only a single supported architecture is used, namely, vCenter Server Appliance.

For more information about each component of vCenter Server, see *vCenter Server and Host Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.vcenterhost.doc/GUID-3B5AF2B1-C534-4426-B97A-D14019A8010F.html.

3-7 vCenter Server Resource Requirements

Hardware requirements for vCenter Server Appliance depend on the size of the vSphere inventory.

Resource	Tiny	Small	Medium	Large	X-Large
	(Up to 10 hosts, 100 VMs)	**(Up to 100 hosts, 1,000 VMs)**	**(Up to 400 hosts, 4,000 VMs)**	**(Up to 1,000 hosts, 10,000 VMs)**	**(Up to 2,000 hosts, 35,000 VMs)**
CPU	2	4	8	16	24
Memory	12 GB	19 GB	28 GB	37 GB	56 GB
Default Storage Sizes	415 GB	480 GB	700 GB	1,065 GB	1,805 GB

For the complete list of system requirements for vCenter Server Appliance, see *vCenter Server Upgrade* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vcenter.upgrade.doc/GUID-9ED7B32A-019F-4A97-BC58-1A9BF7D16C57.html.

3-8 vCenter Server Appliance Features

vCenter Server Appliance has new and innovative features:

- vCenter Server High Availability: A native high availability solution that protects against host and hardware failures, and vCenter Server application failures
- Improved vCenter Server Appliance Management Interface:
  - Resource usage statistics and health data
  - Built-in backup and restore solution
- vSphere Lifecycle Manager for vCenter Server Appliance: Centralized and simplified life cycle management for ESXi hosts
- vCenter Server Appliance migration tool: Single-step migration process for migrating an existing Windows vCenter Server system to vCenter Server Appliance

The installer has a built-in migration tool, providing easy access to vCenter Server Appliance.

With vSphere Lifecycle Manager, you manage ESXi hosts by using either baselines or images. An image is a precise description of the software, drivers, and firmware to run on a host. You set up a single image and apply it to all hosts in a cluster to ensure cluster-wide host image homogeneity.

The vCenter Server Appliance Management Interface exposes configuration data. In addition to CPU and memory statistics, it shows network and database statistics, disk space usage, and health data. This visibility of data reduces dependency on a command-line interface for simple monitoring and operational tasks.

3-9 Design Decision: Determining the Number of vCenter Server Instances

How many vCenter Server instances should you deploy for the vSphere infrastructure?

Consider the following factors:

- Infrastructure size: Hosts and virtual machines cannot exceed the management limits of a single vCenter Server system.
- Requirements of other products: For example, Site Recovery Manager requires a vCenter Server system at both the protected and recovery sites.
- Number of sites: Remote offices might require their own vCenter Server system, and you must configure multiple instances.

Some infrastructures require more than one vCenter Server instance.

For designs that include more than one site, consider the following determining factors:

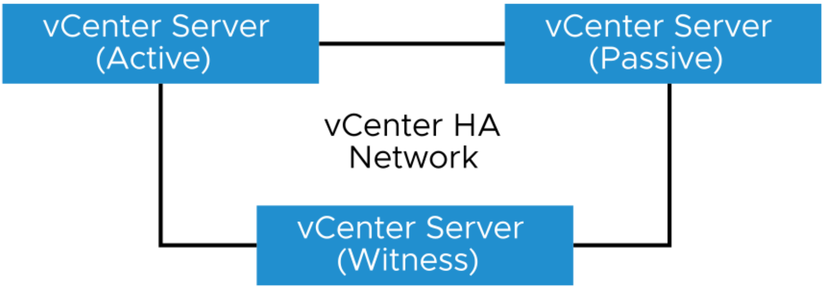
- Management policies
- Security requirements
- Availability requirements
- Size of the remote office deployments
- Type of servers at the remote office: Virtualized servers or user desktops
- Available network bandwidth and latency
- Availability of trained staff at the remote office

For more information about the current configuration maximums, see *Configuration Maximums* at https://configmax.vmware.com/.

3-10 vCenter Server High Availability Considerations

When you use the vCenter Server High Availability feature, the following guidelines apply:

- No shared storage is required.
- Supported vCenter Server and ESXi versions are:
  - ESXi 6.0 or later
  - vCenter Server Appliance 6.5 or later
- The vCenter Server Appliance deployment minimum size must be small but large enough to meet the recovery time objective (RTO).
- Management vCenter Server (if used) must be vCenter Server 6.5 or later.
- Network latency between the active, passive, and witness nodes must be less than 10 milliseconds.
- The vCenter Server High Availability network must be on a different subnet than the management network.



vCenter Server High Availability architecture does not require shared storage, and, as a result, vCenter Server High Availability nodes can be deployed in different geographical areas, if good network connectivity is available.

Your environment can include a management vCenter Server system, or you can set up vCenter Server Appliance to manage the ESXi host on which it runs (self-managed vCenter Server).

Do not select tiny as the vCenter Server Appliance deployment size in production environments.

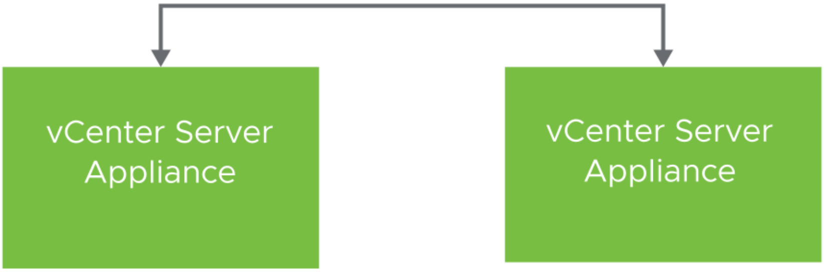
vCenter Server High Availability is supported and tested with vSphere VMFS, NFS, and vSAN datastores.

3-11 Design Decision: Linking Multiple vCenter Server Instances

Should you use Enhanced Linked Mode?

Enhanced Linked Mode is the only way to connect multiple vCenter Server systems. You can use it with or without vSphere HA to deliver the following benefits:

- View and search across all linked systems from a single point.
- View and search the inventories of all linked vCenter Server instances in the vSphere Client.
- Replicate roles, permissions, licenses, policies, and tags.



The design of vCenter Server 7 Enhanced Linked Mode differs significantly from previous vCenter Server versions, and linking is now supported only with the vCenter Server Appliance.

3-12 vCenter Server System Prerequisites

The vCenter Server system has the following infrastructure prerequisites:

- The vCenter Server appliances require a minimum of 2 vCPUs, 12 GB RAM, and 415 GB (10 hosts/100 VMs).
- The vCenter must be able to send data to every managed host and receive data from the vSphere Client
- Static IP address or a DHCP assigned address is supported.
- DNS resolution of the involved hosts must be tested and validated.
  - — The appliance installer also relies on reverse lookup (PTR DNS records).
- Time synchronization (NTP) across all hosts must be tested and validated.

Storage requirements include vSphere Lifecycle Manager. DNS resolution must be working for all system names through fully qualified domain name (FQDN), short name (host name), and IP address (reverse lookup). Time must be synchronized across the environment. You can set up one or more Network Time Protocol (NTP) servers in the vCenter Server configuration.

3-13 Design Decision: Choosing a Single Sign-On Identity Source

Which single sign-on identity provider and source do you use?

vCenter Single Sign-On is an authentication broker and security token exchange service. It provides authentication using one of these providers:

- vCenter Server built-in identity provider: Supports local accounts (vsphere.local), Active Directory (AD), or OpenLDAP
- Identity provider federated authentication: Supports AD Federation Services (ADFS)

  The federated option is new to vSphere 7.

  Use of federated authentication replaces the vCenter Server built-in identity provider.

With the built-in vCenter Server identity provider, you use identity sources to attach one or more domains to vCenter Single Sign-On. A domain is a repository for users and groups that the vCenter Single Sign-On server can use for user authentication. An identity source is a collection of user and group data.

You specify the most common domain as the default domain so that users do not have to qualify their login with a domain when they log in.

After the vCenter Server Appliance is deployed, appropriate identity sources must be configured. Every instance of vCenter Single Sign-On has an internal identity source, for example, `vsphere.local`.

The following types of identity sources can be used with the built-in identity provider in addition to the vCenter Single Sign-On authentication domain (vsphere.local by default):

- Active Directory (integrated Microsoft Windows authentication): This identity source is recommended for use with an Active Directory (AD) environment. This identity source resolves most issues in vSphere 5.x environments because it uses built-in Microsoft Windows API calls to authenticate across forests and domains.
- Active Directory as the LDAP server: This identity source is the legacy mode for vCenter Single Sign-On 6.0. This setup reverts the configuration to the configuration that is used in vCenter Single Sign-On 5.1 and uses Lightweight Directory Access Protocol (LDAP) binds to connect to the directory.
- OpenLDAP: You use this identity source for connectivity to OpenLDAP directories.
- Local OS: This identity source includes local operating system users. Using local operating system users is not recommended.

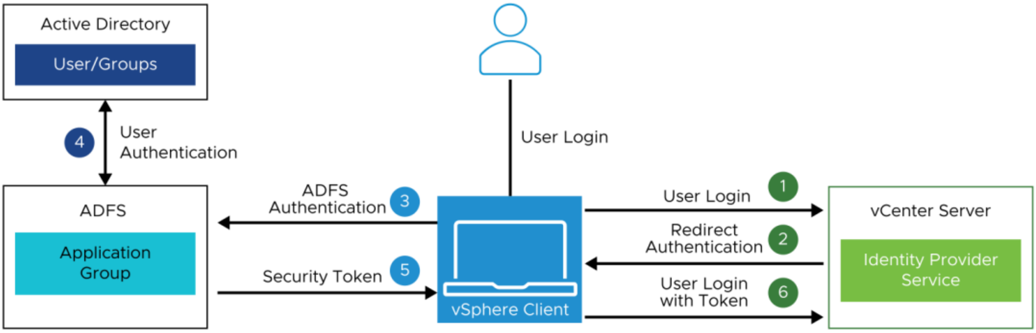
When a user logs in to a vCenter Server system from the vSphere Client, the login behavior depends on whether the user is in the default domain. The default domain is the domain that is set as the default identity source:

- Users who are in the default domain can log in with their user name and password.
- Users who are in a domain that is added to vCenter Single Sign-On as an identity source, but is not the default domain, can log in to vCenter Server. These users must specify the domain in one of the following ways:
  - Include a domain name prefix, for example, MYDOMAIN\user1.
  - Include the domain, for example, user1@mydomain.com.

3-14 Identity Provider Federated Authentication

A vCenter identity provider federation configuration consists of the following components:

- vCenter Server
- An identity provider service configured on vCenter Server
- An ADFS server and associated AD domain
- An ADFS application group
- AD groups and users that map to vCenter Server groups and users



With vCenter Server identity provider federation, you can configure an external identity provider for federated authentication. In this configuration, the external identity provider interacts with the identity source on behalf of vCenter Server.

Starting in vSphere 7, vCenter Server supports federated authentication. In this scenario, when a user logs in to vCenter Server, it redirects the user login to the external identity provider. The user credentials are no longer provided to vCenter Server directly. Instead, the user provides credentials to the external identity provider. vCenter Server trusts the external identity provider to perform the authentication. In the federation model, users never provide credentials directly to any service or application but only to the identity provider. As a result, you federate your applications and services, such as vCenter Server, with your identity provider.

vCenter Server, ADFS, and AD interact as follows:

2. The user logs in to vCenter Server on the usual vCenter Server landing page.
3. vCenter Server redirects the authentication request to ADFS.
4. If needed, ADFS prompts the user to log in with AD credentials.

5. ADFS authenticates the user with AD.
6. ADFS issues a security token with group information from AD.
7. vCenter Server uses the token to log in the user.

If the external identity provider is unreachable, the login process returns to the vCenter Server landing page, showing an appropriate information message. Administrative users can still log in using their local accounts (local OS or vsphere.local identity sources).

3-15 Design Decision: Time Synchronization Configuration

How should the infrastructure be configured to support time synchronization?

Time synchronization must be maintained between VMs, ESXi hosts, and management systems. Whenever possible, apply the following recommendations to the design:

- Use NTP, PTP, or AD to synchronize components to a common time source.
- Make multiple, external NTP or PTP time servers available to the infrastructure.
- In larger infrastructures, configure at least two internal NTP or PTP servers.
- Configure the ESXi hosts and AD servers (if you have AD) as NTP clients.
- If the management servers are not AD clients, configure them as NTP clients.

Time synchronization between the virtual infrastructure components is important for several reasons:

- Yields accurate and useful performance data.
- Provides meaningful time stamps on log file entries.
- Meets application requirements. For example, database transactions might not be processed correctly between multitiered applications running on separate VMs when a time synchronization problem exists.

Use NTP to synchronize all the management and host components to a common time source. Synchronizing with multiple NTP time sources is an NTP best practice because using multiple sources ensures time accuracy and service availability. Larger organizations might consider creating internal NTP servers to synchronize to. Creating internal NTP servers prevents large numbers of internal hosts and VMs from directly accessing (and overloading) an external NTP server.

Synchronize infrastructure components to a common time source by configuring the ESXi hosts and AD servers as NTP clients. If the management servers (for example, the vCenter Server system) are AD clients, their clocks are synchronized by AD.

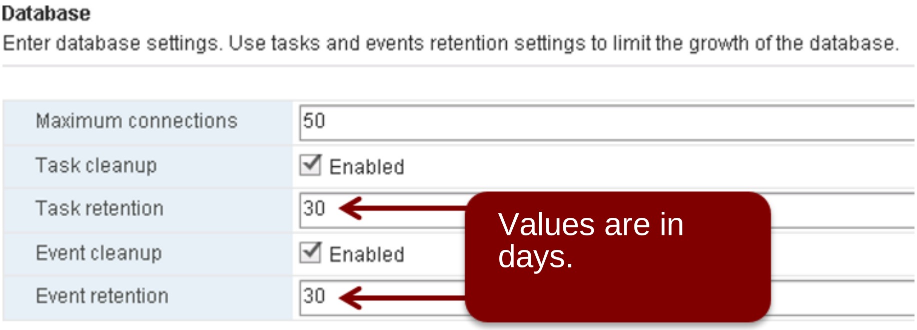
The vSphere 7 platform might need to run workloads that require precise time (to within 1 millisecond accuracy), for example, for financial transaction applications with timestamps. In this scenario, vSphere 7 Precision Clock might be required. Precision Clock is a new virtual device that is presented in virtual hardware and that interfaces with a new VMkernel module called Hyper Clock. The ESXi host must be configured with either NTP or PTP, and only Linux guests with the chrony daemon are supported.

3-16 Design Decision: vCenter Server Tasks and Events Retention Policy

Should vCenter Server tasks and events retention policies be changed?

Consider the following guidelines:

- Check how long you must keep logs and records to adhere to compliance policies and set the retention policy accordingly.
- A longer retention period provides more information for troubleshooting and auditing but requires a larger vCenter Server database.
- If the organization has no specific requirements, use the default settings.



Database

Enter database settings. Use tasks and events retention settings to limit the growth of the database.

Maximum connections	50
Task cleanup	☑ Enabled
Task retention	30
Event cleanup	☑ Enabled
Event retention	30

Values are in days.

vCenter Server collects a log of tasks and events for entities in its inventory. This data is kept indefinitely and can cause substantial growth of the vCenter Server database. Using the tasks and events retention policy, you can configure a threshold after which data is purged.

When determining how much task and event information to retain, you must consider the trade-off. If you decide to keep information for a longer period, your vCenter Server database might require a significant amount of storage resources.

For more information about configuring the vCenter Server database settings, which include task retention and event retention settings, see *vCenter Server and Host Management* at https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.vcenterhost.doc/GUID-3B5AF2B1-C534-4426-B97A-D14019A8010F.html.

3-17 Design Decision: Choosing a Log Collection Method

Which log collection method should you use?

A centralized logging system has several benefits:

- Simplifies log collection, archiving, and troubleshooting
- Captures system-wide events, such as attempted security penetrations
- Allows you to use ESXi hosts with no local storage

If you do not have a usable centralized logging system, consider using a vSphere logging system:

- Forward vCenter Server Appliance log file to a remote host:
  - — vCenter Server and ESXi can stream.
- Use vRealize Log Insight:
  - — Log collection and analytics tool for vSphere environments
  - — Available for purchase

A centralized logging system separates and searches logs quickly.

VMware offers centralized logging tools for vSphere environments:

- You can forward vCenter Server Appliance log files to a remote Syslog server. With this feature, you can further analyze vCenter Server Appliance log files with log analysis products, such as vRealize Log Insight.
- vRealize Log Insight collects and analyzes all types of machine-generated log data, including application logs, network traces, configuration files, messages, performance data, and system state dumps. Administrators can connect vRealize Log Insight to operating systems, applications, storage, firewalls, or network devices for enterprise-wide visibility through log analytics. For more information, see vRealize Log Insight at https://docs.vmware.com/en/vRealize-Log-Insight/index.html.

3-18 Design Decision: Choosing an ESXi Core Dump Collection Method

Which core dump collection method should you use?

VMkernel core dumps are useful debugging tools. Consider the following options for collecting core dumps:

- Default option: Core dumps are saved to the local disk.
- vSphere ESXi Dump Collector:
  - Core dumps are saved on a network server.
  - vSphere ESXi Dump Collector is included with the vCenter Server component.

vSphere ESXi Dump Collector is useful in the following situations:

- Data centers where ESXi hosts might not have any local storage
- Data centers where ESXi hosts have local storage but require an extra location to redirect dumps when critical failures occur.

A core dump is the recorded state of working memory if a host failure occurs.

Use the vSphere Client or vSphere CLI to configure the host to use vSphere ESXi Dump Collector.

In troubleshooting situations, if you cannot access the host with vSphere CLI, you can use esxcli in vSphere ESXi Shell instead.

Authentication and encryption are not required in the file transfer session from a crashed host to the vSphere ESXi Dump Collector.

Configure the vSphere ESXi Dump Collector on a separate VLAN when possible to isolate the ESXi core dump from regular network traffic.

3-19 Design Decision: High Availability Options for vCenter Server Appliance

Which method should you use to make vCenter Server highly available?

Consider the high availability solutions.

Supported High Availability Solutions for vSphere	HA Behavior
vSphere HA	vCenter appliance VM restarts on an alternate host after host or guest failure.
vSphere Fault Tolerance (Tiny and small deployment types only)	Continuous VM operation on failover after host failure (nonstop compute).
VMware Service Lifecycle Manager	A service monitor within the guest OS of appliance restarts failed services.
vCenter Server High Availability	vCenter function fails over to a running replica vCenter appliance.

vSphere HA provides general-purpose protection against hardware and operating system failures for vCenter Server instances running in VMs.

vSphere Fault Tolerance can be used to provide continuous availability for vCenter Server by having identical vCenter Server VMs running on separate hosts. Because vSphere FT supports up to 4 vCPUs, vSphere FT can be used to protect only vCenter Server for the tiny and small deployment types, 2 vCPUs and 4 vCPUs, respectively.

vCenter Server availability is provided by VMware Service Lifecycle Manager. If a vCenter service fails, VMware Service Lifecycle Manager restarts it. VMware Service Lifecycle Manager monitors the health of services and takes preconfigured remediation action when a service failure is detected. A service does not restart if multiple attempts to remediate fail.

vCenter Server High Availability is a native high availability solution for vCenter Server Appliance that protects against hardware failures and also vCenter Server application failures. It monitors the health of the services and takes corrective action if a service fails or is not responding. An advanced failure detection mechanism is constantly monitoring the vCenter Server services, the application, and the database. If a problem is detected, it can fail over to the passive node when necessary.

For more information about vCenter Server high availability, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

For more information on the supported vCenter Server High Availability options, see VMware knowledge base article 1024051 at http://kb.vmware.com/kb/1024051.

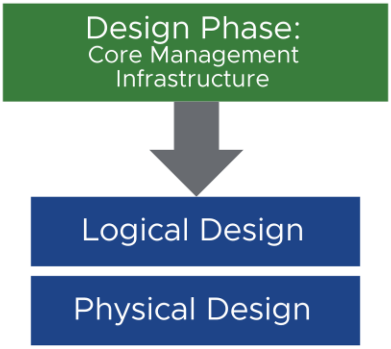
3-20 Lab 3: Designing vCenter Server Architecture

Design the vCenter Server architecture and core management services:

1. Review the Conceptual Design
2. Evaluate vCenter Server Architecture Design Options
3. Diagram the vCenter Server Architecture
4. Create a vCenter Server Physical Design

3-21 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



3-22 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Determine the number of vCenter Server instances to include in a design
- Design a vCenter Server deployment topology that is appropriate for the size and requirements of a data center

3-23 Key Points

- To ensure that a vSphere design meets scalability requirements, always check configuration maximums before deciding on a design.
- vCenter Server Appliance You should be able to meet the following objectives run in a highly available configuration.
- To choose a vCenter Server deployment topology, consider the number of linked vCenter Server instances and redundancy requirements for high availability.
- You specify the most common domain for the vCenter Single Sign-On identity source so that users do not have to qualify their login with a domain.

Questions?

Module 4

Virtual Data Center Infrastructure

4-2 Importance

A vSphere virtual data center helps an organization to achieve its business objectives through several key functions:

- Allocating compute resources to VMs
- Providing user access to data and services
- Protecting the data that is generated by applications and services

Your design for the virtual data center infrastructure determines how well the data center can deliver these important functions.

4-3 Module Lessons

1. Calculating Data Center Resources
2. Cluster Architecture

4-4 Lesson 1: Calculating Data Center Resources

4-5 Learner Objectives

After completing this lesson, you should be able to meet the following objective:

- Calculate the total capacity requirements for a virtual data center design

4-6 Designing the Virtual Data Center Infrastructure

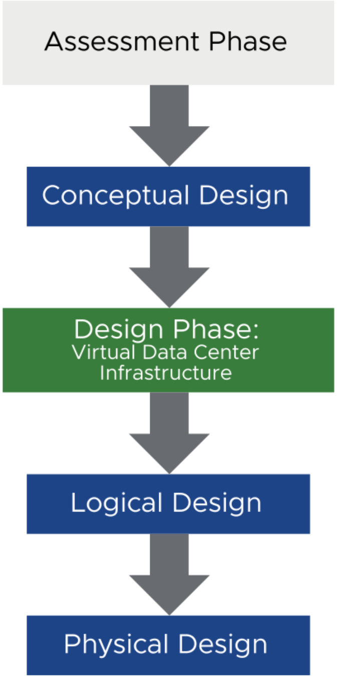
When setting up a virtual data center, you make several design decisions:

- Total CPU and memory requirements for VMs
- Required number of ESXi hosts
- vSphere cluster architecture
- Number of vSphere clusters
- Use of vSphere HA and vSphere DRS
- Use of resource pools

As you make design decisions, validate your logical design decisions against the conceptual design.

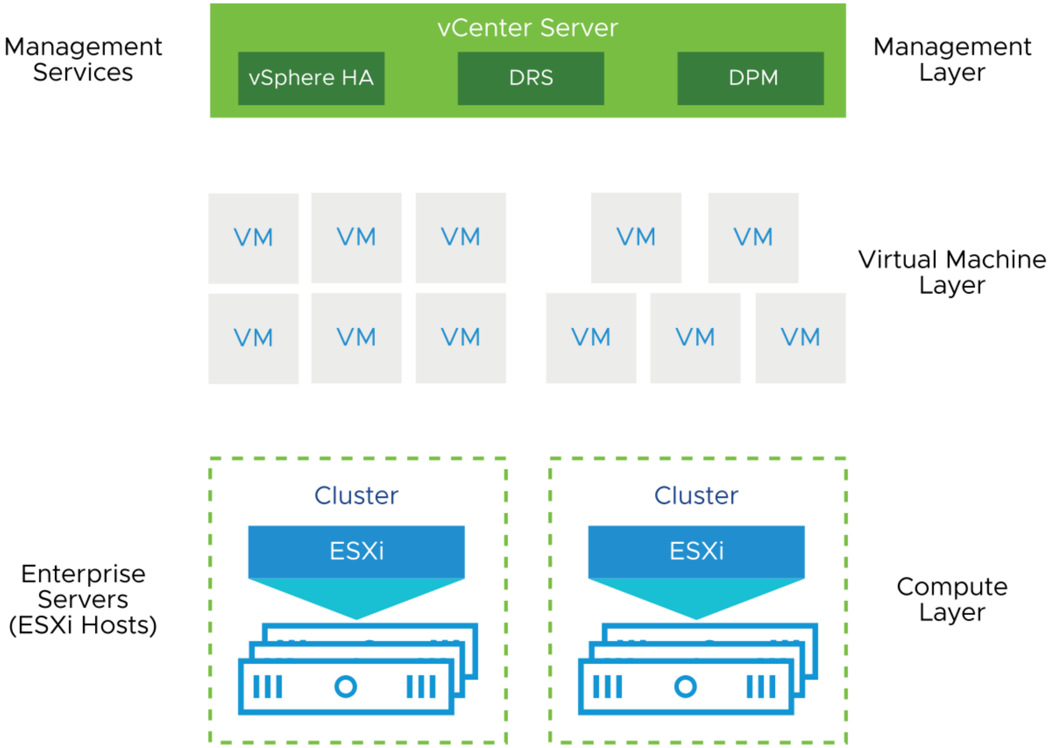
After the conceptual design is created, the design work can begin, starting with the logical design. You consider each of the major vSphere infrastructure components and make the appropriate design decisions.

The design phase is an iterative process. As you make design decisions, always validate these decisions against the organization's goals, requirements, constraints, risks, assumptions, and current best practices.



4-7 Review of the Virtual Data Center Infrastructure

vSphere clusters organize different types of servers into logical groups for failure domains and for configuring management services.



In a vSphere data center, ESXi hosts are typically organized into clusters. Clusters and resource pools group similar hosts into a logical unit of virtual resources, enabling such technologies as:

- vSphere vMotion
- vSphere HA
- vSphere DRS
- vSphere DPM
- vSphere Fault Tolerance

4-8 Basic Process for Designing a vSphere Data Center

To design a vSphere virtual data center:

1. Calculate the capacity requirements.
2. Determine the cluster design.
   - Design the clusters to support the requirements of the workloads and meet the requirements of the management services that will be used for the cluster.
3. Identify the nodes in the cluster, which are the ESXi hosts that provide the resources to support VMs.
4. Add nodes for redundancy, based on the requirements of the organization or specific project.
5. Determine how these nodes connect to one another and to the shared storage where the VM data resides.
6. Add the hosts to the cluster and finish configuring the cluster with the appropriate management services.

Most organizations want to see the results of the capacity analysis. Although you might not always be required to show the raw data to the organization, you should at least consider showing a summary report of the analysis to the organization. You can easily justify the cost of the virtualization project if the capacity analysis report includes the total cost of ownership and return on investment benefits of virtualizing the data center. To calculate these costs, see the VMware TCO Comparison Calculator at http://www.vmware.com/go/tcocalculator.

When you create a cluster, you must configure several settings that determine how the cluster manages the pooled resources. Before you configure settings, identify the cluster's nodes. These nodes are the ESXi hosts that provide the resources to support VMs and that provide failover protection.

Determine how to those nodes should connect to one another and to the shared storage where your VM data resides. After this networking architecture is in place, you can add the hosts to the cluster and finish configuring the cluster with management services, such as vSphere HA, vSphere DRS, and so on.

4-9 Calculating the Total CPU and Memory Requirements

To estimate CPU and memory requirements for the virtual infrastructure:

1. Use the capacity analysis report to determine the required number of VMs for the infrastructure.
2. Adjust the number of VMs based on the following business requirements:
   - Servers to be decommissioned or identified as poor virtualization candidates
   - Growth projections
   - vSphere HA requirements
   - Server isolation requirements
3. Multiply the adjusted total number of VMs by the average resource usage per server by using the CPU and RAM usage data from the capacity analysis report at peak utilization times.
4. Adjust the memory requirement if you decide to enable page sharing because, for security concerns, inter-VM transparent page sharing is disabled by default.

The total number of VMs that are used to calculate total capacity requirements must be adjusted to reflect the requirements of the infrastructure. For example, you do not need to virtualize servers that you plan to decommission. Servers might also use specialized hardware that cannot be virtualized. Increases to the total might be necessary to account for organizational growth projections, and requirements for availability, security, and so on. For example, an organization’s policies for server hardware isolation might require you to add hosts to the design to accommodate different security zones, such as a DMZ network and a production network.

When sizing hosts, use peak utilization levels rather than average utilization. In this way, all systems can run at their observed peak resource levels simultaneously.

Many ESXi workloads present opportunities for sharing memory across VMs and within a single VM. ESXi memory sharing runs as a background activity that scans for sharing opportunities over time. The amount of memory saved varies over time. For a fairly constant workload, the amount generally increases slowly until all sharing opportunities are exploited.

To determine the effectiveness of memory sharing for a given workload, run the workload and run the `resxtop` or `esxtop` command to observe the actual savings. Find the information in the PSHARE field of the interactive mode on the Memory page.

4-10 Calculating the Required Number of Servers

To estimate the number of ESXi hosts that are required to meet total capacity requirements:

1. Determine the available compute resources per host.
   - Use the host specifications from the logical design.
2. Cap use at 10 to 30 percent below capacity to allow for overhead and unanticipated usage, such as:
   - Short-term usage increases tied to business cycles
   - VMkernel overhead
   - Application requirements, such as large memory pages
3. Divide the total requirements by the proposed resources per host:

   a.. Divide the total CPU required by the proposed CPU resources per host.

   b.. Divide the total RAM required by the proposed RAM resources per host.

   c.. Use the higher value of the two values.

When calculating the required number of hosts to support your design, do not plan to use all the available CPU and memory capacity. Always leave at least 10 to 30 percent for overhead and unplanned usage. For example, an organization might have applications that use large memory pages, which can overcommit the active memory. You can avoid performance degradation by providing more physical memory to each host.

After calculating the required number of servers to meet CPU requirements and the required number of servers to meet memory requirements, use the higher of the two values. The higher value must be used because it is the limiting factor. For example, if you find that the CPU workload requires 15 hosts and the memory workload requires 19 ESXi hosts, the number of ESXi hosts required to meet the total capacity is 19.

4-11 Lab 4: Calculating Resource Requirements

Calculate the minimum required number of ESXi hosts:

1. Read the Capacity Planning Assessment Summary
2. Calculate Total CPU and RAM Capacity Requirements
3. Read the ESXi Host Server Specifications
4. Calculate the Minimum Required Number of ESXi Hosts

4-12 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objective:

- Calculate the total capacity requirements for a virtual data center design

4-13 **Lesson 2: Cluster Architecture**

4-14 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create a virtual data center cluster design that meets business and workload requirements
- Evaluate the use of several management services in the virtual data center
- Evaluate the use of resource pools in the virtual data center design

4-15 Design Decision: Choosing a Cluster Architecture

Which cluster architecture type should you use?

The decision to use a scale-out or scale-up cluster architecture depends on budget and infrastructure requirements.

Architecture Type	Benefits	Drawbacks
Scale out: • More hosts • Fewer resources per host	• Less impact during a host failure • Economical host redundancy in large clusters • More efficient bus I/O	• Lower consolidation ratios • Higher power, cooling, and floor space costs • More complex resource management • Easier to reach the limit of hosts per cluster
Scale in: • Fewer hosts • More resources per host	• Simpler resource management • Reduced I/O cabling, power, and real estate • Fewer license costs	• Increased RTO because more objects must recover during host failures • Reduced system bus I/O • Easier to reach the limit of VMs per host

When deciding on a cluster methodology, evaluate the following factors as they relate to the requirements of your design:

- The capital costs of purchasing fewer, larger hosts compared to purchasing more hosts that are smaller. Costs vary between vendors and models.
- The operational costs of managing a few hosts compared to several hosts.
- The maximums for clusters to ensure that you are within the appropriate limits.
- The purpose of the cluster, for example:
  - A virtualized server cluster typically has more hosts with fewer VMs per host.
  - A VMware Horizon cluster typically has fewer hosts with more VMs per host.

4-16 Design Decision: Purpose-Built Clusters

Should you design purpose-built clusters?

Designing clusters based on their purpose might increase capital expenses but reduce operating expenses.

Cluster Type Example	Description
Payload	One or more clusters to run user-workload VMs
Island	A small cluster to segregate a workload with a specific purpose
Management	An island cluster that contains components that manage and monitor the virtual infrastructure and stage and test new VMs
Edge	Island clusters for network gateway devices, such as NSX Manager and NSX Edge devices

A purpose-built cluster design reduces consolidation ratios and increases capital expenses because this design requires more clusters. However, the benefits are significant:

- Licensing and management are simpler when similar environments are grouped.
- Resiliency requirements can be more easily met for workloads that have higher SLAs.
- Resources can be allocated based on the performance needs of the workloads.
- Physical separation gained is often a compliance requirement and is considered secure.

A purpose-built cluster design typically includes some common types of clusters:

- Payload clusters host the end-user VMs. The number of ESXi hosts included in a payload cluster is determined by the resource requirements of the application workload. But the number must be within the support range for the version of vSphere that is used in the design. See the section on cluster and resource pool maximums for ESXi hosts in VMware Configuration Maximums at https://configmax.vmware.com. Additional payload clusters must be added when the maximum for any one component is reached.
- Island clusters host workloads that require segregation for a unique purpose. Island clusters are useful for workloads that have unique licensing or security requirements. These clusters can also be used for segregating resource-intensive workloads that perform best and affect other workloads the least if the workloads are isolated in their

own cluster. The number of hosts included in an island cluster depends on the purpose of the cluster. However, island clusters typically include only two or three hosts.

- A management cluster is a type of island cluster segregated for performance and security. This cluster contains all the components used to manage and monitor the virtual infrastructure, including vCenter Server, Microsoft Active Directory domain controllers, and database servers. The management cluster can also be used for deploying new VMs for staging and testing. The management cluster typically consists of a minimum of three ESXi hosts.
- Edge clusters simplify the physical network switch configuration. These clusters are used to deliver networking services to payload-cluster VMs. Typically, all external networking for user-workload VMs, including corporate and Internet, are accessed through the edge cluster. The minimum cluster size depends on the gateway device requirements.

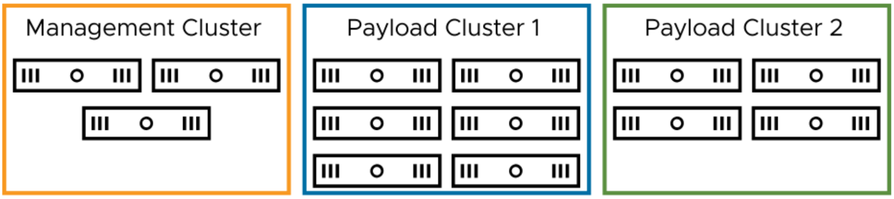
For more information about designing purpose-built clusters, see "Creating Purpose-Built vSphere Clusters" at http://blogs.vmware.com.

4-17 Design Decision: Management Cluster (1)

Should you create a management cluster?

For enterprise environments, consider creating a management cluster:

- Provides resource isolation so that management services can operate at the best possible performance level
- Can satisfy an organization's policy to physically isolate the management hardware from the production hardware
- Can be expensive and therefore is suited for large environments to warrant the cost



Management, monitoring, and infrastructure services are critical functions in a vSphere environment. Placing these functions in their own cluster isolates these functions from applications. Production applications, test applications, and other types of applications cannot use the cluster resources reserved for management, monitoring, and infrastructure services.

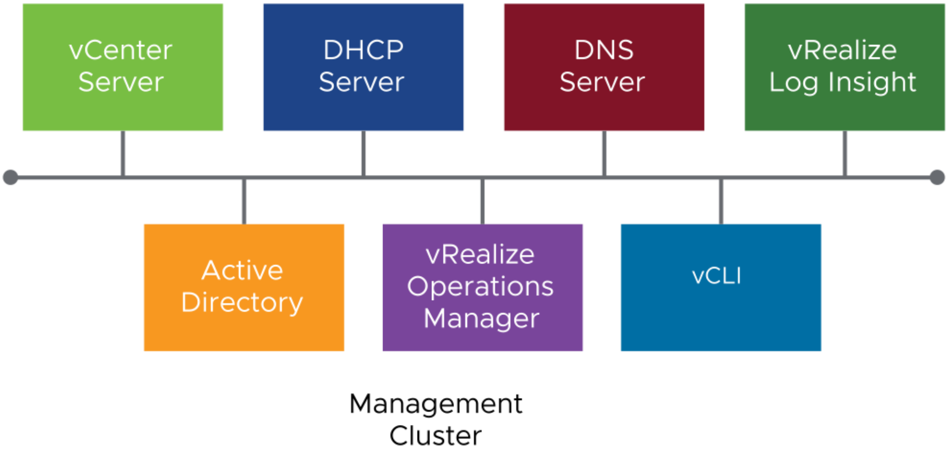
The primary disadvantage of creating a management cluster is cost. A large environment is required to warrant a dedicated management cluster.

For example, a three-host management cluster might proportionally add significant cost to an infrastructure of only 10 hosts. However, a three-host management cluster might proportionally add little cost to an infrastructure of 50 hosts.

4-18 Design Decision: Management Cluster (2)

Consider the following guidelines when designing a management cluster:

- Limit access to this cluster to only vSphere administrators.
- Create at least a three-host cluster that is enabled for vSphere HA and vSphere DRS.
- Run only infrastructure services, such as:
  - vCenter Server
  - DHCP servers
  - Active Directory servers
  - DNS servers
  - Infrastructure services specific to the organization



Placing all management, monitoring, and infrastructure services in a vSphere HA cluster provides higher availability for these critical services.

You can configure permissions on the management cluster to limit access to only vSphere administrators. Access to the VMs running the management, monitoring, and infrastructure services is further protected.

4-19 Design Decision: Evaluating the Use of vSphere HA

Should you enable vSphere HA in the clusters?

vSphere HA clusters provide higher levels of availability for VMs than each ESXi host can provide:

- The decision to create a vSphere HA cluster depends on the workloads being virtualized:
  - With vSphere HA Orchestrated Restart, an application that runs across multiple VMs can recover because the restart order is enforced.
- As a best practice, you should protect all important workloads:
  - If a workload is protected in the current infrastructure, the workload most likely requires protection in the virtual infrastructure.
  - If workloads are already running application-level high availability, vSphere HA might not be needed.
  - vSphere HA is simple to configure and can protect a wide range of workloads and services.

vSphere HA provides high availability for VMs by monitoring hosts in a cluster. If a failure occurs, the VMs on a failed host are restarted on alternate hosts.

vSphere HA Orchestrated Restart improves the current restart priority by creating restart tiers and VM-VM dependencies. You create dependency chains between VMs using VM-to-VM restart rules. These restart rules enforce the restart order for each VM within the dependency chain, increasing the likelihood that an impacted application properly recovers when vSphere HA restarts the VMs.

vSphere HA is simple to configure and easily protects a wide range of workloads. Many services that are not currently protected are typically chosen to be protected in the new virtual infrastructure. When vSphere HA is configured, networking and storage must be configured to support it.

You should protect all workloads with vSphere HA if they are important to the daily business operations. If workloads are already running application-level high availability, vSphere HA might not be needed.

For a more in-depth discussion of vSphere HA features, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

4-20 Requirements for vSphere HA

To configure a cluster for vSphere HA, the cluster and ESXi hosts in the cluster must meet the following requirements:

- The cluster must contain at least two ESXi hosts.
- ESXi hosts in the cluster must have the following characteristics:
  - Be licensed for vSphere HA
  - Be configured with a static IP address
  - Contain at least one management network in common but best practice is two networks
  - Access to the same VM networks and datastores
- For VM Monitoring, VMware Tools must be installed in the VMs that are monitored.
- For Application Monitoring, the application must support VMware Application Monitoring, or you must use the SDK to set up customized heartbeats.
- For VM Component Protection (VMCP), clusters must contain ESXi 6.0 hosts or later.

vSphere HA supports IPv6 network configurations for clusters containing ESXi 6.0 hosts or later.

For a complete list of vSphere HA requirements, which include additional configuration requirements for IPv6, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

4-21 Design Decision: vSphere HA Heartbeat Redundancy

Which vSphere HA heartbeat redundancy methods should you use?

Always configure heartbeat redundancy or explain the risk to key stakeholders and SMEs. If hosts do not receive heartbeats, a cluster-wide isolation response might be triggered.

You can use the following methods to provide heartbeat redundancy:

- Creating management network redundancy
- Using storage that is resilient to failures for your heartbeat datastores

Consider using both network redundancy and datastore heartbeats in your vSphere HA cluster.

Always configure some form of heartbeat redundancy in a vSphere HA environment. vSphere HA uses a heartbeat mechanism to detect host isolation and host failures. Network heartbeating over a management network is the primary heartbeating mechanism. Datastore heartbeating is the secondary heartbeating mechanism.

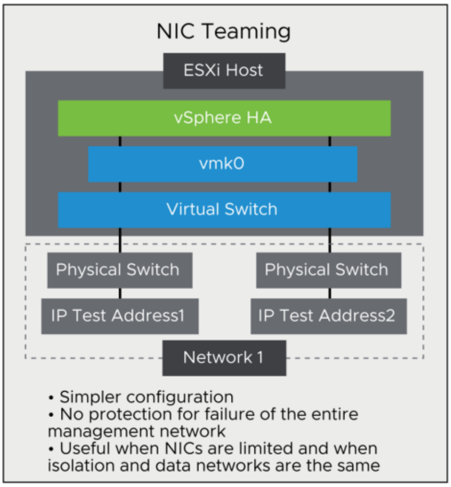
Each host should have multiple paths to the storage that it needs to access. Two or more HBAs or NICs that are connected to separate switches can provide access to redundant storage processors (SPs) and arrays with resilient RAID configurations. HA automatically selects two datastores with the highest number of connected hosts.

You can override this setting and specify datastores. Normally, this step is not necessary, but it might be useful, for example, if certain datastores are supported by more resilient LUNs or storage arrays.

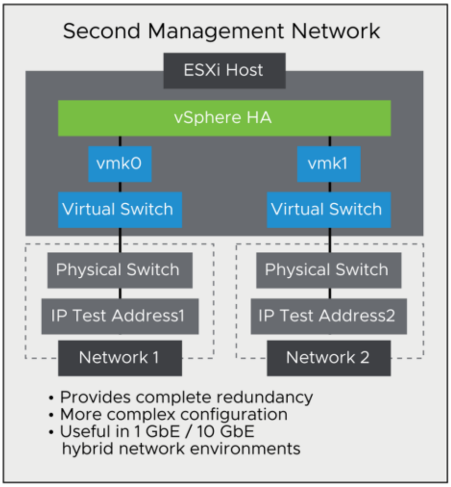
Accessibility issues can occur when heartbeat redundancy does not exist. For example, a physical switch on the management network, or a heartbeat datastore, might become inaccessible. In such a situation, hosts are prevented from receiving heartbeats. Every affected host initiates an isolation response. Depending on the configured isolation response, a large portion of the VMs might shut down although the default isolation response is to leave the VM powered on.

4-22 Choosing a Method for Network Heartbeat Redundancy

Two configuration options are available for creating redundancy for network heartbeats.



- Simpler configuration
- No protection for failure of the entire management network
- Useful when NICs are limited and when isolation and data networks are the same



- Provides complete redundancy
- More complex configuration
- Useful in 1 GbE / 10 GbE hybrid network environments

Network path redundancy between cluster nodes is important for vSphere HA reliability. A single management network can become a single point of failure and can result in failovers, even if only the network fails. Possible failures include network interface card (NIC) failures, network cable failures, network cable removal, and switch resets.

You can create additional heartbeat redundancy by setting up management network redundancy. To set up management network redundancy, create a second network management port on each ESXi host in the cluster. Attach each port on each host to a separate physical management network. The advantage of this method is that two separate and redundant management networks and IP addresses are created. The disadvantage is that this configuration creates a slightly more complex environment to manage. This configuration is a good option for a hybrid network of 1 GB and 10 GB Ethernet environments. The 1 GB controller can host the primary network heartbeat, and the 10 GB controller can be used for the secondary network heartbeat.

Configuring a single management port on a NIC team also provides redundancy. In this configuration, each NIC in the team is connected to a separate physical switch. If either NIC fails, or the path to either NIC fails, an alternate NIC and path are available. This method is a slightly simpler configuration. However, each NIC in the team is still connected to a single management network. If the entire network fails, an isolation response is triggered. This option is most commonly used when NICs are limited and the isolation domains are the same as the data network, for example, in 10 GB Ethernet environments.

Always configure a second isolation test address. A single isolation test address is a single point of failure. If the isolation address is not reachable, an unintended isolation response is initiated. Use the advanced parameter `das.isolationaddress` to add an isolation test address for each additional management network.

4-23 Heartbeat Datastores

vSphere HA uses datastore heartbeating as a secondary heartbeating mechanism:

- Monitors hosts when a management network partition occurs
- Enables vSphere HA to continue to respond to failures

By default, vCenter Server selects two shared datastores for heartbeating.

Consider the following guidelines for configuring datastore heartbeating:

- Give preference to your datastores that are resilient to failures.
- If your design requires more datastores, use the advanced das.heartbeatDsPerHost option to increase the number to up to five datastores.
- If storage and management network traffic are traveling over the same physical NICs, disable datastore heartbeating because it does not provide any benefit in this type of network.

vSphere HA uses datastore heartbeating to distinguish between hosts that have failed and hosts that reside on a network partition. If some datastores are more resilient to failures in the environment, you can configure vSphere HA to give preference to them.

For detailed information about configuring datastore heartbeats, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

4-24 Design Decision: VM Component Protection

Should you enable VMCP?

VMCP is a powerful feature that protects VMs from storage accessibility failures:

- Permanent device loss (PDL): An unrecoverable loss of accessibility that occurs when a storage device reports the datastore is no longer accessible by the host.
- All paths down (APD): A transient or unknown accessibility loss or any other unidentified delay in I/O processing. This issue is recoverable.

VMCP has some limitations:

- Does not support vSphere Fault Tolerance
- Does not protect VMs on vSAN or Sphere Virtual Volumes datastores
- Does not protect against inaccessible raw device mappings (RDMs)

VM Component Protection (VMCP) provides protection against datastore accessibility failures that can affect a VM running on a host in a vSphere HA cluster. When a datastore accessibility failure occurs, the affected host cannot access the storage path for a specific datastore. You can determine the response of vSphere HA to such a failure, ranging from the creation of event alarms to VM restarts on other hosts.

For more information about vSphere HA configuration, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

4-25 Design Decision: Choosing an Admission Control Policy

Which admission control policy should you use?

vSphere HA supports the use of several policies to ensure that sufficient resources in a cluster are reserved for VM recovery. The choice depends on resource fragmentation considerations, flexibility of resource reservation, and the size and heterogeneity of the cluster.

Policy	Benefits	Drawbacks
Cluster resource percentage (default)	Most flexible configuration option. Define the number of host failures to tolerate. Failover capacity is calculated automatically.	Does not address resource fragmentation.
Slot policy (powered-on VMs)	Avoids resource fragmentation by defining a slot.	In heterogeneous clusters, this policy might be too conservative. In small clusters, this policy reserves a large proportion of resources.
Dedicated failover hosts	Eliminates resource fragmentation.	In small clusters, this policy reserves a large proportion of resources.

vSphere HA admission control ensures that sufficient resources in the cluster are reserved for VM recovery if a host failure occurs. The settings determine whether VMs can be started if they violate availability constraints. Configuration options vary by policy.

- For cluster resource percentage, you define the number of host failures to tolerate (FTT), and vSphere HA automatically calculates a percentage of resources to set aside by applying this admission control policy. As hosts are added or removed from the cluster, the percentage is automatically recalculated. This setup is the default configuration, but it is possible to override the automatic calculation.
- For slot policy (powered-on VMs), you select a slot size policy that covers all powered-on VMs or is a fixed size. You can also calculate how many VMs require multiple slots.
- For dedicated failover hosts, you select hosts to use for failover actions.

You can also choose not to reserve failover capacity. If you choose this option, vSphere HA powers on VMs even when they violate availability constraints. But with this option, the expected number of VMs might not be restarted after a failure.

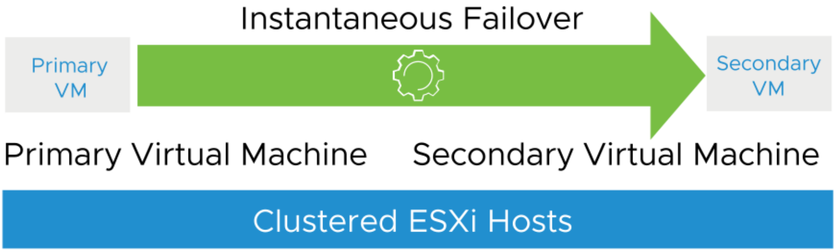
For detailed information about evaluating and configuring admission control policies, see *vSphere Availability* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.avail.doc/GUID-63F459B7-8884-4818-8872-C9753B2E0215.html.

4-26 Design Decision: Evaluating the Use of vSphere Fault Tolerance

Should vSphere FT be used?

vSphere FT ensures higher levels of availability and data protection than vSphere HA and should be considered for the following use cases:

- Continuously available applications:
  - Especially applications that have long-lasting client connections that users want to maintain during hardware failures
  - Workloads that require additional protection during critical periods
- Custom applications with no clustering
- Custom clustering solutions, which are too complicated to configure and maintain



You can use vSphere FT for the most mission-critical VMs. vSphere FT provides continuous availability for a mission-critical VM by creating and maintaining another VM that is identical and continuously available to replace the mission-critical VM if a failover situation occurs.

4-27 Design Decision: Evaluating the Use of vSphere DRS

Should you enable clusters for vSphere DRS?

vSphere DRS provides load balancing of the cluster by migrating workloads from heavily loaded hosts to less used hosts in the cluster. DRS considers CPU and memory use of hosts and VMs.

Earlier releases of vSphere DRS aggregated computing capacity per cluster load.

In vSphere 7, DRS ensures that resource requirements are met per individual VM.

DRS calculates VM scores and triggers VM migrations to satisfy VM resource requirements.

In most cases, include vSphere DRS in the virtual data center design.

vSphere 7 introduces DRS 2.0. Many of the previous DRS features and capabilities are the same in DRS 2.0. The main difference is in the load-balance frequency and DRS per-VM focus. Earlier releases of vSphere DRS were designed to aggregate computing capacity on a cluster load basis. DRS 2.0 now works on an individual VM basis to ensure that resource requirements are met. This new focus on the VMs receiving required resources leads to better VM performance and new information available through the vSphere Client.

The vSphere DRS 2.0 load-balancing algorithm runs every minute. It performs calculations to prepare for VM moves with vSphere vMotion. Internally, it uses a placement decision tree to evaluate current performance, perform a what-if scenario on a different compatible host, check the cost of the VM migration, and perform a vSphere vMotion migration to the new host if the VM runs more efficiently. The new focus of these calculations is for the highest VM DRS score, which is the highest instance at which the VM's resource requirements are met.

The DRS placement algorithm runs every minute and recommends where individual VMs should be moved for maximum efficiency. If the cluster automation level is set high enough, DRS runs the recommendations and migrates VMs to their optimal host based on the underlying calculations done every minute. The load-balancer algorithm also runs every minute.

IMPORTANT: Because of the new focus on the VM's DRS score, a greater number of vSphere vMotion migrations can occur initially, after vSphere DRS 2.0 is enabled. This probability derives from the need to normalize the cluster and ensure that VMs have the resources that they need regardless of which individual host they are placed on.

4-28 Best Practices for Designing vSphere DRS Clusters

Consider the following best practices when designing vSphere DRS clusters:

- Configure vSphere DRS for full automation by using the default migration threshold:
  - Reduces daily monitoring and management requirements
  - Provides sufficient balance without excessive migration activity
- vSphere DRS load balancing benefits from having a larger number of hosts in the cluster (scale-out cluster) rather than a smaller number of hosts.
- Ensure that vSphere DRS affinity and anti-affinity rules are the exception rather than the norm:
  - An affinity rule might be useful in the following situations:
    - VMs on the same network share significant network traffic.
    - Applications share a large memory working set size.
  - An anti-affinity rule might help applications with high-transactional I/O workloads to avoid an I/O bottleneck on the local host.
- Ensure that all vSphere vMotion migration requirements are met by all hosts in the vSphere DRS cluster.

Usually, use the default migration threshold to configure vSphere DRS for full automation. The default migration threshold is typically aggressive enough to maintain workload balance across hosts without creating an excessive CPU workload that is caused by too frequent vSphere vMotion migrations.

Testing shows that vSphere DRS achieves a better balance when more, not fewer, hosts are in the cluster. More hosts offer the vSphere DRS algorithms more choices for migration. As a result, a scale-out cluster is preferred if the cluster is enabled for vSphere DRS.

Although vSphere DRS affinity and anti-affinity rules are helpful in certain situations, you should limit the total number of such rules. Affinity and anti-affinity rules limit the number of available migration choices and collectively might have a negative effect on the ability of DRS to achieve the optimal workload balance:

- Affinity rules are useful for VMs on the same network whose applications generate significant network traffic between them. This network traffic traverses only the virtual switch, not the physical network links. Having the network traffic traverse only the virtual switch can reduce congestion on the physical network links and components.
- Anti-affinity rules should be configured when appropriate to increase availability. For example, identically configured DNS server VMs can be maintained on separate hosts to eliminate the possibility that a single host failure results in DNS service failure. (If vSphere HA is also configured, the DNS server VMs might be restarted anyway.) In rare cases, anti-affinity rules might be configured for performance reasons. Application performance might improve if applications with high-transactional I/O workloads are not allowed to run on the same host.

For more information about vSphere DRS, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html.

4-29 Design Decision: Evaluating the Use of Predictive DRS

Should you use predictive DRS?

Predictive DRS integrates with vRealize Operations Manager:

- Uses both current and forecasted usage information to make the best VM placement decision
- Acknowledges predicted workload spikes and proactively balances (migrates) VMs before the spikes occur
- Leads to better performance for all workloads in a vSphere cluster

vSphere DRS	Predictive DRS
Resolves unexpected resource demand	Predicts future resource demand
Minimal overhead	Minimal overhead
Only moves VMs that must be moved	Only moves VMs that must be moved
Moves VMs when contention starts	Moves VMs before contention starts

Historically, vSphere DRS has been reactive. It reacts to any changes in VM workloads and migrates the VMs to distribute load across different hosts. From vSphere 6.5, with vCenter Server working together with vRealize Operations Manager, DRS can act on predicted future changes in workloads. This action helps DRS migrate VMs proactively and makes room in the cluster for future workload demand.

For example, if your VMs' workload is going to spike at 9:00 AM every day, predictive DRS detects this pattern based on historical data from vRealize Operations Manager. It can then prepare the cluster resources by using either of the following techniques:

- Migrating the VMs to different hosts to accommodate the future workload and avoid host overcommitment
- Bringing back a new host from standby mode using vSphere DPM to accommodate the future demand

4-30 Design Decision: Evaluating the Use of Proactive HA

Should Proactive HA be enabled?

Proactive HA is a vSphere DRS feature that proactively reduces the need for vSphere HA by anticipating host failures and avoiding unnecessary downtime:

- Unnecessary downtime might be related to issues with power supply, memory, network, storage, and even fan failures.
- vSphere DRS uses the host’s health statistics to make recommendations about VM placement to prevent VM outages if a host is degraded.

Proactive HA requires support from plug-ins and Proactive HA solutions from certified vendors. Check the VMware website for the list of certified vendors.

Proactive HA is based on the observation that some symptoms of potential future host hardware failure can occur without an administrator's knowledge. Some indications of future trouble can be detected minutes, hours, or even days before the failure occurs. The risk of workload failure can be mitigated by taking actions when these indications are detected.

In vSphere 6.5, a new API set allows third-party server vendors to provide information and diagnosis about server health states. The third-party tool monitors the host sensor data and determines whether particular failures jeopardize the health state of a host. The tool sends alerts to vCenter Server informing it of the host’s state: healthy (green), moderately degraded (yellow), severely degraded (red), or unknown (gray).

Proactive HA uses vSphere DRS to migrate virtual machines away from a degraded host.

The migration reduces data loss and avoids potential vSphere HA restarts.

Proactive HA integrates with the server vendor’s monitoring software, through a vSphere Client plug-in, which passes on detailed server health status and alerts to vSphere DRS.

vSphere DRS reacts based on the health state of the host’s hardware.

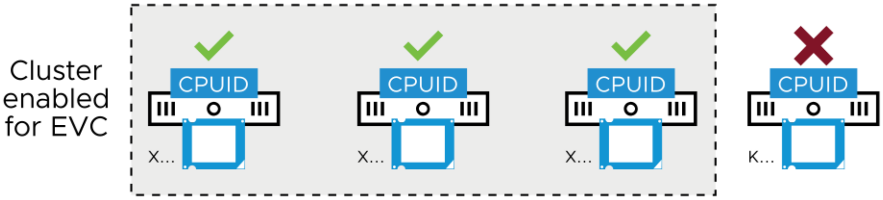
4-31 Design Decision: Evaluating the Use of Enhanced vMotion Compatibility

Should Enhanced vMotion Compatibility be used?

Use this feature to prevent vSphere vMotion migrations from failing because of incompatible CPUs. Set the EVC mode to the highest level possible with the current CPU in use:

- Hosts with newer CPUs can be added later without disruption.
- Rolling hardware upgrades can be performed with zero downtime.

You can also enable, disable, or change the EVC mode at the VM level to simplify cross vCenter vSphere vMotion.



vSphere vMotion requires that the CPUs in each host be similar to ensure that live migrations can occur. Enhanced vMotion Compatibility ensures that all hosts in a cluster present the same CPU feature set to VMs, even if the actual CPUs on the hosts differ.

Enhanced vMotion Compatibility masks only processor features that affect vSphere vMotion compatibility. Enabling this feature does not prevent a VM from taking advantage of faster processor speeds, increased numbers of CPU cores, or hardware virtualization support that might be available on newer hosts.

To check Enhanced vMotion Compatibility support for a specific processor or server model, see the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility. Search for the server model or CPU family and click the entry in the CPU Series column to display the compatible EVC modes.

The per-VM Enhanced vMotion Compatibility feature facilitates the migration of the VM beyond the cluster and across vCenter Server systems and data centers that have different processors. The EVC mode of a VM is independent from the EVC mode defined at the cluster level. The cluster-based EVC mode limits the CPU features a host exposes to VMs.

The per-VM EVC mode determines the set of host CPU features that a VM requires to power on and migrate.

By default, when you power on a newly created VM, it inherits the feature set of its parent Enhanced vMotion Compatibility cluster or host. However, you can change the EVC mode for each VM separately. You can raise or lower the EVC mode of a VM. Lowering the EVC mode increases the CPU compatibility of the VM. You can also use the API calls to customize the EVC mode further.

The way that Enhanced vMotion Compatibility works at the host cluster level is different from how it works at the VM level:

- Unlike with cluster-based Enhanced vMotion Compatibility, you can change the per-VM EVC mode only when the VM is powered off.
- With cluster-based Enhanced vMotion Compatibility, when you migrate a VM out of the Enhanced vMotion Compatibility cluster, a power cycle resets the EVC mode of the VM. With per-VM Enhanced vMotion Compatibility, the EVC mode becomes an attribute of the VM.

  A power cycle does not affect the compatibility of the VM with different processors.
- When you configure Enhanced vMotion Compatibility at the VM level, the per-VM EVC mode overrides cluster-based EVC. If you do not configure per-VM EVC, when you power on the VM, it inherits the EVC mode of its parent Enhanced vMotion Compatibility cluster or host.
- If a VM is in an Enhanced vMotion Compatibility cluster and the per-VM EVC is also enabled, the EVC mode of the VM cannot exceed the EVC mode of the EVC cluster in which the VM runs.

  The baseline feature set that you configure for the virtual machine cannot contain more CPU features than the baseline feature set applied to the hosts in the EVC cluster.

4-32 Design Decision: Evaluating the Use of vSphere DPM

Should vSphere DPM be used?

Consider using vSphere DPM so that a vSphere DRS cluster can reduce its power consumption:

- vSphere DPM is useful in the following situations:
  - Environments where workloads vary significantly over time
  - Configuration of vSphere DPM to run only during nonbusiness hours
- In most cases, the default automation and threshold levels are sufficient to realize savings without increasing administrative management overhead.

vSphere DPM requires one of the following technologies:

- ESXi hosts that support either the HP Integrated Lights-Out (iLO) interface or Intelligent Platform Management Interface (IPMI)
- NICs for the vSphere vMotion interfaces that support Wake on LAN (WOL).

vSphere DPM is a feature of vSphere DRS. During periods of inactivity, vSphere DRS migrates current workloads to other hosts and places the host in standby. When demand increases in the cluster, vSphere DRS sends instructions to power on the host.

vSphere DPM is similar to vSphere DRS but has different modes of operation. These modes are independent of the vSphere DRS settings:

- Off: vSphere DPM is not used.
- Manual: vSphere DPM provides recommendations on hosts to power off, but the recommendations must be manually initiated by an administrator.
- Automatic: vSphere DPM recommendations for host shutdown and restarts are triggered.

Migration thresholds from Aggressive to Conservative behave in similar ways to vSphere DRS.

vSphere DPM can use one of three power management protocols to bring a host out of standby mode. Each protocol requires its own hardware support and configuration. If a host does not support any of these protocols, vSphere DPM cannot put it into standby mode. If a host supports multiple protocols, they are used in the following order: Intelligent Platform Management Interface (IPMI), HP Integrated Lights-Out (iLO), and Wake on LAN (WOL).

For additional information about vSphere DPM, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html.

4-33 Design Decision: Evaluating the Use of Resource Pools

Should you use resource pools?

Use resource pools to dedicate and isolate resources for the workloads that require them.

Avoid putting resource pools and VMs as siblings in a hierarchy: Share values for resource pools and VMs might not compare appropriately, potentially resulting in unexpected performance.

Benefits	**Drawbacks**
Flexible assignment of CPU and memory resources	Can cause poor use of compute resources if not constantly analyzed or balanced
Resource isolation to prevent unfair use of resources	Creates another management layer that must be monitored
Prevention of overutilization of resources	
Prioritized, automated access to a pool of resources	

Resource pools can be useful for dedicating CPU, memory, and network resources to departments or projects. Using resource pools to dedicate resources might also be useful for chargeback purposes, if the organization cannot purchase a chargeback product.

Resource pools help improve the management and troubleshooting of performance problems. For performance reasons, each level in the resource pool hierarchy should contain only resource pools or only VMs, but not a combination of both.

4-34 Design Decision: Evaluating the Use of Scalable Shares

Should scalable shares be enabled for DRS resource pools?

Use the scalable shares setting on resource pools in DRS clusters in the following situations:

- High design priority to maintain the share ratio between peer resource pools
- Requirement to support an unbalanced scale-out, where one or more resource pools scale out the number of VMs considerably greater than peer resource pools
- Where analyzing and changing share allocations later based on scale-out might be an unwanted administrative overhead

vSphere 7 provides a new option for resource pools, called scalable shares. With this option, the number of shares set on a resource pool becomes a computed value, based on the original resource pool value multiplied by the sum of the VM shares in that resource pool.

In other words, even as the number of VMs in a resource pool scales out, the intended effect of shares set at the parent resource pool level is always retained.

For example, RP-1 should have double the number of shares allocated to RP-2.

If the number of VMs in RP-1 starts to increase, that scale-out effect diminishes the available resource allocation to VMs in RP-1.

However, by dynamically increasing the shares on the parent pool by a computed value based on the number of VMs, the behavior of the system is preserved as per the design, even after many more VMs are added.

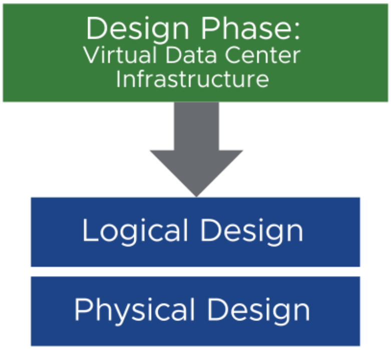
4-35 Lab 5: Designing the Virtual Data Center Infrastructure

Calculate capacity requirements and create the vSphere cluster design:

1. Review the Conceptual Design and Capacity Planning Assessment Data
2. Evaluate vSphere Cluster Design Options
3. List vSphere Clusters
4. Create a vSphere Cluster Physical Design

4-36 Review of Lab: Designing the Virtual Data Center Infrastructure

The instructor facilitates a class discussion of the design decisions of one or more teams.



4-37 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Create a virtual data center cluster design that meets business and workload requirements
- Evaluate the use of several management services in the virtual data center
- Evaluate the use of resource pools in the virtual data center design

4-38 Key Points

- Do not plan on fully using all host CPU or memory resources.
- Use one or more clusters that are enabled for vSphere HA and vSphere DRS to increase availability and scalability.
- Always provide management network redundancy for vSphere HA clusters.
- Whenever possible, enable Enhanced vMotion Compatibility on a cluster and configure it for the highest level possible with the current CPU in use.
- vSphere FT is a good option for applications that must be always available.
- vSphere DPM is a good option for environments where workloads vary significantly over time.
- Configure resource pools only for workloads that require dedicated and isolated resources.

Questions?

Module 5

Compute Infrastructure

5-2 Importance

ESXi hosts are the fundamental compute building blocks of the virtual data center. These hosts are aggregated to build clusters of highly available pools of compute resources.

You must design a compute infrastructure that meets calculated consolidation ratios, protects against system failures using component redundancy, and supports all necessary vSphere features.

5-3 Learner Objectives

After completing this module, you should be able to meet the following objectives:

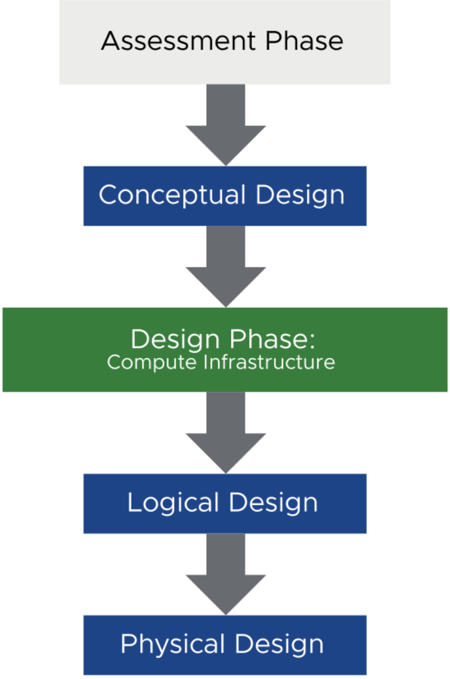
- Create a compute infrastructure design that includes the appropriate ESXi boot, installation, and configuration options
- Select the ESXi host hardware for the compute infrastructure

5-4 Designing the Compute Infrastructure

Design decisions for the compute infrastructure relate to the following areas:

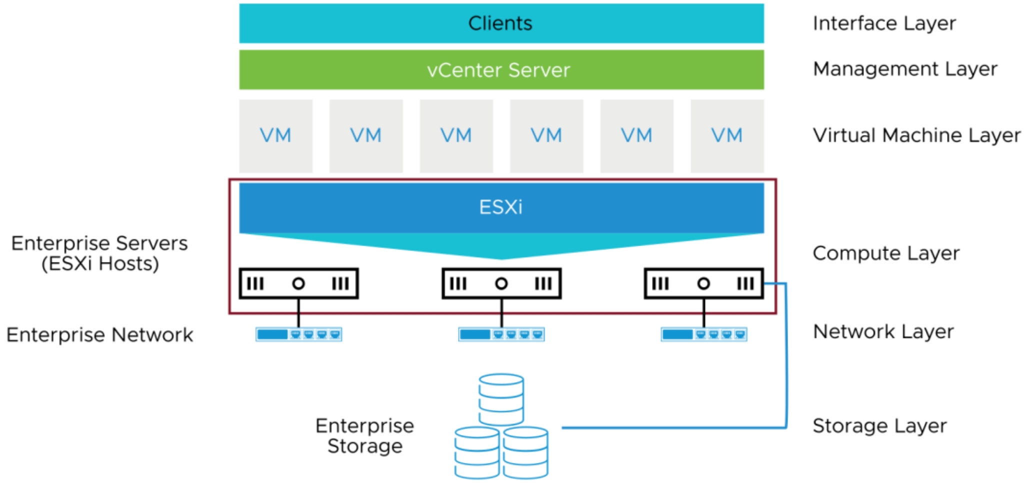
- ESXi server platform
- ESXi installation method
- ESXi configuration:
  - Boot disk location
  - IP addresses
  - Naming conventions

The design phase is an iterative process. As you make Design decisions, always validate these decisions against the organization's goals, requirements, constraints, risks, assumptions, and current best practices.



5-5 Review of the Compute Layer in a vSphere Infrastructure

ESXi hosts provide resources for virtual machines and run the VMs.



The compute layer of the architecture design encompasses the CPU, memory, and hypervisor technology components.

Compute resources are provided by enterprise servers. Enterprise servers are industry-standard x86 servers that run ESXi on the bare metal. Each server is called a standalone host in the virtual environment.

You can group several similarly configured x86 servers with connections to the same network and storage subsystems to provide an aggregate set of resources in the virtual environment. The aggregate set of resources is a vSphere cluster.

5-6 Design Decision: Hardware Platform Consistency

Should all ESXi hosts share a consistent configuration?

Whichever hardware platform is selected, design a consistent platform configuration, especially in vSphere clusters, for the following components:

- CPU type
- Memory capacity and memory slot assignment
- Network interface card
- Host bus adapter types
- PCI slot assignments

Consistent hardware and configuration have the following benefits:

- Simplify capacity planning
- Simplify automated installation and configuration
- Simplify troubleshooting
- Preserve vSphere vMotion compatibility

Standardizing the entire physical configuration of the compute infrastructure is critical to providing an easily manageable and supportable infrastructure. You achieve consistency by purchasing a single-server model and configuring each server in the same way.

Verify host functionality because you can find hosts that have components that might produce intermittent failures. Placing a host that is nearly, but not completely, functional in production can result in unnecessary downtime.

5-7 Choosing an ESXi Server Platform

Use only supported server platforms from the VMware Compatibility Guide.
An ESXi 7 host must have the following hardware and resources:

- 64-bit x86 processors released after September 2006
- At least two cores
- NX/XD bit enabled for the CPU in the BIOS
- Minimum of 4 GB physical RAM, but provide at least 8 GB
- Hardware virtualization (Intel VT-x or AMD-V) enabled on CPUs
- One or more 1 GbE or faster (10 GbE, 25 GbE, 40 GbE) Ethernet controllers
- SCSI disk or a local, non-network RAID LUN with unpartitioned space for VMs
- For Serial ATA (SATA), a disk connected through supported SAS controllers or supported on-board SATA controllers

Servers that you use in the infrastructure must be listed in the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility. VMware supports only servers that are listed in the guide. Supported hardware is tested for operational functionality and is safer to deploy than unsupported systems.

Although the minimum requirement is 4 GB RAM, provide at least 8 GB of RAM to take full advantage of ESXi features and run virtual machines in typical production environments.

5-8 ESXi Booting Requirements

Consider the following boot requirements:

- ESXi hosts support boot from Unified Extensible Firmware Interface (UEFI).
- Boot systems from hard drives, CD-ROM drives, or USB media.
- vSphere Auto Deploy supports UEFI.
- Starting with vSphere 6.5, ESXi supports secure boot if it is enabled in the hardware.
- For vSphere 7, use high-endurance media for the boot partition.

ESXi 7 has the following boot disk requirements:

- At least 8 GB for USB or SD devices.
- 32 GB for other device types such as HDD, SSD, or NVMe.
- A boot device must not be shared between ESXi hosts.

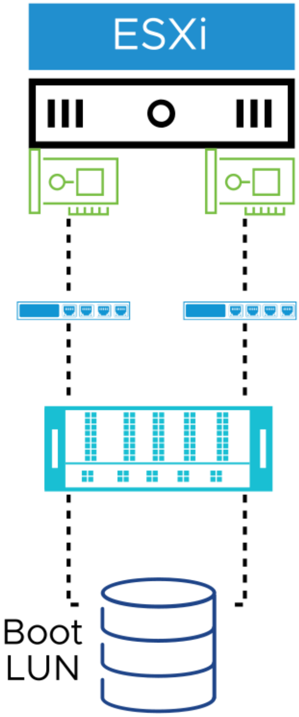
ESXi 7 supports installing on and booting from the following storage systems:

- SATA disk drives
- Serial-attached SCSI (SAS) disk drives
- Dedicated SAN disks on Fibre Channel or iSCSI
- USB devices (8 GB or more recommended)
- Software Fibre Channel over Ethernet (FCoE)

The recommended ESXi 7 installation options are as follows:

- An 8 GB USB or SD and an extra 32 GB local disk. The ESXi boot partitions reside on the USB or SD, and the ESX-OSData volume resides on the local disk.
- A local disk with a minimum of 32 GB. The disk contains the boot partitions and ESX-Osdata volume.
- A local disk larger than 128 GB. The disk contains the boot partitions, ESX-OSData volume, and VMFS datastore.

5-9 Design Decision: Evaluating the Use of ESXi Boot from SAN



Should ESXi hosts boot from SAN?

Boot from SAN is a good option in the following situations:

- Maintenance of local storage is undesirable.
- Hosts have diskless hardware, for example, blade systems.

Boot from SAN offers the following benefits:

- Cheaper servers
- Easier server replacement
- Less wasted space
- Easier backup processes
- Improved management
- Better reliability

Requirements and configuration tasks vary by storage protocol.

If you use boot from SAN, the following benefits are available for your environment:

- Servers can be more dense and run cooler without internal storage.
- You can replace servers and have the new server point to the old boot location.
- Servers without local disks often take up less space.
- You can back up the system boot images in the SAN as part of the overall SAN backup procedures. You can also use advanced array features, such as snapshots, on the boot image.
- Creating and managing the operating system image is easier and more efficient.
- You can access the boot disk through multiple paths, which protect the disk from being a single point of failure.

Boot from SAN can provide numerous benefits to your environment. However, sometimes you should not use boot from SAN for ESXi hosts. Before you set up your system for boot from SAN, decide whether boot from SAN is appropriate for your environment.

For environments that boot from a SAN or use vSphere Auto Deploy, the ESX-OSData volume for each ESXi host must be set up on a separate SAN LUN. However, if /scratch is configured not to use ESX-OSData, you do not need to allocate a separate LUN for /scratch for each host. You can colocate the scratch regions for multiple ESXi hosts onto a single LUN. The number of hosts assigned to any single LUN should be weighed against the LUN size and the I/O behavior of the virtual machines.

For detailed considerations to implement this solution, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-8AE88758-20C1-4873-99C7-181EF9ACFA70.html.

5-10 Design Decision: Selecting Blade or Rack-Mounted Servers

What hardware type should be selected, blade or rack-mount servers?

Both blade and rack-mount servers are viable virtualization solutions.

Server Type	Benefits	Drawbacks
Blade	• Modularity makes it easier to install, replace, or scale. • Reduced cabling and wiring because midplane helps reduce costs. • Sharing components such as power supplies, fans, DVD drives, and consoles can decrease capital expenses.	• Modularity locks in a vendor. • The form factor makes it difficult to apply a scale-up model. • Blade chassis redundancy comes at a higher cost because ROI can be achieved with only a fully populated chassis.
Rack-Mount	• More I/O expansion options provide for scaling within a box. • Huge capacity provides better consolidation ratios compared to a single blade.	• With higher consolidation ratios, the RTO increases. • A bigger footprint increases housing costs, especially in large implementations. • Sharing components can become a single point of failure.

The choice between a blade or rack-mount server might not be a purely technical decision. For example, the organization might already have a relationship with a specific vendor. If a specific blade or rack-mount server meets the requirements of the design, this solution is a viable virtualization solution.

Both options have advantages and disadvantages.

Blade server modularity makes replacing a blade easy, particularly if it has no local disk. The only configuration after replacement might be to configure the BIOS. The ease of adding blades to a chassis is also useful for quickly adding processing power to a cluster that is built on blade servers. However, blade servers from different vendors are not interchangeable, which locks you into a single vendor.

Buying a blade chassis is cost-efficient only to fill the chassis with more blades. Blade servers offer attractive compute density and large amount of computing power can be housed in a small space, which reduces data center real estate costs. Shared components such as power supplies, fans, and DVD drives can reduce the overall hardware costs. However, blade servers are typically more expensive. So you must purchase enough blade servers to realize the cost savings. Shared components can also become a single point of failure. This single point of failure includes the chassis itself. This risk can be mitigated by using features like vSphere HA.

5-11 Design Decision: ESXi Host Installation

Which method should be used to install and configure ESXi hosts?

When installing ESXi on five or more hosts, consider using the following automated methods:

- Native scripted installation utility
- vSphere Auto Deploy
- Third-party automated installers that support ESXi, which, if already in use, can be adapted for ESXi.

Whenever possible, use automated installers to install ESXi. Use interactive installations for a small installation of fewer than five hosts. If the organization already has a third-party installer and the installer can be adapted to install ESXi, use that installer. ESXi includes a native scripted installation utility if a third-party installer is not available.

For managing large deployments efficiently, use vSphere Auto Deploy.

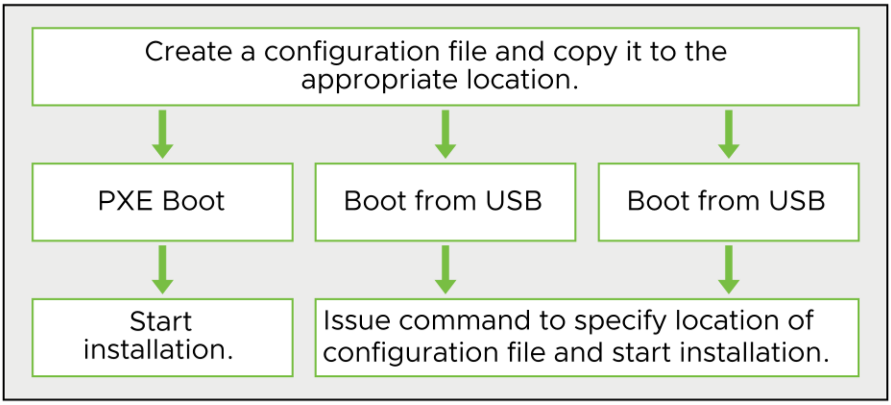
By using vSphere Auto Deploy, experienced system administrators can manage large deployments efficiently. Unless stateless caching or a stateful install is used, vSphere Auto Deploy does not store the ESXi configuration or state on the host disk. Instead, the state is managed through an image profile and other host attributes are managed through host profiles.

5-12 Evaluating the Use of Native Scripted Installations

Scripted installations are an efficient way to deploy multiple similar ESXi hosts and the native scripted installation utility is a good option for performing unattended ESXi installations.

The installation configuration file must use the supported commands, and administrators must be comfortable working with command-line constructs.
The installation configuration file must reside in one of the following locations:

- FTP server
- HTTP/HTTPS server
- NFS server
- USB flash drive
- CD-ROM drive


Create a configuration file and copy it to the appropriate location.
PXE Boot
Boot from USB
Boot from USB
Start installation.
Issue command to specify location of configuration file and start installation.

Scripted installations using the native utility supports preboot execution environment (PXE) booting the ESXi installer or booting it from a CD/DVD or USB drive.

For information about scripted installations and the support commands to use in the scripts, see VMware vSphere documentation at https://docs.vmware.com/en/VMware-vSphere/index.html.

5-13 Evaluating the Use of vSphere Auto Deploy

Using vSphere Auto Deploy has the following benefits:

- Large numbers of ESXi hosts can be deployed quickly and easily.
- A standard ESXi image can be shared across many hosts.
- Two deployment modes are available.

Mode	Host Image Storage	Operation
Stateless Caching	Provision host without configuration or state on the host disk.	An image profile defines the image that the host is provisioned with. This image is cached locally. Other host attributes are managed through host profiles. A host that uses Auto Deploy for stateless caching still must connect to the vSphere Auto Deploy Server and the vCenter Server.
Stateful installs	Provision host with the image stored to disk.	On subsequent boots, the host boots from the disk.

With vSphere Auto Deploy, many hosts can be rapidly deployed. vSphere hosts are network booted from a central vSphere Auto Deploy server where the ESXi software is installed directly into the server's memory. After installation, a host profile is used to configure the host. After configuration, the host is connected to vCenter Server where it is available to host virtual machines.

vSphere Auto Deploy simplifies ESXi host management by eliminating the necessity of maintaining a separate boot image for each host. You can share ESXi software images with all hosts running on matching hardware. Centralized image management eliminates the need to patch and update individual ESXi hosts. You can perform a single update to the shared image profile and reboot the vSphere hosts.

When a host is provisioned using vSphere Auto Deploy, the host image is decoupled from the physical server. The host can be recovered without recovering the hardware or restoring from a backup.

vSphere Auto Deploy stateless caching PXE boots the ESXi host and loads the image in memory. However, when the host profile is applied to the ESXi host, the image running in memory is copied to a boot device. The saved image acts as a backup if the PXE infrastructure or vSphere Auto Deploy server is unavailable. If the host must reboot and cannot contact the DHCP, TFTP, or vSphere Auto Deploy server, the network boot times out and the host reboots using the cached disk image. With stateless caching or stateful deployments, the host can be booted from the local disk even if the vSphere Auto Deploy server fails.

Stateless caching does not guarantee that the image is current or that the vCenter Server system is available after the boot. The primary benefit is that stateless caching provides a means to boot the host to troubleshoot and resolve problems that prevent a successful PXE boot. Unlike stateless ESXi hosts, stateless caching requires a dedicated boot device to be assigned to the host.

5-14 vSphere Auto Deploy Dependencies

vSphere Auto Deploy has a high infrastructure overhead because it has many dependencies:

- PXE boot infrastructure requires a DHCP and TFTP server.
- Approximately 400 MB of storage is required for each image profile, and best practice is to allocate 2 GB.
- ESXi host profiles.
- A remote Syslog server.
- vSphere ESXi Dump Collector.
- Authentication proxy server.

You create image profiles and configure vSphere Auto Deploy with the vSphere Client.

When a vSphere host is powered on, a PXE-based network boot is started. The PXE boots require a Dynamic Host Configuration Protocol (DHCP) and Trivial File Transfer Protocol (TFTP) server, which must be set up outside vSphere Auto Deploy.

The vSphere Auto Deploy server uses a repository to store the necessary data. The data includes the rules and rule sets that you create and the VIBs and image profiles that you specify in your rules. The best practice is to allocate 2 GB to have enough room for four image profiles and some extra space. Each image profile requires approximately 400 MB. Determine how much space you must reserve for the vSphere Auto Deploy repository by considering how many image profiles you expect to use.

A rules engine on the vSphere Auto Deploy server uses information that is provided by the host during the PXE boot. The rules engine uses this information to determine the proper image and host profiles must configure the host and the location in the vCenter Server system to place the host.

Setting up logging on a remote host is important for hosts provisioned with vSphere Auto Deploy that have no local storage. You can optionally install the vSphere Syslog Collector to collect logs from all hosts. Likewise, you should use vSphere ESXi Dump Collector to keep core dumps on a network server for use during debugging.

An authentication proxy server is especially helpful when used with vSphere Auto Deploy. You set up a reference host that points to the authentication proxy server and create a rule that applies the reference host’s profile to any ESXi host provisioned with vSphere Auto Deploy.

For more information about implementing vSphere Auto Deploy, see *VMware ESXi Upgrade* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.esxi.upgrade.doc/GUID-65B5B313-3DBB-4490-82D2-A225446F4C99.html.

5-15 ESXi Host Configuration

The Host Profiles feature has the following benefits:

- Faster than manual configuration
- Less error-prone than manual configuration
- Can be used for compliance monitoring
- Reduces the need for developing custom postinstallation scripts
- Does not require knowledge of scripting interfaces, such as PowerCLI or vSphere CLI

Consider using postinstallation scripts or vRealize Orchestrator workflow for features or components that cannot be configured by using Host Profiles.

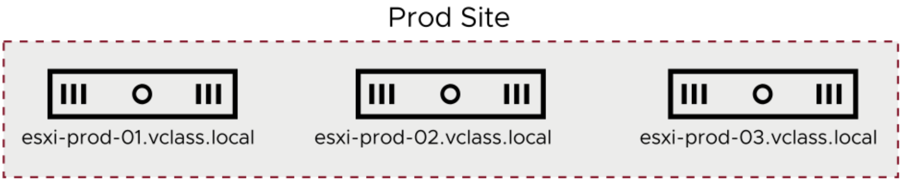
Host profiles eliminate per-host, manual, or UI-based host configuration and maintain a configuration consistency and correctness across the data center. Host profiles policies capture the blueprint of a known, validated reference host configuration, including the networking, storage, security, and other settings.

With host profiles, the time required to set up, change, audit, and troubleshoot configurations decreases by using centralized configuration and compliance checking. The use of host profiles reduces labor costs and minimizes the risk of downtime for applications because of misconfigured systems.

5-16 Assigning IP Names and Addresses to ESXi Hosts

Use a simple, descriptive naming convention such as `esxi-`*`<locationcode>-<##>.<domain_name>.`*
Assigning a static IP address and name to each host helps avoid management connection problems:

- Locally configured IP addresses are best.
- If the design requires using reserved DHCP IP addresses, ensure that the DHCP server is highly available.
- Configure ESXi hosts as DHCP clients if the organization plans to use vSphere Auto Deploy to provision hosts.


Prod Site
esxi-prod-01.vclass.local
esxi-prod-02.vclass.local
esxi-prod-03.vclass.local

A static host name and IP address are important for maintaining management access. If the IP address of the host changes, the vCenter Server system must be reconfigured to regain management control. You might also be required to configure or reissue certificates.

Locally configured host names and IP addresses are best because this configuration removes any dependence on DHCP. However, a reserved IP address can be assigned by a DHCP server. If a DHCP server is used, it should be highly available.

The ESXi host design should also include a host naming convention that uses simple, descriptive, and easy to understand host names. A standardized naming convention results in easier use, management, and troubleshooting, which all reduce operational complexity.

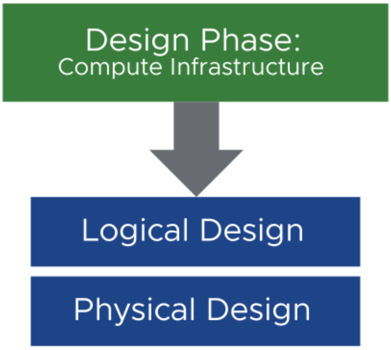
5-17 Lab 6: Designing the Compute Infrastructure

Create a compute infrastructure design:

1. Review the Conceptual Design
2. Evaluate Compute Infrastructure Design Options
3. Document the ESXi Host Physical Design

5-18 Review of Lab

The instructor facilitates a class discussion of the Design decisions of one or more teams.


Design Phase:
Compute Infrastructure
Logical Design
Physical Design

5-19 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Create a compute infrastructure design that includes the appropriate ESXi boot, installation, and configuration options
- Select the ESXi host hardware for the compute infrastructure

5-20 Key Points

- Design a consistent hardware configuration.
- The default hardware BIOS settings on servers might not always be the best choice for optimal performance.
- Verify that hosts are fully functional before installing ESXi.
- Use automated methods to install ESXi.
- Configure a static IP address and host name.

Questions?

Module 6

Storage Infrastructure

6-2 Importance

A storage platform design includes hardware specifications and configuration instructions to support the storage requirements of the vSphere infrastructure.

You must design a storage infrastructure that performs well and achieves the following objectives:

- Prevents unauthorized access to business data
- Protects data from hardware and software failures
- Protects data from malicious or accidental corruption

6-3 Module Lessons

1. Storage Platform
2. Calculating Storage Capacity Requirements
3. Storage Management Features

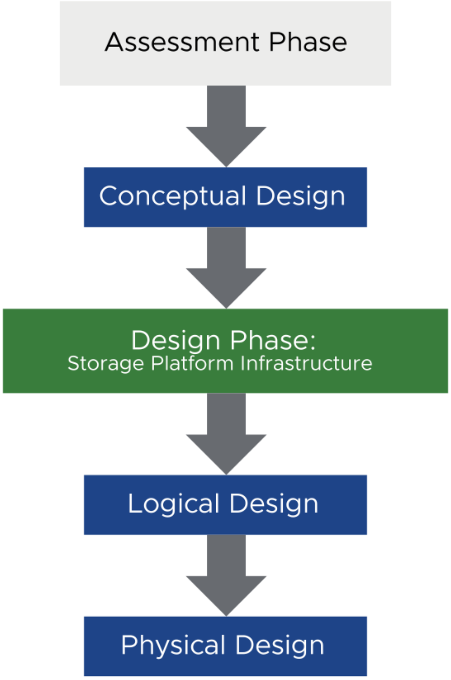
6-4 Lesson 1: Storage Platform

6-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of different storage platform solutions
- Design a storage platform architecture that meets the needs of the vSphere environment

6-6 Designing the Storage Platform Infrastructure



When you create a design for the storage platform infrastructure, you make decisions in the following areas:

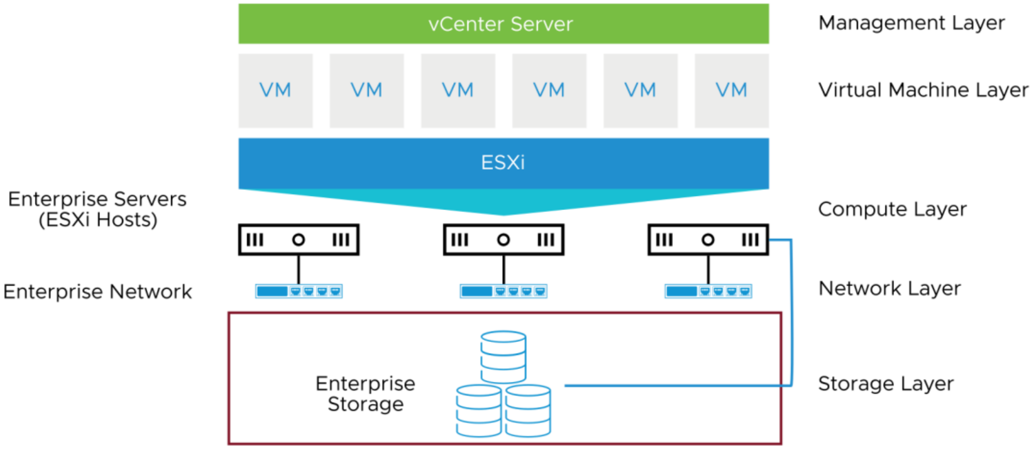
- Storage types to use
- Storage path redundancy
- Use of SSDs
- Use of persistent memory (PMem) and remote direct memory access (RDMA)
- Total storage capacity requirements
- Datastore size
- Number of datastores

When creating the logical design, you look at each of the major vSphere infrastructure components and make the appropriate design decisions. You always consider the organization's goals, requirements, and constraints, and current best practices.

Customer involvement during all phases of the process is key to the success of the project. During the assess and design phases, stakeholders must understand vSphere and virtualization concepts so that they can provide inputs into the design process.

6-7 Review of the Storage Layer in a vSphere Infrastructure

In a vSphere infrastructure, storage arrays are connected to and shared between groups of servers through storage area networks.



With shared storage, you can aggregate the storage resources and improve flexibility, efficiency, and reliability. Using ESXi with a SAN, also supports centralized management, failover, and load-balancing technologies.

6-8 High-Level Storage Design Guidelines

Consider the following guidelines whenever possible:

- Design storage to meet the diverse needs of applications, services, administrators, and users.
- Strategically align business applications and the storage infrastructure to reduce costs, boost performance, improve availability, provide security, and enhance functionality.
- Build a modular storage solution that considers both capacity and performance.
- Design and configure redundancy for storage network components, storage paths (multipathing), storage processors, and LUN configurations (RAID).
- Avoid oversubscription by distributing workloads.

Verify that storage devices and arrays are listed in the VMware Compatibility Guide.

A good storage design should meet the following objectives:

- The storage design should help to reduce costs. This objective is one of the simpler goals to achieve in the vSphere infrastructure. Server and storage consolidation alone reduce storage costs by reducing the number of required storage logical unit numbers (LUNs). Features such as vSphere Thin Provisioning help in reducing costs.
- The design should not impede performance. Performance is achieved by providing the necessary storage bandwidth, which reduces contention and its accompanying latency. Providing necessary storage bandwidth might mean adding additional paths to storage, adding additional back-end disks to storage LUNs, configuring more cache, and so on.
- The design should improve availability and storage availability greatly affects applications. Server and storage consolidation mean that availability is even more critical. A storage failure affects more services and applications. If a single point of failure cannot be removed, list it as a risk.
- The design must be secure. Security in a design is achieved by access control and isolation, as required.
- The design should help to enhance the functionality of the infrastructure. Configuring storage to support vSphere features such as vSphere vMotion, vSphere HA, and vSphere FT.

- The storage design should be as modular as possible. A modular storage design means that you can expand the design by adding storage components to the existing storage components. A module storage solution considers scaling components such as back-end disks, array front end, and the storage network infrastructure.
- Oversubscription leads to contention and contention leads to higher latencies. Oversubscription can occur on the host, the storage paths, or the storage array. Storage oversubscription is reduced by distributing the workload across multiple components, paths, and LUNs. You can also reduce latency by reducing the number of hops in the storage path. Configure the storage network so that a host traverses fewer switches to reach the storage array.

6-9 Design Decision: Choosing a Storage Type

Which storage types should be used?

The key to performance is proper design and configuration.

Several technologies can support the storage needs of a VM:

- PMem
- RDMA
- Fibre Channel
- iSCSI
- NAS
- vSAN
- vSphere Virtual Volumes

The decision to implement a technology should be based on the following considerations:

- The organization’s current in-house expertise and installation base
- The cost, including both capital and long-term operational expenses
- The organization’s current relationship with a storage vendor

The key to performance is the specific design and configuration of the storage components. Sometimes, the choice of technology is determined by nontechnical factors such as cost or existing in-house expertise. The key to designing storage by using these technologies is reducing contention and latency.

6-10 Comparing Storage Types

Type	Protocols	Transfers	Interface
RDMA	Infiniband, RoCE	Direct memory access	RDMA Host Channel Adapter (HCA)
Fibre Channel	FC/SCSI	Block access or LUN	FC HBA
Fibre Channel over Ethernet	FCoE/SCSI	Block access or LUN	Converged network adapter (hardware FCoE) NIC with FCoE support (software FCoE)
iSCSI	IP/SCSI/iSER	Block access or LUN	RDMA over Converged Ethernet (RoCE) iSCSI HBA or iSCSI enabled NIC (hardware iSCSI) Network adapter (software iSCSI)
NAS	IP/NFS	File (no LUN access)	Network adapter
vSAN	IP	Block access	Network adapter
vSphere Virtual Volumes	FC/FCoE/iSCSI/NFS/SCSI	Storage transports for the protocols	Same as protocol used

The characteristics of different storage types affect your storage design decisions.

6-11 vSphere Features Compared by Storage Type

Support for some vSphere functionality depends on the storage type.

Type	Boot VM	vSphere vMotion	Datastore	RDM	HA/DRS	Storage APIs Data Protection
Local storage	Yes	Yes	VMFS	No	No	Yes
Fibre Channel	Yes	Yes	VMFS	Yes	Yes	Yes
iSCSI	Yes	Yes	VMFS	Yes	Yes	Yes
NAS over NFS	Yes	Yes	NFS	No	Yes	Yes
vSAN	Yes	Yes	vSAN	No	Yes	Yes
vSphere Virtual Volumes	Yes	Yes	vVol	No	Yes	Yes

6-12 Fibre Channel Storage Design Considerations (1)

ESXi supports different types of storage arrays:

- Active-active
- Active-passive
- Asymmetric Logical Unit Access (ALUA)

For performance in a Fibre Channel environment, use single-initiator zoning:

- Prevents RSCN messages from crossing zone boundaries and affecting normal I/O traffic.
- Prevents problems and misconfigurations on the SAN.

When using Fibre Channel storage, a Host Bus Adapter (HBA) is required. The HBA must be supported by both the array and the VMware Hardware Compatibility list.

ESXi supports different storage systems and arrays:

- Active-active: This storage system allows access to the LUNs simultaneously through all the storage ports that are available without significant performance degradation. All the paths are active always, unless a path fails.
- Active-passive: In this storage system, one storage processor actively provides access to a given LUN. The other processors act as backup for the LUN and can actively provide access to other LUN I/O. I/O can be successfully sent only to an active port for a given LUN. If access through the active storage port fails, one of the passive storage processors can be activated by the servers accessing it.
- Asymmetrical Logical Unit Access (ALUA) compliant: This storage system provides different levels of access per port. With ALUA, hosts can determine the states of target ports and prioritize paths. The host uses some of the active paths as primary, while using other paths as secondary.

A single-initiator zone is created by configuring a single HBA in a zone with one or more target devices. In a multi-initiator zone, multiple initiators exist in a zone, with one or more target devices. Single-initiator zones are the preferred choice. Single-initiator zones are more secure because initiators cannot communicate with one another. Single-initiator zones also eliminate registered state change notification (RSCN) messages from one initiator interrupting the normal I/O operations of another initiator. With multi-initiator zoning, multiple initiators entering or leaving the fabric interrupt the normal transfer of data by other initiators.

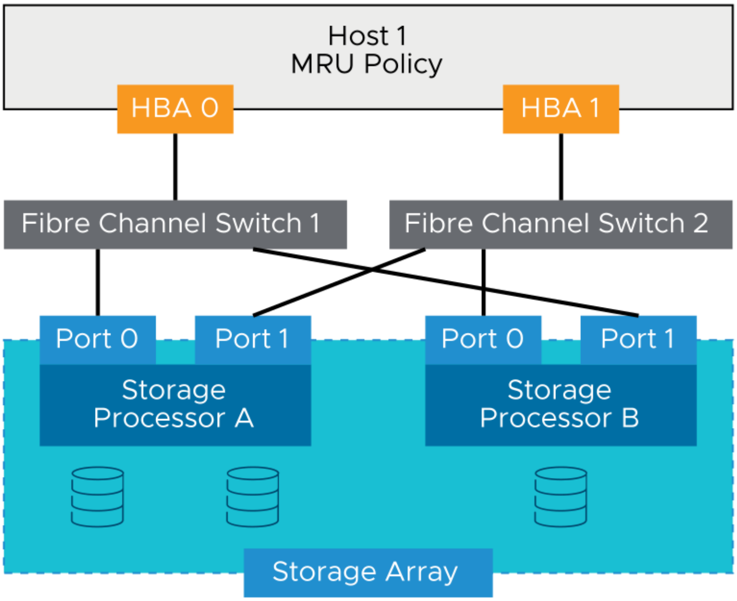
For additional information about Fibre Channel storage, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-8AE88758-20C1-4873-99C7-181EF9ACFA70.html.

6-13 Fibre Channel Storage Design Considerations (2)

You can use different methods to control access to Fibre Channel datastores. The following methods are listed from the most general to the most specific:

- Fibre Channel zoning
- LUN masking that is configured at the array is the most reliable
- Datastore permissions in the vCenter Server inventory

To ensure performance and availability in the zone configuration, include a path from each initiator on a host to each storage processor on the array. This configuration provides availability with redundancy.



Several methods are used to control access to Fibre Channel datastores:

- Fibre Channel zoning restricts access to data stored on a specific array. Devices outside a zone are not visible to the devices inside the zone. Typically, zones are created for each group of servers that access a shared group of storage devices and LUNs.
- Zoning is configured at the Fibre Channel switch, and LUN masking is configured at the array, the ESXi host, or both. LUN masking is also called selective storage presentation, access control, and partitioning, depending on the vendor. LUN masking makes a LUN

invisible when a target is scanned. The administrator configures the disk array so that each server or group of servers can see only certain LUNs. Masking capabilities for each disk array are vendor-specific, as are the tools for managing LUN masking. LUN masking configured on an array is enforced more consistently than masking done on ESXi hosts. Masking done at the ESXi host level is often based on controller, target, and LUN numbers, all of which can change with the hardware configuration.

- Datastore permissions provide the most specific access control. Configuration is done by using the standard vCenter Server permissions mechanism. Permissions prevent users from accessing only the datastores, not the ESXi hosts.

6-14 iSCSI Storage Design Considerations (1)

For ESXi to discover iSCSI storage, you must have iSCSI adapters:

- Adapter cost is usually the deciding factor.
- In general, hardware adapters provide better performance because they offload traffic processing.

Consider the following performance guidelines when designing the iSCSI configuration:

- Use dedicated LUNs and RAID groups for ESXi hosts.
- Assign LUNs on a cluster basis to each of the ESXi hosts in the cluster.

iSCSI Adapter	Description
Independent hardware iSCSI adapter	Third-party adapter that offloads iSCSI and network processing and management from host
Software iSCSI adapter	Uses standard NICs to connect your host to a remote iSCSI target
Dependent hardware iSCSI adapter	Third-party adapter that depends on VMware networking and iSCSI configuration and management interfaces
VMware iSER adapter	Uses an RDMA-capable network adapter to connect your host to a remote iSCSI target

When logical unit numbers (LUNs) are assigned, each shared LUN can be accessed through all hosts that have been granted access. Therefore, LUNs and RAID groups used for ESXi hosts should be dedicated.

Assign LUNs on a cluster basis to each of the ESXi hosts that are in the cluster. All hosts can then run all VMs on the assigned storage. This approach avoids problems with features such as vSphere vMotion and vSphere HA.

Certain types of iSCSI adapters depend on VMkernel networking. These adapters include the software or dependent iSCSI adapters and the VMware iSCSI over RDMA (iSER) adapter.

iSER can be configured on ESXi hosts so that the iSCSI framework on the host can use the RDMA transport instead of TCP/IP.

For additional information about iSCSI storage, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-8AE88758-20C1-4873-99C7-181EF9ACFA70.html.

6-15 iSCSI Storage Design Considerations (2)

The following methods, from most general to specific, can be used to control access to iSCSI datastores:

- Physical or VLAN network segmentation
- Challenge Handshake Authentication Protocol (CHAP)
- Datastore permissions in the vCenter Server inventory

The following methods are available to control access to iSCSI datastores:

- Host access to an array can be controlled by using network segmentation. Network segmentation can be accomplished by using either physical networks or VLANs. This level of access control allows or denies access to the whole array.
- More specific access control is available by using the Challenge Handshake Authentication Protocol (CHAP). Per-target access control is available by assigning different CHAP passwords to different targets. Hosts that do not have the correct CHAP password cannot establish a connection to the target.
- The most specific access control uses the datastore permissions in the vCenter Server inventory. Permissions to access a datastore can be granted per user or per group.

6-16 VMFS Design Considerations

Place only one VMFS datastore on each LUN:

- A SCSI reservation on the LUN might be required by a host when updating VMFS metadata on a partition:
  - LUN access by other hosts is not allowed while the SCSI reservation exists.
  - Metadata updates on multiple partitions can increase latency and reduce performance.
- For storage devices that support T10 standard VAAI specification, ESXi uses more efficient hardware assisted locking

Size the VMFS datastore to achieve acceptable latency on the VMs:

- Determine latency for a specific array configuration using the projected IOPS and MBps.
- Performance is affected by the path components.
- Performance is heavily dependent on the storage array vendor. Consult the vendor documentation.

The datastore size also depends on the RTO for the data. You must be able to recover data within the RTO limits.

The main factor that determines the size of a VMFS datastore is how many virtual machines can run in the datastore with acceptable latency. Determining acceptable latency requires information about the application workloads.

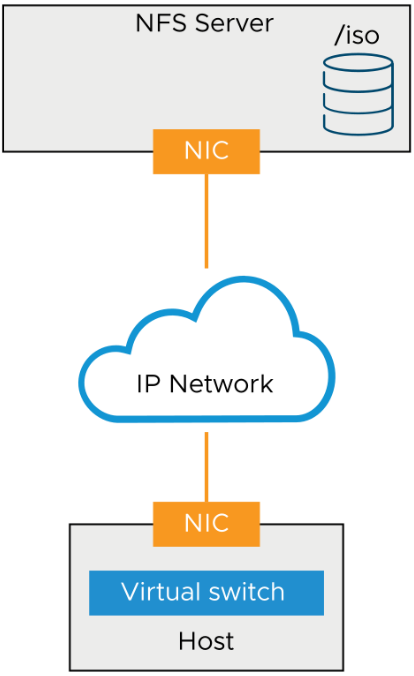
For example, if 10 VMs can access the datastore without exceeding the established average latency limit, and each VM requires an average of 10 GB of storage space, the datastore size must be at least 100 GB.

A general guideline is 10 to 15 virtual machines per VMFS datastore. However, updates to VMFS and storage typically increase this value over time. The number of virtual machines per LUN can be much higher if linked clones are used.

In a shared storage environment, when multiple hosts access the same VMFS datastore, specific locking mechanisms are used. These locking mechanisms prevent multiple hosts from concurrently writing to the metadata and ensure that no data corruption occurs. Depending on its configuration and the type of underlying storage, a VMFS datastore can use different types of locking mechanisms. It can exclusively use the atomic test and set locking mechanism (ATSonly), or use a combination of ATS and SCSI reservations (ATS+SCSI).

For storage devices that support T10 standard-based VAAI specifications, VMFS provides ATS locking, also called hardware assisted locking. The ATS algorithm supports discrete locking per disk sector. All newly formatted VMFS5 and VMFS6 datastores use the ATS-only mechanism if the underlying storage supports it, and never use SCSI reservations. When you create a multi-extent datastore where ATS is used, vCenter Server filters out non-ATS devices. With this filtering, you use only those devices that support the ATS primitive. In certain cases, you might need to turn off the ATS-only setting for a VMFS5 or VMFS6 datastore

6-17 NFS Storage Design Considerations



NFS storage requires a standard Ethernet network. The higher the speed for the Ethernet device, the better the performance when accessing a datastore.

ESXi supports NFS protocols version 3 and 4.1.

NFS 4.1 does not support:

- vSphere Storage DRS
- Site Recovery Manager

The following methods, from most general to specific, can be used to control access to NFS datastores:

- Physical or VLAN network segmentation
- Not mounting an NFS datastore to a host
- Datastore permissions in the vCenter Server inventory

Determining the size for an NFS datastore is accomplished in the same way as for VMFS datastores.

When using NFS, you require a standard Ethernet network. Customized HBAs or adapters are not required.

Several methods are available to control access to an NFS datastore. Choose a method based on how specific the access control must be:

- Network segmentation provides the least specific access control. This method prevents an entire host from accessing the NFS array. The administrator of the host cannot mount or unmount a datastore from the array.
- If network segmentation is not used, the administrator can mount or unmount specific NFS datastores from the array. However, all hosts in a VMware cluster typically need access to the same storage resources for VM migration and failover to work properly.
- Datastore permissions provide the most specific access control. An administrator can use datastore permissions in the vCenter Server inventory to control access per user or group.

Site Recovery Manager 8.2 released in May 2019 does not support NFS 4.1 datastores.

For additional information about NFS datastores, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-8AE88758-20C1-4873-99C7-181EF9ACFA70.html.

6-18 Storage Path Redundancy Considerations

Multipathing provides both availability and load balancing for performance.

Consider these guidelines when designing multipathing:

- Always configure storage multipathing.
- For availability, use a minimum of two active paths to a LUN:
  - Four paths are better because this configuration accommodates more types of failures.
- Configure paths using two HBAs or NICs, two switches, and two array storage processors:
  - If using HBAs, use two single-port storage HBAs rather than a single dual-port storage HBA.
- Do not change the path policy that the system sets unless the implications of the change are recommended by the storage vendor.
- Use vendor recommendations for third-party multipathing plug-ins.

Providing redundancy in the storage design is essential to increase availability, scalability, and performance. Configure the following components for storage redundancy:

- Storage network components
- Storage paths
- Storage processors
- LUN configurations (RAID)

Configure a minimum of four paths to a LUN. With all paths correctly configured, the failure of an HBA, switch, or array storage processor does not prevent access to data.

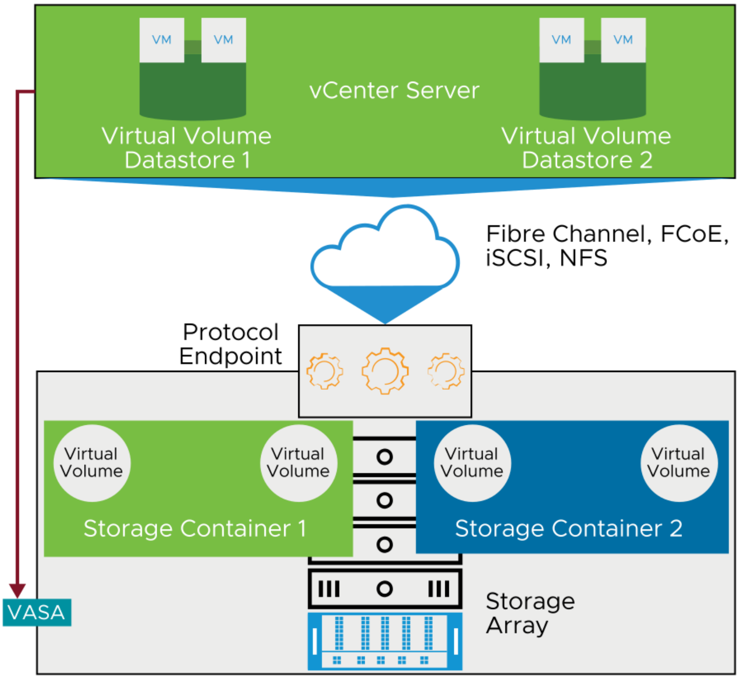
Always consult array documentation for specific multipathing policy support. However, in general, configure multipathing policies as follows:

- Most Recently Used (MRU) for active-passive arrays to avoid path thrashing
- Fixed or round robin as an option for active-active arrays
- MRU or round robin as an option for ALUA arrays
- MRU for virtual port storage arrays

6-19 vSphere Virtual Volumes Considerations

vSphere Virtual Volumes define virtual disk containers that are independent of the underlying storage hardware with the following characteristics:

- Storage is VM-centric.
- Requires a storage array that supports virtual volumes and can integrate with vSphere API for Storage Awareness.
- NFS 3, NFS 4.1, iSCSI, Fibre Channel, and FCoE are supported.
- IPv6 is supported end to end.
- From vSphere 7, SCSI-3 persistent reservations are supported for Microsoft WSFC clusters.



vSphere Virtual Volumes is a virtual machine disk management and integration framework that exposes virtual disks as the primary unit of data management for storage arrays. With vSphere Virtual Volumes, the virtual disk becomes the primary unit of data management at the array level. This setup turns the virtual datastore into a VM-centric pool of capacity.

You can run storage operations with virtual disk granularity and provision native array-based data services, such as compression, snapshots, deduplication, encryption, and so on, to individual VMs. As a result, you can provide the right storage service levels to each individual virtual disk within a VM.

vSphere Virtual Volumes requires a compatible storage array system. In most cases, a software solution such as a virtual storage appliance from one of the supporting vendors is supported for testing management workflows, operations, and functionalities.

Depending on the vendor-specific implementation, storage array system might or might not require a firmware upgrade to support vSphere Virtual Volumes. Check with your storage vendor for detailed information and configuration procedures.

SCSI-3 reservations support shared disks or volumes between VMs across nodes or hosts. This configuration is often used for Microsoft WSFC clusters. With this new enhancement, RDMs can be removed.

Check the VMware Compatibility Guide for up-to-date information on storage arrays that support vSphere Virtual Volumes.

6-20 vSphere Virtual Volumes Benefits

vSphere Virtual Volumes offers several benefits:

- Faster storage operations:
  - Intensive operations such as snapshot, cloning, and replication are offloaded to the storage array.
- Simplified storage operations:
  - Operational dependencies between the vSphere administrator and storage administrator are eliminated.
  - The storage administrator retains control of the storage resources, and the vSphere administrator consumes the published storage array capabilities.
- Service-level delivery:
  - vSphere administrators predefine VM storage policies to meet the various needs of each application.
- Flexible storage consumption:
  - The vSphere administrator can change policies, and the necessary infrastructure changes are configured through automation.

For both the vSphere administrator and storage administrator, vSphere Virtual Volumes simplifies management over the existing operational model. With vSphere Virtual Volumes, you can separate the provisioning and consumption of storage for VMs.

With vSphere Virtual Volumes, the classes of service are no longer physical pre-allocations but are logical entities controlled and automated entirely by software and interpreted through policies. By associating one or more VMs to the right policy, the provisioning and instantiation of storage service levels are automated for those VMs. Automated policy enforcement simplifies the monitoring process and ensures the compliance of storage service levels by the application.

With policy-driven automation, agile storage consumption for VMs results in faster provisioning for new applications.

Environments that have dynamic resource requirements can benefit from vSphere Virtual Volumes and storage-policy-based management. For example, during a seasonal workload increase, an automated process might update a given storage policy to deliver higher levels of performance.

Multitiered applications can benefit from using vSphere Virtual Volumes. Each VM of the multitiered application can have an appropriate policy attached to it at the time of deployment, and each VM is instantiated as a virtual volume with the required storage characteristics, even when deployed to the same container.

For more information about working with vSphere Virtual Volumes, see *vSphere Storage* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.storage.doc/GUID-8AE88758-20C1-4873-99C7-181EF9ACFA70.html.

6-21 vSAN Considerations for Storage Configuration

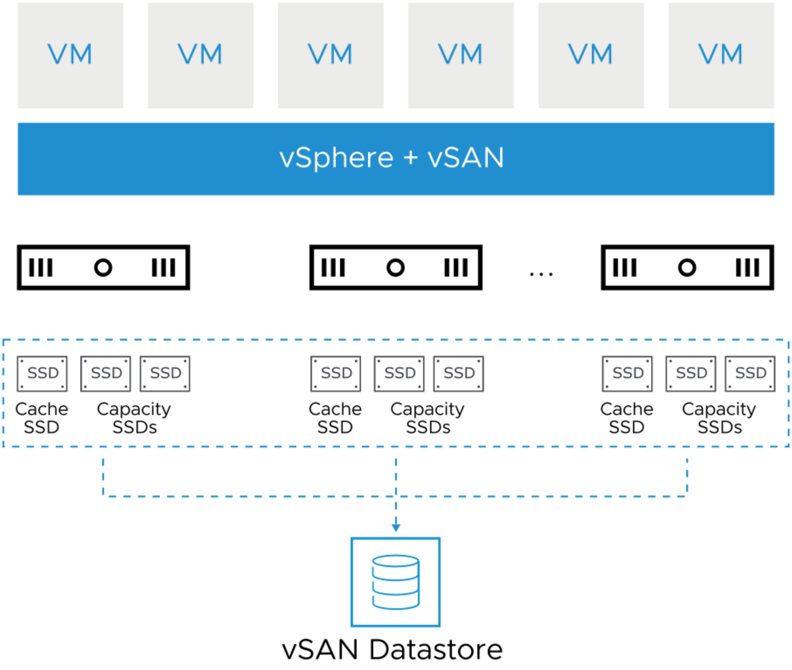
vSAN provides a simple and cost-effective storage configuration:

- An all-flash device configuration is used to create storage pools.

  A hybrid configuration of magnetic drives and flash devices can also be used.

- Internal or external DAS storage can be used.
- With the vSAN iSCSI target service, hosts and physical workloads that reside outside of the vSAN cluster can access the vSAN datastore.
- All-flash configurations support the following functions:
  - — Deduplication and compression
  - — RAID 5 or RAID 6 erasure coding
- vSAN has a default VM storage policy.

  You create custom storage policies to ensure that VMs get the appropriate level of service.



vSAN virtualizes the local physical storage resources of ESXi hosts and turns them into pools of storage. These pools of storage can be assigned to VMs and applications according to their quality-of-service requirements. vSAN removes the need for external shared storage and is implemented directly in the ESXi hypervisor.

All-flash configurations can take advantage of features that are not available in hybrid configurations:

- Deduplication and compression are space-savings features that can reduce the amount of storage consumption by as much as seven times. Actual reduction numbers vary. For example, spreadsheet files compress well, but video files do not compress well.
- Erasure coding provides the same levels of redundancy as mirroring (RAID 1), while consuming less storage capacity. Use of erasure coding reduces capacity consumption by as much as 50 percent versus mirroring at the same fault tolerance level.

The vSAN iSCSI target service provides native iSCSI support within vSAN. The main use cases are supporting physical servers and also Microsoft Clustering Technologies that require shared disks. One can create iSCSI targets and LUNs on vSAN and use the iSCSI initiator to access the storage.

For information about vSAN, see VMware vSAN at
https://storagehub.vmware.com/t/vmware-vsan/.

6-22 vSAN Considerations for Workloads

vSAN remains a platform of choice for a wide variety of workloads:

- Large or small
- Test and development
- Business-critical applications
- Virtual desktop infrastructure (VDI)

vSAN provides the following features:

- Disaster and downtime avoidance: Stretched clusters protect VMs across data centers, not only racks.
- Remote office/branch office (ROBO): Two-node clusters are used for this scenario.
- Disaster recovery: You can use Site Recovery Manager with vSAN and vSphere Replication to orchestrate the recovery of multiple VMs.

vSAN is enterprise-class storage for hyperconverged infrastructure (HCI). vSAN simplifies storage configuration and VM provisioning activities. Seamless integration with vSphere and the entire VMware stack makes it the simplest storage platform for business-critical workloads, virtual desktop infrastructure (VDI), remote office IT, disaster recovery, and DevOps infrastructures. Customers of all industries and sizes use vSAN to run their most important applications.

For information about vSAN use cases, see VMware vSAN at
https://storagehub.vmware.com/t/vmware-vsan/.

6-23 vSAN Considerations for Clusters

A vSAN cluster has the following requirements:

- vSAN components must be listed in the VMware Compatibility Guide.

  vSAN ReadyNodes are typically the easiest, most flexible approach when considering deployment methods.

- The cluster must have a minimum of three hosts that contribute capacity.

  For stretched clusters and two-node clusters, you need two hosts and a witness node

Follow vSAN best practices to improve performance and throughput:

- Use a dedicated or shared 10 GbE or 25 GbE network:
  - Place vSAN traffic on a distributed switch and configure Network I/O Control to guarantee bandwidth.
  - For all-flash configurations, use a dedicated or shared 10 GbE NIC.
- Provision one additional physical NIC as a failover NIC.

vSAN does not support:

- vSphere DPM
- vSphere Storage I/O Control

Hardware must be chosen from the vSAN compatibility list. Although hardware might be on the vSphere compatibility list, it also must be qualified for vSAN. Hardware choice is important for performance and supportability by VMware.

vSAN ReadyNodes are x86 servers, available from all the leading server vendors that have been preconfigured, tested, and certified for vSAN. Each vSAN ReadyNode is optimally configured for vSAN with the required amount of CPU, memory, network, I/O controllers, and storage devices.

vSAN does not use external storage arrays and does not support array-based replication. All replications must be done at the host, at the individual virtual machine level.

vSAN can be managed by vSphere administrators, no specialized storage skills are required to administer vSAN systems. However, if the storage team must manage the vSAN systems, the tools used by the vSphere administration team must be used.

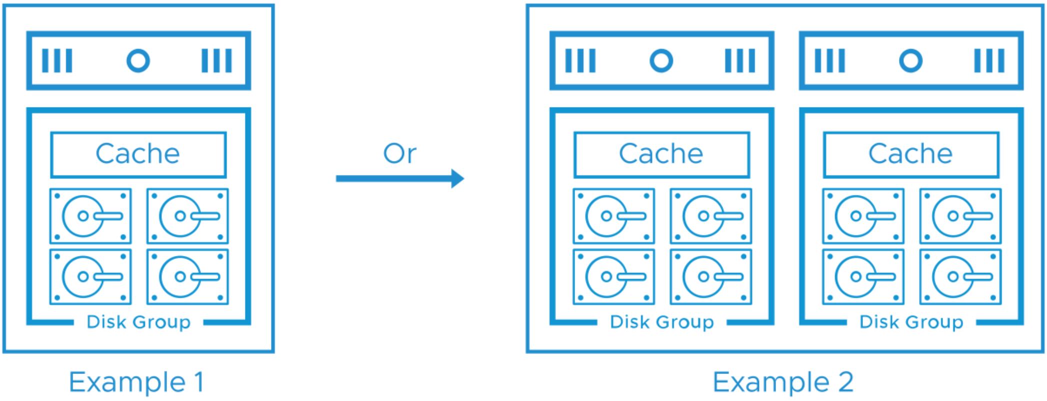
When working with vSAN, consider the following limitations:

- vSAN does not support hosts participating in multiple vSAN clusters. However, a vSAN host can access other external storage resources that are shared across clusters.
- vSAN does not support SE Sparse disks.
- vSAN does not support RDM, VMFS, diagnostic partition, and other device access features.

6-24 vSAN Considerations for Disk Devices

Storage design is controlled through disk groups and host count:

- Redundancy: In example one, a single disk group can result in a host failure.
- Flexibility: Scale up or scale out are both viable options.
- Performance: More cache devices mean more I/O paths.



Consider design example 1 in terms of availability: What happens if the cache device fails? Do my VMs go offline? Have I lost data?

Unless you changed the default policy to RAID 0, the default storage policy is RAID 1 with FTT=1, which means that you can tolerate one failed device. Your disk group goes offline, but an additional copy of the data on those devices exists in a separate failure domain.

Configuring additional disk groups is not a way to ensure availability. Desired availability is determined by using storage policy based management (SPBM). Until the failed device is replaced (and the disk group rebuilt), the objects on the six capacity devices remain offline and are impacted. The VM is uninterrupted because it uses the additional copy.

Design example 2 reduces the failure domain because a failed cache device only impacts the disk group associated capacity devices, instead of all disks in the host. This design also reduces the time to rebuild these impacted objects.

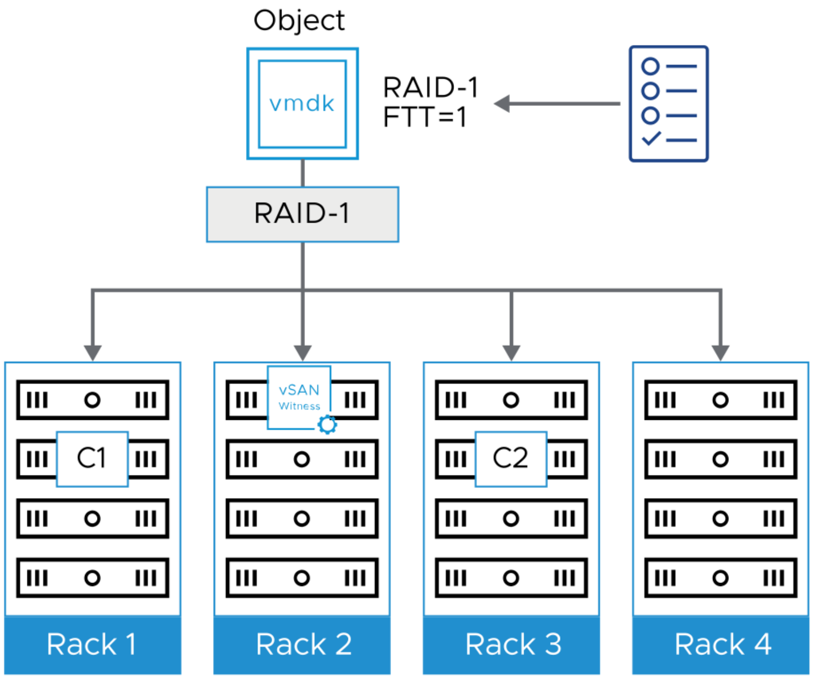
From vSAN 6.7 and later, vSAN enhances the way background I/O operations (such as rebuild) are handled, ensuring applications receive the necessary I/O in times of contention.

In terms of performance, creating two disk groups containing one cache device and capacity devices is advantageous. Testing reveals that moving disk groups to separate storage controllers can significantly boost the performance of a vSAN environment.

6-25 vSAN Considerations for Hardware Failure

vSAN supports the creation of explicit fault domains that have the following benefits:

- Increase availability
- Protect against rack and other types of failures
- Ensure that redundancy components of the same object do not live in the same rack



Using fault domains, you can protect against rack or chassis failure if your vSAN cluster spans across multiple racks or blade server chassis. You can create fault domains and add one or more hosts to each fault domain.

A fault domain consists of one or more vSAN hosts grouped according to their physical location in the data center. When fault domains are configured, vSAN can tolerate failures of entire physical racks and failures of a single host, capacity device, network link, or a network switch dedicated to a fault domain.

The Primary level of failures to tolerate policy setting for the cluster depends on the number of failures that a VM is provisioned to tolerate. When a VM is configured with the Primary level of failures to tolerate set to 1 (PFTT = 1), vSAN can tolerate a single failure of any kind and of any component in a fault domain, including the failure of an entire rack.

When you configure fault domains on a rack and provision a new VM, vSAN ensures that protection objects, such as replicas and witnesses, are placed in different fault domains. For example, if a VM's storage policy has the Primary level of failures to tolerate set to N (PFTT = n), vSAN requires a minimum of 2*n+1 fault domains in the cluster. When VMs are provisioned in a cluster with fault domains using this policy, the copies of the associated VM objects are stored across separate racks.

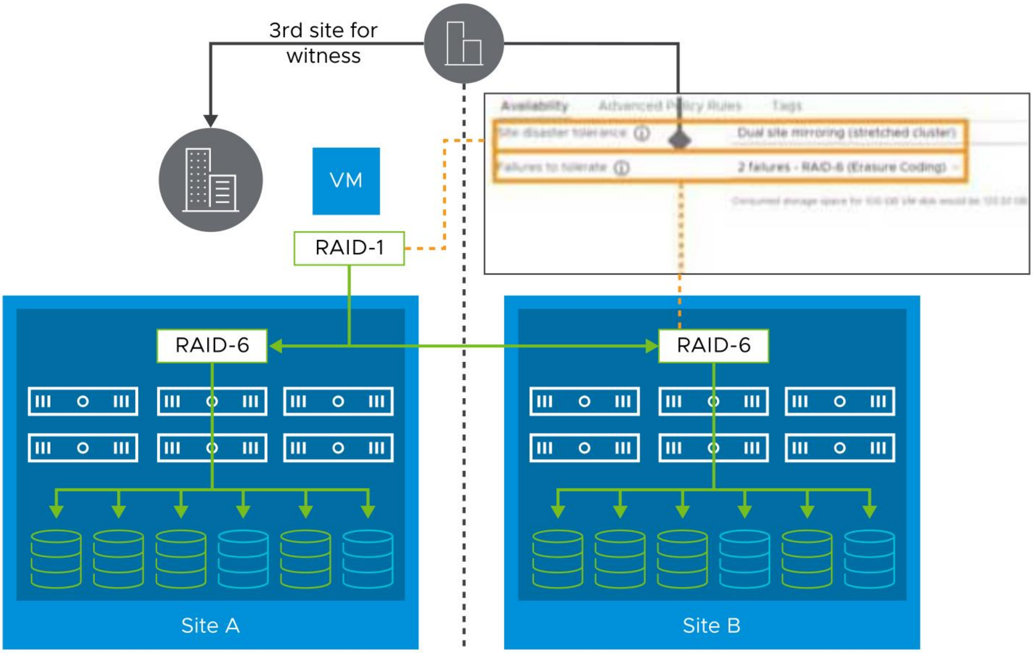
A minimum of three fault domains are required to support PFTT=1. For best results, configure four or more fault domains in the cluster. A cluster with three fault domains has the same restrictions that a three-host cluster has, such as the inability to reprotect data after a failure and the inability to use the Full data migration mode.

6-26 vSAN Local and Remote Protection

With its integrated stretched clustering technology, vSAN provides redundancy locally and across sites.

When a site failure occurs, vSAN maintains availability with local redundancy in the surviving site, and no change occurs in the stretched cluster configuration steps.

In a stretched cluster configuration, vSAN provides optimized site locality logic to minimize I/O traffic across sites.



Site disaster tolerance: Dual site mirroring (stretched cluster)

Failures to tolerate: 2 failures - RAID-6 (Erasure Coding)

Consumed storage space for [illegible]

vSAN uses fault domains to protect an environment from any downtime resulting from a site failure. The ability of vSAN to provide a fully active-active, stretched cluster is valuable for data centers. vSAN provides the capability for storage redundancy within a site and across sites at the same time. This redundancy helps deliver effective, affordable protection against entire site outages, and host outages within a site. This level of protection is what many of customers require. Lower total cost of ownership is the result because you do not need to purchase any additional hardware or software. You can address protection requirements with software that you already know.

For more information, see *vSAN Planning and Deployment* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.vsan-planning.doc/GUID-194D9B44-7593-4D3E-A9DD-A6646C1DCC05.html.

6-27 vSAN File Service

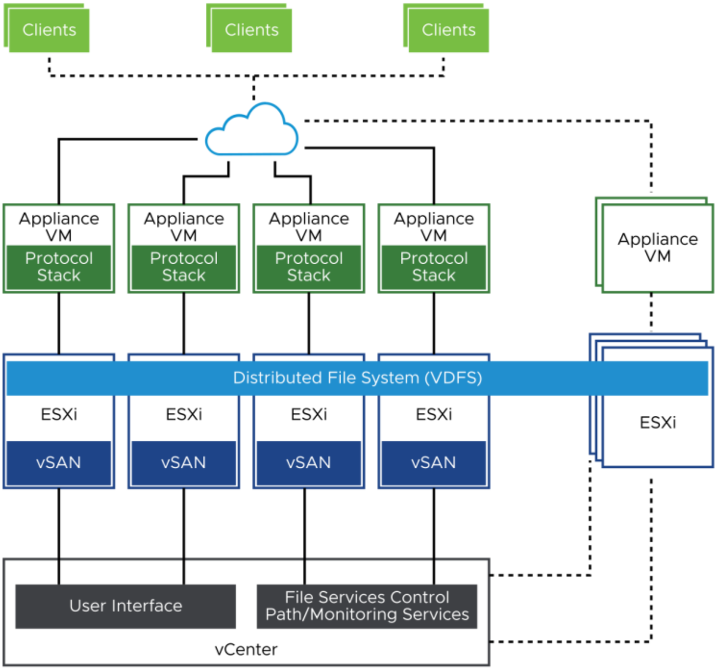
The vSAN file service eases the burden of managing a separate file server component:

- File shares are created in the vSAN datastore and accessed by the client workstations or VMs.
- Data is stored in a file share that can be accessed from any device that has access rights.
- The vSAN file service provides a single storage solution for block and file storage.

vSAN file service is a layer that sits on top of vSAN to provide SMB and NFS file shares:

- vSAN Virtual Distributed File System
- Storage services platform (VM appliances per host)

File shares are integrated into the existing vSAN storage policy based management on a per-share basis.



When you configure the vSAN file service, vSAN creates a single VDFS distributed file system for the cluster. A file service VM (FSVM) is placed on each host. The FSVMs manage file shares in the vSAN datastore. Each FSVM contains an NFS file server.

An IP pool enabled with the vSAN file service provides multiple NFS endpoints to balance NFS requests across the FSVMs. The primary IP is used as the single point of access to vSAN file shares. To provide computing resources that help manage access requests, the number of IP addresses must be equal to the number of hosts in the vSAN cluster.

6-28 Design Decision: Use of Solid-State Disks and Flash Devices (1)

Should you use solid-state disks (SSDs) and flash devices?

vSphere can use SSDs and flash devices in multiple ways:

- Regular datastores
- vSAN for cache and capacity devices
- ESXi boot disk
- ESXi core-dump device
- ESXi logging device
- ESXi host swap cache

SSDs and flash devices can be used in a vSphere environment in the following ways:

- Regular datastore: A (local) SSD is used instead of a hard disk drive. This usage model is supported for SATA and SCSI connected SSDs.
- vSAN: This usage module is supported for SATA and SCSI SSDs.
- ESXi boot disk: This usage model is supported for USB flash devices and SCSI/SATA connected devices.
- ESXi core-dump device: This usage model is supported for boot USB flash devices and for any SATA or SCSI connected SSD that is local. This usage model also applies to autodeployed hosts which have no boot disk.
- ESXi logging device: This usage model is supported for any SATA or SCSI connected SSD that is local.
- ESXi host swap cache: This usage model is supported for SATA and SCSI connected SSDs. USB and low-end SATA or SCSI flash devices are not supported. The workload is heavily influenced by the degree of host memory overcommitment.

6-29 Design Decision: Use of Solid-State Disks and Flash Devices (2)

SSDs offer much higher throughput and much lower latency than traditional magnetic hard disks. Flash devices, such as USB and SATADOM, can also be appropriate for some use cases.

SSDs and flash device storage have less endurance than magnetic disks. Devices vary based on workload type and factors such as drive capacity.

For each use case, the amount and frequency of data written to the SSD or flash device determines the minimum requirement for performance and endurance by ESXi.

SSDs offer much higher throughput and much lower latency than traditional magnetic hard disks.

While offering lower throughput and higher latency, flash devices such as USB or SATADOM can also be appropriate for some use cases.

The potential drawback to using SSDs and flash device storage is that the endurance can be less than traditional magnetic disks.

Also, SSDs and flash devices can vary based on the workload type and factors such as the drive capacity, underlying flash technology, and so on.

Performance and endurance are critical factors when selecting SSDs. In general, SSDs can be deployed in all the use cases mentioned, but (low-end) flash devices, including SATADOM, can be deployed only in some use cases.

For information about SSD and flash device selection requirements, see vSphere Flash Device Support at https://storagehub.vmware.com.

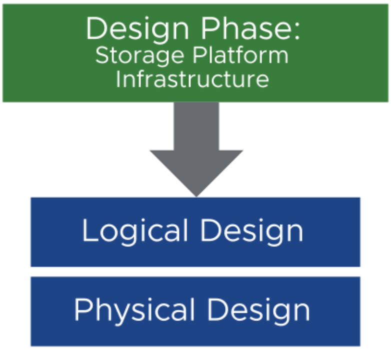
6-30 Lab 7: Designing the Storage Platform Infrastructure

Create the storage platform infrastructure design:

1. Review the Conceptual Design
2. Evaluate Storage Platform Design Options
3. Diagram the Storage Platform Architecture
4. Create the Storage Platform Physical Design

6-31 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



6-32 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of different storage platform solutions
- Design a storage platform architecture that meets the needs of the vSphere environment

6-33 Lesson 2: Calculating Storage Capacity Requirements

6-34 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Calculate the storage capacity required for VMs
- Calculate the number of required datastores

6-35 Calculating Total Storage Capacity Requirements

When determining total storage capacity requirements, you must account for several items.

Item	Notes
All VMs	Use server inventory data from the capacity analysis. Account for new VMs: Management VMs and forecasted VMs (year-over-year growth).
VM swap space	Average VM memory size multiplied by total number of VMs.
VM snapshots	Guideline: 15 percent of total VM capacity.
VM growth	Guideline: 15 percent of total VM capacity.
vSphere Auto Deploy repository for image profiles	Each image profile requires approximately 350 MB of storage.
Templates and ISO images	Sizes vary.

After calculating the base datastore capacity requirements from the capacity analysis data, you must factor in the required spare capacity for vSphere feature requirements.

The size of the swap file depends on the amount of RAM that is configured for the VM. For example, if the average VM memory size is 3 GB, 3 GB of swap space is required for each VM.

The retention policy for active snapshots determines the overall space that is required. However, 15 percent is a good guideline. Make the retention policy for active snapshots as short as possible. If the retention policy exceeds two days, store the snapshots on a second-tier datastore.

If your design includes ESXi boot from vSphere Auto Deploy, ensure that you have enough storage for the vSphere Auto Deploy repository. Consider the number of image profiles that you expect to use.

The size of template and ISO storage depends only on the size and number of templates and ISO images in the library. Templates do not run and do not have snapshots. So extra datastore overhead for snapshots is not necessary.

6-36 Calculating Datastore Size

To size a datastore, you must consider the following information:

- Number of VMs to allocate to each datastore:
  - The maximum value should be calculated based on the VM workload (IOPS).
  - If you do not know the number of VMs to allocate to a datastore, use the guideline of 10 to 15 VMs per datastore.
- Average size of all VMDKs for a VM:
  - Number of VMs x Average size of VM's VMDK files = VM capacity
- The total datastore capacity accounts for the additional storage requirements:
  - Total datastore capacity = VM capacity + Swap space capacity + Snapshot reserve capacity + VM growth capacity
  - + (Optional) First Class Disks

The total capacity required for a datastore includes not only VM capacity but also capacity for VM swap space and VM snapshot overhead, and capacity for future growth.

Sizing for average utilization always introduces the risk of workloads peaking at the same time. During peak utilization, the environment might be constrained for resources.

6-37 Design Decision: Determining the Number of Datastores (1)

How many datastores should be created?

You can determine the minimum number of datastores by dividing the total required storage capacity by the sizes of the individual datastores.

More than the minimum number of datastores might be required, depending on the following factors:

- Policies that require separate storage for individual organizations or departments.
- Policies that require separate storage for different security zones.
- Storage performance requirements: The datastore should be able to handle the IOPS requirements for the VMs on the datastore.

In a vSphere HA enabled cluster, you can have a maximum of 2,048 powered-on VMs on a single datastore.

You might be required to add datastores to the design because of other factors. For example, an organization’s policies regarding hardware isolation between departments might force you to add datastores.

You can determine the storage performance requirement for each server by calculating the average IOPS for each server. This measurement can also help you determine the number of datastores required and how many VMs are on each datastore.

6-38 Design Decision: Determining the Number of Datastores (2)

Separate datastores can be implemented for the following use cases:

- Individual projects
- Production, test, and development environments
- ISO images and templates managed by a content library
- A VM with exceptionally large storage requirements, for example, a VM that uses 1.5 TB of storage

Different use cases might also force you to add datastores. For example, a development environment might have a more lenient VM snapshot policy than a production environment. In this case, you might need separate datastores for development and production to accommodate different snapshot policies. You might also need a datastore for ISO images and a datastore for templates.

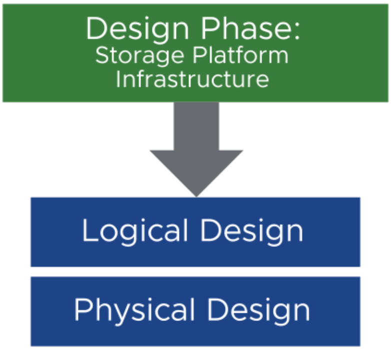
6-39 Lab 8: Calculating Storage Capacity

Create the storage platform infrastructure design:

1. Review the Capacity Planning Assessment Findings for Storage Usage
2. Calculate Storage Capacity Requirements

6-40 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



6-41 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Calculate the storage capacity required for VMs
- Calculate the number of required datastores

6-42 Lesson 3: Storage Management Features

6-43 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

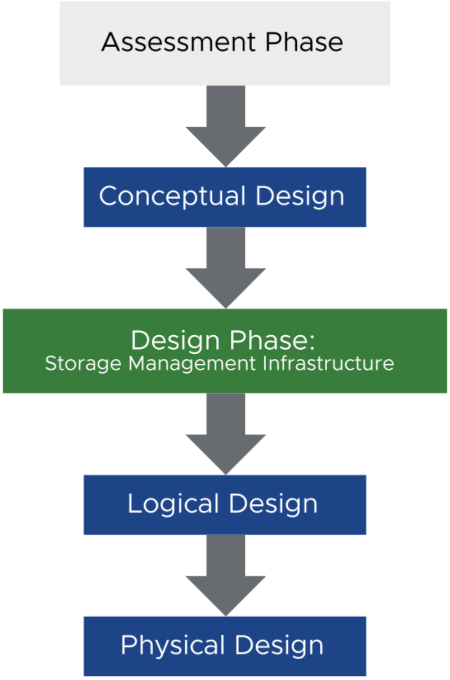
- Evaluate the use of storage management solutions
- Design a storage management architecture that meets the needs of the vSphere environment

6-44 Designing the Storage Management Infrastructure

You must decide how to manage storage during normal operations.

Design decisions for the storage management infrastructure relate to:

- Storage tiering
- Use of vSphere Storage DRS
- Use of Storage I/O Control



After the conceptual design is created, the design work can begin, starting with the logical design. The designer looks at each of the major vSphere infrastructure components and makes the appropriate design decisions.

The design phase is an iterative process. As you make design decisions, always verify these decisions against the organization's goals, requirements, constraints, risks, assumptions, and current best practices.

6-45 Design Decision: Evaluating the Use of Tiered Storage

Should storage tiers be created? If so, how do you use storage tiering?

Using storage tiers, you can provide the appropriate levels of performance and availability for your application requirements.

You can create storage tiers by using the following methods:

- VM storage policies: You create storage policies so that administrators can assign VMs to the appropriate storage tiers.
- Array-based autotiering: LUN segments are migrated to different disk types, based on the use pattern.

  Migration policies and thresholds depend on array type and vendor.

Enable storage tiering at either the storage array or in vSphere, but not in both places at the same time.

Align your design decision with the storage vendor’s best practices.

Storage, server, and application administrators must select the correct storage configuration for each application that is deployed in the environment. All application workloads are not the same. Storage tiering caters for these differences by creating multiple levels of storage with varying degrees of performance, reliability, and cost, depending on the application workload needs.

Automated storage tiering is a common feature in today’s enterprise-class storage arrays. This feature migrates LUN segments (chunks) to different disk types, based on the use pattern. Hot (frequently accessed) segments typically move to faster disks, whereas cold segments move to slower disks.

6-46 Tiered Storage Example

This example of tiered storage shows the storage type and use case for each tier in the hierarchy.

Tier	Storage Type	Speed	Capacity	Cost	Use Case
2	SSD	High	Medium	High	Transactional databases and workloads that require many small I/O transactions
3	SAS	Medium	Medium	Medium	Email and web servers
4	SATA	Low	High	Low	Noncritical data requests and streaming media

Enterprise-class storage arrays contain multiple drive types and protection mechanisms. All application workloads are not the same. Storage tiering accounts for these differences by creating multiple levels of storage with varying degrees of performance, reliability, and cost, depending on the application workload requirements.

When you optimize performance for your VMs, storage location is an important factor. You must make a trade-off between expensive storage that offers high performance and storage with lower cost and lower performance.

SSD has higher performance than SAS, and SAS has higher performance than SATA. However, trade-offs between speed, storage capacity, and cost must also be considered.

6-47 Storage Tiering with VM Storage Policies

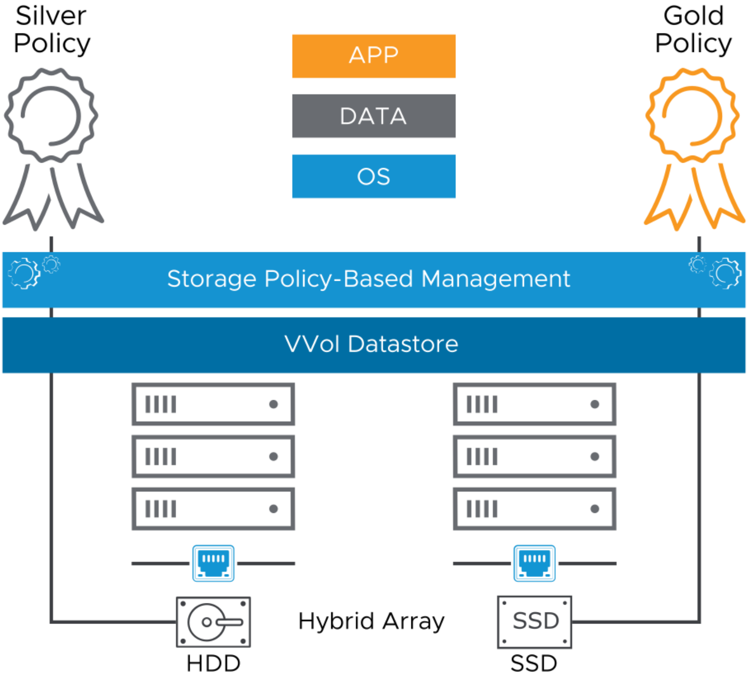
VM storage policies ensure that VMs run on storage with the correct performance and availability characteristics.

Storage subsystem capabilities are identified in the following ways:

- By storage vendors, using vSphere API for Storage Awareness
- By the vSphere administrator (user-defined)

You can use storage policies to place a database server's disks on appropriate storage. For example:

- Assign the Silver policy, which uses only HDDs, to the operating system disk.
- Assign the Gold policy, which uses only SSDs, to the data disk.



You create a VM storage policy to identify capabilities and characteristics of the storage subsystem. When a VM is created, cloned, or migrated, it can be associated with a VM storage policy.

When you select a VM storage policy, the vSphere Client shows the datastores that are compatible with the capabilities of the policy, and a datastore or datastore cluster can be selected.

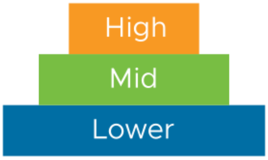
When a datastore is selected and it does not match the VM storage policy, the vSphere Client shows that the VM is using noncompliant storage.

vSAN uses storage policies that allow the user to define the characteristics that VMs must have when running. The policy can be defined on an individual disk level rather than at the volume level.

VM storage policies can be associated to VMs and periodically checked for compliance so that the VM is running on storage with the correct performance and availability characteristics.

For information about storage policy based management, see "Understanding Storage Policy-Based Management" at https://blogs.vmware.com.

6-48 Designing Storage Tiers



Storage tiers do not need to be complicated to be effective. To implement storage tiers:

1. For each application or service, determine the storage characteristics:
   - IOPS
   - MBps
   - Capacity
   - Availability
   - Latency
2. Design storage tiers with characteristics that match the application and service requirements listed in the service-level agreement (SLA).
3. Assign applications and services to the appropriate tiers. Data might move between tiers during the information life cycle.

When deciding where to place a VM, ask the following questions:

- How critical is the VM?
- What are its PiT (point-in-time) restoration requirements?
- What are its backup requirements?
- What are its replication requirements?

A VM might change tiers throughout its life cycle because of changes in criticality or changes in technology that push higher tier features to a lower tier. Criticality is relative and might change for various reasons, including changes in the organization, operational processes, regulatory requirements, disaster planning, and so on.

6-49 Design Decision: Evaluating the Use of vSphere Storage DRS

Should you use a datastore cluster that is enabled for vSphere Storage DRS?

vSphere Storage DRS automates the management of datastores based on latency and utilization.

Consider the following guidelines when using vSphere Storage DRS:

- Enable vSphere Storage DRS in a datastore cluster whenever possible.
- All vSphere Storage vMotion prerequisites must be in place.
- All datastores should use the same version of VMFS and be on the same storage subsystem.
- Turn on balancing based on IOPS but turn it off during maintenance cycles.
- Minimize the use of affinity-based rules because they add complexity.
- If array-based autotiering is used, configure vSphere Storage DRS in manual mode with I/O metrics disabled.

When a datastore cluster is created, vSphere Storage DRS can be used to manage storage resources. When vSphere Storage DRS is enabled on a datastore cluster, vSphere automates the process of initial VM file placement and balances storage resources across the cluster to avoid bottlenecks. Datastore space use and I/O load are considered when vSphere Storage DRS makes migration recommendations.

When configuring vSphere Storage DRS, the default setting is to balance use when a datastore becomes 80 percent full. Consider leaving more available space if snapshots are used often or multiple snapshots are frequently used.

Consider turning on balancing based on IOPS. The feature balances datastores based on sustained usage over time. Turn off balancing based on IOPS during maintenance cycle, such as when VMs on the datastores are being backed up.

An example of using affinity and anti-affinity rules is to improve the performance of an application by keeping the application disk on a datastore separate from the operating system disk.

For more information about datastore clusters and vSphere Storage DRS, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html.

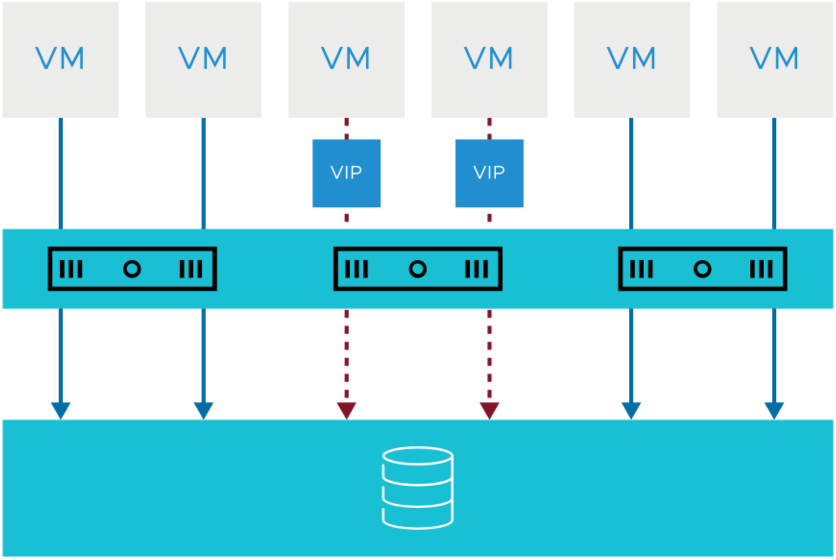
6-50 Design Decision: Evaluating the Use of Storage I/O Control

Should Storage I/O Control be used?

Using Storage I/O Control, you can configure shares on a VM disk to prioritize storage I/O.

After a latency threshold is reached, Storage I/O Control distributes the resources based on the virtual disk share values.

Consider configuring higher share values for critical applications and default share values for other applications.



With vSphere Storage I/O Control, you can configure shares on a VM disk. When datastore latency exceeds a threshold (30 ms), Storage I/O Control can balance the I/O based on the share value. If the threshold is never reached, Storage I/O Control remains idle.

Configuring Storage I/O Control is a two-step process:

4. Enable Storage I/O Control for the datastore.

5. Set the number of storage I/O shares and the upper limit of I/O operations per second (IOPS) to be allowed for each VM.

Storage I/O Control is disabled by default. But if vSphere Storage DRS is configured, Storage I/O Control is enabled.

Storage I/O Control delivers better workload consolidation and helps reduce extra costs associated with overprovisioning. Applications critical to an organization that are sensitive to latency are good candidates for higher share values.

For more information, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html.

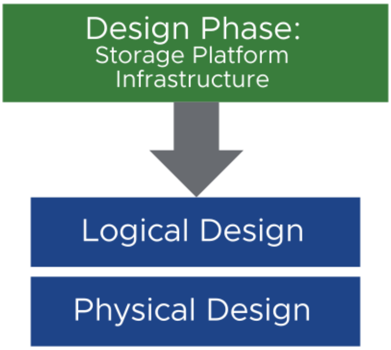
6-51 Lab 9: Designing the Storage Management Infrastructure

Create the storage management infrastructure design:

1. Review the Conceptual Design
2. Evaluate Storage Management Design Options
3. (Optional) Document the Datastore Cluster Physical Design

6-52 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



Design Phase:
Storage Platform
Infrastructure

Logical Design

Physical Design

6-53 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of storage management solutions
- Design a storage management architecture that meets the needs of the vSphere environment

6-54 Key Points

- For Fibre Channel, NFS, and iSCSI storage, design storage to reduce latency and increase availability.
- vSphere Virtual Volumes is a VM-centric solution where storage operations are simplified, and intensive storage operations, such as snapshots and replication, are offloaded to the storage array.
- vSAN provides a simple and cost-effective storage configuration, where an all-flash device configuration can be used to create pools of storage.
- SSDs offer higher throughput and lower latency, but less endurance, than magnetic hard disks.
- Using storage tiers, you can provide the appropriate levels of performance and availability based on application requirements.

Questions?

Module 7

Network Infrastructure

7-2 Importance

Effective network design can help your organization meet its business goals. With proper network design, you can ensure authorized and timely access to business data, and, at the same time, prevent unauthorized access to data.

7-3 Module Lessons

1. Network Components

2. Network Management and Monitoring

7-4 Lesson 1: Network Components

7-5 Learner Objectives

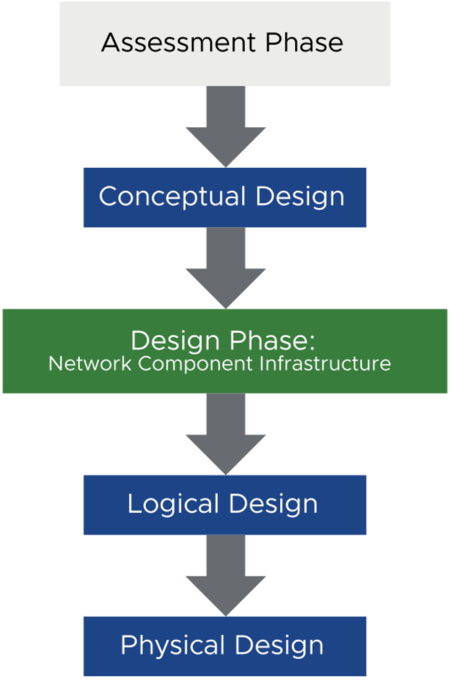
After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of different network component solutions
- Design a network component architecture that includes information about network segmentation and virtual switch types

7-6 Designing the Network Component Infrastructure

When you design the network component infrastructure, you make decisions in the following areas:

- Number of networks
- Network segmentation
- Micro-segmentation
- Multiple TCP/IP stacks
- Use of jumbo frames
- Virtual switch types
- Number of virtual switches

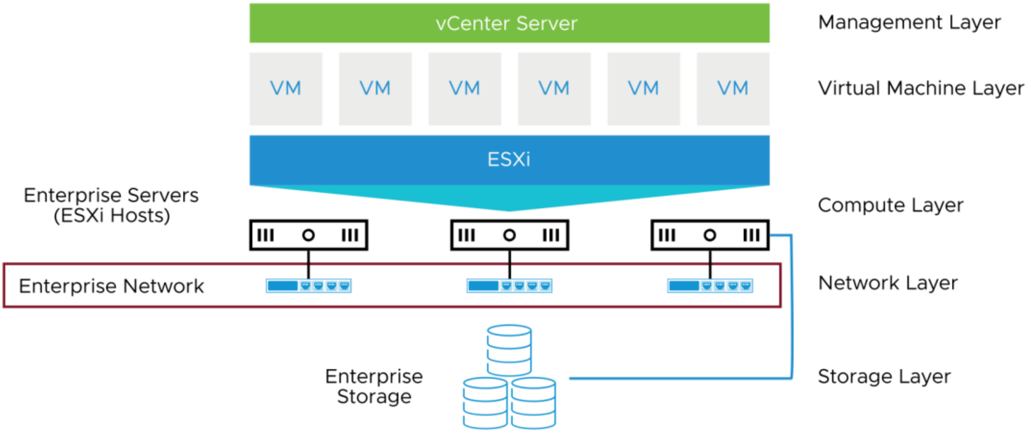


When creating the logical design, the designer looks at each of the major vSphere infrastructure components and makes the appropriate design decisions.

The design phase is an iterative process. As you make design decisions, always validate these decisions against the organization’s goals, requirements, constraints, risks, assumptions, and current best practices.

7-7 Network Layer in a vSphere Infrastructure

In a vSphere infrastructure, virtual switches connect virtual machines and VMkernel services to the physical network.



The network layer of the architecture design is built around the virtual switch. You can choose between standard switches and distributed switches. More than one network can coexist on the same virtual switch or networks can exist on separate virtual switches. The decision partly depends on the layout of your physical networks.

Physical NICs are assigned at the virtual switch level, so all ports and port groups defined for a particular switch share hardware.

7-8 High-Level Network Design Guidelines

Consider the following design guidelines for your network:

- Meet the diverse needs of applications, services, administrators, and users.
- Reduce costs, boost performance, improve availability, provide security, and enhance functionality.
- Separate network services to achieve greater security and better performance.
- Deploy firewalls to protect the most sensitive virtual machines.
- Design network security around applications.
- Consider service-oriented firewalling and micro-segmentation to reduce attack surfaces.
- Prevent network connectivity problems with the following configurations:
  - For multiple network adapters on a virtual switch, ensure that each network adapter is connected to the same physical network.
  - Configure all VMkernel network adapters on a distributed switch to use the same MTU.

Reducing costs is one of the simpler goals to achieve in the vSphere infrastructure. Server consolidation alone reduces network costs by reducing the number of required network ports and NICs. But you should explore ways to create a more efficient network design. For example, configuring two 10 Gb NICs with VLANs might be more cost-efficient than configuring a dozen 1 Gb NICs on separate physical networks.

You can improve or maintain network performance by providing sufficient bandwidth, which reduces contention and latency. Network availability is typically achieved by providing network redundancy. Network security can be achieved through controlled access where required and isolation where necessary.

The network design should help enhance the functionality of the infrastructure. You can achieve enhanced functionality by configuring the network to support vSphere features such as vSphere vMotion, vSphere HA, and vSphere FT.

Use the information gathered from the current-state analysis and the information gathered from the key stakeholder and subject matter expert (SME) interviews for guidance in creating the network design.

7-9 Network Segmentation Considerations

Separating different types of network traffic is recommended for the following reasons:

- Reduces contention and latency and improves performance, which is important for IP storage and vSphere FT logging networks
- Enhances security by limiting network access, for example, by separating management traffic and VM traffic

Use interview information from the key stakeholders and SMEs to determine which workloads and networks are sensitive to high latency and which networks must be secured.

High latency on any network can negatively affect performance. However, some components are more sensitive to high latency than others. For example, reducing latency is important on IP storage networks and the vSphere FT logging network. Latency on these networks can negatively affect the performance of multiple VMs.

Depending on the application or service, high latency on specific VM networks can also negatively affect performance. Use information gathered from the current-state analysis and key stakeholder and SME interviews to determine the existence of specific workloads and networks that are especially sensitive to high latency.

Separate networks are also required for access security. Use information gathered from key stakeholder and SME interviews to determine the specific access requirements for users and services, and design the network to isolate them appropriately.

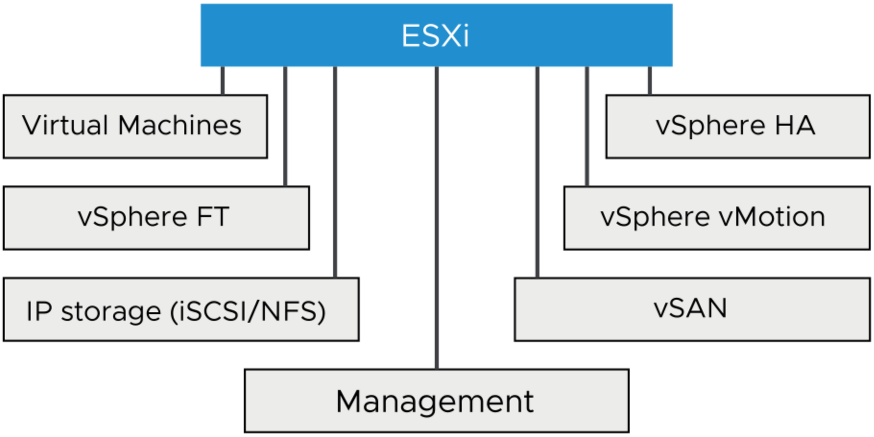
7-10 Design Decision: Types of Networks

What types of networks should be created?

The number of networks depends on an organization's business needs and determines the types of traffic required:

- vSphere operation traffic
- Organizational service and application traffic

The diagram shows the vSphere components that should be on separate networks.



The number of required networks depends on the different types of traffic that are necessary in the infrastructure. For example, the environment might benefit from a separate virtual machine backup network.

The number of networks, and whether VLANs are configured, affects the number of physical network components required. Physical network components include the number of NICs and the number of switches, ports, and cables.

7-11 Design Decision: Network Segmentation: Physical Networks or VLANs (1)

Should network traffic be segmented with physical networks or VLANs?

The decision to use VLANs instead of physical networks depends on the number of networks required and the number of available physical ports:

- If the number of required networks is less than or equal to the number of physical ports, the use of VLANs is optional. However, you should consider redundancy as well.
- If the number of networks required is greater than the number of physical ports, you have the following choices:
  - Configure VLANs to create separate virtual networks.
  - Purchase hardware that supports a greater number of network ports.

VLANs provide for logical groupings of stations or switch ports, allowing communications as if all stations or ports were on the same physical LAN segment. Physical Ethernet adapters serve as bridges between virtual and physical networks.

Physical network segmentation and redundancy require a host with significant I/O expansion capability to provide the required number of physical ports. This type of host might be more expensive, require more rack space, more cabling, and a greater number of physical switch ports.

7-12 Network Segmentation: Physical Networks or VLANs (2)

The following factors might determine the choice between physical or VLAN network segmentation:

- Physical port speed. For example, a 10 Gb Ethernet port often offers sufficient bandwidth to support several VLANs.
- Whether the physical network infrastructure supports VLANs.
- An organization's current use of VLANs.
- Network security policies that might require the physical separation of networks.

You can consider several interrelated factors to help determine whether network segmentation should be accomplished through physical networks or VLANs.

Some organizations might have network policies against the use of VLANs. But the primary factor is whether enough network ports are available to accommodate the required number of networks.

When determining the number of networks, consider not only the number of networks but also whether any of the networks should be made more highly available through redundancy.

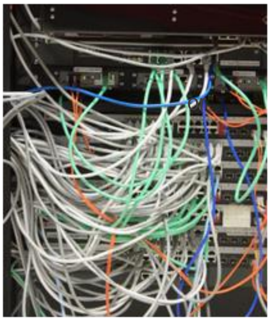
7-13 VLAN Design Considerations

VLANs have the following benefits:

- Reduce the number of required physical ports and cabling
- Provide easier reconfiguration of networks and servers
- Reduce hardware and administration costs

When designing a VLAN configuration, consider the following guidelines:

- Build network redundancy into the VLAN environment:
  - If a trunk port fails, multiple VMs and services can become unavailable.
- Do not exceed the bandwidth limitations of the physical network:
  - A combination of NIC teaming and 10 Gb or 40 Gb Ethernet adapters can be used to address bandwidth and redundancy issues.



More Ports/Cables

Fewer Ports/Cables

With VLANs, the use of fewer physical ports reduces the number of cables and even the number of switches required. The use of fewer cables and switches also lowers hardware and administrative costs and complexity. When reconfiguring networks with VLANs, networks can be created or changed without the need to add or move cables.

In a VLAN environment, the bandwidth limitations of the physical network can be easily exceeded. Because many virtual networks share a single physical path, the aggregated network traffic can exceed the capabilities of the physical network.

Network redundancy is important in a VLAN environment. NIC teaming is a good solution but requires additional network ports. These additional network ports require more I/O expansion capability on the host, more network switch ports, and more cables. A 10 Gigabit Ethernet solution requires fewer hosts, switch ports, and cabling than a 1 Gigabit Ethernet NIC team solution.

Combining NIC teaming with 10 Gigabit Ethernet or even 40 Gigabit Ethernet is perhaps the best solution because both bandwidth and redundancy issues are addressed.

7-14 Design Decision: Evaluating the Use of Multiple TCP/IP Stacks

Should multiple TCP/IP stacks be used?

Use multiple TCP/IP stacks to isolate management traffic from system traffic, which improves security:

- Each TCP/IP stack is assigned a dedicated default gateway, routing table, and DNS configuration for its traffic.
- A separate set of buffers and sockets are used by each stack.
- Routing table conflicts (which might appear when many features use a common stack) are avoided.

vSphere provides the following TCP/IP stacks:

- Default for management traffic
- Provisioning for provisioning traffic, such as VM cold migrations and cloning
- vMotion for vSphere vMotion traffic:
  - With a separate stack for vSphere vMotion, migrations can occur across long distances, as well as between vCenter Server instances.

Take appropriate security measures to prevent unauthorized access to the management and system traffic in your vSphere environment. For example, isolate the vSphere vMotion traffic in a separate network that includes only the ESXi hosts that participate in the migration. Isolate the management traffic in a network that only network and security administrators can access.

After you create a VMkernel adapter on the vSphere vMotion TCP/IP stack, you can use only this stack for vSphere vMotion migration on this host. The VMkernel adapters on the default TCP/IP stack are disabled for the vSphere vMotion service. The same rule applies to the Provisioning TCP/IP stack.

Keep the vSphere vMotion connection on a separate network. When a vSphere vMotion migration occurs, the contents of the guest operating system’s memory is transmitted over the network. You can keep the vSphere vMotion connection on a separate network by using VLANs to segment a single physical network or by using separate physical networks (the latter is preferable).

7-15 Design Decision: Evaluating the Use of Jumbo Frames

Should jumbo frames be enabled?

Jumbo frames provide an increase in performance for workloads, if configured correctly.

Consider the following guidelines:

- If you have NFS or iSCSI storage, enable jumbo frames to improve storage throughput by decreasing network protocol overhead.
- Enable jumbo frames if the workload requires them:
  - The Geneve overlay in NSX-T Data Center requires an MTU of 1,600 bytes or greater
- Jumbo frames must be configured end to end on all devices on the network:
  - VMs require specific NIC support.
  - If jumbo frames are not configured end to end, performance can suffer because the larger packets must be broken up to be transmitted.

A jumbo frame is an Ethernet frame with a maximum transmission unit (MTU) that is bigger than 1,500 bytes and up to 9,000 bytes.

With jumbo frames, increasing the per-frame payload from 1,500 bytes to 9,000 bytes increases the efficiency of data transfer. Workloads such as iSCSI might benefit from configuring jumbo frames. If the workload consistently transfers large amounts of network data, then configure jumbo frames if possible.

For NSX virtual switches, the MTU for each switch must be set to 1,600 or higher for the Geneve overlay. By default, it is set to 1,600. The Geneve overlay provides the overlay capability in NSX-T Data Center to create isolated, multitenant broadcast domains across data center fabrics. Customers can create elastic, logical networks that span physical network boundaries.

A common issue is that jumbo frames must be configured end to end on all devices on the network. Otherwise, performance can suffer because the larger packets must be broken up to be transmitted. End-to-end configuration might be easily accomplished in a LAN and might be more difficult for WAN configurations. Often, the act of configuring jumbo frames end to end is not easily justifiable because of the cost of configuration versus the expected gains.

The virtual machine operating system and the virtual machine NIC must also support jumbo frames. For information about which virtual NICs provide jumbo frame support, see VMware knowledge base article 1001805 at http://kb.vmware.com/kb/1001805.

7-16 Design Decision: Choosing a Virtual Switch Type (1)

Which virtual switch types should be used?

Consider the benefits and drawbacks of each type.

Type	Benefits	Drawbacks
Standard Switch	• Simple to configure • No dependency on vCenter Server	• Manual configuration per ESXi host • Lacks many advanced features supported by distributed switches
Distributed Switch	• Ensures consistent network configuration across hosts and clusters • Enhanced network monitoring and troubleshooting capabilities • Supports advanced vSphere networking features	• Cannot be managed when vCenter Server is not available

A standard switch models a physical Ethernet switch. One network adapter of a virtual machine can be connected to each port. Each uplink adapter associated with a standard switch uses one port. Each logical port on the standard switch is a member of a single port group. Each standard switch can also have one or more assigned port groups.

A distributed switch centralizes network management in the vCenter Server system when the vCenter Server system is running.

7-17 Design Decision: Choosing a Virtual Switch Type (2)

Consider the following guidelines when choosing between virtual switch types:

- Only use standard switches if the organization does not require distributed switch features and does not want to pay the additional licensing cost.
- Use distributed switches if possible:
  - A distributed switch requires a vSphere Enterprise Plus license.
  - If you deploy NSX-T Data Center with logical switching, a distributed switch is required.

You use standard switches to provide network connectivity to hosts and VMs. A standard switch can only bridge traffic internally between VMs in the same VLAN and link to external networks.

A distributed switch provides centralized management and monitoring of the networking configuration of all hosts that are associated with the switch. You set up a distributed switch on a vCenter Server system, and its settings are propagated to all hosts that are associated with the switch

For more information about distributed switches, see VMware vSphere Distributed Switch Best Practices at https://www.vmware.com/techpapers.html.

7-18 Standard Switch and Distributed Switch Feature Comparison

Standard and distributed switches share some features. But distributed switches have several features that standard switches do not have.

Feature	Standard Switch	Distributed Switch
Layer 2 switch	Yes	Yes
VLAN segmentation	Yes	Yes
IPv6 support	Yes	Yes
802.1Q tagging	Yes	Yes
NIC teaming	Yes	Yes
Outbound traffic shaping	Yes	Yes
Inbound traffic shaping	No	Yes
Configuration back up and restore	No	Yes
Private VLANs	No	Yes
Link Aggregation Control Protocol	No	Yes

A distributed switch offers features that are not available on standard switches. These features might be useful or required by applications and services.

For a comparison of standard switch and distributed switch features, see VMware knowledge base article 1010555 at http://kb.vmware.com/kb/1010555.

7-19 Design Decision: Choosing the Number of Virtual Switches

How many virtual switches should be used?

The number of virtual switches depends on the organization’s preferences and the available hardware configuration.

In general, keep virtual switches to a minimum:

- To simplify configuration and monitoring, configure a single virtual switch with a port group for each type of network traffic.
- One virtual switch with VLANs works in environments with a limited number of physical network ports.
- Connect all hosts in a cluster enabled for vSphere HA and vSphere DRS to a single distributed switch so that VMs can easily reconnect after migration.
- If the organization has a policy that VM to VM traffic must pass through a physical firewall, multiple virtual switches are required.

To simplify configuration, administration, and monitoring, create a single virtual switch if possible.

Different types of network traffic can be isolated using multiple port groups on a virtual switch. Using a combination of VLANs and port groups is a good solution, particularly in situations where a limited number of network ports are available.

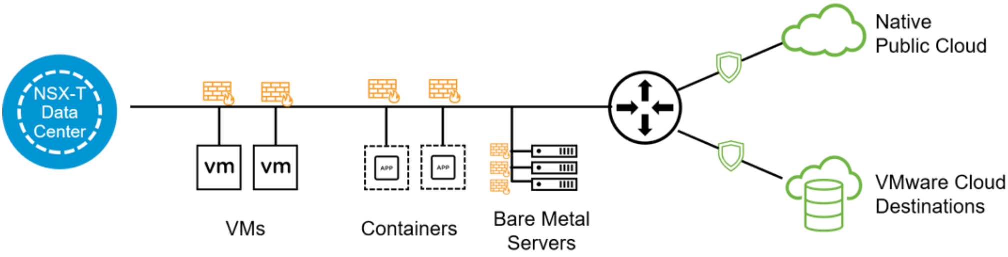
7-20 About NSX-T Data Center

NSX-T Data Center is a software layer providing connectivity from data center to cloud to edge infrastructure, with data visibility and security.

NSX-T Data Center provides consistent networking and security across the entire IT environment.

You deploy NSX-T Data Center nondisruptively on top of the existing physical infrastructure.

NSX-T Data Center reproduces the entire network model in software so that any network topology can be created and provisioned in seconds.



Network virtualization is conceptually similar to server virtualization. Network virtualization is the functional equivalent of a network hypervisor that reproduces network functions in software.

Network virtualization provides an abstraction layer that sits between underlying network and virtual networks. It provides an operational model of the VM for the network. Whether you are running VMs, containers, or a mixture, you can spin up their networks in seconds. With micro-segmentation enabled by default, these networks are isolated and secured from each other.

7-21 Evaluating the Use of NSX-T Data Center

NSX-T Data Center has the following benefits:

- Provides micro-segmentation capabilities to lock down critical applications, reduce the attack surfaces, and achieve zero-trust security
- Streamlines multicloud operations with consistent networking and security
- Accelerates application delivery by automating the provisioning and management of networking and security services
- Provides integrated, full-stack networking and security for containerized applications and microservices

Use the micro-segmentation capabilities in NSX to lock down critical apps, create a logical DMZ in software, and reduce the attack surface of a virtual desktop environment. Zero-trust security is now attainable and efficient in private and public cloud environments.

Streamline your multicloud operations with consistent networking and security, and, in the process, enable multicloud use-cases ranging from seamless data center extension, to multi data center pooling, to rapid workload mobility.

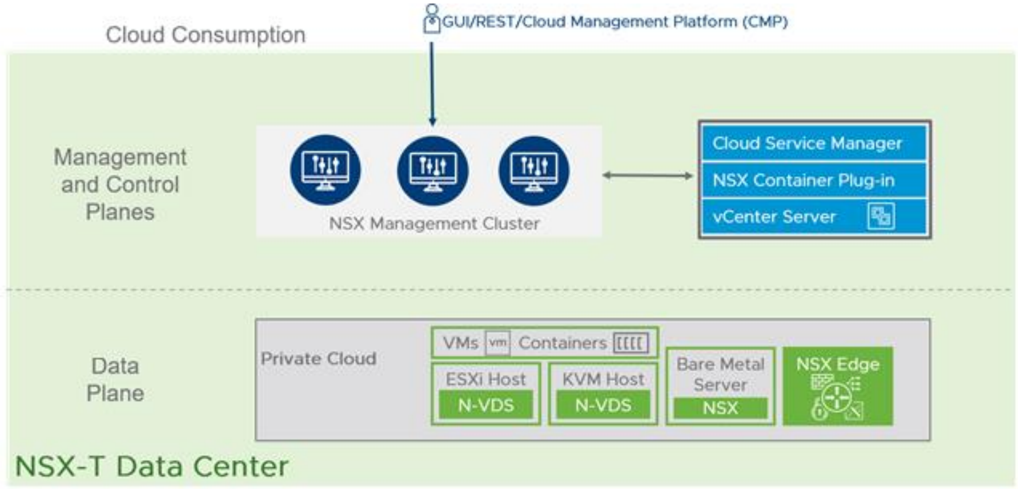
Accelerate application delivery by using blueprints to automate the provisioning and management of networking and security services consistently across all sites and clouds, or by exposing the infrastructure as code.

Provide integrated, full-stack networking and security for your containerized apps and microservices, including native container networking for Kubernetes, micro-segmentation, and end-to-end observability for microservices.

7-22 NSX-T Data Center Components

An NSX-T Data Center deployment consists of the following planes:

- Management plane to handle the user management input
- Control plane to handle the implementation
- Data plane, which includes the NSX Virtual Distributed Switch (N-VDS)



Each plane has its own components:

- The management plane is designed with advanced clustering technology, whereby the platform can process large-scale concurrent API requests. NSX Manager provides the REST API and a web-based UI interface entry point for all user configurations.
- The control plane includes a three-node controller cluster, which is responsible for computing and distributing the runtime virtual networking and security state of the NSX-T Data Center environment. The control plane is separated into a central control plane and a local control plane. This separation significantly simplifies the work of the central control plane, and the platform can extend and scale for various endpoints. With NSX-T Data Center 2.4, the management plane and control plane are converged. Each manager node in NSX-T Data Center is an appliance with converged functions, including management, control, and policy.
- The data plane includes a group of ESXi or kernel-based virtual machine (KVM) hosts, and NSX Edge nodes. The group of servers and edge nodes prepared for NSX-T Data Center are called transport nodes. Transport nodes are responsible for the distributed

forwarding of network traffic. Rather than relying on the distributed virtual switch, the data plane includes a host switch called the NSX Virtual Distributed Switch (N-VDS), which decouples the data plane from the compute manager, such as vCenter Server, and normalizes networking connectivity.

Although the consumption plane is not part of NSX-T Data Center, it provides integration into virtually any cloud management platform (CMP) through the REST API and integration with VMware cloud management planes such as vRealize Automation:

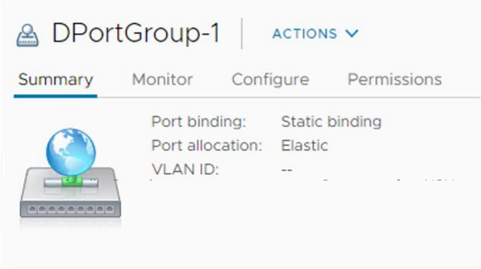
- The consumption of NSX-T Data Center can be driven directly through the NSX Manager user interface (UI).
- Typically, end-users tie network virtualization to their cloud management plane for deploying applications.
- Integration is also available through OpenStack, Kubernetes, and Pivotal Cloud Foundry.
- All operations are performed from the management plane. These operations include create, read, update, and delete (CRUD).

7-23 NSX-T Data Center Port Groups on VDS

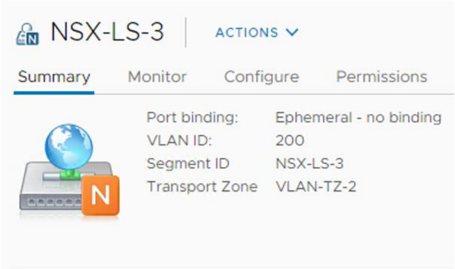
In vSphere 7, NSX-T Data Center logical switches (segments) behave as virtual distributed port groups.

Using these NSX port groups, you can perform the following tasks in vSphere:

- Bind VMkernel adapters to NSX port groups.
- Assign VM network adapters to NSX port groups.
- Migrate VMs to and from NSX port groups.



Port binding: Static binding
Port allocation: Elastic
VLAN ID: --



Port binding: Ephemeral - no binding
VLAN ID: 200
Segment ID NSX-LS-3
Transport Zone VLAN-TZ-2

In vSphere 7, you can now view and use NSX-T Data Center logical switches in the vSphere Client as transparent networks.

In the vSphere Client, you can easily distinguish NSX port groups from traditional distributed port groups. vCenter Server requires that networking interfaces use unique names. For example, two or more distributed port groups with the same name cannot exist in the same system. However, NSX allows multiple logical switches to use the same name. The same NSX port group can span multiple distributed switches.

To avoid confusion between seemingly identical NSX port groups in the vSphere Client, additional properties are provided for NSX port groups to differentiate them from virtual distributed port groups. Also, the icon that represents NSX port groups is different from the icon for virtual distributed port groups.

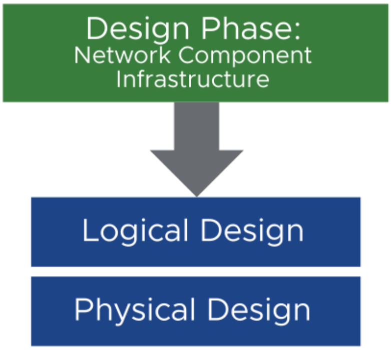
7-24 Lab 10: Designing the Network Component Infrastructure

Create the network component design:

1. Review the Conceptual Design
2. Evaluate Network Component Design Options
3. Create a Network Component Relationship Diagram
4. Document the Network Component Physical Design
5. Document the ESXi Host Networking Physical Design

7-25 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



7-26 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of different network component solutions
- Design a network component architecture that includes information about network segmentation and virtual switch types

7-27 Lesson 2: Network Management and Monitoring

7-28 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

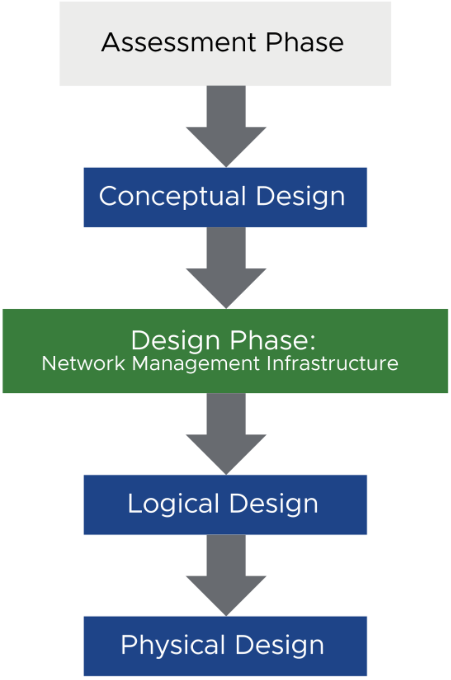
- Evaluate the use of different network management solutions
- Design a network management architecture that meets the needs of the vSphere environment

7-29 Designing the Network Management Infrastructure

After the network component architecture is designed, you can decide how to build redundancy and performance into the network design.

For the network management infrastructure, you make decisions in the following areas:

- NIC teaming design
- Use of vSphere Network I/O Control
- Use of traffic filtering and QoS tagging
- Use of CDP and LLDP
- Use of NetFlow and port mirroring
- Exporting distributed switch configurations
- DNS design
- Use of IPv6

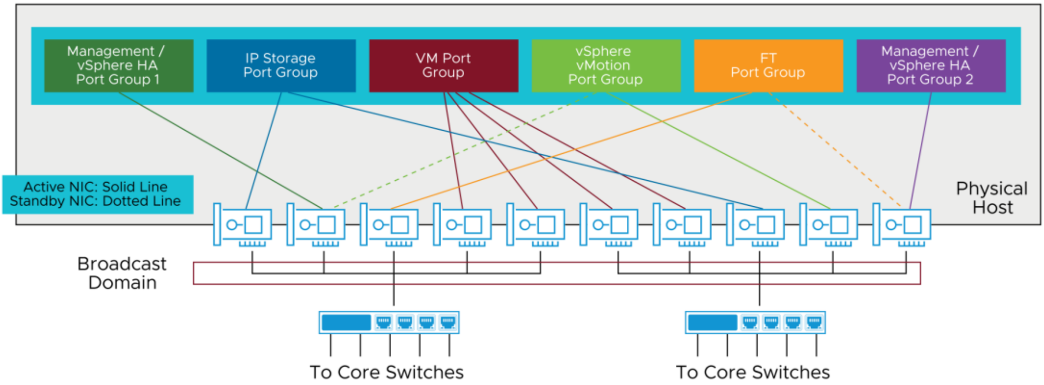


When creating the logical design, the designer looks at each of the major vSphere infrastructure components and makes the appropriate design decisions.

The design phase is an iterative process. As you make design decisions, always validate these decisions against the organization’s goals, requirements, constraints, risks, assumptions, and current best practices.

7-30 Review of NIC Teaming

NIC teams are used to increase the network bandwidth that is available in a network path and provide redundancy to avoid single points of failure.



NIC teaming is done by assigning two or more physical NICs to a virtual switch. Any NIC can be used. However, if the team is built by using ports from multiple NICs and motherboard interfaces, the single points of failure are further reduced and protection is increased.

NIC teaming requires that at least two NICs be exclusively assigned to a virtual switch. Also, all NICs in the same port group must be in the same layer 2 broadcast domain.

The virtual machine port group has four active NIC ports. If a port or physical switch fails, alternate paths are available.

The IP storage port group has two active NIC ports. If either port or physical switch fails, an alternate path is available.

Two separate management port groups provide management and vSphere HA network heartbeat redundancy. If either NIC port or physical switch fails, an alternate path is available.

The vSphere vMotion port group has an active NIC port. If the vSphere vMotion port group or the physical switch fails, an alternate standby path is available.

The vSphere FT port group has an active NIC port. If the vSphere FT port group or the physical switch fails, an alternate standby path is available.

7-31 Design Decision: NIC Teaming Design Considerations

Should NIC teaming be used?

NIC teaming is recommended in most situations.

When configuring NIC teaming, consider the following guidelines:

- Reduce single points of failure by creating a virtual switch with teamed NICs across separate physical switches.
- Ensure better availability by separating physical NIC ports across physical cards.

NIC teaming has the following requirements:

- At least two NICs must be exclusively assigned to a virtual switch.
- All NICs in the same port group must be in the same layer 2 broadcast domain.

To further avoid a single point of failure, configure the NIC team across multiple physical switches when possible. Whether this configuration is possible depends on the following factors:

- The available number of physical switches
- The capabilities of the physical switches
- The load-balancing policy configured on the virtual switch

7-32 Design Decision: Choosing a Load-Balancing Policy

Which load-balancing policy should be used for NIC teaming?

For distributed switches, consider using load-based teaming to gain the following benefits:

- Ensure that a distributed switch's uplink capacity is optimized.
- Avoid the weakness of other teaming policies where some uplinks can be idle while others are saturated.

Otherwise, configure the load-balancing policy for either source port ID or IP hash, depending on hardware support.

Load-Balancing Policy	Chosen Uplink for VM Is Based On
Route based on originating virtual port	The VM port IDs on the virtual switch
Route based on source MAC hash	VM MAC address
Route based on IP hash	The source and destination IP address of each packet
Route based on physical NIC load (load-based teaming)	Current load on the physical NICS
Use explicit failover order	Highest order uplink from the active list

When using Network I/O Control, use load-based teaming as a distributed switch teaming policy to maximize the networking capacity use. Load-based teaming moves a flow only when the mean send or receive use on an uplink exceeds 75 percent of capacity over a 30-second period. Load-based teaming does not move flows more often than every 30 seconds.

If you intend to configure the members of a NIC team across multiple physical switches, using the IP address load-distribution policy requires support in the physical switches. The physical switches must be managed from a single interface as a single switch, where the ports on separate switches can be aggregated into a single logical port. Such physical switches are sometimes called stacked switches or multichassis EtherChannel switches.

Such physical switch capability is not required for the source port ID or MAC address load-distribution policies.

Source MAC address was the default setting for older versions of ESXi. Explicit failover order requires manual configuration.

7-33 Activity: Load-Balancing Policy Discussion (1)

Discuss the advantages and disadvantages of policies for the distribution of traffic, resource consumption, physical switch support, and awareness of load.

Load-Balancing Policy	Advantages	Disadvantages
Route based on originating virtual port		
Route based on source MAC hash		
Route based on IP hash		
Route based on physical NIC load (load-based teaming)		
Use explicit failover order		

7-34 Activity: Load-Balancing Policy Discussion (2)

Load-Balancing Policy	Advantages	Disadvantages
Route based on originating virtual port	Simple. No physical switch configuration.	Does not consider actual load. Risks heavy workloads on same pNIC.
Route based on source MAC hash	Simple. No physical switch configuration. Slightly more random distribution.	Does not consider actual load. Risks heavy workloads on same pNIC.
Route based on IP hash	Most likely to achieve a fair distribution of traffic across pNICs.	Requires correct physical switch configuration.
Route based on physical NIC load (load-based teaming)	Actual traffic is considered. VMs are assigned based on load.	Distributed switches only.
Use explicit failover order	Simple. No physical switch configuration.	Not a load distribution mechanism.

7-35 Design Decision: Choosing Failure Detection Type

Which failure detection type should be used for NIC teaming?

You can choose from the failure detection types and their features:

- Link Status Only:
  - Relies only on the link status that the network adapter provides
  - Detects failures, such as removed cables and physical switch power failures
  - Does not detect blocked physical switch port
  - Does not detect a pulled cable that connects a physical switch to another networking device
- Beacon Probing:
  - Sends out and listens for Ethernet broadcast frames that physical NICs send to detect link failure in all physical NICs in a team
  - Intended for use with three or more NICs in a team that are connected to multiple external switches
  - Can indicate a false failure if a switch is configured to block beacon packets
  - Is unable to detect upstream network failures

The link-state tracking feature allows some ports on the same switch to be defined as upstream ports and other ports to be defined as downstream ports. After the upstream and downstream ports are associated with each other, upstream port link failures are automatically communicated to downstream ports on the same switch.

Beacon probing is most useful to detect failures in the physical switch closest to the ESXi host, where the failure does not cause a link-down event for the host.

The beacon-probing algorithm needs at least three NIC ports to determine where the failure has occurred. With fewer than three NIC ports, beacon probing cannot determine which link has the failure. As a result, the VMkernel floods traffic to both NIC ports to guarantee packet delivery.

For information about beacon probing, see VMware knowledge base article 1005577 at http://kb.vmware.com/kb/1005577.

7-36 Review of Network I/O Control Version 3

Network I/O Control version 3 allocates network bandwidth over distributed switches by using network resource pools:

- Allocate traffic bandwidth using shares, reservations, and limits.
- Reserve no more than 75 percent of the bandwidth of a physical network adapter.
- Allocate bandwidth to individual VMs in the network resource pool for VMs:
  - Use shares, reservations, and limits on a VM’s network adapter.
  - Configure shares, reservations, and limits to a pool associated with a distributed port group.

System Traffic	**Bandwidth Reservation** **(40 GbE NIC)**
vSphere Fault Tolerance	2.0 Gbps
Management	2.0 Gbps
NFS	0.0 Gbps
Virtual Machine	4.0 Gbps
Virtual SAN	3.0 Gbps
iSCSI	3.0 Gbps
vSphere vMotion	2.0 Gbps
vSphere Data Protection Backup	1.0 Gbps
vSphere Replication	1.0 Gbps
Unreserved	22.0 Gbps

Network I/O Control version 3 can be used to allocate network bandwidth to business-critical applications and to resolve situations where several types of traffic compete for common resources.

The amount of bandwidth available to a system traffic type is determined by its relative shares and by the amount of data that the other system features are transmitting.

Reserved bandwidth that is unused becomes available to other types of system traffic. However, Network I/O Control version 3 does not redistribute the capacity that system traffic does not use to virtual machine placement. For example, you configure a reservation of 2 Gbps for iSCSI. The distributed switch might never impose this reservation on a physical adapter because iSCSI uses a single path. The vacant bandwidth is not allocated to virtual machine system traffic. Network I/O Control version 3 can safely meet a potential need for bandwidth for system traffic, for example, for a new iSCSI path where you must provide bandwidth to a new VMkernel adapter.

To enable bandwidth allocation for virtual machines by using Network I/O Control version 3, configure the virtual machine system traffic. Network I/O Control version 3 allocates bandwidth for virtual machines across the entire distributed switch and on the physical adapter carrying the virtual machine's traffic.

Network I/O Control version 3 allocates bandwidth to each type of system traffic by using shares, reservations, and limits:

- Shares: The relative priority, from 1 to 100, of the traffic through the virtual machine's network adapter against the capacity of the physical adapter that is carrying the virtual machine's traffic to the network
- Reservations: The minimum bandwidth, in Mbps, that the virtual machine network adapter must receive on the physical adapter
- Limit: The maximum bandwidth on the virtual machine network adapter for traffic to other virtual machines on the same or on another host

The bandwidth reservation for virtual machine traffic is also used in admission control. When you power on a virtual machine, admission control verifies that enough bandwidth is available. In the example shown, the amount of bandwidth reserved is less than 30 Gbps, which is less than the 75% limit for this 40 GbE NIC.

To guarantee bandwidth, Network I/O Control version 3 implements a traffic placement engine that becomes active if a bandwidth reservation is configured in a virtual machine. The distributed switch tries to place the traffic from a virtual machine network adapter to the physical adapter that can supply the required bandwidth and is in the scope of the active teaming policy.

The actual limit and reservation also depend on the traffic shaping policy on the distributed port group that the adapter is connected to. For example, if a virtual machine network adapter asks for a limit of 200 Mbps and the average bandwidth configured in the traffic shaping policy is 100 Mbps, then the effective limit becomes 100 Mbps.

A network resource pool represents a part of the bandwidth aggregated for network resource pools across the uplinks of a distributed switch.

The bandwidth quota that is dedicated to a network resource pool is shared among the distributed port groups associated with the pool. A virtual machine receives bandwidth from the pool through the distributed port group that the virtual machine is connected to.

7-37 Design Decision: Evaluating the Use of Network I/O Control

Should Network I/O Control be used?

Consider using Network I/O Control for the following use cases:

- You have multiple traffic types following through a single, high-bandwidth network adapter, for example, a 10 GbE, 25 GbE or 40 GbE NIC.
- You want to prioritize critical traffic over the physical network when contention for those resources occurs.

Consider these guidelines when implementing Network I/O Control:

- Use hard reservations and limits sparingly.
- For a latency-sensitive traffic flow, keep the shares value for this resource set to high.

As multiple traffic types flow through a single physical network interface, the traffic must be effectively managed. You do not want critical application flows to suffer because of a burst of low-priority traffic.

Network I/O Control has the following benefits:

- Maintains NFS and iSCSI storage performance in the presence of other network traffic, such as vSphere vMotion migrations and bursty virtual machines
- Provides network service-level guarantees for critical virtual machines
- Ensures sufficient bandwidth for management traffic
- Ensures adequate bandwidth for vSphere FT logging
- Ensures predictable vSphere vMotion performance and duration
- Provides required network resources for the replication process that is used with Site Recovery Manager
- Facilitates any situation where a minimum or weighted level of service is required for a particular traffic type, independent of other traffic types

7-38 Design Decision: Evaluating the Use of Traffic Filtering and QoS Tagging

Should traffic filtering and QoS tagging be used?

For business-critical and latency-sensitive application, consider using traffic filtering and tagging with Network I/O Control.

Traffic filtering has the following benefits:

- Builds on the Network I/O Control qualifiers and is used to provide port-level security
- Allows the distributed switch to reserve bandwidth for the most important traffic for the best possible performance

QoS tagging has the following benefits:

- If the network is congested, the highly tagged traffic does not get dropped, which provides the traffic with a higher QoS.
- When used with Network I/O Control, end-to-end QoS and SLA requirements can be met.

Use the traffic filtering and marking policy to create a set of rules for security and QoS tagging of packets flowing through distributed switch ports.

The traffic filter and marking policy are supported only on distributed switches. The distributed switch can apply network traffic rules on the data path between the virtual machine network adapter and the distributed port. The distributed switch can also apply network traffic rules between the uplink port and the physical network adapter.

Traffic filtering allows an administrator to qualify the following options:

- The type of traffic, such as vSphere vMotion, Management, vSphere FT, and so on
- The destination and source MAC address or the destination and source IP address, protocol, and port

These options allow a filter for inbound, outbound, or both inbound and outbound. Use of these options can provide substantial benefit in a shared traffic environment by enabling segmentation of specific traffic.

Physical network devices use QoS tags to identify important traffic types and provide QoS based on the value of the tag. When using business-critical and latency-sensitive applications, you must consider this technology with Network I/O Control.

7-39 Design Decision: CDP and LLDP Considerations

Should CDP or LLDP be used to monitor the network?

Cisco Discovery Protocol (CDP) and Link Layer Discovery Protocol (LLDP) are data link layer protocols used to discover capabilities of network devices.

Weigh the advantages and disadvantages of enabling CDP or LLDP and consider the organization's requirements and constraints:

- Advantage: Additional network information can be useful for troubleshooting issues.
- Disadvantage: CDP and LLDP create a potential security issue by advertising switch information that should be protected.

If you must use CDP or LLDP, control which interfaces run these protocols because the network topology information can be used by an unauthorized user to breach security.

Using Cisco Discovery Protocol (CDP) and Link Layer Discovery Protocol (LLDP), administrators can automate the deployment and configuration process in a complex network switching environment. These protocols also help avoid downtime because of the misconfiguration of network devices.

Having CDP and LLDP advertising on the host ports is a potential security problem because CDP or LLDP advertises switch information that should be protected.

7-40 Design Decision: NetFlow and Port Mirroring Considerations

Should NetFlow and port mirroring be used to monitor the network?

Drawing on business requirements, decide whether to use NetFlow, port mirroring, or both.

You can use NetFlow and port mirroring to provide visibility into network traffic:

- Security administrators have visibility into the network packets transmitted by the server workloads, if a security breach occurs.
- Network administrators can analyze packets to debug network issues in the virtual infrastructure.

Both NetFlow and port mirroring are available on distributed switches, not standard switches.

Determine whether the infrastructure management design should allow network and security administrators to monitor network traffic of certain virtual machines.

NetFlow capability, with a NetFlow collector tool, has the following benefits:

- Helps monitor application flows and measures flow performance over time
- Helps in capacity planning and ensuring that I/O resources are used properly by different applications, based on their needs

Port mirroring capability duplicates network packets of a switch port to another port and has the following benefits:

- Helps network administrators in debugging network issues in a virtual infrastructure
- Helps administrators fine-tune the traffic that is sent for analysis (ingress, egress, or all traffic)

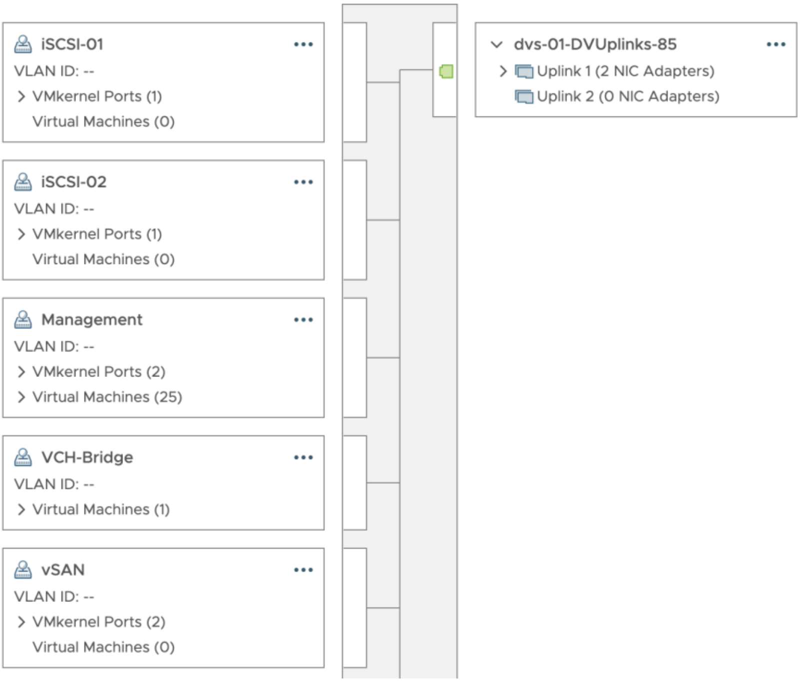
7-41 Design Decision: Exporting Distributed Switch Configurations

Should distributed switch configurations be exported?

The vSphere Distributed Switch configuration is managed through vCenter Server and all virtual network configuration details are stored in the vCenter Server database.

Consider exporting your distributed switch configurations for the following use cases:

- Making a backup of the configuration
- Creating a template of the configuration
- Creating a revision control system for your distributed switch configuration



iSCSI-01
VLAN ID: --
> VMkernel Ports (1)
Virtual Machines (0)

iSCSI-02
VLAN ID: --
> VMkernel Ports (1)
Virtual Machines (0)

Management
VLAN ID: --
> VMkernel Ports (2)
> Virtual Machines (25)

VCH-Bridge
VLAN ID: --
> Virtual Machines (1)

vSAN
VLAN ID: --
> VMkernel Ports (2)
Virtual Machines (0)

dvs-01-DVUplinks-85
> Uplink 1 (2 NIC Adapters)
Uplink 2 (0 NIC Adapters)

You export the distributed switch and distributed port group configuration to a file on the system that is running the vSphere Client. The file preserves valid network configurations, which can be distributed distribution to other deployments.

In addition to your regular vCenter Server backups, you can make periodic backups of your distributed switch configuration with the export function.

You can use the template created from the export function to create similar distributed switch configurations on other vCenter Server systems.

You can keep revisions by saving the distributed switch configuration after each change. By keeping revisions, you can restore the current configuration to an older configuration if necessary.

7-42 DNS Design Considerations

Consider the following guidelines when configuring your DNS servers:

- Configure DNS servers to resolve long and short names and perform forward and reverse lookups for servers used in the vSphere environment:
  - ESXi hosts
  - vCenter Server systems
  - Other services that might be involved in the infrastructure:
    - vRealize Operations
    - vRealize Log Insight
    - NSX-T Manager
    - vRealize Automation
  - IP storage addresses
- Configure DNS servers to be highly available to prevent service outages caused by DNS lookup failures.

To simplify the management of vSphere, all servers involved in vSphere management should be configured in DNS. These servers include servers that host products such as NSX-T Data Center, vRealize Suite components including Log Insight, vRealize Operations, and vRealize Automation.

Although IP storage addresses are not a management component, you should also configure them in DNS.

To increase the availability of DNS, configure redundant servers. Configure each host and the management server systems, such as vCenter Server Appliance with redundant DNS servers. Also, configure virtual machines with redundant DNS servers.

7-43 Design Decision: IPv6 Design Considerations

Should IPv6 addresses be used in the design?

Keep IPv6 enabled only if it is required by the organization because IPv6 increases network management and monitoring overhead.

vSphere supports IPv6 for communication between ESXi hosts, virtual machines, vCenter Server Appliances, NFS storage, iSCSI storage and so on.

The following vSphere features do not support IPv6:

- vSphere Auto Deploy
- vSphere Virtual Volumes (data plane)
- vCenter Server Appliance connected to Active Directory
- ESXi hosts connected to Active Directory
- NFS 4.1 storage with Kerberos

The VMware Compatibility Guide lists the I/O devices that support IPv6.

If you use IPv4 in your environment, no compelling reason exists to enable IPv6, unless IPv6 is a requirement of your organization. Although improvements are made to IPv4, such as integrated network security and an increased address space, enabling IPv6 increases management overhead. Use IPv6 only if necessary.

For information about IPv6 connectivity of vSphere features, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.networking.doc/GUID-35B40B0B-0C13-43B2-BC85-18C9C91BE2D4.html.

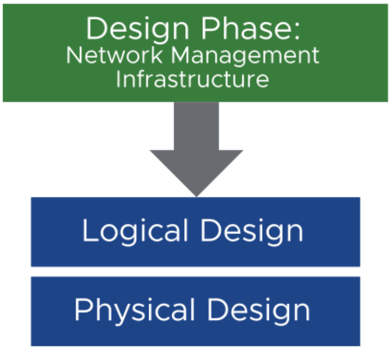
7-44 Lab 11: Designing the Network Management Infrastructure

Create a network management design:

1. Review the Conceptual Design
2. Evaluate Network Management Design Options
3. Document the Network Management Physical Design

7-45 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



7-46 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Evaluate the use of different network management solutions
- Design a network management architecture that meets the needs of the vSphere environment

7-47 Key Points

- Separate different types of network traffic to reduce network contention and latency, and to enhance security.
- Use distributed switches if possible because they centralize network management and offer features that are not available on standard switches.
- NIC teaming reduces single points of failure and is recommended in most situations.
- Use Network I/O Control to prioritize critical traffic that flows through a single, high-bandwidth adapter.

Questions?

Module 8

Virtual Machine Design

8-2 Importance

Virtual machines run the applications and services that support individual users and entire lines of business. You must design, provision, and manage VMs so that applications and services run efficiently. To make the best decisions, you should understand the needs of the applications in your vSphere environment.

8-3 Learner Objectives

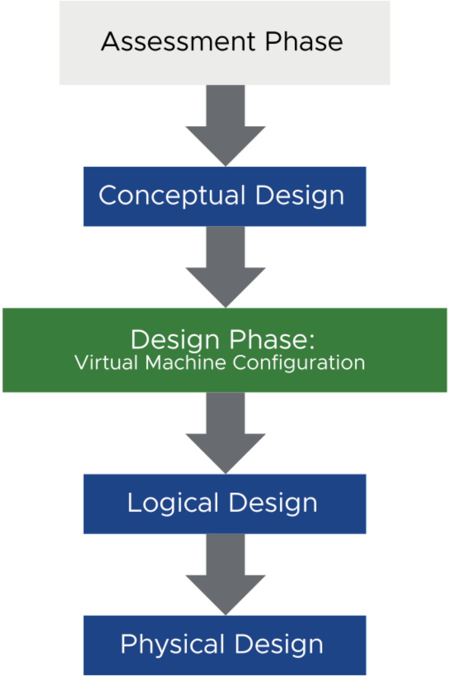
After completing this module, you should be able to meet the following objectives:

- Make virtual machine design decisions, including decisions for resources
- Design VMs that meet the needs of the applications in the vSphere environment and that follow VMware best practices

8-4 Designing Virtual Machines (1)

For VM configurations in your vSphere design, consider:

- Virtual hardware version
- Number of vCPUs and memory size
- Use of shares, reservations, and limits
- Single or multiple virtual disks
- Virtual disk location
- Use of thin-provisioned disks
- VM swap file location



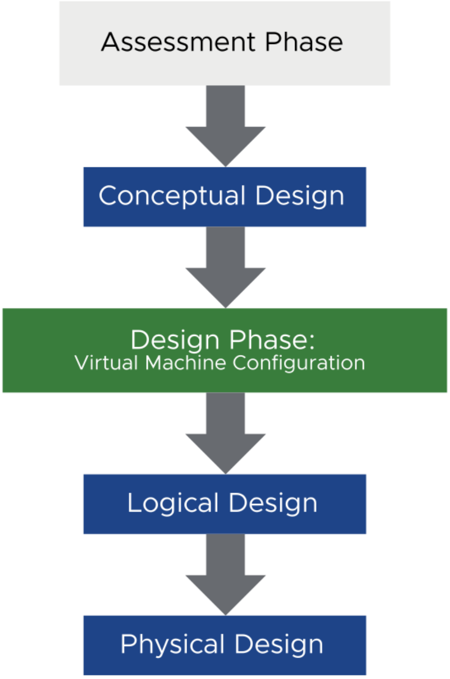
When creating the logical design, the designer looks at each of the major vSphere infrastructure components and makes the appropriate design decisions.

The design phase is an iterative process. As you make design decisions, always validate these decisions against the organization's goals, requirements, constraints, risks, assumptions, and current best practices.

8-5 Designing Virtual Machines (2)

For VM configurations in your vSphere design, consider:

- Virtual SCSI HBA type
- Virtual NICs
- Virtual GPUs
- Use of persistent memory (PMem)
- Use of PVRDMA
- Use of Precision Clock or watchdog timer



8-6 Evaluating Application Requirements

Evaluate application requirements and determine whether the vSphere environment meets the following requirements:

- Hardware and software requirements:
  - Check the VM configuration maximums.
  - Check the guest operating system support.
- Service dependencies
- Functional requirements
- Nonfunctional requirements

The key to optimal performance of an application is to ensure that the correct decisions are made when configuring a VM's resources.

Verify that the vSphere environment meets application hardware and software requirements.

To check VM maximums, see Configuration Maximums at https://configmax.vmware.com/. To verify guest operating system support by a particular ESXi version, see the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility.

Determine whether services that the application needs are accessible in the vSphere environment. Determine whether items in the application’s service-level agreement can be satisfied in the vSphere environment, such as availability, manageability, performance, recoverability, and security.

Nonfunctional requirements include the following operations:

- Availability
- Manageability
- Performance
- Recoverability
- Security

For information about virtualizing business-critical applications, see the following resources:

- https://blogs.vmware.com/apps/
- https://blogs.vmware.com/vsphere/tag/business-critical-apps

8-7 High-Level VM Design Guidelines

Consider the following guidelines whenever possible:

- Use only supported guest operating systems that include VMware Tools.
- Always install VMware Tools on the guest operating system:
  - Some of the primary benefits of VMware Tools include:
    - Efficient memory management by using vmmemctl (the balloon driver)
    - High-performance network drivers, such as VMXNET3
    - Time synchronization, if needed
  - As a best practice, run the latest version of VMware Tools.
- Rightsize your VMs using the following guidelines:
  - Start with the minimum guest resource requirements and then increase resources to improve application performance.
  - Use vRealize Operations Manager to rightsize your VMs and make efficient use of host resource capacity.
  - Decrease resources to reduce any wasted capacity, where possible.

Although a guest operating system can run without VMware Tools, you lose important features and convenience if you do not run it. VMware Tools improves management of the VM by replacing operating system drivers with VMware drivers tuned for virtual hardware.

Rightsizing VMs is the process of optimizing the infrastructure resources that are used by the VMs. This process helps you to allocate sufficient CPU and memory resources to handle the cumulative workload of all VMs.

Rightsizing should be performed regularly to ensure maximum performance of your workloads and efficient use of your underlying hardware. Use a tool such as vRealize Operations Manager to rightsize your VMs.

8-8 Design Decision: Virtual Hardware Version Compatibility

Which virtual hardware version should be used?

Consider the following options before you choose the virtual hardware version:

- Choice of virtual hardware version determines which hosts can run the VM
- Use the latest version unless you have a good reason to use an older version
- Reverting to an older version of hardware is not easy

You can upgrade to a newer version with vSphere Lifecycle Manager.

Compatibility	Virtual Hardware Version
ESXi 7	17
ESXi 6.7 U2 and later	15
ESXi 6.7 and later	14
ESXi 6.5 and later	13
ESXi 6.0 and later	11
Virtual hardware versions 12 and 16 are specific to Workstation and Fusion Pro.	

With each new version of ESXi, new features and capabilities are added to the VM hardware. Depending on the virtual hardware level, and the required feature set, the VM might not run on older hardware. If newer versions of virtual hardware are used, but older hosts exist in the clusters, compatibility problems might arise. A VM cannot run on older hardware if using a newer hardware version.

An organization that has an existing vSphere environment might choose to remain on an older VM hardware version because they cannot upgrade all VMs because of the number of VMs that are deployed. Also, upgrading to a newer VM hardware version can be time-consuming and requires planned outages because the VMs must be turned off. An organization might not need to use the latest features available in the newer VM hardware version.

The VM maximum memory (6128 GB) and maximum logical processors (256) remain the same in virtual hardware 15. USB 3.1. SuperSpeed+ standard is now supported in virtual hardware version 17 as is 4 GB 3D graphics memory.

The Configuration Maximums guide for a particular vSphere release lists the maximum values for the VM hardware version associated with it. To view the configuration maximums, see Configuration Maximums at https://configmax.vmware.com/.

8-9 Design Decision: Determining the Number of vCPUs

How many vCPUs should be configured for each virtual machine?

The number of virtual CPUs (vCPUs) required for a VM depends on the operating system, application, and workload:

- Configure a single vCPU unless the need for Virtual SMP is clear.
- Only workloads that support parallelization can really benefit from Virtual SMP.

If you configure a multiple-vCPU VM, using fewer vCPUs provides a potential performance improvement for both the VM and the host:

- Unused vCPUs consume timer interrupts.
- Maintaining a consistent memory view among multiple vCPUs consumes resources.
- The guest scheduler might migrate a single-threaded workload among multiple vCPUs which loses cache locality.

vSphere Virtual SMP enhances VM performance so that a single VM can use multiple physical processor cores simultaneously. The biggest advantage of a virtual SMP system is the ability to use multiple processors to execute multiple tasks concurrently, which increases throughput. Read the application vendor’s documentation to determine if the application is multithreaded (supports parallelization) and how many CPUs are necessary or possible.

If the workload requires multiple vCPUs, configure as few as possible. The more vCPUs a VM has, the more CPU and memory overhead are necessary to manage it. This additional overhead needlessly wastes resources if the additional vCPUs are not required.

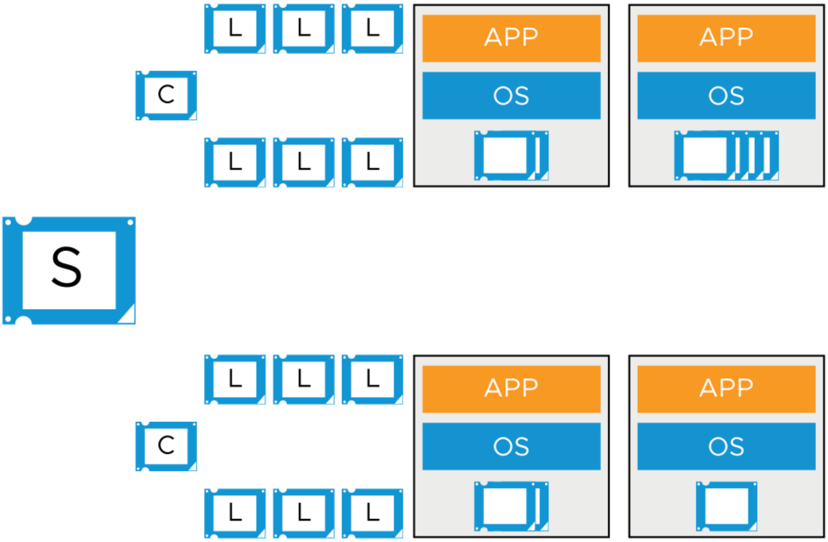
For a sample of the amount of overhead memory on VMs, see *vSphere Resource Management* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-98BD5A8A-260A-494F-BAAE-74781F5C4B87.html.

8-10 Virtual SMP Considerations

When configuring a VM with multiple vCPUs, follow these requirements:

- The guest operating system must support symmetric multiprocessing (SMP).
- The application must be multithreaded.
- The number of vCPUs cannot exceed the number of physical CPU cores (or logical CPUs, if hyperthreading is enabled) on the ESXi host.

  A VM with this configuration cannot be scheduled to run.
- For physical-to-virtual migrations, VM hardware must match the physical machine’s hardware abstraction layer (HAL).



S = Socket
C = Core
L = Logical CPU

If more vCPUs are required, the operating system and application must support them. If the application is not multithreaded, use a scale-out strategy instead. Install multiple VMs, each providing the same service or application to different users.

The VMkernel uses a relaxed co-scheduling algorithm to schedule processors. With this algorithm, when the VM is scheduled to run, every vCPU does not have to be scheduled to run on a physical processor at the same time. The number of vCPUs that run at once depends on the operation being performed at that moment.

For information about CPU scheduling, see *The CPU Scheduler in VMware vSphere 5.1* at http://www.vmware.com/files/pdf/techpaper/VMware-vSphere-CPU-Sched-Perf.pdf.

With some operating systems, scheduling inefficiencies negatively affect performance if a multiprocessor hardware abstraction layer (HAL) runs on a uniprocessor VM. For operating system information related to the HAL, see the guest operating system documentation.

8-11 NUMA and vNUMA Considerations

For applications that are NUMA-aware, understand how NUMA affects your applications.

The ESXi scheduler attempts to maintain good NUMA locality.
For best performance, consider these guidelines:

- Keep the VM size to within the NUMA node size.
  - The number of vCPUs should be less than or equal to the number of cores in the physical node.
- For wide VMs, align vCPUs to physical NUMA boundaries and enable vNUMA:
  - A wide VM has more vCPUs than the available cores in a NUMA node.
  - vNUMA requires virtual hardware version 8 or later, and is enabled by default when the number of vCPUs is greater than eight.

You can obtain maximum performance benefits from vNUMA if your vSphere cluster contains hosts with matching NUMA architecture.

In selecting a home NUMA node for a VM, the scheduler attempts to keep both the VM and its memory on the same node to maintain good NUMA locality.

You must size VMs with the NUMA node size in mind. For example, in a system with four quad-core processors and 128 GB memory, sizing the VM to four virtual CPUs and 32 GB means that the VM does not have to span multiple NUMA nodes. Keeping the VM size to within the NUMA node provides the best performance.

vSphere supports large VMs (up to 128 vCPUs and 6 TB of memory), and these large VMs span NUMA nodes. Ensure that you align vCPUs to physical NUMA boundaries. For example, if a NUMA node has six cores, size your wide VMs with a multiple of six vCPUs: 6 vCPUs, 12 vCPUs, 18 vCPUs, and so on.

Wide VMs are assigned two or more NUMA nodes and are preferentially allocated memory local to those NUMA nodes. Because vCPUs in these wide VMs might sometimes need to access memory outside their own NUMA node, the VM might experience higher average memory access latencies than VMs that fit entirely within a NUMA node.

Virtual NUMA exposes NUMA technology to guest operating systems, which can provide significant performance benefits, although the benefits depend heavily on the level of NUMA optimization in the guest operating system and applications.

8-12 Design Decision: Virtual Memory Considerations

How much memory should be configured for each virtual machine?

Proper sizing of memory is highly dependent on the application:

- Web servers are great candidates for higher consolidation ratios.
- Search and SQL services consume more memory resources than other applications and might benefit from memory reservations.
- Some applications might not perform as expected with vSphere memory management techniques.

To maximize VM memory performance, you must keep a VM's active memory in physical RAM:

- Limit host memory overcommitment or configure VM reservations, or do both.
- Run the `esxtop` command or use the vSphere Client or vRealize Operations Manager to get information about CPU and active memory.

Proper sizing of memory for a VM is based on many factors and depends on the application’s characteristics. For example, web servers have periods where resources can be reclaimed and are great candidates for higher VM-to-host consolidation ratios. For VMs running Search and SQL services, memory reservations can guarantee that those services have the resources that they require while allowing for high consolidation ratios of other VM workloads.

Always ensure that the memory resources are used effectively to help maximize the technology investment in the infrastructure hardware and ensure peak performance.

Active memory is the amount of guest memory that is used by the guest operating system and its applications. Active memory is displayed as active guest memory in the `esxtop` command-line utility, the vSphere Client performance charts, or vRealize Operations Manager.

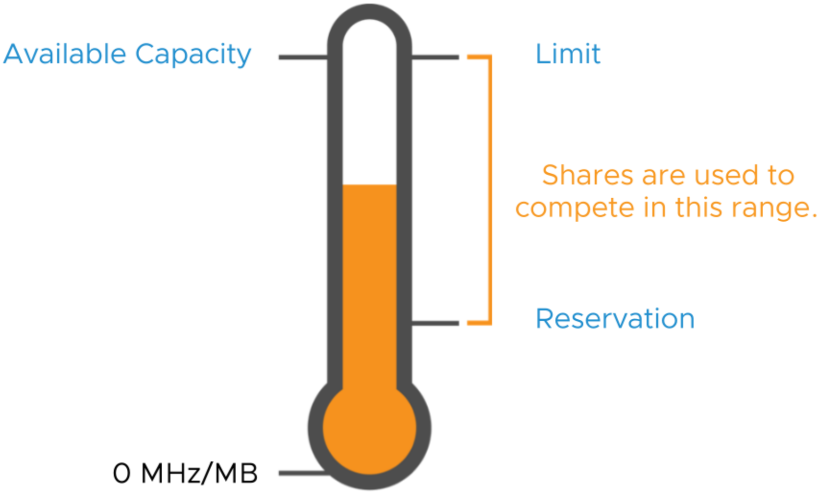
ESXi uses five memory management mechanisms: page sharing, ballooning, memory compression, swap to host cache, and regular swapping to dynamically reduce the amount of machine physical memory required for each VM. Application performance can be significantly affected by these techniques, and the architect should be familiar with the techniques.

8-13 Design Decision: CPU and Memory Shares, Reservations, and Limits

Should shares, reservations, and limits be configured?

Configure shares, reservations, and limits only if necessary. Consider these guidelines:

- Use the default settings to simplify provisioning, administration, and troubleshooting.
- Use shares, reservations, and limits for critical applications or services that must continuously receive CPU and memory resources, even during periods of resource contention.
- Use resource pools as an alternative to reservations. For example, if a department or group paid for a certain amount of CPU or memory resources, use resource pools.



The use of shares, reservations, and limits make CPU and memory configurations for VMs highly configurable.

Using shares, an administrator can set a relative priority for VM access to CPU or memory. If the host’s memory is overcommitted and a mission-critical VM is not achieving an acceptable performance level, the administrator can adjust the VM's shares to escalate the relative priority. The hypervisor allocates more host memory to the mission-critical VM.

Using reservations, an administrator can guarantee CPU or memory to a VM. For example, if a customer-facing or mission-critical application needs a guaranteed memory allocation, you can configure a memory reservation for that VM. If a VM is configured for latency sensitivity, CPU and memory resources are reserved automatically for the VM.

Limits define the maximum amount of physical resources that a VM can consume. Configure this value thoughtfully. If this value is misconfigured, users might experience application performance issues although the host has plenty of resources available.

In general, CPU and memory resources are best reserved or limited using resource pools in a DRS-enabled cluster and detailed per-VM resource configurations used for exceptions.

8-14 Guidelines for Setting Reservations

Reservations must be specified carefully because they can affect the performance of other VMs and significantly reduce the consolidation ratio.

Consider the following factors when configuring CPU or memory reservations:

- Set the memory reservation slightly above the VM's average active memory size.
- Reservations increase administrative overhead, so it might be better to design a consolidation ratio that does not overcommit host memory.
- Instead of creating a large VM swap file, consider setting a high reservation.
- Reservations can increase the resources needed for failover capacity in a vSphere HA cluster.

When you create a large swap file (for example, larger than 100 GB), the amount of time taken for the VM to power on can increase significantly. To avoid this scenario, set a high reservation for large VMs.

The size of the swap file depends on the amount of virtual RAM and memory reservations that are configured.

The swap file size is determined by subtracting the amount of memory reserved (if any) from the amount of memory allocated to the VM. For example, if a VM is configured with 8 GB of RAM and has a memory reservation of 2 GB, the swap file size is 6 GB. The higher the memory reservation value, the smaller the swap file size.

8-15 Design Decision: Virtual Disk Considerations

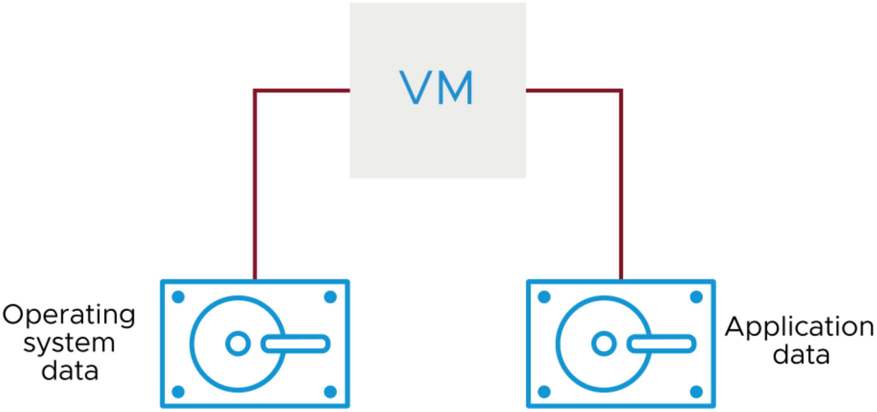
Should single or multiple virtual disks be provisioned?

Consider the following factors when deciding whether to provision a single virtual disk or multiple virtual disks:

- Separating application data and system data helps to distribute the I/O load.
- A separate application data disk can easily be increased in size if necessary.
- A separate application data disk simplifies backups.
- A separate system disk simplifies provisioning.

Use the following guidelines when planning virtual disks:

- Configure one partition per virtual disk.
- If you are using large capacity disks (disks greater than 2 TB), not all vSphere functions are supported such as vSphere FT



A separate application data disk can easily be increased in size without affecting the running operating system on the system disk.

An application data disk can be regularly backed up. If the VM fails, a new VM can be provisioned using a template. After a new VM is running again, the application data can be restored from backup media.

Whether the VM is a Windows or Linux system, separate system and application disks can simplify provisioning. A template can be used to provision a VM with a system disk. If a VM requires an application data disk, you can later add an application data disk that is the appropriate size for that VM.

VMs with large capacity disks have certain conditions and limitations. For example, the guest operating system must support large capacity virtual hard disks. Also, vSphere FT is not supported with large capacity disks. BusLogic Parallel controllers are also not supported.

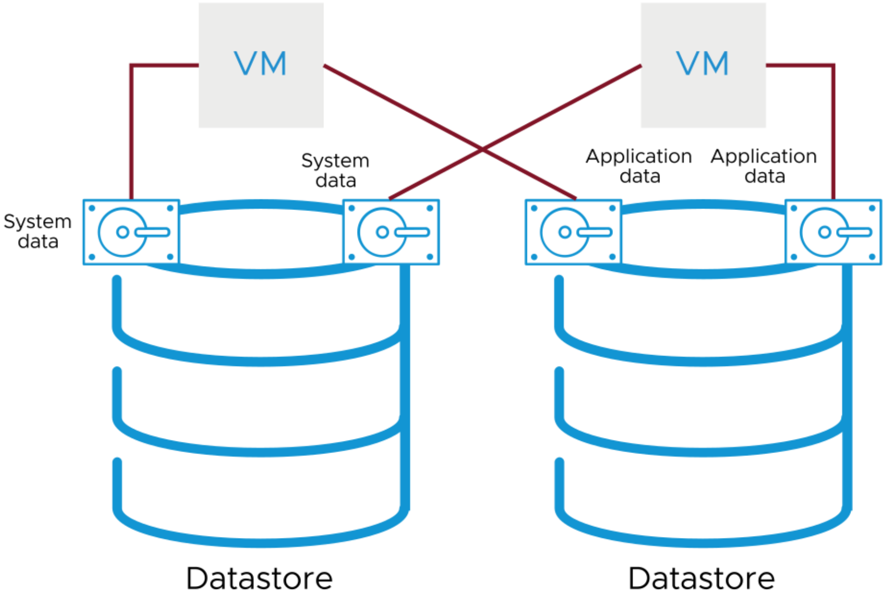
For more information about large capacity virtual disk conditions and limitations, see *vSphere Virtual Machine Administration* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.vm_admin.doc/GUID-55238059-912E-411F-A0E9-A7A536972A91.html.

8-16 Design Decision: Virtual Disk Location

Where should virtual disks be located?

Consider the following guidelines when deciding on where to locate a VM's disks:

- Keep the system and data disks together on a single datastore if they require the same I/O characteristics (RAID level, storage bandwidth, latency):
  - Simplifies array-based replication
  - Simplifies VM snapshot creation
- Keep the VMDKs on shared storage:
  - Supports migration, availability, and load balancing
  - Simplifies administration
- If a VMDK is exceptionally large, put this disk on its own datastore or use a raw device mapping (RDM).



If a VM has multiple disks, placing the virtual disks on different datastores can improve performance. Keeping the disks on the same datastore simplifies configuration and administration. For example, if you plan to replicate the VM's files to another logical unit number (LUN) or data center for availability, replicating a single LUN within the storage array software replicates all the VM’s files.

In most cases, store VMDKs together on shared storage rather than on local storage. Nearly all the more important benefits of vSphere, such as vSphere vMotion, vSphere HA, and vSphere DRS, depend on shared storage.

Local storage might be required. For example, local storage might be more secure. Local storage might also be less expensive and be used by smaller organizations. More commonly, local storage is used by vSAN to provide software-defined shared storage.

8-17 Design Decision: Evaluating the Use of Thin Provisioning

Should thin-provisioned disks be used?

Understand the tradeoffs between thick-provisioned and thin-provisioned virtual disks.

Consider thin-provisioned disks if data growth is slow or static.

	Benefits	Drawbacks
Thick Provisioning	• Simpler storage configuration and monitoring	• Less efficient space use • Higher storage costs
Thin Provisioning	• Decreases the cost of storage • Eliminates storage underutilization • Eliminates the need to dedicate full capacity up front	• Increases the risk of running out of datastore space • Must create alarms and monitor for overcommitment and disk use

Thin provisioning is a method that optimizes storage use by allocating storage space in a flexible on-demand manner. Thin provisioning contrasts with the traditional model, called thick provisioning. With thick provisioning, large amounts of storage space are provided in advance and to anticipate future storage needs. However, the space might remain unused causing underutilization of storage capacity.

With thick provisioning, you have less need to configure the alarms and alerts necessary to ensure that you do not unexpectedly run out of space in the datastore. You also have less need to increase the size of the VMFS datastore in response to VMDK file growth. However, vSphere Storage DRS usage balancing features can be used when one datastore in a datastore cluster starts to run out of space, possibly as a result of thin-provisioned disks expanding in size.

Thin provisioning might not be suitable for environments where virtual machine disk sizes have a rapid rate of growth. Virtual machine disks that rapidly approach their maximum size reduce the benefit of thin provisioning.

8-18 Thin Provisioning Considerations

Consider the following guidelines if you choose thin-provisioned disks:

- Two models of thin provisioning are available:
  - Array-level thin provisioning, with vSphere Storage APIs - Array Integration
  - Virtual disk-level provisioning, with vSphere Thin Provisioning
- Both models can reduce costs, but the tradeoff is increased management complexity.
- You can use both models together as long as they are managed correctly.
- If the operating system initial disk format operation writes zeroes to all sectors, the disk is prematurely inflated, which negates the benefits of virtual disk-level thin provisioning.
- Fragmentation on thin-provisioned disks has negligible effects on disk performance.
- NFS datastores are thin-provisioned by default.
- For VMFS and NFS, select a storage array configuration that supports adding physical storage space.

You can use virtual disk-level thin provisioning and array-side thin provisioning together in your vSphere environment. However, you should carefully monitor and manage thin-on-thin storage systems.

In an NFS environment, the monitoring and management of thin provisioning is done at the NFS server and not through the vSphere Client. The vSphere Client lacks complete visibility into the NFS server storage and cannot report on the thin provisioning done in the NFS server store.

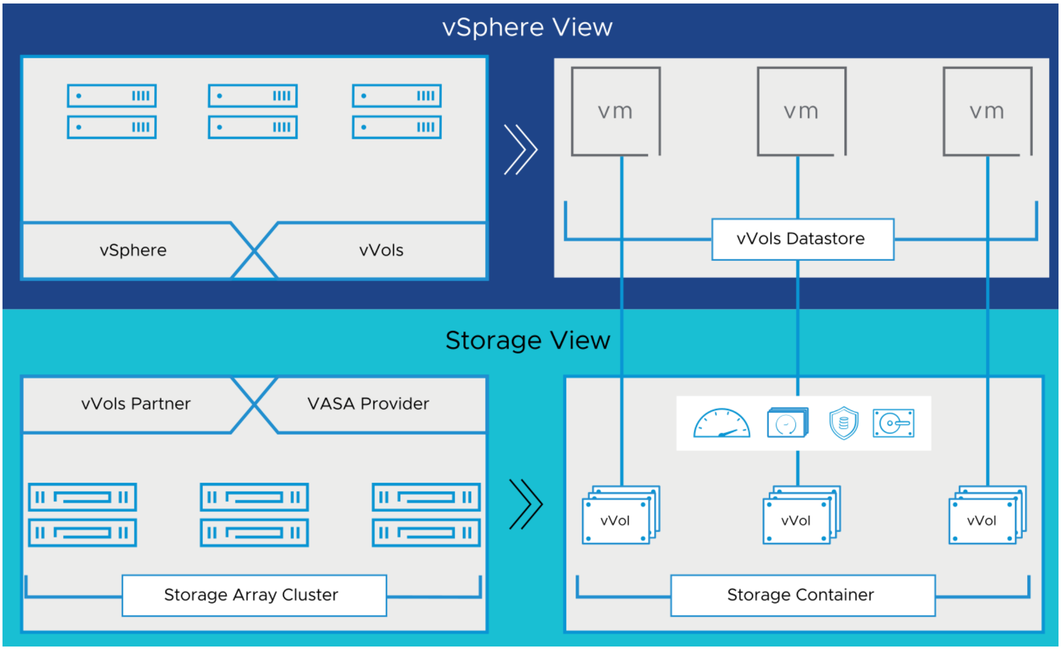
As the size of the thin-provisioned virtual disk files increase over time, additional space might need to be added to the datastore. Ensure that you understand the time involved for ordering, adding, and configuring additional physical storage space. Also ensure that you have configured the appropriate alarms and alerts so that you do not unexpectedly run out of storage space.

8-19 Design Decision: Evaluating the Use of vSphere Virtual Volumes

Should vSphere Virtual Volumes be used?

LUNs are rigid and fixed in their capabilities. Consider using vSphere Virtual Volumes for the following design requirements:

- Array-level operations scoped to vSphere Virtual Volumes
- Simplification of consumption of array capabilities
- Multiple capabilities in a single datastore
- Policy driven storage from vSphere



LUNs or volumes are rigid and fixed in their capabilities. They also require a file system such as VMFS or NFS. vSphere Virtual Volumes has no file system. A file system is created in each virtual volume. That file system depends on the OS in the data virtual volume or the vSphere Virtual Volumes function, whether it is a config, memory, swap, or other function. In terms of capabilities, a LUN or volume is generally provisioned with a specific set of capabilities that are not easily changed without creating a LUN or volume and migrating the data.

With vSphere Virtual Volumes, changing the capability is simple. You can change a storage policy based management (SPBM) policy and the capability is changed. vSphere Storage vMotion or data migration is not required. You can change from spinning media to all-flash with a policy change, depending on the array. The array handles the performance and data migration, and no administration action is required. Another key aspect of vSphere Virtual Volumes over traditional storage is the array. vSphere has complete insight into the data and I/O requirements. The array can manage the I/O and data accordingly.

vSphere Virtual Volumes provides the following functionality:

- Native representation of VMDKs on SAN/NAS: No LUNs or volume management.
- Works with existing SAN/NAS systems.
- A new control path for data operations at the VM/VMDK level.
- Snapshots, replications, and other operations at the VM level on external storage.
- Automates control of per-VM service levels.
- vSphere API for Storage Awareness protocol endpoint provides standard protocol access to storage.
- Storage containers can span an entire array.

8-20 Design Decision: VM Swap File Considerations

Where should virtual machine swap files be located?

By default, the swap file is stored in the same directory as the VM’s other files.

Consider the following guidelines when placing VM swap files:

- Place swap files on solid-state drives to reduce performance issues caused by actively using swap files.
- Place swap files on nonreplicated datastores.
- Do not place a swap file on a thin-provisioned LUN.

Relocating the swap file can affect the following:

- Ease of administration and provisioning
- vSphere vMotion performance
- Datastore replication performance

When a VM is powered on, the system creates a VMkernel swap file to serve as a backup for the RAM contents of the VM. The default swap file is stored in the same location as the configuration file of the VM. This process simplifies the configuration, but can cause an excess of replication traffic.

The size of the swap file depends on the amount of RAM and reservations that are configured for the VM.

Place swap files on a nonreplicated datastore to eliminate the additional replication traffic and prevent slow vSphere vMotion migrations.

If you run a VM with a swap file that is stored on a thin-provisioned LUN, that LUN can cause swap file growth failure and the VM might end.

8-21 VM Swap File Location

Swap file location depends on an organization's business requirements.

Swap File Location	VM Files Location	Benefits	Drawbacks
Shared storage	Shared storage	• Simplest configuration to administer • vSphere vMotion migration proceeds at best possible rate	• Can increase the required replication bandwidth for products that replicate datastores
Local storage	Shared storage	• Reduces the need for more expensive network storage • Reduces bandwidth for products that replicate the datastore • Does not affect VM performance	• Slows vSphere vMotion migration and vSphere DRS operations • Increases administrative overhead
Dedicated datastore	Separate datastore	• Can improve replication performance for products that replicate datastores	• Increases administrative overhead
Local storage	Local storage	• Reduces the need for more expensive network storage	• Prohibits operations that use shared storage

Swap file location can affect the performance of vSphere vMotion and vSphere DRS. During a migration with vSphere vMotion, if the swap file location specified on the destination host differs from the swap file location specified on the source host, the swap file is copied to the new location. This activity can result in slower migrations with vSphere vMotion.

Swap file location can make administration and provisioning more difficult. If a swap file is not located in the same directory as the VM's other files, administration is more difficult than if all the VM's files are in a single directory.

Swap file location can affect datastore replication performance. Replicating swap data does not make sense because swap files are recreated when a VM powers on. So placing the swap file separately from the rest of the VM's files benefits replication by reducing the bandwidth requirements.

Placing all the VM's files on a local disk is normally not used in a data center solution. This configuration might be appropriate for remote offices with only one or two ESXi hosts.

8-22 Design Decision: Virtual SCSI HBA Type

Which virtual SCSI HBA type should be used?

Follow these guidelines when choosing the virtual SCSI HBA type:

- Use the default choice unless this choice does not support a required feature.

  For example, Microsoft Windows Server 2012 cluster services require a Serial Attached SCSI device.

- Consider using the VMware Paravirtual SCSI (PVSCSI) adapter in environments where hardware or applications drive a high amount of I/O throughput:

  — Suitable for SAN environments.

  — Does not provide any performance benefits for the following use cases:

    - Local disk storage environments
    - VMs with snapshots

  — You must use a supported guest operating system.

If you consistently configure a nondefault choice, create a template to simplify VM provisioning.

VMware Paravirtual SCSI (PVSCSI) Host Bus Adapters (HBAs) increase throughput while lowering CPU overhead. PVSCSI is helpful for workloads such as performance-critical database applications.

Disks with snapshots might not experience performance gains when used on Paravirtual SCSI adapters if memory on the ESXi host is overcommitted.

For platform support for VMware Paravirtual SCSI controllers, see the VMware Compatibility Guide at http://www.vmware.com/resources/compatibility.

8-23 Design Decision: Virtual NIC Considerations

Which virtual NIC should be used?

Consider the following guidelines when choosing a virtual NIC:

- Consider using VMXNET3 adapters for their enhanced performance and feature set.
- If VMXNET3 is not an option, choose the newest driver that the guest operating system supports.
- In all cases, use a virtual network adapter that the guest operating system and VMware Tools support.

When configuring a virtual NIC, set the operating system correctly when creating a virtual machine.

The type of network adapters available to the VM depends on the following factors:

- The version of the VM hardware, which depends on which host created it or most recently updated it
- Whether the VM is updated to the latest version for the current host
- The guest operating system

The VMXNET3 adapter is a paravirtualized NIC designed for performance and is not related to VMXNET or VMXNET2. The VMXNET3 adapter offers all the features available in VMXNET2 and adds several features such as multiqueue support (also known as Receive Side Scaling in Windows), IPv6 offloads, and MSI/MSI-X interrupt delivery. Because operating system vendors do not provide built-in drivers for this card, you must install VMware Tools to have a driver for the VMXNET3 network adapter available.

For more information about network adapter types, see *vSphere Networking* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.networking.doc/GUID-35B40B0B-0C13-43B2-BC85-18C9C91BE2D4.html.

8-24 Design Decision: Virtual GPU Considerations (1)

Should virtual GPUs be used for graphics acceleration?

Virtual GPUs add support for graphics acceleration. Consider adding virtual GPUs, if you have one of the following use cases:

- Rich 2D and 3D graphics
- Graphics-intensive applications, for example, applications used by architects and engineers
- VMware Horizon virtual desktops

vSphere supports virtual GPU graphics acceleration options.

Virtual GPU Type	GPU Partner	ESXi Version
vSGA	NVIDIA, AMD*	ESXi 6.5 and later
vDGA	NVIDIA, Intel, AMD*	ESXi 6.0 and later
NVIDIA GRID vGPU	NVIDIA*	ESXi 6.0 and later

** See the VMware Compatibility Guide for specific models.*

Virtual Shared Graphics Acceleration (vSGA) provides hardware-accelerated 3D graphics by allowing multiple virtual machines to share physical graphics processing units (GPUs) installed locally in the ESXi hosts.

Virtual Dedicated Graphics Acceleration (vDGA) graphics adapters installed in the underlying host are assigned to virtual machines using vSphere DirectPath I/O. Assigning a discrete GPU to a virtual machine dedicates the entire GPU to it.

With GRID vGPU, a single GPU can be shared among multiple virtual desktops. GRID vGPU uses native NVIDIA drivers and supports newer releases of OpenGL and DirectX.

When configuring the virtual GPU, select the appropriate rendering mode for the virtual machine's needs: Automatic, hardware, or software. vSphere vMotion is available as long as the physical hardware exists on the destination host.

For more information about vSGA and vDGA, see Graphics Acceleration for View Virtual Desktops in Horizon 7 at https://www.vmware.com/techpapers.

For more information about NVIDIA GRID vGPU, see VMware Horizon and NVIDIA GRID vGPU at http://www.vmware.com/files/pdf/products/horizon/vmware-nvidia-grid-vgpu-FAQ.pdf.

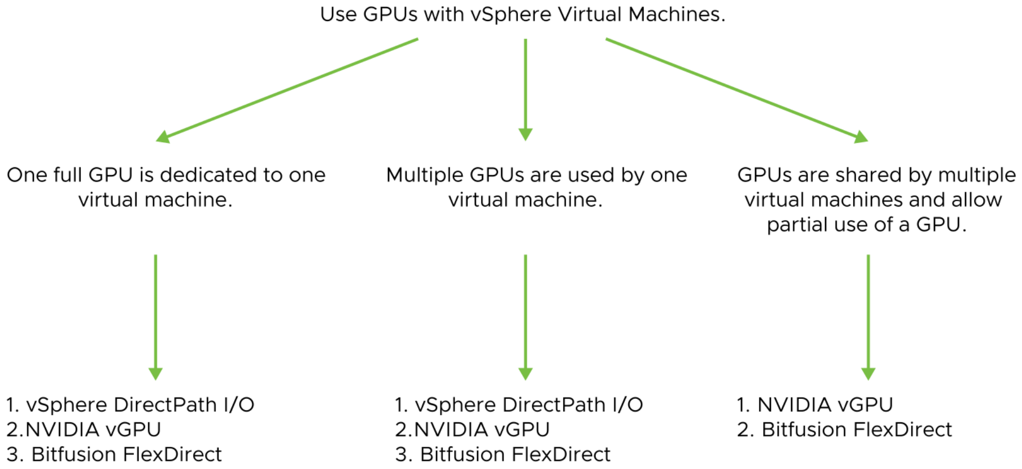
8-25 Design Decision: Virtual GPU Considerations (2)

Should virtual GPUs be used for machine learning inference or other data science applications?

GPUs for compute use cases can be connected to VMs in three ways.

The choice is a function of use case and GPU dedication to the VM.

Compute Workloads: Decision Tree for Different Use Cases



Use GPUs with vSphere Virtual Machines.

One full GPU is dedicated to one virtual machine.

Multiple GPUs are used by one virtual machine.

GPUs are shared by multiple virtual machines and allow partial use of a GPU.

1. vSphere DirectPath I/O
2.NVIDIA vGPU
3. Bitfusion FlexDirect

1. vSphere DirectPath I/O
2.NVIDIA vGPU
3. Bitfusion FlexDirect

1. NVIDIA vGPU
2. Bitfusion FlexDirect

Your company's data scientists, machine learning practitioners, or developers ask you to provide them with a GPU-capable machine setup to do their work. They want to execute workloads that need GPU compute power. The data scientist describes the workloads as machine learning training, inference, or development.

The reason your end users need GPU capability is for faster time to results. Machine learning models involve large matrix multiplications, and GPUs are designed to compute these operations much faster than CPUs.

One of your early decision points as a system administrator is to determine how the GPUs are used in your environment. VMs can use GPUs in different ways. The approach that you take depends on the type of users and applications that use the GPUs for their applications:

- A full GPU dedicated to one VM
- Multiple GPUs dedicated to one VM
- Sharing a GPU across multiple VMs, including partial GPU use

Some use cases are provided by VMware partners' products such as the NVIDIA Virtual GPU, also called NVIDIA vGPU technology. This family was formerly named NVIDIA Grid. The NVIDIA vGPU set is a family of software products that includes the NVIDIA Virtual Compute Server (vCS) software product as well as others, such as NVIDIA Quadro vDWS.

The Bitfusion FlexDirect software increases the flexibility with which you might use your GPUs on vSphere. It does so by allowing your physical GPUs to be allocated in part or as a whole to applications running in VMs. Those consumer VMs might be hosted on servers that do not themselves have physical GPUs attached to them. Bitfusion FlexDirect uses techniques for remote execution of the CUDA instructions to other servers to achieve this use of the physical GPUs. Bitfusion was acquired by VMware in 2019.

8-26 Guest Operating System Considerations

Consider the following guidelines when choosing guest operating systems:

- Simplify administration and troubleshooting by keeping the variations minimal for each guest operating system.
- Use standardized templates for the installation of each key application.
- Simplify administration, troubleshooting, and chargeback by using standard sizing for virtual machines based on your business requirements.

Standard sizing exists for small, medium, large, and extra-large virtual machines.

Item	Small VM	Medium VM	Large VM	Extra-Large VM
CPU	2	8	16	64
RAM	4 GB	16 GB	256 GB	1 TB
Disk	100 GB	300 GB	1 TB	4 TB

vSphere supports many different guest operating systems to be run at the same time, replacing legacy hardware.

The table illustrates an example of creating virtual machines in standard sizes. An organization can choose to use standard sizing, using sizing values that are specific to the organization's environment. An alternative might be for the organization to develop a more dynamic sizing policy. This dynamic policy might be to start with the minimum resource requirements for the virtual machine, then increase the resources as needed.

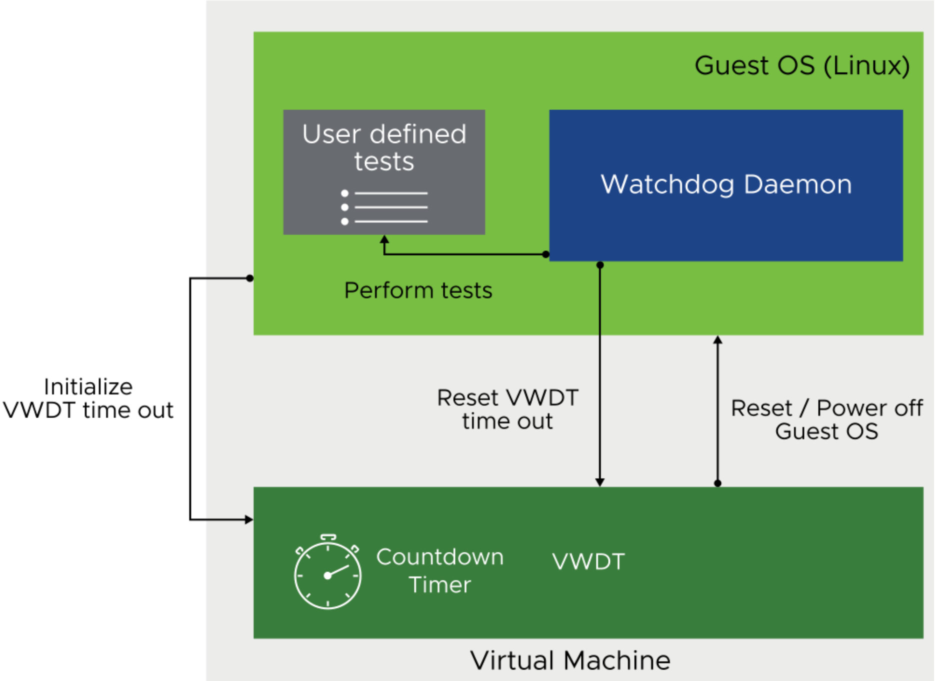
8-27 Design Decision: Evaluating the Use of the Virtual Watchdog Timer Device

Should a virtual watchdog timer device be used?

The virtual watchdog timer (VWDT) is used to detect and recover from OS problems.

The VWDT is useful in the following scenarios:

- With high availability solutions such as Red Hat High Availability and the MS SQL failover cluster
- On VMware Cloud and in hosted environments for implementing custom failover logic to reset or power off VMs



This feature is based on Microsoft specifications: Watchdog Resource Table (WDRT) and Watchdog Action Table (WDAT).

The virtual watchdog timer (VWDT) detects and recovers from OS problems. The device's timeout can be initialized with VM power on or explicitly by the guest OS. The guest OS or an application in the guest OS can reset the device's timeout based on a predefined logic. If the VWDT is not reset before its timeout, it resets or powers off the guest OS.

How the guest OS or the application in the guest OS interacts with the VWDT differs depending on the guest OS and the application that interacts with the virtual device. Configuring the guest OS or the application that interacts with the VWDT is beyond the scope of this lesson.

If a failure occurs, the VWDT helps recovery:

- The guest OS or application can reset the timeout.
- The responses of VM power off or VM reset are appropriate for the workload.

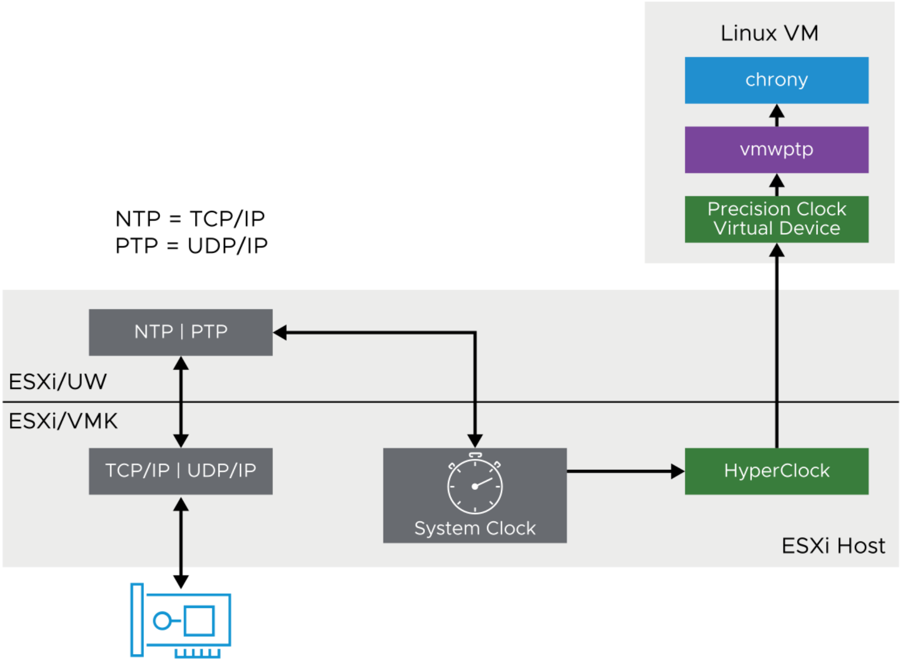
8-28 Design Decision: Evaluating the Use of the Precision Clock Virtual Device

Should the Precision Clock virtual device be used?

Precise timekeeping is a requirement for many applications, such as financial services applications.

Consider Precision Clock in the following scenarios:

- Your Linux-based application requires timekeeping within 1 second accuracy.
- The ESXi host is set for NTP or PTP time synchronization.
- The chrony time-sync agent can be deployed in the Linux guest.
- The VM can only run on vSphere 7 and above.



In financial services applications, the time stamps on transactions and records must be precise. The financial sector must adhere to many regulations to time stamp records and transactions. These time stamps must be accurate within a 1 second range.

The timing model shows the following configurations:

- Precision Clock is a new virtual device in vSphere 7.
- HyperClock is a new VMkernel module in vSphere 7.
- vmwptp is a Linux device driver for Precision Clock.
- Chrony is a Linux time-synchronization agent.
- Timing packets move through the ESXi host network interface to NTP or PTP.
- NTP or PTP synchronizes the system clock.
- Hyper Clock gets the time from the system clock.
- Hyper Clock pushes this time to Precision Clock.
- Chrony consumes time that is provided by Precision Clock through vmwptp, to synchronize the guest OS time.

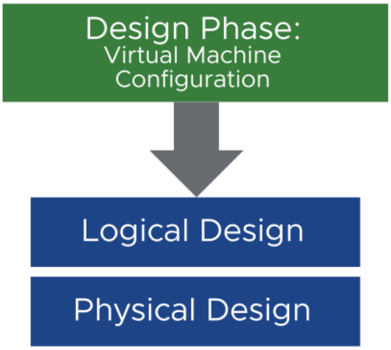
8-29 Lab 12: Designing Virtual Machines

Create a virtual machine design:

1. Review the Conceptual Design and Capacity Planning Assessment Data
2. Evaluate Virtual Machine Design Options

8-30 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.


Design Phase:
Virtual Machine Configuration

Logical Design

Physical Design

8-31 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Make virtual machine design decisions, including decisions for resources
- Design VMs that meet the needs of the applications in the vSphere environment and that follow VMware best practices

8-32 Key Points

- The key to optimal performance of an application is to ensure that the correct decisions are made when configuring a virtual machine's resources.
- Use tools such as vRealize Operations Manager to rightsize your VMs.
- Use only supported guest operating systems that include VMware Tools.
- Default to one vCPU unless a clear need for more vCPUs exists.
- Maximize memory performance by keeping a virtual machine's active memory in physical RAM.
- Keep the system and data disks together on a single datastore if they require the same I/O characteristics.
- Use VMXNET3 adapters for its enhanced performance and feature set.

Questions?

Module 9

Infrastructure Security

9-2 Importance

vSphere components are secured by several features, such as certificates, authorization, limited access, and so on.

You can modify the default setup in many ways, for example, by setting permissions on vCenter Server objects, opening firewall ports, and enabling additional security features. vSphere gives you flexibility in securing vCenter Server systems, ESXi hosts, and virtual machines.

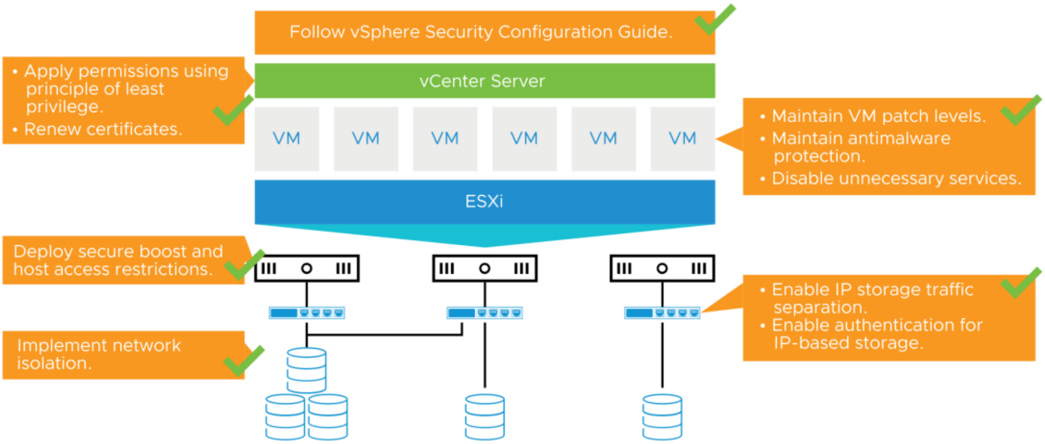
9-3 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Make security design decisions for various layers in the vSphere environment
- Design a security strategy that meets the needs of the vSphere environment and follows VMware best practices

9-4 vSphere Infrastructure Security Best Practices

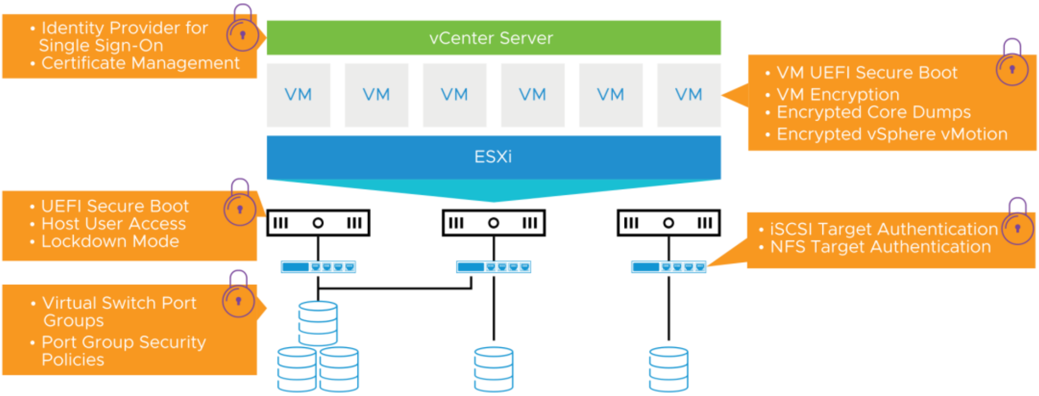
Implement security best practices at the various levels of the vSphere infrastructure.


Follow vSphere Security Configuration Guide.
vCenter Server
VM
VM
VM
VM
VM
VM
ESXi
• Apply permissions using principle of least privilege.
• Renew certificates.
• Maintain VM patch levels.
• Maintain antimalware protection.
• Disable unnecessary services.
Deploy secure boost and host access restrictions.
Implement network isolation.
• Enable IP storage traffic separation.
• Enable authentication for IP-based storage.

From a technical standpoint, the vSphere virtualization platform is an inherently secure environment. vSphere has a minimal hypervisor footprint, provides APIs for monitoring that eliminate the need for third-party software on the host, uses secure logging with Syslog, integrates with Active Directory (AD), and more.

9-5 vSphere Infrastructure Security Features

Security features are available for the various levels of the vSphere infrastructure.


• Identity Provider for Single Sign-On
• Certificate Management
vCenter Server
VM
VM
VM
VM
VM
VM
• VM UEFI Secure Boot
• VM Encryption
• Encrypted Core Dumps
• Encrypted vSphere vMotion
ESXi
• UEFI Secure Boot
• Host User Access
• Lockdown Mode
• iSCSI Target Authentication
• NFS Target Authentication
• Virtual Switch Port Groups
• Port Group Security Policies

Secure the different layers of the vSphere infrastructure in your security strategy by following these guidelines:

- Your vCenter Server system and associated services are protected by authentication through vCenter Single Sign-On and by authorization through the vCenter Server permissions model.
- To secure your virtual machines, keep the guest operating systems patched and protect your environment in the same way that you protect a physical machine.
- The ESXi hypervisor is secured out of the box. You can further protect ESXi hosts by using lockdown mode and other built-in features.
- vSphere includes the full array of features necessary for a secure networking infrastructure. You can separately secure each element of the infrastructure, such as virtual switches, distributed virtual switches, virtual network adapters, and so on.
- Follow best practices for storage security as outlined by your storage security provider.

9-6 General Security Design Guidelines

Consider the following guidelines whenever possible:

- Ensure that all systems use the same relative time source, such as an NTP server:
  - Simplifies the task of tracking and correlating an intruder's actions when reviewing log files
- Enforce strong passwords and password policies for all vSphere users.
- Regularly check for security alerts that might impact your environment:
  - Sign up to receive new and updated VMware security advisories in email.

Ensure that all systems use the same relative time source including the relevant localization offset. You must also ensure that the relative time source can be correlated to an agreed-on time standard, such as Coordinated Universal Time (UTC). Incorrect time settings can make it difficult to inspect and correlate log files to detect attacks and can make auditing inaccurate.

As a best practice, require all ESXi users, vCenter Server users, and vCenter Single Sign-On users to create strong passwords. Password restrictions, lockout, and expiration in your vSphere environment depend on the system that the user targets, who the user is, and how policies are set.

VMware security advisories document remediation for security vulnerabilities that are reported in VMware products. Sign up to receive security advisories through email at http://www.vmware.com/security/advisories.

9-7 Security Design Resources

Design your vSphere security strategy by using the VMware security resources:

- vSphere Security:
  - Provides operational guidelines, or best practices, for securing a vSphere environment.
  - The IT administrator and security teams should discuss how to apply these guidelines to the environment.
- vSphere Security Configuration Guide:
  - Formerly known as the vSphere Hardening Guide.
  - Spreadsheet provides security guidelines that can be checked with APIs, CLIs, and other tools, for attestable values.

vSphere Security provides operational guidelines, whereas the vSphere Security Configuration Guide provides programmatic guidelines. Operational guidelines are generally open to interpretation and can be mitigated in various ways. These guidelines usually cannot be automatically tested and verified. Instead, individual signing off is required after each guideline is met.

You can check programmatic guidelines by using a script or program. For example, one of the guidelines in the security configuration guide is to configure Network Time Protocol (NTP) time synchronization. The guide provides the command line, which you can use in a script, to check that NTP time synchronization is configured.

For more information about security, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-52188148-C579-4F6A-8335-CFBCE0DD2167.html.

9-8 About the vSphere Security Configuration Guide

The vSphere Security Configuration Guide is a spreadsheet that helps you to configure your vSphere environment according to operations security best practices:

- Over time, vSphere has become more secure by default.
- As a result, this guide is less about hardening and more about ensuring that best practices are followed.

The spreadsheet categorizes each setting based on one or more characteristics.

Characteristic	Description	Example of setting
Hardening	If set to true, this setting is an actual hardening setting.	VM.disable-console-gui-options: Explicitly disable VM copy and paste operations.
Site specific	If set to true, this setting is one that VMware cannot provide default values for.	ESXi.set-password-policies: Establish a password policy for password complexity.
Audit	If set to true, check setting regularly for changes to the default value.	vNetwork.reject-mac-changes-dvportgroup: Ensure that the MAC Address Changes policy is set to reject.

The vSphere Security Configuration Guide replaces the vSphere Hardening Guide, which is available for vSphere 6.0 and earlier. vSphere already meets the Common Criteria standards of security without any changes applied from the vSphere Security Configuration Guide. This guide helps you to maintain a high standard of security for vSphere.

The vSphere Security Configuration Guide is provided in a spreadsheet format, with metadata to allow for guideline classification and risk assessment. The guide also includes script examples for enabling security automation.

The Hardening, Site specific, and Audit characteristics in the table can be found in the spreadsheet in columns V, W, and X, respectively.

For information about the Common Criteria standards, see https://www.commoncriteriaportal.org.

For configuration information, see the vSphere Security Configuration Guide at http://www.vmware.com/security/hardening-guides.

9-9 Design Decision: Accessing vCenter Server

How should vCenter Server access be controlled?

Strictly control access to different vCenter Server components.

Consider the following guidelines:

- Require that all users log in to vCenter Server using the vSphere Client:
  - Login events are logged by vCenter Server and can be audited if necessary.
  - Do not allow anyone to log in directly to vCenter Server Appliance.
- Create named vCenter Server users and groups in your existing directory service (Active Directory, OpenLDAP, and so on):
  - For example, do not use the domain administrator user account or the Domain Administrators group account.
- Use of the local operating system users and groups from the vCenter Linux appliance is not recommended.
- Configure an additional vCenter Single Sign-On administrator account to access the system in case an account is locked out.
- Evaluate the use of any existing ADFS single sign-on with vCenter Server identity federation.

Avoid allowing users to log in directly to vCenter Server Appliance. Allow only those users who have legitimate tasks to log in to the appliance and confirm that these events are audited.

Ensure that vCenter Server and the ESXi hosts are connected to a directory service. After completion, users and groups in the directory service should be created to simplify user and group management and to present a consistent user and group view to any interface managing the environment.

Instead of using a generic group, such as Domain Administrators, create a specific group for vCenter Server system administration. Creating a specific group for vSphere administrators reduces the risk of Windows-based administrators not trained in vSphere from gaining privileged access to the vCenter Server system.

9-10 Role-Based Access Control

Consider the following vCenter Server access control guidelines:

- Apply the principle of least privilege to all vCenter Server users.
- Grant full vCenter Server administrative rights only to those administrators who need it.
- Do not grant vCenter Server roles to generic groups, for example, if using AD authentication source then do not use Domain Administrators

vCenter Server Role	AD or LDAP Group	vSphere Privileges	Inventory Level for Permissions	Description
Enterprise vSphere Administrators	VI-Admins	All	Data center and all child objects	Administrative rights to the entire vSphere infrastructure
vSphere Storage Administrators	SAN-Admins	Datastore privileges, storage view privileges	Datastores and its child objects only	Administrative rights to vSphere storage components only
Operators	Admin-1	Read-only privileges	Data center and all child objects	Used for monitoring purposes

The principle of least privilege gives users the minimal amount of access that is enough for them to do their jobs. Applying this principle simplifies vCenter Server administration and enhances security by reducing the attack surface.

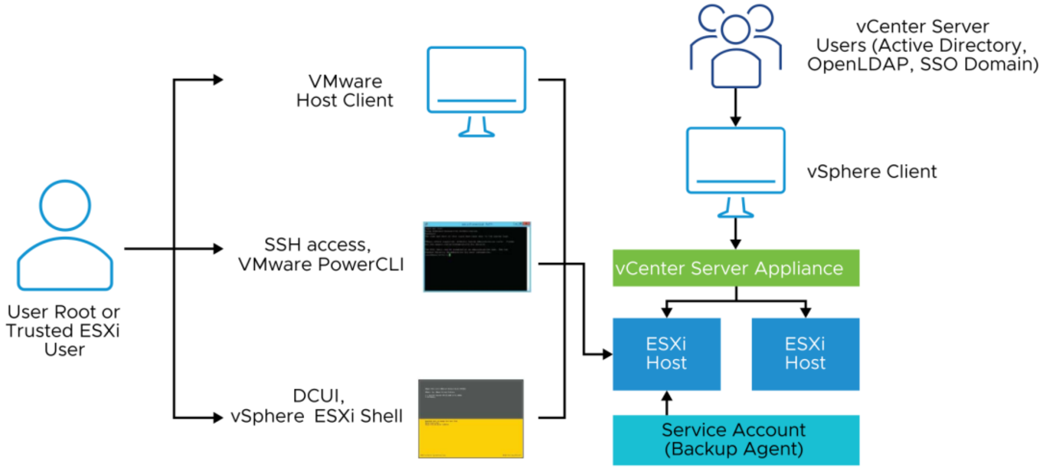
Restrict who gets full vCenter Server administrative rights. Do not grant this privilege to any group whose membership is not strictly controlled.

Do not add Windows special identity groups (such as Everyone) to vCenter Server roles. Instead, create specific groups for specific vSphere management and assign the appropriate user permissions.

In most cases, enable propagation when you assign permissions to an object. Propagation ensures that when new objects are inserted in to the inventory hierarchy, they inherit permissions and are accessible to users.

9-11 Review of ESXi Host Access

Most users access ESXi hosts by using the vSphere Client. Trusted users, including the user root, and service accounts can access ESXi hosts through direct methods.



User Root or Trusted ESXi User
VMware Host Client
SSH access, VMware PowerCLI
DCUI, vSphere ESXi Shell
vCenter Server Users (Active Directory, OpenLDAP, SSO Domain)
vSphere Client
vCenter Server Appliance
ESXi Host
ESXi Host
Service Account (Backup Agent)

Most of the ESXi host management tasks can be performed with the vSphere Client. However, a vSphere administrator might have to access an ESXi host directly to perform maintenance or troubleshooting tasks.

Using the following tools, you can access an ESXi host directly:

- VMware Host Client
- SSH session (for example, by using PuTTY)
- vSphere CLI
- PowerCLI
- Direct Console User Interface (DCUI)
- vSphere ESXi Shell (To access, press Alt+F1 from the DCUI login screen.)

A service account that performs a specific task, such as a host backup, can have direct access to an ESXi host.

9-12 Design Decision: Restricting ESXi Host User Access

How should ESXi host access be controlled?

To protect the host against unauthorized intrusion and misuse, consider the following guidelines:

- Use the vSphere Client to connect to vCenter Server to centrally administer all ESXi hosts.

  Only revert to VMware Host Client when central administration unavailable.

- Keep the SSH and vSphere ESXi Shell services disabled.
- If you enable vSphere ESXi Shell, enforce strict policies:
  - Enable vSphere ESXi Shell temporarily and restrict use to troubleshooting tasks only.
  - Give vSphere ESXi Shell access to trusted users only.
  - Set timeout values appropriately:
    - vSphere ESXi Shell availability timeout (in seconds)
    - vSphere ESXi Shell idle timeout (in seconds)
    - DCUI timeout (in minutes): Ensure that this value is not changed from 10 to 0.

For hosts that are managed by vCenter Server, avoid accessing managed hosts directly with the VMware Host Client, and avoid making changes to managed hosts from the host's DCUI. If you manage hosts with a scripting interface or API, do not target the host directly. Instead, target the vCenter Server Appliance that manages the host and specify the host name.

For vSphere ESXi Shell, you can set the availability timeout and the idle timeout from the vSphere Client, VMware Host Client, and DCUI. By default, these timeouts are disabled (value of 0). The availability timeout forces a user to log in within a specific time period. After the availability timeout period passes, the service is disabled, and users are not allowed to log in. The idle timeout is the amount of time that can elapse before the user is logged out of an idle, interactive session. The idle timeout does not affect existing sessions. Changes to the idle timeout apply the next time a user logs in to vSphere ESXi Shell.

Ensure that the DCUI timeout value is not changed from 10 to 0. The DCUI is used for directly logging in to the ESXi host and carrying out host management tasks. You must close the idle connections to DCUI to avoid any unintended use of the DCUI originating from a leftover login session.

9-13 Design Decision: Enabling Lockdown Mode

Should ESXi lockdown mode be enabled?

Consider lockdown mode for increased ESXi host security.

Mode	UI Access to ESXi Host? *	DCUI Access to ESXi Host?	Direct Access to ESXi Host? **
Normal	Yes	Yes, but only to user accounts in the DCUI.Access list, or the Exception Users list	Yes, but only to user accounts in the Exception Users list
Strict	Yes	No	Yes, but only to user accounts in the Exception Users list, and only if the vSphere ESXi Shell and SSH services are enabled

* UI access refers to access by the vSphere Client.

** Direct access refers to access by VMware Host Client, vSphere ESXi Shell, or SSH.

Enabling lockdown mode disables direct access to an ESXi host requiring the host to be managed remotely from vCenter Server. The roles and access controls implemented in vCenter Server are always enforced and users cannot bypass them by logging in to a host directly. By forcing all interaction to occur through vCenter Server, the risk of someone inadvertently accessing elevated privileges or performing tasks that are not properly audited is greatly reduced.

When lockdown mode is set to Normal, the DCUI service is not stopped. If the connection to the vCenter Server system is lost and access through the vSphere Client is not available, only users on the DCUI.Access list or the Exception Users list can log in to the ESXi host’s DCUI and exit lockdown mode.

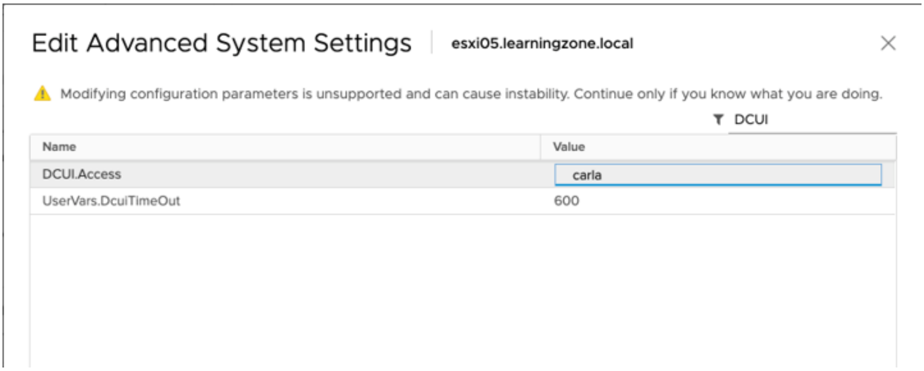
When lockdown mode is set to Strict, the DCUI service is stopped. If the connection to vCenter Server is lost, and the vSphere Client is not available, the ESXi host becomes unavailable unless vSphere ESXi Shell and SSH services are enabled and users on the Exception Users list are defined. If you cannot restore the connection to the vCenter Server system, you have to reinstall the host.

The vSphere ESXi Shell and SSH services are independent of lockdown mode. For lockdown mode to be an effective security measure, ensure that the vSphere ESXi Shell and SSH services are also disabled. Those services are disabled by default.

9-14 Lockdown Mode Considerations

For the DCUI.Access list, consider the following guidelines:

- Add at least one trusted, ESXi host user with nonadministrative privileges to the list.
- Remove root from the list.



Edit Advanced System Settings | esxi05.learningzone.local

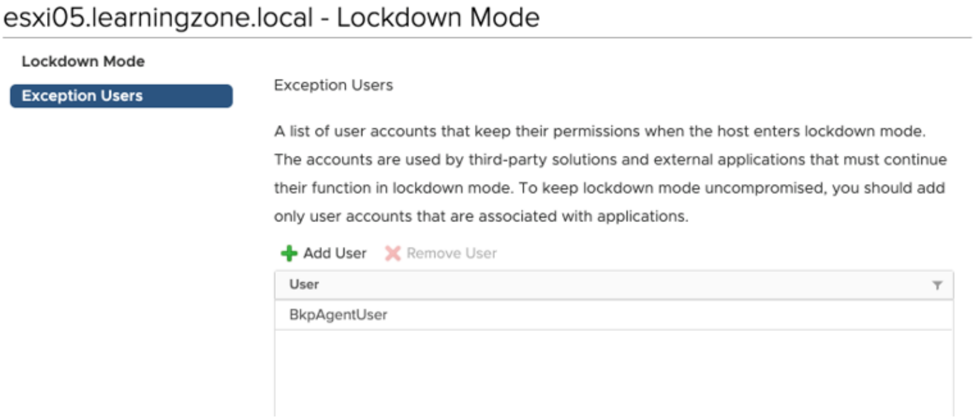
⚠ Modifying configuration parameters is unsupported and can cause instability. Continue only if you know what you are doing.

Name	Value
DCUI.Access	carla
UserVars.DcuiTimeOut	600

For the Exception Users list, consider the following guidelines:

- Do not add users with administrative privileges to the Exception Users list.

  Adding administrator users to this list defeats the purpose of lockdown mode.
- Periodically audit the Exception Users list for unauthorized users.



esxi05.learningzone.local - Lockdown Mode

Exception Users

A list of user accounts that keep their permissions when the host enters lockdown mode. The accounts are used by third-party solutions and external applications that must continue their function in lockdown mode. To keep lockdown mode uncompromised, you should add only user accounts that are associated with applications.

User
BkpAgentUser

The DCUI.Access list should contain only trusted users. These users do not require administrative privileges on the host. These users are on the DCUI.Access list to override lockdown mode in an emergency.

The Exception Users list is for service accounts that perform specific tasks, such as host backups, and is not for administrators. Periodically audit this list. Verify that the list of users who are exempted from losing permissions is legitimate and as needed for your environment.

9-15 Additional ESXi Host Security Considerations

Consider these additional ESXi security guidelines:

- Configure ESXi to use a directory service such as Active Directory to manage users.
- Do not use any third-party agents on the ESXi host.
- Use only VMware sources to upgrade or patch ESXi hosts.
- Maintain consistent, permanent logs in either of these ways:
  - Configure persistent logging for all ESXi hosts.

    Temporary (nonpersistent) storage of log files complicates auditing.
  - Perform remote logging to a central log host.
- Configure the ESXi host firewall to restrict access to services running on the host:
  - Close the ports that are not required for management access to the host.
- Starting with vSphere 6.5, ESXi supports secure boot if it is enabled in the hardware.

Creating local user accounts on each host presents challenges because you must synchronize account names and passwords across multiple hosts. Consider joining ESXi hosts to an Active Directory domain to eliminate the need to create and maintain local user accounts. Using Active Directory for user authentication simplifies the ESXi host configuration and reduces the risk for configuration issues that can lead to unauthorized access.

ESXi hosts run only services that are essential for functioning. A limited subset of vendors has hardware agents that can run. Using unsupported third-party agents or products creates a serious security risk and should be avoided.

VMware does not support upgrading ESXi hosts from any source other than a VMware source. If a download or patch is used from another source, management interface security or functions might be compromised.

From vSphere 7, the contents of what was the scratch partition are now in a dedicated partition called ESX-OSData created on high-endurance storage media such as HDD or SDD. Where high endurance storage media is not available, the host’s log files are stored on an in-memory file system. As a result, only a single day’s worth of logs are stored at any time. In addition, log files are reinitialized on each reboot. Storing log files on an in-memory file system presents a security risk because user activity logged on the host is stored only temporarily and does not persist across reboots. Monitoring events and diagnosing issues might also be difficult. ESXi host logging should always be configured to a persistent datastore.

Remote logging to a central log host provides a secure, centralized store for ESXi logs. By gathering host log files on a central host, you can more easily monitor all hosts with a single tool. You can also do aggregate analysis and searching to look for coordinated attacks on multiple hosts. Logging to a secure, centralized log server helps prevent log tampering and also provides a long-term audit record.

Unrestricted access to services running on an ESXi host can expose a host to outside attacks and unauthorized access. Reduce the risk by configuring the ESXi firewall to only allow access from authorized networks.

By default, all ports not required for management access to the host are closed. Ports must be opened if additional services are required.

Secure boot is part of the Unified Extensible Firmware Interface (UEFI) firmware standard. With secure boot enabled, a machine refuses to load any UEFI drivers unless the operating system bootloader is cryptographically signed. When ESXi boots, the UEFI firmware validates the bootloader at the VMkernel. Then, the secure boot VIB verifier verifies every VIB package that is installed on the system. If the security verifications pass during the boot sequence, then the entire system is booted up.

9-16 Design Decision: Using VMware CA Certificates

Should VMware certificates be used?

If you do not currently replace VMware certificates, VMware CA can handle all certificate management for you.

If your company policy requires using certificates that are signed by a third-party or enterprise certificate authority, you have the following options:

- Replace the VMware CA root certificate with a CA-signed certificate.

  The VMware CA certificate becomes an intermediate certificate for this third-party CA.

- If your company policy does not allow intermediate certificates, explicitly replace certificates.

VMware CA can be used as is, or as an intermediary certificate authority. You can replace the VMware CA root certificate with a certificate that is signed by an enterprise CA or third-party CA. VMware CA signs the custom root certificate each time it provisions certificates, making VMware CA an intermediate CA.

If you do not want to use VMware CA, you can replace the existing certificates signed by VMware CA with custom certificates. If you use this approach, you are responsible for all certificate provisioning and monitoring. You can perform different types of certificate replacement depending on the company policy and requirements for the system that you are configuring. You can perform each replacement with the vSphere Client or manually by using the CLIs included with your installation.

You can also use a hybrid model. For example, because solution user certificates are used only to authenticate to vCenter Single Sign-On, VMware CA can provision those certificates. Replace the machine SSL certificates with custom certificates to secure all SSL traffic.

For information about replacing certificates, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-52188148-C579-4F6A-8335-CFBCE0DD2167.html.

9-17 Design Decision: Storage Security Considerations

How should storage be secured?

In general, follow best practices for storage security, as outlined by your storage security provider.

Consider these guidelines for providing iSCSI storage security:

- Ensure that the iSCSI name is unique and traffic is only allowed for expected initiators.
- Enable bidirectional CHAP authentication (also known as mutual CHAP):
  - Ensure the uniqueness of CHAP secrets.
- Use a VLAN, or dedicated storage-only switches, to isolate iSCSI traffic.

Consider these guidelines for providing NFS storage security:

- To encrypt traffic, enable Kerberos with v4.1 NFS connectivity.
- Use a VLAN, or dedicated storage-only switches, to isolate NFS traffic.

ESXi supports unidirectional Challenge Handshake Authentication Protocol (CHAP) for all types of iSCSI initiators and bidirectional (mutual) CHAP for software and dependent hardware iSCSI.

Before configuring CHAP, check whether CHAP is enabled at the iSCSI storage system and check the CHAP authentication method that the system supports. Enable CHAP for your initiators and ensure that the CHAP authentication credentials match the credentials on the iSCSI storage.

Ensure uniqueness of CHAP secrets. The mutual authentication secret for each host should be different. If possible, the secret should be different for each client authenticating to the server. Uniqueness ensures that if a single host is compromised, an attacker cannot create another arbitrary host and authenticate to the storage device. With a single shared secret, compromise of one host can allow an attacker to authenticate to the storage device.

9-18 Design Decision: Network Security Considerations

Which networks should be isolated for security reasons?

Network security in the vSphere environment shares many characteristics with securing a physical network environment.

Isolation of network traffic is essential to a secure vSphere environment. Isolation prevents snooping.

Isolate networks for specific applications or functions:

- iSCSI
- NFS
- Virtual SAN
- Management

Network segmentation can be implemented by using separate physical network adapters or by setting up VLANs.

Maintaining separate physical network adapters for isolating traffic is probably the most secure method and is less prone to misconfiguration after the initial segment creation. However, VLANs provide almost all the security benefits inherent in implementing physically separate networks without the hardware overhead. As a result, VLANs offer a viable solution that can save you the cost of deploying and maintaining additional devices, cabling, and so forth.

Ensure that IP-based storage traffic is isolated. IP-based storage includes iSCSI and NFS. IP-based storage is not frequently encrypted, so anyone with access to this network can view it. VMs might share virtual switches and VLANs with the IP-based storage configurations. This type of configuration might expose IP-based storage traffic to unauthorized VM users.

Configure the IP-based storage adapters on separate VLANs or network segments from the production traffic and management network to limit unauthorized users from viewing the traffic.

9-19 Design Decision: VM Network Security Considerations

How should virtual machine networking be secured?

VM network security is built into the infrastructure:

- A VM is isolated from other VMs if it does not share the same virtual switch.
- A VM is isolated from physical networks if no physical network adapter is configured for a virtual switch.

VM network security can be enhanced in several ways:

- Keeping different VM trust zones within a host on different network segments:
  - Lowers the chances of packet transmissions between VM zones
  - Prevents sniffing attacks that require sending network traffic to the victim
- Adding firewall protection to the virtual network

The network can be one of the most vulnerable parts of any system. The virtual machine network requires as much protection as its physical counterpart. If the same safeguards, such as firewalls or antivirus, are used to protect a VM, the VM is as secure as a physical machine.

A trust zone is loosely defined as a network segment within which data flows relatively freely. Data flowing in and out of the trust zone is subject to stronger restrictions. If VM trust zones on their own network segments are isolated, the risks of data leakage from one VM zone to the next are minimized. Segmentation prevents various threats, including Address Resolution Protocol (ARP) spoofing.

With ARP spoofing, an attacker manipulates the ARP table to remap MAC and IP addresses. This remapping allows access to network traffic to and from a host. Attackers use ARP spoofing to generate man-in-the-middle (MITM) attacks and denial of service (DoS) attacks, hijack the target system, and disrupt the virtual network.

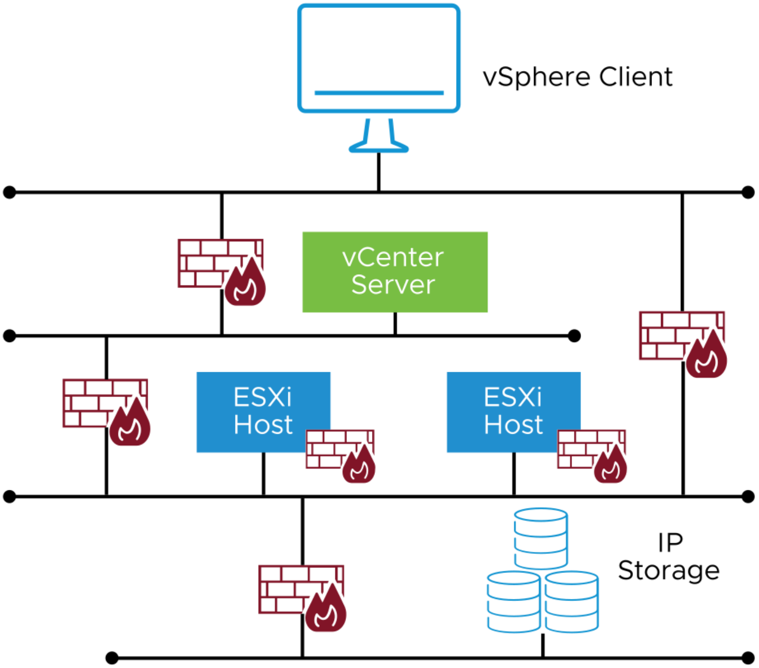
9-20 Protection with Firewalls

Although firewalls might add complexity to the environment, firewalls are a necessary component in securing the environment.

These firewall rules are common to ESXi implementations:

- Between the clients and vCenter Server
- Between the clients and ESXi hosts
- Between vCenter Server and ESXi hosts
- Between the ESXi hosts in your network
- Between ESXi hosts and network storage

Consider using one or more of these firewall rules, based on your security policy.



Firewalls provide basic protection for the network. During normal operation, vCenter Server listens for data from its managed hosts and clients on designated ports. vCenter Server also assumes that its managed hosts listen for data from vCenter Server on designated ports. If a firewall is present between any of these elements, you must verify that the firewall has open ports to support data transfer.

Firewalls might also be included at various access points in the network, depending on how the network is used and the level of security that various devices require. Select the locations for firewalls based on the security risks that have been identified for network configuration.

You might plan to clone virtual machines or perform vSphere vMotion migrations. The correct ports must be opened in any firewall that divides the source host from the target hosts so that the source and targets can communicate. The ports required for NFS and iSCSI storage are not specific to VMware and these ports can be configured according to the specifications of the network.

Additional firewalls increase the complexity of administration of the networking elements. You must secure the network with other means, such as VLANs, rather than an overcomplicated firewall configuration.

9-21 Securing Virtual Switch Ports

Both standard and distributed switches have security policies, which can be configured to restrict different types of traffic.

For increased network security, keep the default (Reject) for all network security.

Network Security Policy	Distributed Switch Defaults	Standard Switch Defaults	Use Cases for Reject	Use Cases for Accept
Promiscuous mode	Reject	Reject	To prevent a VM administrator from viewing traffic destined for other guest operating systems	To run network intrusion detection software or packet sniffers
MAC address changes	Reject	Reject	To protect against MAC address impersonation	To use Microsoft Network Load Balancing in unicast mode
Forged transmits	Reject	Reject	To protect against MAC address impersonation	To use Microsoft Network Load Balancing in unicast mode

Promiscuous mode eliminates any reception filtering that the virtual machine adapter performs so that the guest operating system receives all traffic observed on the wire. Although Promiscuous mode can be useful for tracking network activity, it is an insecure mode of operation. An adapter in promiscuous mode can access the packets even if some of the packets are received only by a particular network adapter. An administrator or root user within a virtual machine can potentially view traffic destined for other guest or host operating systems.

The MAC address changes option affects traffic that a virtual machine receives. When the MAC address changes option is set to Accept, ESXi accepts requests to change the effective MAC address to a different address than the initial MAC address. When the MAC address changes option is set to Reject, ESXi does not honor requests to change the effective MAC address to a different address than the initial MAC address. This setting protects the host against MAC impersonation.

The Forged transmits option affects traffic that a virtual machine transmits. When the Forged transmits option is set to Accept, ESXi does not compare source and effective MAC addresses. To protect against MAC impersonation, you can set the Forged transmits option to Reject. If you set this option to Reject, the host compares the source MAC address that is transmitted by the guest operating system with the effective MAC address for its virtual machine adapter. If the addresses do not match, the ESXi host drops the packet.

The guest operating system does not detect that its virtual machine adapter cannot send packets by using the impersonated MAC address. The ESXi host intercepts any packets with impersonated addresses before they are delivered and the guest operating system might assume that the packets are dropped.

9-22 Design Decision: VM Security Considerations (1)

How should virtual machines be secured?

Secure virtual machines as you might secure physical machines.

Consider the following VM guidelines:

- Keep all security measures up-to-date, including applying appropriate patches:
  - Use a patch management tool.
- Protect the guest operating system from viruses by installing antivirus software.
- Create VMs from VM templates that are secured and hardened.
- Minimize the use of the VM console:
  - Use vCenter Server roles to limit access to the VM console.
- Disable unnecessary devices and features in VMs:
  - Unnecessary hardware devices
  - Unused display features
  - Unexposed features

Ensure that antivirus software, antispyware, intrusion detection, and other protection are enabled for every VM in your virtual infrastructure. Also ensure that you have enough space for the VM logs and track updates for VMs that are powered off.

Stagger the schedule for virus scans in deployments with many VMs. The performance of systems in your environment degrades significantly if you scan all VMs simultaneously.

By using a template to capture a hardened base operating system image with no applications installed, you can ensure that all virtual machines are created with a known baseline level of security. Use templates that can contain a hardened, patched, and properly configured operating system to create other, application-specific templates.

The VM console provides the same function for a VM that a monitor on a physical server provides. Users with access to the VM console can access VM power management and removable device connectivity controls, which might allow a malicious attack on a VM.

For information about virtual machine security best practices, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-52188148-C579-4F6A-8335-CFBCE0DD2167.html.

9-23 Design Decision: VM Security Considerations (2)

Consider the following virtual machine guidelines:

- Prevent denial of service (DoS) attacks by performing the following VM configurations:
  - Limit informational messages from the VM to `VMX` file to avoid filling the datastore.
  - Disable the ability to shrink virtual disks:
    - If you shrink a virtual disk repeatedly, the disk can become unavailable.
  - Use resource reservations and limits to protect VMs from performance degradation that results if another VM consumes excessive resources.
- Consider UEFI secure boot if the following prerequisites are met:
  - The VM uses EFI firmware.
  - The virtual hardware is version 13 and later.
  - The guest operating system supports UEFI secure boot.

When one VM consumes most of the host resources so that other VMs on the host cannot perform their intended functions, a DoS might occur. To prevent a VM from causing a DoS, use host resource management features, such as setting shares, reservations, limits, and resource pools.

A DoS can occur when you do not control the size of a virtual machine's `VMX` file, and the amount of information exceeds the datastore's capacity. The `tools.setInfo.sizeLimit` advanced option has a default value of 1 MB. This capacity is often sufficient, but you can change this value if necessary. For example, you might increase the limit if large amounts of custom information are being stored in the configuration file.

Nonadministrative users in the guest operating system are able to shrink virtual disks. Shrinking a virtual disk reclaims the disk's unused space. To prevent reclaims, set the `isolation.tools.diskWiper.disable` and `isolation.tools.diskShrink.disable` advanced options to `true`.

Any service running in a virtual machine provides the potential for attack. By disabling unnecessary system components that are not necessary to support the application or service that is running on the system, you reduce the number of components that can be attacked. Virtual machines do not usually require as many services or functions as physical servers. When you virtualize a system, evaluate whether a particular service or function is necessary.

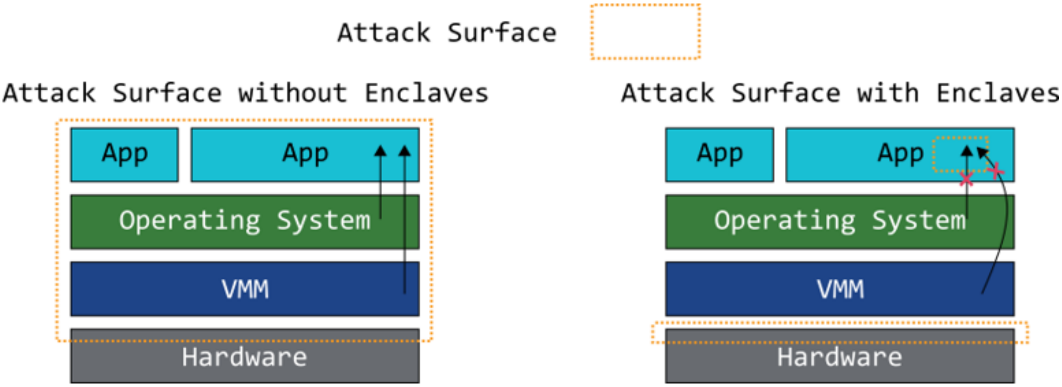
To properly support UEFI secure boot in a virtual machine, see the guest operating system vendor documentation for the proper way to configure and use EFI firmware.

For a list of operating systems that support UEFI secure boot, see VMware Compatibility Guide at http://www.vmware.com/resources/compatibility.

9-24 Intel Software Guard Extensions

With Intel Software Guard Extensions (SGX) technology, software programs can create private memory regions called enclaves.

- Data in enclaves can be accessed only by the intended program.
- The enclave region is isolated from other programs, operating systems, hypervisors, and so on.
- Enclaves are used by the software program to secure data.
- Virtual SGX (vSGX) exposes Intel's SGX to VMs running in vSphere.



Applications frequently work with private information such as passwords, encryption keys, bank account numbers, and so on. This private data should be accessed only by the designated recipient. The job of the operating system is to protect such private data from other applications and users.

Operating systems and applications often employ safeguards to protect this private data. Despite these protections, most computer systems are still vulnerable. Malware with administrative privileges can access all operating system resources including all applications running on that system. Malware can target an application's protected data to extract encryption keys and secret data directly from memory.

With Intel Software Guard Extensions (SGX), applications can create enclaves, which protect the confidentiality and integrity of applications from privileged malware. Privileged malware is malware that gets privileges by manipulating the underlying system.

9-25 Using the Virtual SGX Device

vSGX exposes Intel's SGX to VMs running in vSphere.

Consider the use of vSGX for the following situations:

- You have applications such as secure remote computing, secure web browsing, and digital rights management.
- You are deploying VMware virtual appliance components.
- You are running applications that must conceal proprietary algorithms and encryption keys.
- You do not want cloud service providers to inspect client's code and data.

Virtual SGX is implemented as part of the vSphere core virtualization stack. The vSGX implementation occurs between the VMkernel, VMM, and the management layer (vCenter Server, hostd, VMX processes).

VMkernel is responsible for initializing SGX support on the ESXi host. It also performs the initial hardware compatibility checks. Host level Enclave Page Cache (EPC) memory allocation and management are implemented by VMkernel. Enclave Page Cache (EPC) is a static portion in RAM that is used for storing running enclaves.

VMM handles the core virtualization of Intel's SGX instructions, EPC memory, and launch control configuration. vCenter Server, hostd, and VMX processes coordinate to perform the VM compatibility checks for power-on and initial DRS placement. They also implement feature restrictions and perform basic life cycle management of the vSGX features when a VM is powered on, reset, or powered off.

9-26 Using Virtual Trusted Platform Module Devices

vTPMs perform the same functions as a physical TPM, but it performs cryptographic coprocessor capabilities in software:

- When a vTPM is added to a VM, the guest operating system can create and store keys that are private.
- These keys are not exposed to the guest operating system.
- The VM attack surface is reduced.
- Usually, compromising the guest operating system compromises its secrets, but enabling a vTPM greatly reduces this risk.
- These keys can be used only by the guest operating system for encryption or signing.
- With an attached vTPM, a third party can remotely attest to (validate) the identity of the firmware and the guest operating system.

A hardware TPM can store information securely in what is called nonvolatile secure storage. On a physical TPM, this type of storage is a hardware-based component of the TPM device itself. It is part of a tamper resistant integrated circuit.

A virtual TPM does not have a hardware-based component. When the data to be secured is written to the nonvolatile secure storage by the guest OS, it is written using the in-guest TPM 2.0 APIs to the `.nvram` file, which is encrypted using VM Encryption. This feature provides for a software-controlled method for storage of data and is tamper-resistant because it is encrypted with VM Encryption. This software control does not live in the VM, it is part of the virtual device presented to the VM and is controlled by the ESXi hypervisor.

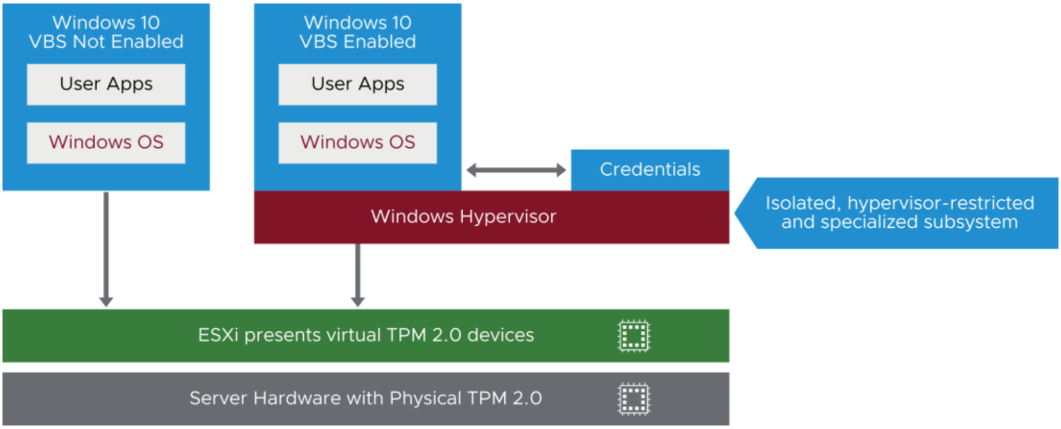
You can add a vTPM to either a new or an existing VM. A vTPM depends on VM encryption to secure vital TPM data. When you configure a vTPM, VM encryption automatically encrypts the VM files but not the disks. You can choose to add encryption for the VM and its disks.

The specific use case for a vTPM on vSphere is to support Windows 10 and 2016 security features. The HTML5 UI is designed for the security features.

9-27 Using Virtualization-Based Security for Windows VMs

Microsoft VBS uses hardware and software virtualization to enhance system security by permitting you to use the following Windows security features to harden your system:

- Credential Guard to isolate and harden key system and user secrets against compromise
- Device Guard features that work together to prevent and eliminate malware from running on a Windows system
- Configurable Code Integrity to ensure only trusted code runs from the boot loader onward


Windows 10
VBS Not Enabled
User Apps
Windows OS
Windows 10
VBS Enabled
User Apps
Windows OS
Credentials
Windows Hypervisor
Isolated, hypervisor-restricted and specialized subsystem
ESXi presents virtual TPM 2.0 devices
Server Hardware with Physical TPM 2.0

Microsoft VBS is a feature of Windows 10 and Windows Server 2016 and 2019 operating systems. It uses hardware and software virtualization to enhance system security by creating an isolated, hypervisor-restricted, and specialized subsystem.

After you enable VBS for a VM through vCenter Server, you enable VBS within the Windows guest operating system. Windows configures and enforces VBS through a Group Policy Object (GPO). With the GPO, you can turn off and on the various VBS services, such as Secure Boot, Device Guard, and Credential Guard. Certain Windows versions also require that you perform the additional step of enabling the Hyper-V platform. You do not require a vTPM to enable VBS.

Use the following Intel hardware for VBS:

- Haswell CPU or later. For best performance, use the Skylake-EP CPU or later.
- The Ivybridge CPU is acceptable.
- The Sandybridge CPU might cause some slow performance.

Create a VM that uses hardware version 14 or later and one of the following supported guest operating systems:

- Windows 10 (64 bit)
- Windows Server 2016 (64 bit)
- Windows Server 2019 (64 bit)

Not all VBS functionality is available on AMD CPUs. For more information VBS services, see VMware knowledge base article 54009 at https://kb.vmware.com/s/article/54009.

9-28 Design Decision: VM Encryption (1)

Should virtual machines be encrypted?

Use VM encryption if you want to secure confidential data on a VMDK so that the data is unreadable:

- For example, VM encryption reduces the risk of someone downloading a VM to a removable storage device and then leaving the company with the data.

VM encryption uses keys generated from a key management server (KMS):

- The KMS must be compatible with KMIP 1.1.
- vCenter Server must be able to communicate with the KMS.
- Protect the KMS by creating a key management cluster.

With VM encryption, you can secure sensitive data on a virtual machine's disk so that it is readable only with a digital key that was used to encrypt the disk. The key is not readable in any file, but secured in an extra layer of encryption. VM encryption protects not only the data on disk, but the swap file, as well as any guest-specific information that might be contained in the VMX or NVRAM files.

vSphere can be integrated with various third-party key management server (KMS) vendors and products. For vSphere, the KMS must be a server that communicates using the KMIP protocol. vCenter Server implements a KMIP client, which can issue KMIP-formatted commands to request keys using specific identifiers. A KMIP server can return the key values to vCenter Server using a secure channel.

Because the KMS plays a critical role in protecting VMs, create a key management cluster to provide some kind of failover protection.

For more information about the impact of using VM encryption on a VM's I/O performance, see *VMware vSphere Virtual Machine Encryption Performance* at https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/vm-encryption-vsphere65-perf.pdf.

9-29 Design Decision: VM Encryption (2)

By default, the vCenter Server Administrator role has cryptographic privileges:

- The No Cryptography Administrator role should be assigned to administrators who do not require encryption privileges.

If you use VM encryption, and if an error occurs on the ESXi host, the resulting core dump is encrypted to protect customer data.

For encrypted VMs, migration with vSphere vMotion always uses VMware vSphere Encrypted vMotion:

- Encrypted vSphere vMotion protects sensitive data that is transferred over a shared network.
- vSphere vMotion migration across vCenter Server systems is only supported when a shared KMS exists between the vCenter Servers

You can limit access to cryptographic operations by enforcing roles and permissions for users. Only users with the required permissions can perform cryptographic tasks, such as encrypting or decrypting a virtual machine.

A core dump is useful when debugging system failures and particularly host purple diagnostic screens. If the VMkernel fails, the failure usually results in a host purple diagnostic screen. This type of screen is the most common type of core dump analyzed by VMware technical support. It is important to protect core dumps, even from VMware technical support personnel who are trying to help you debug your system failure.

VMware vSphere Encrypted vMotion secures confidentiality, integrity, and authenticity of data that is transferred with vSphere vMotion. Encrypted vSphere vMotion migrations support all vSphere vMotion migrations for unencrypted VMs, including migration across vCenter Server systems.

9-30 VM Encryption Considerations (1)

Consider the following best practices:

- Do not encrypt any vCenter Server Appliance instances.
- Encryption is CPU-intensive:
  - Enable AES-NI in your BIOS because AES-NI significantly improves encryption performance.
- Consider storage trade-offs:
  - Deduplication and compression might not be effective for encrypted VMs.
  - Encrypting existing VMs is more time consuming than encrypting a VM during creation.
- Ensure that the KMS is available.
  - If the KMS is not available, VM operations that require that vCenter Server request a key from the KMS are not possible.
  - Most KMS solutions include high availability features.

The encryption process encrypts data on the host before it is written to storage. Back-end storage features, such as deduplication and compression, might not be effective for encrypted virtual machines. Therefore, consider storage trade-offs when using VM Encryption.

If the KMS is not available, virtual machine operations that require that vCenter Server request the key from the KMS are not possible. That means running virtual machines continue to run, and you can power on, power off, and reconfigure those virtual machines. However, you cannot relocate the virtual machine to a host that does not have the key information.

9-31 VM Encryption Considerations (2)

Establish a policy for encrypted core dumps:

- Core dumps are encrypted because they contain sensitive information, such as keys.
- Always use a password when you collect a vm-support bundle.

By default, encrypted vSphere vMotion can be used for unencrypted VMs if the source and destination ESXi hosts use version 6.5 and later.

Certain features do not work with VM Encryption:

- vSphere Fault Tolerance
- Content library
- Raw device mappings
- Some backup solutions that use vSphere Storage API - Data Protection
- Some linked cloning operations
- vSphere ESXi Dump Collector
- Multiwriter or shared disks (MSCS, WSFC, or Oracle RAC)

Establish a policy regarding core dumps. If you decrypt a core dump, consider it sensitive information. ESXi core dumps might contain keys for the ESXi host and for the VMs on it. Consider changing the host key and re-encrypting encrypted VMs after you decrypt a core dump.

You can specify a password when you generate the support bundle from the vSphere Client or using the `vm-support` command. The password re-encrypts core dumps that use internal keys to use keys that are based on the password. You can use the password in the future to decrypt any encrypted core dumps that might be included in the support bundle. Unencrypted core dumps or logs are not affected.

Site Recovery Manager supports the protection and recovery of encrypted virtual machines with storage policy protection groups.

Not all backup solutions that use vSphere Storage API - Data Protection for virtual disk backups are supported:

- vSphere Storage API - Data Protection SAN backup solutions are not supported.
- vSphere Storage API - Data Protection hot-add backup solutions are supported if the vendor supports encryption of the proxy VM that is created as part of the backup workflow.

  The vendor must have the privilege `Cryptographic Operations.Encrypt Virtual Machine`.
- vSphere Storage API - Data Protection NBD-SSL backup solutions are supported.

  The vendor application must have the privilege `Cryptographic Operations.Direct Access`.

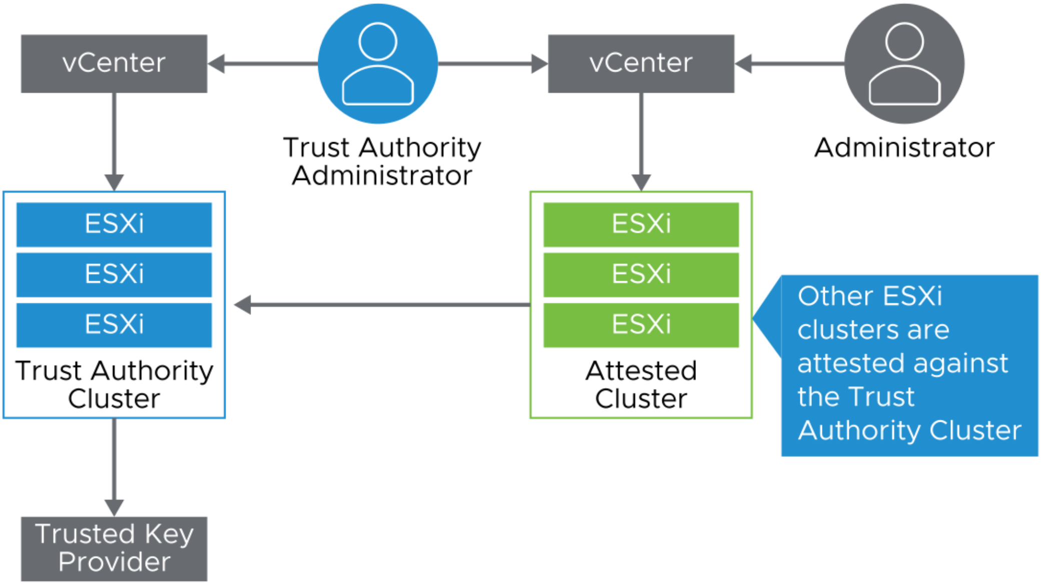
For more information on VM encryption, see *vSphere Security* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.security.doc/GUID-52188148-C579-4F6A-8335-CFBCE0DD2167.html.

9-32 Using vSphere Trust Authority to Secure Compute Infrastructure

vSphere Trust Authority comprises a vSphere Trust Authority Cluster and a trusted key provider.

Consider use of vSphere Trust Authority to perform security tasks if you have the following use cases:

- Providing ESXi hosts with a hardware root of trust and remote attestation capabilities
- Performing cryptographic operations on VMs but with an enhanced level of encryption key management
- Restricting encryption key management by releasing keys only to attested ESXi hosts



In vSphere 7, you can create a trusted computing base, which consists of a secure, manageable set of ESXi hosts. vSphere Trust Authority implements a remote attestation service for the ESXi hosts that you want to trust. Furthermore, vSphere Trust Authority improves on TPM 2.0 attestation support (added to vSphere beginning in the 6.7 release), to implement access restrictions on encryption keys and so better protect VM workload secrets. In addition, vSphere Trust Authority allows only authorized Trust Authority administrators to configure vSphere Trust Authority services, and configure Trust Authority hosts. The Trust

Authority administrator can be the same user as the vSphere administrator user, or a separate user.

Several concepts apply to vSphere Trust Authority:

- Trust Authority Cluster: A set of restricted ESXi hosts with a known good configuration.
- Administered by a smaller number of trusted administrators.
- Trusted key provider (also called key management server [KMS]): A key provider that only the Trust Authority Cluster should know.
- Attested cluster: Also known as a Trusted Cluster. The attestation report for a cluster must be validated by the Trust Authority Cluster.

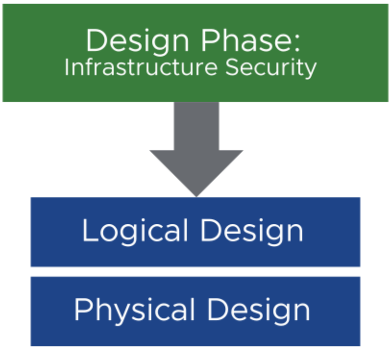
9-33 Lab 13: Designing Infrastructure Security

Create an infrastructure security design:

1. Review the Conceptual Design
2. Evaluate Security Design Options
3. Document vCenter Server Users, Groups, and Roles

9-34 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



9-35 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Make security design decisions for various layers in the vSphere environment
- Design a security strategy that meets the needs of the vSphere environment and follows VMware best practices

9-36 Key Points

- Consult the vSphere Security Configuration Guide and determine what risk profile to apply to your vSphere environment.
- Apply the principle of least privilege to all vCenter Server users.
- If you do not currently replace VMware certificates, VMware CA can handle all certificate management for you.
- Use a VLAN, or dedicated storage-only switches, to isolate different types of network traffic.
- Secure virtual machines as you secure physical machines.
- Use modern platform-based security capabilities including Intel SGX, vTPM, and Windows Virtualization Based Security to reduce attack surfaces
- Use VM encryption if you want to secure confidential data on a VMDK so that the data is unreadable without a digital key.

Questions?

Module 10

Infrastructure Manageability

10-2 Importance

The vSphere virtual infrastructure must run efficiently day after day to support an organization’s efforts to reach its business goals.

To maintain this type of efficiency, you must start with a strong design for managing and monitoring components. Achieving and maintaining infrastructure manageability requires skills in lifecycle management, infrastructure scalability, and capacity planning.

10-3 Module Lessons

1. Lifecycle Management
2. Scalability and Capacity Planning

10-4 **Lesson 1: Lifecycle Management**

10-5 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

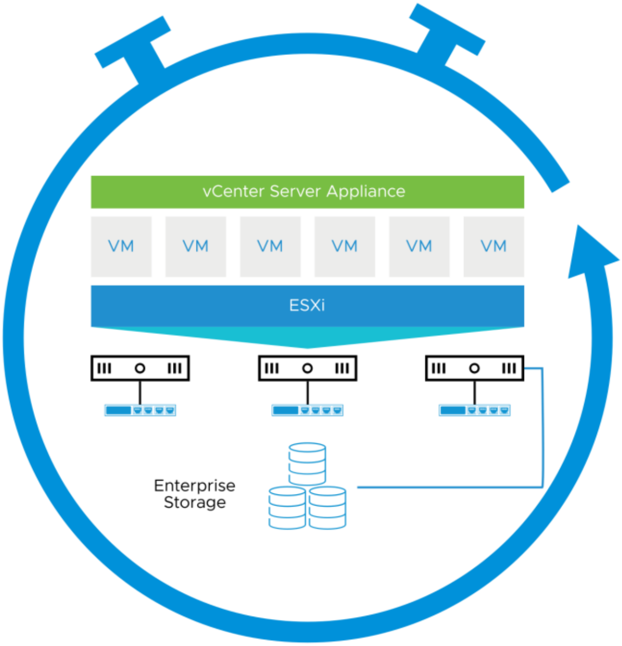
- Make lifecycle management design decisions that adhere to business requirements
- Design a lifecycle management strategy that meets the needs of the vSphere environment and follows VMware best practices

10-6 About Lifecycle Management

Lifecycle management involves managing the following operations tasks:

- Patching
- Updating firmware and drivers
- Managing VM templates
- Managing snapshots
- Managing vApps
- Managing VM hardware versions and VMware Tools
- Updating operating systems and applications
- Upgrading VM appliances
- Updating datastores
- Updating distributed virtual switches

You must consider the design requirements for each of these functions.

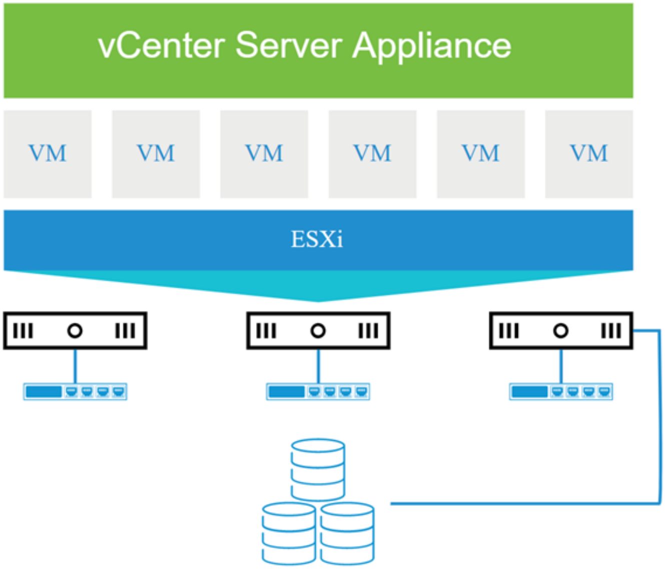


10-7 Patch Management

A vSphere design should specify how the following components are kept up to date with the latest patches and upgrades:

- ESXi hosts
- VMs and virtual appliances (vApps):
  - VMware Tools
  - Virtual hardware versions
- Guest operating systems and applications
- VMware virtual appliances:
  - vCenter Server Appliance
  - vRealize Operations Manager
  - vRealize Log Insight appliance
  - NSX Manager appliances

The design should also maintain version interoperability across VMware software components.



Patch management is not only a critical component of a management and security policy, it requires definition in your vSphere design. You must consider how the environment manages the life cycle of all the components, from hosts to virtual hardware and guest operating systems. Automation makes the task of maintaining patching policy significantly easier.

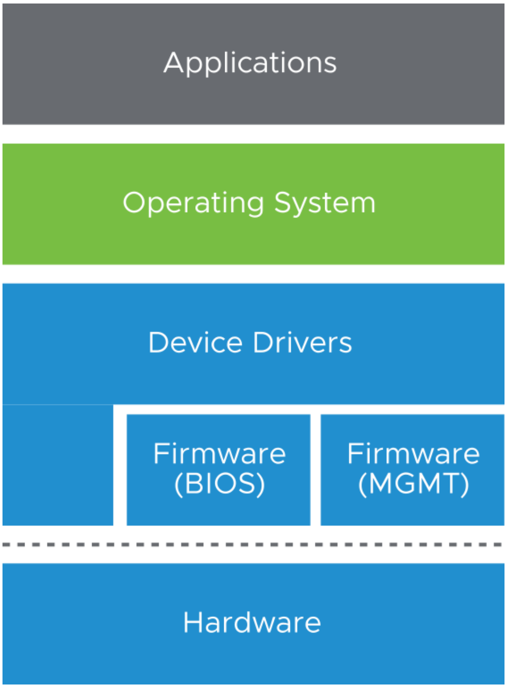
You must establish a process for ensuring continued interoperability. With architectural design frameworks such as VMware Validated Design, the bill of materials for the interoperability of VMware software components is defined in the VMware Validated Design version. In VMware Cloud Foundation, automation ensures that the patching sequence across VMware software components follows correct upgrade paths and maintains version interoperability support.

10-8 ESXi Host Firmware and Drivers

Management of ESXi host patching applies to more than the software layer.

Architects must consider the lifecycle management of hardware, firmware, and device drivers:

- Use identical server hardware when building clusters because homogeneous clusters use a single image for all hosts, making hardware management easier.
- Ensure models and firmware levels of server hardware and I/O adapters are consistent:
  - Use hardware vendors tools for centralizing and updating firmware.
  - Design joint-engineered solutions to maintain firmware and drivers in line with VMware software components.



Eliminating vulnerabilities and improving performance characteristics are not goals that are unique to software components. The firmware for the server hardware and the firmware of the I/O adapters are continually patched in their support lifetime by hardware vendors.

For example, interoperability support might dictate that a particular level of firmware is required for an I/O adapter. This requirement is especially relevant for storage I/O adapters used with vSAN.

The architect must consider the initial selection of server hardware and I/O adapters, with awareness that homogeneous server configurations provide the greatest ease of management by minimizing variance in operational support and patching. The architect should consider if one or more chosen hardware vendors can provide automation tooling to manage their firmware through its life cycle, or if the hardware vendor firmware can be maintained from within vSphere.

Jointly engineered solutions provide unique integration with VMware Cloud Foundation components that creates a user experience that uses and extends existing integrated system features and operations processes into VMware Cloud Foundation. These features and processes include, but are not limited to, lifecycle management of the hardware and software subsystems using native SDDC Manager orchestrated workflows. Dell EMC VxRail is an example of a jointly engineered solution that enables updates of Dell EMC device drivers from within VMware products.

10-9 Design Decision: VM Template Management

How should VM templates be managed?

With templates, VMs can be provisioned faster with fewer errors and reduced administration overhead.

Consider the following guidelines for managing templates:

- Keep the total number of templates to a minimum for easier maintenance of templates.
- Create templates that meet the needs of as many applications on a given operating system as possible.
- Create a template for each operating system and each key application.
- Create a versioning and release process for template updates.

Creating a template is a more straightforward process than maintaining the template. You must apply the latest patches and software upgrades, manage configuration drift, and so on. In addition, successive updates can introduce new issues into the template.

When creating templates based on guest operating systems, include the necessary software that is compatible with all the applications that must run on that operating system. Leave out anything that is not compatible with all the applications for that operating system. Instead, deploy the VM machine from the template and install these less-compatible items when you install the application into the VM. Include VMware Tools, service packs, most recent patches, and so on.

If you manage multiple versions of a template, consider creating a process for controlling the version and release of your templates. At the least, use a naming convention for your templates that includes the version number as part of the name. Your organization’s security policy can help determine how often to update templates.

For more information about managing templates, see "Manage Your Templates!" at https://blogs.vmware.com.

10-10 Design Decision: Using Content Libraries

Should content libraries be used in a single-site or multisite deployment?

You can use content libraries to manage content such as VM templates, vApp templates, and ISO images.

The content library can be stored on a local vCenter Server file system, on an NFS share, or on a VMFS datastore:

- Consider creating a content library using less expensive storage.

With content libraries, you can distribute the content to multiple sites, which results in the following benefits:

- You can manage identical VMs at different sites from a single, local template.
- You can add ISO images to the local library and synchronize them across all sites.
- You can keep all sites up to date with recent security patches or other software updates.

The content library helps manage templates and other file types across multiple vCenter Server instances. Organizations might have multiple vCenter Server instances in data centers around the globe. Templates can be stored at a central location and published, enabling other content libraries to subscribe and download content. The content library keeps content up to date by periodically synchronizing with the publisher, ensuring that the latest version is available.

10-11 Content Library Considerations

For the best performance, place the content library on a datastore that supports vSphere Storage APIs - Array Integration.

Operations, such as creating new VMs from templates, occur on the datastore, while freeing CPU and I/O resources on the host.

Decide whether a subscribed library should either download all content immediately or on demand as items are used.

Scenario for Data Transfer between Published and Subscribed Libraries	Performance
Datastore to datastore with Enhanced Linked Mode	1 (fastest)
NFS file system to NFS file system	2
NFS file system to datastore	3
Datastore to datastore without Enhanced Linked Mode	4 (slowest)

Datastores can be created on storage that supports vSphere Storage APIs – Array Integration. Many content library operations are performed on these datastores. For example, the creation of new VMs from content library templates takes place largely on the datastore. This setup improves performance, and CPU and I/O resources are freed on the hosts.

The contents of the content library can be shared between different vCenter Server systems. Sharing is achieved through the synchronization operation, which clones a published library by downloading all the content to a subscribed library.

A VMware performance study compares the performance of the following shared content library scenarios:

- Enhanced Linked Mode is used between vCenter Server systems, and content is transferred between datastores.

  In this scenario, the contents of a published library residing under one vCenter Server system can be synchronized to a subscribed library residing under another vCenter Server system. File transfer occurs by directly copying the files from the source datastore to the destination datastore if the ESXi hosts connected to those datastores have direct network connectivity. This method significantly shortens the path taken by the data, improving performance and reducing load on the vCenter Server systems.

- Enhanced Linked Mode is not used between vCenter Server systems and content is transferred between datastores.
- Content is transferred between NFS file systems.
- Content is transferred between an NFS file system and a datastore.

In scenarios 2 through 4, the contents of a published library must be streamed through the content library Transfer Service components, which reside on each vCenter Server system.

For information about this performance study, see "How to Efficiently Synchronize, Import, and Export Content in VMware vSphere Content Library" at https://blogs.vmware.com.

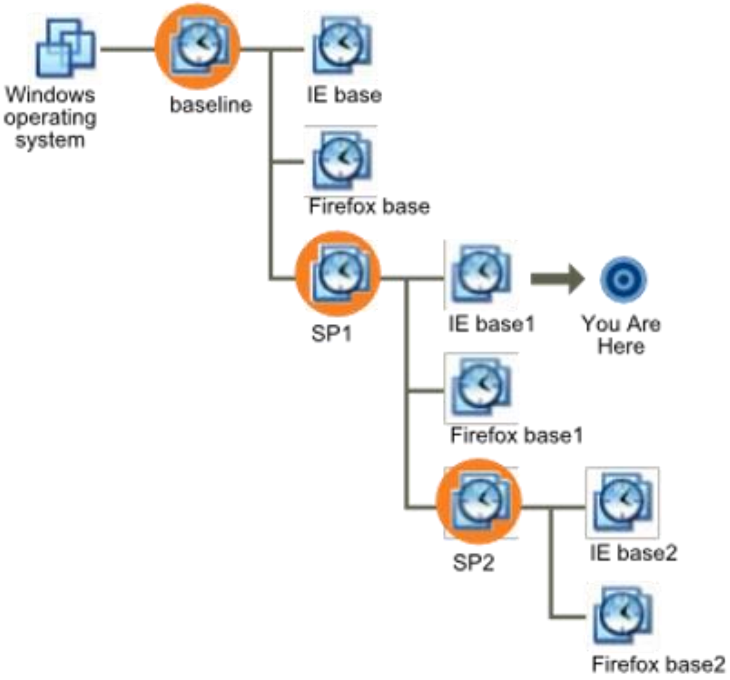
10-12 Design Decision: VM Snapshot Management

How should virtual machine snapshots be managed?

Virtual machine snapshots are used as a tool to revert changes, such as applying patches to an application.

Consider these guidelines when managing snapshots:

- Do no use snapshots as a backup solution.
- Create a management policy for snapshot creation.
- Create different management policies for different environments:
  - Production
  - Test
  - Development
- Add the snapshot creation policy to your change management procedures



Snapshots are a useful tool for backing out of changes made, such as when you apply patches to an operating system or application. With snapshots, you can test the changes before committing to them.

Create a snapshot management policy for the company’s different environments. For example, a production environment might require a different policy than a development and testing environment. Because the performance of a VM with multiple snapshots might be an issue, a production environment policy should allow only a single snapshot and for a short time.

However, a test and development environment policy might allow multiple snapshots for a longer time.

Snapshot creation can also be included in a change management policy. For example, creating a snapshot might require management approval and might not be the decision of a single administrator.

10-13 Design Decision: VM Management Using vApps

Should vApps be used?

Using vApps, you can perform resource management and other management activities for multiple VMs as a single unit:

- Power on VMs in a particular order.
- Clone-related VMs as a single unit.

vApps are a good solution for many use cases.

Use Case	Example
Multitiered applications	Web server, application server, and database server
Interrelated applications	Mail server, DNS server, and domain controller
Preconfigured environments	Consistent training environment for each student
Multitenant cloud or multitenant data center	Hosting provider

A vApp is a container for one or more VMs or vApps. You can use vSphere as a platform for running applications, such as multitiered applications. A vApp shares functionality with VMs. A vApp can power on and power off and it can be cloned. The distribution format for a vApp can be either Open Virtualization Format (OVF) or Open Virtualization Appliance (OVA).

10-14 Design Decision: Updating VM Hardware and VMware Tools

How should you maintain virtual hardware and VMware Tools?

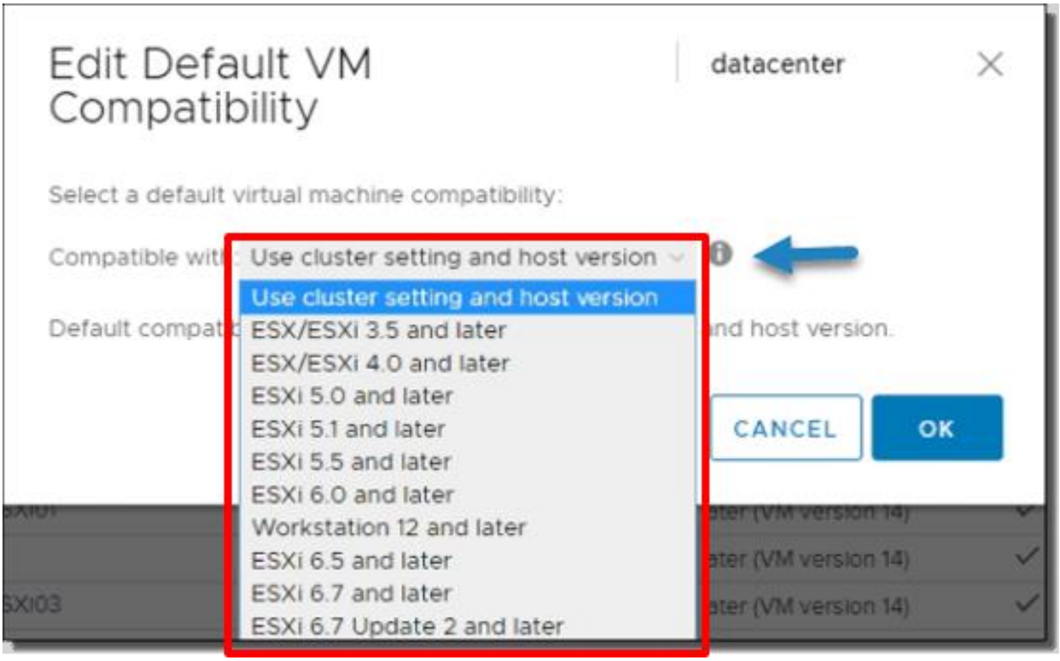
You can upgrade VMs to a higher level of compatibility and a higher version of VMware Tools.

After upgrade, VMs benefit from new hardware options and new features but cannot run on down-level hosts.

Consider the mobility requirements of the VMs:

- Migration to down-level hosts, for example, in a Cross vCenter vMotion
- Creation with latest or down-level compatible virtual hardware

The design must specify tooling to ease the management effort required to upgrade existing VMs.



Edit Default VM Compatibility | datacenter

Select a default virtual machine compatibility:

Compatible wit Use cluster setting and host version

Default compat nd host version.

When upgrading virtual hardware, consider these points:

- Modification of the virtual hardware version on the vCenter Server Appliance is not supported. Likewise, you should not manually install any additional version of VMware Tools inside the guest OS of the vCenter Server Appliance.
- Upgrading a VM to the latest hardware version is the physical equivalent of swapping the drive out of one system and placing it into a new one. Its success depends on the resiliency of the guest operating system in the face of hardware changes.
- Before you upgrade the virtual hardware version of a VM, create a snapshot or backup of the VM in case issues occur after the upgrade.
- To automate this process, consider using vSphere Lifecycle Manager to upgrade VMs.

  For more information about vSphere Lifecycle Manager, see *Managing Host and Cluster Lifecycle* at https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere-lifecycle-manager.doc/GUID-46CC2BE8-C2EC-4535-A8C8-5E6AD04A62AB.html.
- vSphere Lifecycle Manager takes automatic snapshots before performing VM upgrades.
- When you upgrade the virtual hardware, no downtime is required for vCenter Server or ESXi hosts. For VMs, the only significant downtime is the time to shut down and restart the guest operating systems.

From ESXi 5.5 onward, you can schedule hardware version upgrades for the next guest OS restart.

10-15 Design Decision: Updating Guest Operating Systems and Applications

How should guest operating systems and applications be patched and updated?

Use existing, established management tools and processes:

- Less disruption occurs to the organization’s processes.
- The tools might need slight modification to work in a virtual environment.

An alternative approach to patching is to employ a cloud design pattern of disposable VMs.

Refer to the operating system and application documentation for recommendations and best practices.

If the automated software tools and processes for updating operating systems and applications are working well and can scale with the new virtual infrastructure, continue to use these tools. The choice to continue to use these tools results in less disruption to the organization.

Employing a cloud design pattern of disposable VMs is an alternative approach to patching. If the VMs are operating as part of a load-balanced pool, for example, a fleet of web servers, then rather than patching individual VMs, a VM is removed from the pool and replaced. The new VM added to the load-balanced pool is built from an updated template. This process is a VM replacement strategy where the VM itself is not holding persistent data.

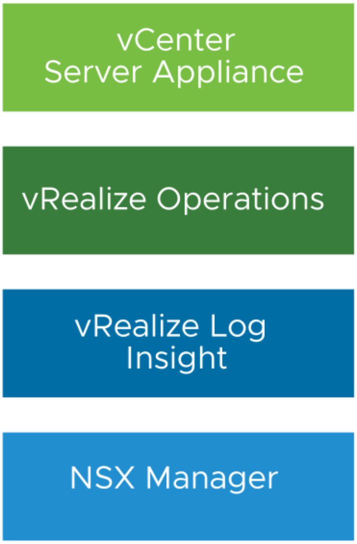
10-16 Design Decision: Upgrading VMware Virtual Appliances

How should VMware virtual appliances be patched and upgraded?

The method to patch and upgrade VMware virtual appliances varies:

- Avoid updating individual components or services within the appliance.
- Avoid changing any of the underlying operating system settings.
- Apply only VMware-released updates and patches to the appliance.

Always follow the documentation and release notes of the specific appliance for instructions on applying patches or performing an upgrade.



vCenter Server Appliance

vRealize Operations

vRealize Log Insight

NSX Manager

Patching of all VMware virtual appliances must be performed as a VMware life cycle operation where the whole appliance is updated using the update procedure relevant for that particular appliance. For example, the virtual appliance management interface (VAMI) interface of the vCenter Server Appliance provides an appliance update interface for the download and application of updates.

10-17 Design Decision: Updating VMFS and vSAN Datastores

How should VMFS and vSAN datastores be upgraded?

Upgrading VMFS datastores:

- ESXi 7.0 supports VMFS5 and VMFS6.
- Cannot upgrade VMFS5 to VMFS6, must migrate VMs from a VMFS5 to a new VMFS6 datastore.

Upgrading vSAN datastores:

- vSAN has several on-disk format versions for the different versions and upgrade history of the cluster.
- Because some vSAN features are tied to the on-disk format version, the format version must be accounted for when determining interoperability.
- For best results and access to all vSAN features, upgrade the objects to use the latest on-disk format:
  - After you upgrade the on-disk format, you cannot roll back software on the hosts.
  - Disk groups are upgraded one at a time.

You cannot upgrade a VMFS5 datastore to VMFS6. If you have a VMFS5 datastore in your environment, create a VMFS6 datastore and migrate virtual machines from the VMFS5 datastore to VMFS6.

vSAN has several different on-disk format versions available depending on the version and upgrade history of the cluster. Some on-disk format versions are transient although some are intended for long-term production.

Depending on the size of disk groups, the disk format upgrade can be time-consuming because the disk groups are upgraded one at a time. For each disk group upgrade, all data from each device is evacuated and the disk group is removed from the vSAN cluster. The disk group is then added back to vSAN with the new on-disk format.

For more information on vSAN on disk format interoperability, see VMware knowledge base article 2145267 at https://kb.vmware.com/s/article/2145267.

10-18 Design Decision: Updating Distributed Virtual Switches

How should distributed switches be upgraded?

You can upgrade vSphere Distributed Switch 6.x to a later version in the vSphere Client:

- Hosts do not require maintenance mode for this operation.
- Brief network downtime is experienced during the upgrade.
- Upgrade when all hosts participating in the switch are at the new host version

Option	Description
Distributed Switch 7.0.0	Compatible with ESXi 7.0 and later
Distributed Switch 6.6.0	Compatible with ESXi 6.7 and later Features released with the latest vSphere distributed versions are not supported.
Distributed Switch 6.5.0	Compatible with ESXi 6.5 and later Features released with latest vSphere distributed versions are not supported.

You can upgrade vSphere Distributed Switch 6.x to a later version. With the upgrade, the distributed switch can use features that are available only in the later version.

The upgrade of a distributed switch causes the hosts and the VMs attached to the switch to experience a brief downtime. For more information about upgrading to distributed virtual switch, see VMware knowledge base article 52621 at https://kb.vmware.com/s/article/52621.

To be able to restore the connectivity of the VMs and VMkernel adapters if the upgrade fails, back up the configuration of the distributed switch. If the upgrade is not successful, you can import the switch configuration file to recreate the switch with its portgroups and connected hosts.

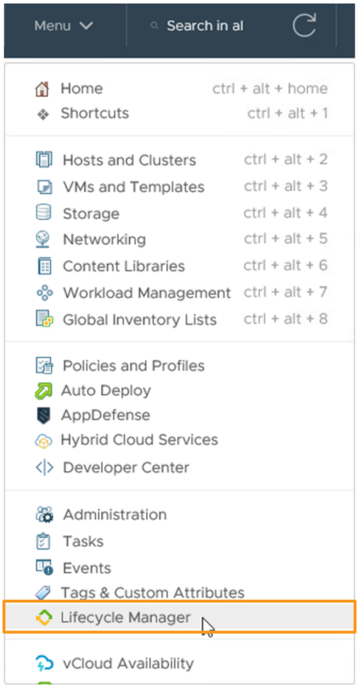
After you upgrade vSphere Distributed Switch, you cannot revert to an earlier version. You also cannot add ESXi hosts that are running an earlier version than the new version of the switch.

10-19 Introduction to vSphere Lifecycle Manager

vSphere Lifecycle Manager centralizes automated patch and version management for clusters, ESXi, drivers and firmware, VM hardware, and VMware Tools.

vSphere Lifecycle Manager features include:

- Upgrading and patching ESXi hosts
- Installing and updating third-party software on ESXi hosts
- Standardizing images across hosts in a cluster
- Installing and updating ESXi drivers and firmware
- Managing VMware Tools and VM hardware upgrades



In general terms, lifecycle management refers to the processes of installing software, maintaining it through updates and upgrades, and finally decommissioning it. In the context of maintaining a vSphere environment, lifecycle management refers to tasks such as installing ESXi and firmware on new hosts and updating or upgrading the ESXi version and firmware where required, vSphere Lifecycle Manager is a service that runs in a vCenter Server. On deploying vCenter Server Appliance, the vSphere Lifecycle Manager user interface becomes automatically enabled in the vSphere Client.

You can use vSphere Lifecycle Manager in a secured network without access to the Internet. In such cases, you use the vSphere Update Manager Download Service (UMDS) to download.

10-20 Design Decision: Configuring vSphere Lifecycle Manager for Host Updates

How should vSphere Lifecycle Manager be configured for host updates?

vSphere Lifecycle Manager supports two methods for updating and upgrading ESXi hosts:

- Only one method is supported per cluster at a time.
- After you switch a cluster to using images, you cannot revert that cluster to using baselines.

Managing Using Baselines	Managing Using Images
Compares ESXi hosts against an ESXi major version, group of patches, or set of extensions.	Compares ESXi hosts against a customized image that includes a base ESXi image, one or more add-on components, one or more vendor add-on components, firmware and drivers.
Supports all versions of ESXi from 6.5 and later.	Supports ESXi 7.0 and later.
Baselines attach to individual ESXi hosts.	Hosts in a cluster are managed collectively, with one ESXi host image per cluster.
ESXi upgrades through ISO images	ESXi upgrades through image depots (ZIP files).
ESXi updates or patches are bundled into baselines.	ESXi updates or patches are bundled and distributed as new ESXi versions.

ESXi images are commonly used to upgrade ESXi hosts from one version to another, for example, from ESXi 6.7 to ESXi 7.0. Before uploading ESXi images, obtain the image files from the VMware website or another source. ESXi images come from VMware or VMware partners. You can also create custom ESXi images that contain third-party VIBs by using vSphere Client.

Baselines and baseline groups are collections of patches that can be assigned to a vSphere cluster or ESXi hosts. The contents of a dynamic baseline are based on available patches that meet the specified criteria. As the set of available patches changes, dynamic baselines are also updated.

vSphere Lifecycle Manager provides the following baselines by default:

- Critical Host Patches
- Non-Critical Host patches
- Host Security Patches

10-21 Using vSphere Lifecycle Manager for VM Patch Management

Before upgrading the VM's virtual hardware, you first upgrade VMware Tools.

Use vSphere Lifecycle Manager to upgrade the VM hardware version and VMware Tools of your VMs.

For the VM settings, vSphere Lifecycle Manager is configured to take snapshots of VMs before applying updates.

Consider the following guidelines when managing snapshots:

- Snapshots kept indefinitely might consume a large amount of disk space and degrade VM performance.
- Having no snapshots save space, ensures the best VM performance, and might reduce the amount of time it takes to complete remediation, but it limits the availability of a rollback.
- Snapshots kept for a set period use less disk space and offers a backup for a short time.

vSphere Lifecycle Manager automates the process of upgrading and patching VMs, ensuring the steps occur in the correct order. This functionality is independent of whether you use baselines or images to manage ESXi hosts.

You can use vSphere Lifecycle Manager to directly upgrade:

- Virtual Machine Hardware
- VMware Tools
- Virtual Appliances

You can also patch and update third-party software running on the VMs and virtual appliances

VMware offers the following vSphere Client methods for upgrading VMs:

- Manual: Requires that you perform the VM upgrade one step at a time, but does not require vSphere Lifecycle Manager
- Automatic: Uses vSphere Lifecycle Manager to ensure that the upgrade and patching steps occur in the correct order, across many VMs

10-22 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Make lifecycle management design decisions that adhere to business requirements
- Design a lifecycle management strategy that meets the needs of the vSphere environment and follows VMware best practices

10-23 Lesson 2: Scalability and Capacity Planning

10-24 Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Make scalability and capacity planning design decisions that adhere to business requirements
- Design a scalability and capacity planning strategy that meets the needs of the vSphere environment and follows VMware best practices

10-25 Infrastructure Scalability

The design must consider how and when the environment scales.

Type of Scaling	Description
Scale Up (Vertical Scaling)	• Adds resources such as CPU or memory to existing hosts. • Host count remains the same, but hosts get bigger.
Scale Out (Horizontal Scaling)	• Hosts are typically uniform in size and more of the same are added. • Provides simple units of compute elasticity with linear increase in both storage and compute

Consider the impact of scale choices:

- Power and cooling differences.
- Fabric port count, with more hosts, more ports consumed
- Constraints related to capital and operating expenditures can influence design.
- vSphere and vSAN are licensed by CPU socket count.

You need instrumentation to determine when it is appropriate to scale.

When scaling the vSphere design for compute and memory, you can make servers bigger by adding RAM, CPU, network I/O bandwidth or processing power, and storage I/O bandwidth or processing power. You can also scale the vSphere design for compute and memory by simply adding more servers.

When scaling the design, you might need to address licensing considerations. vSphere editions are licensed by CPU sockets, so having more sockets equals greater license cost. Further, each socket license entitles the customer to up to 32 cores per socket, so vertical scaling can still result in more socket licenses being required.

When horizontally scaling a vSAN cluster, you must consider whether new hosts contribute local storage. Where scale out of compute is required, but not scale out of storage capacity, hosts that do not contribute to the vSAN datastore capacity can be added to the cluster but yet can still run VMs stored on the vSAN datastore.

vSAN can also be vertically scaled up by increasing the local storage on the hosts. For example, additional disk devices are added to create additional disk groups.

Instrumentation is required to provide metrics on utilization and ideally to predict resource exhaustion or impact to existing workloads. vRealize Operations provides this functionality.

10-26 About Capacity Management

At the VM level, the application team performs capacity management:

- Determines the size of the VM and buy the capacity from the infrastructure team.
- Adjusts the size because oversizing leads to performance issues.

At the physical infrastructure level, the infrastructure team performs capacity management:

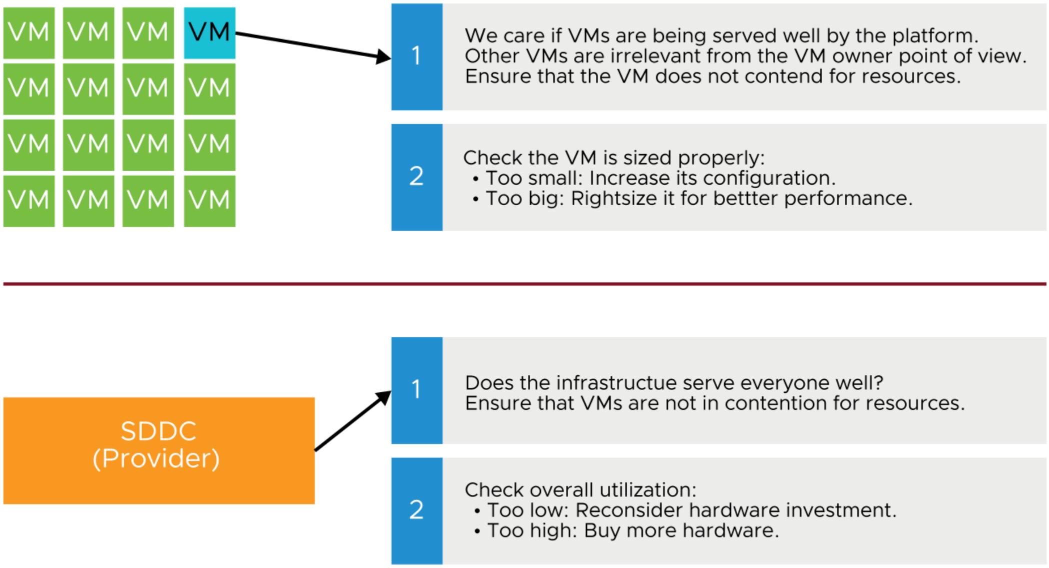
- Joint team manages computing, networking, storage, and disaster recovery.
- Must change mindset from a system builder to a service provider.

The changes to capacity management can be grouped into the following areas:

- Operations:
  - Virtualization changes the IT infrastructure team from system builder to service provider. The application team no longer owns the physical infrastructure because it is now a shared infrastructure. This change creates two-tier capacity management:
    - At the VM level, capacity management is done by the application team. This team determines the size of the VM and buys the capacity from the infrastructure team. The application team must adjust the size because oversizing leads to performance issues.
    - At the physical infrastructure level, the infrastructure team must perform capacity management as a single team. The joint team must take care of computing, networking, storage, and disaster recovery. The mindset must change from system builder to service provider.
- Architecture:
  - The infrastructure moves from a bespoke system to standardized hardware.
  - The application team no longer needs to dictate the specifications of the hardware:
    - For example, the team does not specify a server brand, model, and CPU frequency.
    - The team can only specify how many virtual CPUs are required.
    - Sometimes, especially in a large environment, the application team can only choose small, medium, or large vCPUs, which are all preconfigured.

10-27 Two-Tier Capacity Management

The VM and physical infrastructure levels have different important capacity management concerns.



1 We care if VMs are being served well by the platform.
Other VMs are irrelevant from the VM owner point of view.
Ensure that the VM does not contend for resources.

2 Check the VM is sized properly:

- Too small: Increase its configuration.
- Too big: Rightsize it for bettter performance.

1 Does the infrastructue serve everyone well?
Ensure that VMs are not in contention for resources.

2 Check overall utilization:

- Too low: Reconsider hardware investment.
- Too high: Buy more hardware.

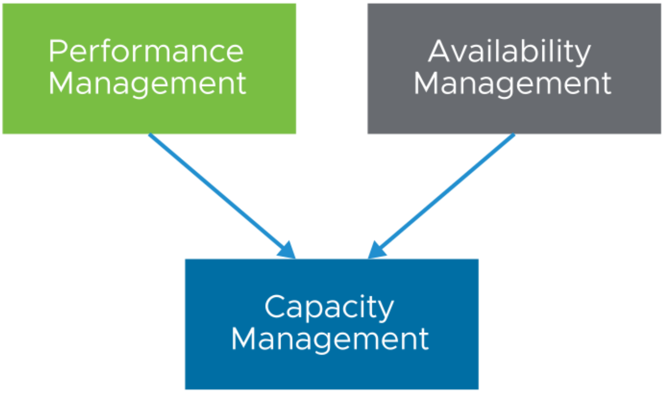
The diagram shows the two tiers and the important capacity management items for each tier.

10-28 Capacity Management Policy

Capacity management policy is interlinked to the performance management policy and availability management policy.

For both performance and availability policies, different service tiers and most enterprises use a three-tier model:

- Tier 1, the highest and most important tier, is for mission-critical VMs.
- Tier 2, the middle tier, is for production VMs.
- Tier 3, the lowest tier, is for test and development VMs.
- The availability of a mission-critical VM is much higher than a development VM.
- You cannot accept any form of resource contention for a mission-critical VM, but you can allow contention in the development environment.



Both performance and availability management drive your capacity management policy.

Avoid having more than three tiers, even in a large environment, for example, more than 100,000 VMs. Keep only three tiers. The more tiers, the more confusing it is for your customers and the application team. The positioning of each tier must be clear. Having too many tiers blurs this positioning.

10-29 Defining a Performance Service Level Agreement

A performance management policy can define service levels in a three-tier model based on resources.

Service Tier (Cluster)	CPU	RAM	Storage	Network
1 (highest)	0% CPU contention All hosts in this cluster have identical specifications.	0% RAM contention	10 ms latency Thick provisioned	No packet drop
2	5% CPU contention All hosts in this cluster have identical specifications.	5% RAM contention	20 ms latency Thin provisioned	No packet drop
3	10% CPU contention All hosts in this cluster might not have identical specifications.	10% RAM contention	30 ms latency Thin provisioned	No packet drop

For the three-tier model example shown, consider the following oversubscription comments for the CPU and RAM resources:

- An oversubscription ratio, such as 1.5x CPU oversubscription or 2x RAM oversubscription, is not defined. Oversubscription is an incomplete policy, and it fails to take into account utilization.
- VMware Professional Services observe that customers experience higher tiers not performing as well as the lower tiers. If you oversubscribe, you cannot guarantee consistent performance.

- Oversubscription leads to resource contention, even if your overall ESXi utilization is not high:
  - For example, more concurrent VM requests for CPU scheduling result in ready time for some VMs.
  - The host CPU level does not need to be near fully used for contention to occur at the individual resource level when concurrency of request for the same oversubscribed resource occurs.
  - Avoid oversubscription if you want to be able to predict performance.
- You can use contention to quantify the SLA. The potential of contention increases in the lower tiers.
- The Tier 1 cluster has no oversubscription because the design ensures enough CPU and RAM is available for every VM in the host. No VM needs to wait or contend for resources and the impact of the design is that reservation is not required.
- Tier 2 and Tier 3 permit oversubscription.

You can specify that all the hosts in the Tier 1 cluster are identical to ensure that the CPU generation and speed are identical. This setup makes performance predictable.

The Tier 3 cluster does not guarantee identical hosts. The cluster might start with four identical hosts, but, over time, it might expand to 16 hosts and not be identical in terms of performance because the new hosts have faster CPUs.

Enhanced vMotion Compatibility does not make a difference to the performance SLA when expanding a cluster to 16 hosts. Enhanced vMotion Compatibility assists only in masking CPU features to ensure VM mobility across different CPU hardware generations. Enhanced vMotion Compatibility does not change performance, for example, newer hosts are likely to have larger layer 1 cache size or greater GHz clock speed.

For storage resources, the performance SLA is set at 10 milliseconds. You can use a 5 minute average for a good balance. In Tier 1, the disk is thick provisioned, so no performance penalty occurs in the first write. The same service quality is not provided in the lower tier.

10-30 Defining an Availability Service Level Agreement

An availability management policy can defines service levels in a three-tier model based on required redundancy.

Service Tier	vSphere HA	Max Cluster Size	Max Number of VMs
1 (highest)	Two spare ESXi hosts	8 nodes	10 VMs per host 50 VMs per cluster
2	One spare ESXi host	12 nodes	20 VMs per host 200 VMs per cluster
3	One spare ESXi host	16 nodes	30 VMs per host 400 VMs per cluster

Mission-critical VMs must be better protected than development VMs. If a failure occurs, you want as small a blast radius as possible. The more hosts you have, with lower density of VMs, the less damage.

In the example, defining two spare ESXi hosts in a cluster that can have a maximum of 8 nodes ensures a high percentage of redundant capacity availability. Similarly, defining one spare ESXi host in a 12-node cluster provides less redundant capacity, and so on.

10-31 Performance Monitoring

Use vSphere best practices with the organization's service-level agreements (SLAs) to determine what to monitor. For example, you might monitor performance thresholds for CPU, memory, storage, and networks.

Select appropriate tools for monitoring performance:

- vCenter Server alarms
- vSphere Client performance charts
- Command-line management tools, such as `esxtop`
- vRealize Operations
- Third-party management software currently in use

Beyond normal vSphere monitoring, determine if certain services or applications require specific monitoring.

vCenter Server is configured with a set of predefined alarms that monitor data centers, clusters, hosts, VMs, datastores, and networks. Consider the vSphere features to be configured in the infrastructure and the information gathered from the key stakeholder and SME interviews to determine which alarms should be configured and at which thresholds. For example, a thin-provisioned environment requires that you configure alarms that are not necessary in a thick-provisioned environment.

vRealize Operations alerts notify you about many of the events that trigger vCenter Server alarms.

Automate performance monitoring whenever possible. Create alarms to notify you when performance thresholds are exceeded. You can also enable alerts if you use vRealize Operations.

Evaluate the organization's current monitoring tools and determine whether they can be used in the virtual infrastructure.

10-32 Design Decision: Management Tools

Which management tools should be used?

Choose appropriate tools for managing your vSphere environment:

- Day-to-day management:
  - vSphere Client
  - vRealize Operations
  - vRealize Log Insight
- PowerCLI for automating management tasks

Install these tools in VMs in a vSphere HA cluster, for example, in the management cluster, to ensure availability.

vSphere can also be automated using either SOAP web services API or REST API.

The vSphere Client has a full range of administrative functionality and an extensible plug-in based architecture. Typical users are virtual infrastructure administrators, help desk, network operations center operators, and virtual machine owners.

Although most of the management tasks are performed with the vSphere Client, VMware Host Client might be required on occasion.

The vRealize Operations console provides a central point of management for your vSphere environment. The operations console has all the capabilities needed for monitoring, troubleshooting issues, and administering the environment.

Automation is integrated into all aspects of the vSphere platform. PowerCLI is a command-line tool with which you can automate all aspects of vSphere management, including network, storage, VM, and guest operating system. PowerCLI is distributed as a Windows PowerShell snap-in.

PowerCLI includes several command-line utilities for automating, managing, and configuring vSphere hosts. PowerCLI can be installed on both Windows and Linux platforms.

10-33 Additional Management Solutions

Consider using other VMware products to enhance your management infrastructure.

Management Solution	Description	Available with
vRealize Orchestrator	Provides a library of workflows for creating and running automated processes to manage the vSphere infrastructure	Base functionality included with vCenter Server license and supports vSphere
vRealize Automation	Accelerates the deployment and management of applications and compute services	Part of vCloud Suite and vRealize Suite
vRealize Network Insight	Provides recommendations for implementing micro-segmentation and operational views to manage and scale an NSX deployment	Packaged separately, as an add-on to NSX

For information about vRealize Orchestrator, see the product page at http://www.vmware.com/products/vrealize-orchestrator.

For information about vRealize Automation, see the product page at http://www.vmware.com/products/vrealize-automation.

For information about vRealize Network Insight, see the product page at http://www.vmware.com/products/vrealize-network-insight.html.

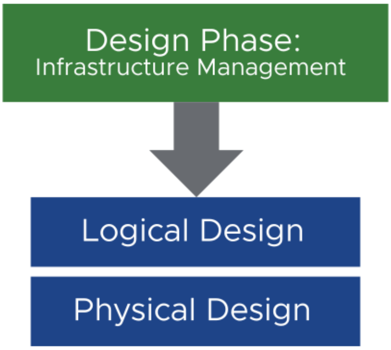
10-34 Lab 14: Designing Infrastructure Manageability

Create an infrastructure manageability design:

1. Review the Conceptual Design
2. Evaluate Manageability Design Options

10-35 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



10-36 Review of Learner Objectives

After completing this lesson, you should be able to meet the following objectives:

- Make scalability and capacity planning design decisions that adhere to business requirements
- Design a scalability and capacity planning strategy that meets the needs of the vSphere environment and follows VMware best practices

10-37 Key Points

- vSphere Lifecycle Manager is bundled with vCenter Server Appliance as an optional service.
- A management cluster provides resource isolation so that management services can operate at the best possible performance level.
- Using the content library, you can distribute content to multiple sites, and you can keep all sites up to date and synchronized across all sites.
- Consider using vRealize Operations to provide visibility and insights into the performance, capacity, and health of the vSphere environment.

Questions?

Module 11

Infrastructure Recoverability

11-2 Importance

A business continuity and disaster recovery (BCDR) solution is critical for any environment. By designing a BCDR solution, you can ensure that your environment is backed up and restored appropriately if a failure occurs.

11-3 Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Make infrastructure recoverability design decisions that adhere to business requirements
- Design an infrastructure recoverability strategy that meets the needs of the vSphere environment and follows VMware best practices

11-4 vSphere Infrastructure Recoverability Overview

Recoverability is the ability to restore your environment after a failure occurs. Typically, your recovery strategy includes a combination of processes.

Recovery Process	Description
Data protection	Perform regular backups Recover data from backups when data is corrupted or lost.
High availability	Respond (reactively) to unplanned outages at the host hardware, vSphere, and VM levels. Restart VMs as appropriate.
Disaster avoidance	Proactively avoid service outages before an impending disaster (such as a hurricane). Brief outages followed by an orderly restart at a recovery site.
Downtime avoidance	Proactively migrate workloads between systems and sites with no downtime and no loss of data.
Disaster recovery	Rapidly recover from unplanned outages that drop services so that local recovery within an acceptable time period is unlikely.

Data protection is the process of backing up your data regularly and being able to successfully recover entire systems (including operating system, applications, and data) within your recovery time objectives.

Highly available systems recover services after an unplanned outage. High availability technologies reduce the period of outage of services during failure, resulting in a rapid recovery of VMs. Traditional high availability clusters also provide automated fault recovery.

With disaster avoidance technologies, services can continue with minimum interruption. Usually, disaster avoidance involves brief outages to services at a site, followed by an orderly restart at a recovery site. A minimum outage sustained under controlled circumstances is typically considered acceptable as an alternative to sustaining an uncontrolled and extended outage associated with a true disaster.

Downtime avoidance differs from disaster avoidance in that downtime avoidance technologies migrate the workloads between systems or sites with no downtime and no loss of data. For example, vSphere vMotion and vSphere Storage vMotion move VMs or VM storage with no interruption of the services that the VMs provide.

In disaster recovery scenarios, the goal is to rapidly return services to operational status, usually in a different data center in a safe location. Disaster recovery solutions help automate a return to operations of services that have stopped because of catastrophic failure of infrastructure.

11-5 Designing Infrastructure Recoverability

Determine each application's SLA requirements for data protection and disaster recovery and apply the appropriate tools and technologies to meet those requirements.

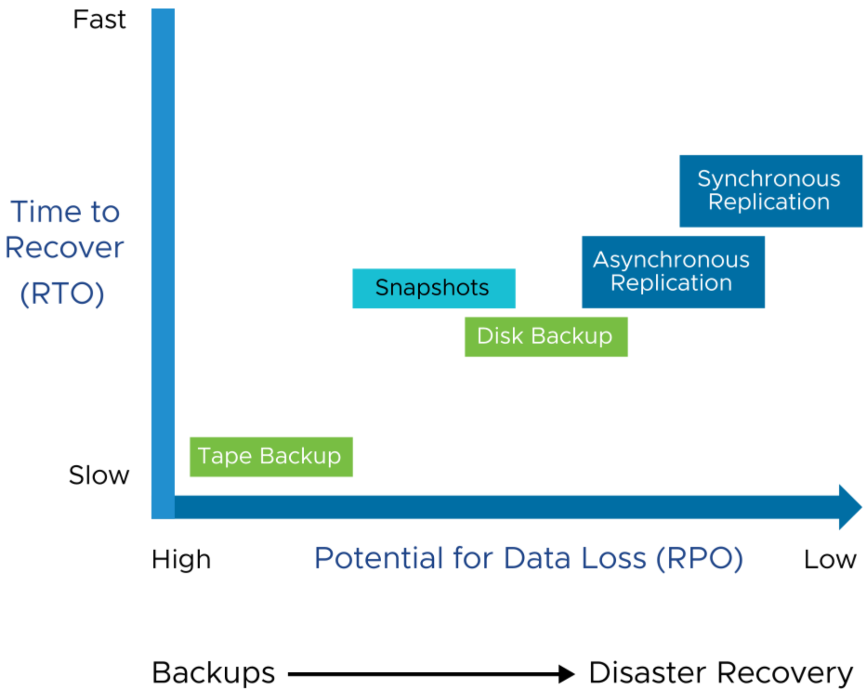
Recovery Process	Tools and Technologies
Data protection	vCenter native backup and recovery VM third-party data protection vSphere Storage APIs - Data Protection
High availability	vCenter Server High Availability vSphere HA vSphere FT
Disaster avoidance	Site Recovery Manager
Downtime avoidance	vSphere vMotion vSphere Storage vMotion
Disaster recovery	vSphere Replication Site Recovery Manager vSAN stretched cluster

VMware provides products on which an administrator can back up and restore data, provide high availability, and provide disaster recovery solutions. Use these recoverability options in the design, if they adhere to business policies and SLAs.

11-6 Using Recovery Objectives

Use recovery time objectives (RTOs) and recovery point objectives (RPOs) to determine the appropriate data protection tools and technologies:

- RTO: Targeted length of time that an application can be down after a failure
- RPO: Maximum tolerable period in which data might be lost because of a disruption



An infrastructure should be recoverable from any kind of outage. Design an efficient recovery plan that minimizes the recovery time for your environment. Backup and data protection policies affect your design. For example, aggressive recovery time objective (RTO) and recovery point objective (RPO) values might force you to add array-based replication or snapshots to your design to augment the tape backup.

RTO is the amount of time taken to restore the services of the architecture after a disaster or disruption. RPO is the point in time that data must be recovered to after a disaster or disruption. Both metrics can be defined in an SLA.

Replication can be synchronous or asynchronous. Synchronous replication is used in disaster recovery deployments that have RPO requirements of zero data loss. Synchronous replication writes data to storage in the primary site and recovery site simultaneously.

However, asynchronous replication writes data to storage in the primary site first and then commits the data to be replicated to memory. Typically, the data is written on a scheduled basis to the storage at the recovery site.

11-7 Using RTOs and RPOs

Depending on the RTOs and RPOs, recovery can be as simple as applying backups to implementing disaster recovery procedures.

Method	Typical RTO	Typical RPO	Retention	Use Cases
Backup	Hours	Daily, weekly, monthly, yearly	Months+	• Corruption/data loss • File recovery • Long-term retention • Compliance
VM snapshots	Seconds to minutes	Minutes to hours	Days	• Test/dev environment • Cloning • Patching/maintenance • Corruption/data loss
Asynchronous replication	Minutes to hours	Minutes to hours	Days	• Hardware failure • Site/disaster recovery • Planned migration • Corruption/data loss
Synchronous replication	Subminute/instant	Zero data loss	None	• Disaster recovery

The frequency of taking backups can impact the ability to meet RTOs and RPOs. With backups taken more frequently, the data is more current. As a result, the RTO and RPO are lower than for backups that are taken less frequently.

The RPO for asynchronous replication can range from a few minutes to hours depending on the customer requirements. This replication is usually constrained by the bandwidth between the primary and recovery sites.

You should not use VM snapshots as a backup solution. Snapshots are a useful tool to back out of changes, such as when you apply patches to an application or update an operating system. You can test the changes before committing the changes.

When determining the maximum tolerable downtime of the service or application, account for the time it takes to bring the VM back online and get the service or application running within the VM.

11-8 General Guidelines for Protecting Data with Backups

Consider the following general guidelines for protecting data with backups:

- Solutions that use vSphere Storage APIs - Data Protection back up the entire VM.
- Existing backup solutions require an agent to back up the guest operating system.
- Backup type and recovery time depend on what is backed up, for example, entire VM or files in the guest operating system.
- Balance cost with need because having a backup agent in every VM can add additional licensing fees.
- Create a backup schedule that minimizes the downtime of an application or service.
- If necessary, provide additional procedures to ensure the availability of business applications after recovering the VM from failure.

For a list of some of the backup solutions that support vSphere Storage APIs – Data Protection, see VMware knowledge base article 1021175 at http://kb.vmware.com/kb/1021175.

vSphere Data Protection was a VMware data protection solution for backing up and restoring VMs, virtual appliances, and vCenter Server instances that are now discontinued. vSphere 6.5 is the last release that includes vSphere Data Protection. Existing vSphere Data Protection installations continue to be supported until the End of General Support (EOGS) date.

VMware continues to support the vSphere Storage APIs – Data Protection framework, which is used by the vSphere partner backup ecosystem.

11-9 Design Decision: Protecting vCenter Server Appliance

How should you protect vCenter Server?

To protect vCenter Server Appliance, use the following methods:

- For data protection, use the appliance's native backup and restore tool:
  - Always includes the vCenter Server database and system configuration files in the backup operation
  - Includes an option for encryption
  - Requires an FTPS, HTTP, HTTPS, SFTP, FTP, NFS, or SMB server with sufficient disk space to store the backup
  - Can automate backups with vCenter Server Appliance Management API scripts
- For protection against host and hardware failures, use the vCenter Server High Availability feature for automatic failover:
  - RTO: Target limit is five minutes.
  - RPO: Zero.

All configuration information about the vSphere inventory objects, roles, alarms, and host profiles are kept in the vCenter Server database. The database also contains a history of tasks and events and a history of performance information. If the database is lost or corrupted, all this information must be restored from backup.

The vCenter Server Appliance backup operation always includes the vCenter Server database and system configuration files, so that a restore operation has all the data for recreating an operational appliance. Optionally, you can specify that a backup operation should include statistics, events, and tasks from the current state of the data center. Current alarms are always included in a backup.

The vCenter Server Appliance backup process collects key files into a tar bundle and compresses the bundle to reduce network load. To minimize storage impact, the transmission is streamed without caching in the appliance. To reduce the total time required to complete the backup operation, the backup process handles the different components in parallel.

You can encrypt the compressed file before transmission to the backup storage location. When you choose encryption, you must supply a password that can be used to decrypt the file during restoration.

For information about how to automate vCenter Server Appliance backups using scripts, see *VMware vCenter Server Appliance Management Programming Guide* at https://code.vmware.com/docs/6529/vmware-vcenter-server-appliance-management-programming-guide.

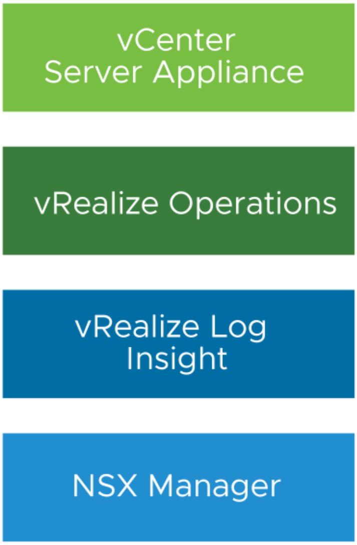
11-10 Design Decision: Backing Up VMware Virtual Appliances

How should you back up VMware virtual appliances?

The method to back up VMware appliances is specific to the appliance.

Always follow the documentation and release notes of the specific appliance for instructions on performing backups:

- Prerequisites and backup guidelines might be provided.
- Some appliances might have their own backup API:
  - For example, vCenter Server Appliance or NSX Manager.
- Often, the documentation recommends using the backup tool of your choice:
  - For example, you can use a backup tool integrated with vSphere Storage APIs – Data Protection.



vCenter Server Appliance

vRealize Operations

vRealize Log Insight

NSX Manager

11-11 Design Decision: Backing Up ESXi Hosts

How should you back up ESXi hosts?

Consider the following options:

- Reinstall the ESXi hosts instead of backing them up:
  - ESXi host software configuration is minimal, and reinstalling the hosts might be faster.
  - For many ESXi hosts, consider using vSphere Auto Deploy for rapid deployment.
- Back up the ESXi hosts by using the vSphere CLI. The command `vicfg-cfgbackup` backs up the host's configuration data.
- Use host profiles to create a baseline of the host configuration:
  - You can apply a standard configuration to multiple hosts.
  - You can check hosts for compliance.
- Use host profiles to recover an ESXi host configuration.

Periodically back up your ESXi host configuration, and, in particular, after you change the configuration or upgrade the ESXi image. However, the installation and configuration time for an ESXi host is minimal. So the time taken to restore an ESXi host from a backup is often longer than reinstalling the software.

You can use vSphere Auto Deploy to deploy large numbers of ESXi hosts quickly and easily. vSphere Auto Deploy simplifies ESXi host management because you do not need to maintain a separate boot image for each host. You can share ESXi software images with all hosts running on matching hardware. Centralized image management eliminates the need to patch and update individual ESXi hosts. You can perform a single update to the shared image profile and reboot the vSphere hosts.

To make reinstalling the ESXi host easier, you can use host profiles to create a baseline of the host configuration. Host profiles eliminate per-host, manual, or user interface based host configuration and maintain configuration consistency and correctness across the data center through policies.

For information about the `vicfg-cfgbackup` command, see vSphere Command-Line Interface Documentation at https://www.vmware.com/support/developer/vcli.

11-12 Design Decision: Backing Up Virtual Machines

How should you back up virtual machines?

Back up VMs in accordance with the business policies and SLAs

Generally, VMs are backed up by using either using a VM backup or by an in-guest agent.

Both backup methods can employ full or incremental types of backup.

vSphere Storage APIs – Data Protection backups have the following benefits:

- Nondisruptive: Do not require any downtime for VMs
- Do not need extended backup windows

Method	Backup Contents	Advantage	Disadvantage
VM backup	• VM config • VM disks	VMs easily restored to exact state of VM at backup	Requires significant amounts of space
In-guest backup	Guest OS Files	Uses less space than full VM backups	More difficult to back up and restore because the operating system must be running before the recovery

You can use third-party backup solutions to protect system, application, and user data in your VMs. When using vSphere Storage APIs - Data Protection, third-party software can perform backups without loading ESXi hosts with the processing of backup tasks.

Third-party products using vSphere Storage APIs - Data Protection can perform the following backup tasks:

- Run a full, differential, and incremental image backup and restore of VMs.
- Run a file-level backup of VMs that use supported Windows and Linux operating systems.
- Ensure data consistency by using Microsoft Volume Shadow Copy Services (VSS) for VMs that run supported Microsoft Windows operating systems.

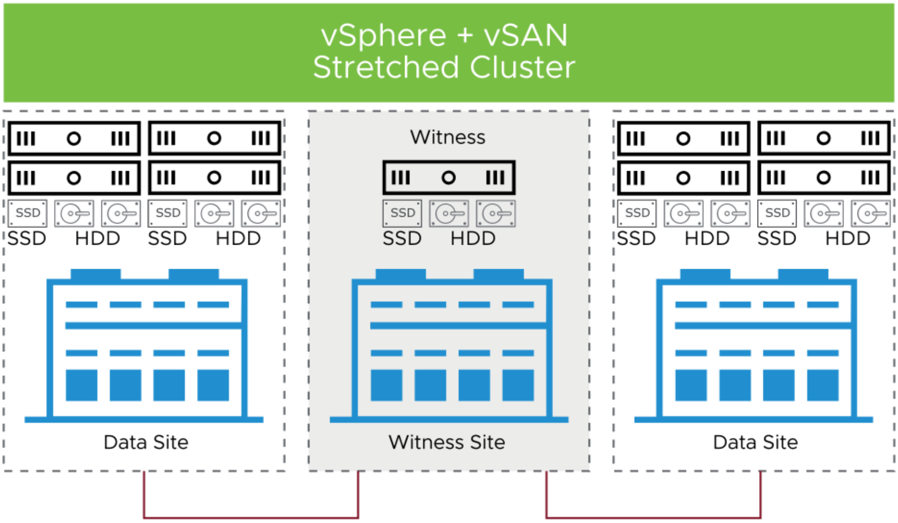
Because vSphere Storage APIs - Data Protection uses the snapshot capabilities of VMFS, backups that you can perform do not require downtime for VMs.

11-13 Design Decision: Disaster and Downtime Avoidance with vSAN (1)

Should you implement vSAN stretched clusters?

Use vSAN stretched clusters in environments where disaster and downtime avoidance is a key requirement:

- Protects VMs across data centers (sites), not just racks. Failover is automatic:
  - RTO: Depends on the number of VMs to restart
  - RPO: Zero
- Can also be used for the planned maintenance of one site without any service downtime.



A stretched cluster can be used for planned and unplanned downtime of data sites. For planned downtime, you can migrate the VMs from one data site to the other data site. You can then perform maintenance on the evacuated data site without causing any VM downtime.

If one data site might be affected by an impending disaster, such as a hurricane, the VMs running at this site can be migrated to another data site and can continue to run.

If a failure occurs at one data site, the stretched cluster can automatically restart VMs at the other data site.

11-14 Design Decision: Disaster and Downtime Avoidance with vSAN (2)

Consider the following stretched cluster guidelines:

- Minimum of 3 hosts and a maximum of 31 hosts
- Network configuration:
  - Low latency (<= 5 milliseconds) and high bandwidth between data sites.
  - All three sites must be connected to the management network and the vSAN network.
  - The data sites must be connected to the vSphere vMotion network.
  - VM network connectivity between the data sites.
- Only the data sites contribute to storage and CPU, the witness host does not run any VMs.

The minimum number of hosts in a stretched cluster is three: One host in each data site plus the witness host in the witness site.

The maximum number of hosts in a stretched cluster is 31: 15 hosts in each data site plus the witness host in the witness site.

For stretched clusters, geographical distances are not a support concern. The key requirement is the actual latency numbers between sites.

The vSAN witness host is an appliance that is available as an Open Virtual Appliance (OVA) from VMware. The witness host is uniquely designed with the sole purpose of providing cluster quorum services during failure events. The witness host does not contribute to compute and storage capacities. The size of the witness host depends on the size of the stretched cluster deployment. For guidelines on sizing the witness host, see Witness Host Sizing at https://storagehub.vmware.com.

11-15 Design Decision: Disaster Recovery with vSphere Replication (1)

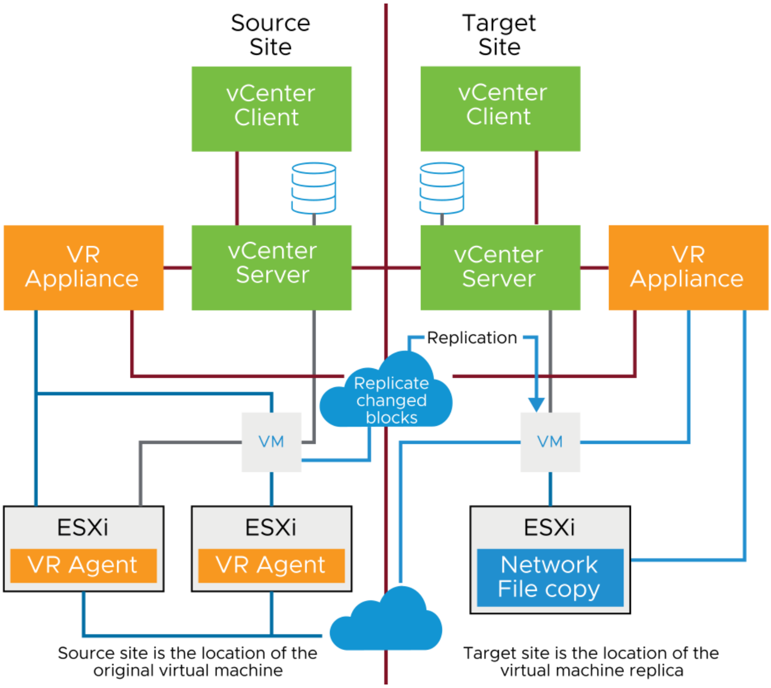
Should you implement vSphere Replication?

vSphere Replication is an extension to vCenter Server that provides hypervisor-based asynchronous replication.

vSphere Replication can be used in the following scenarios:

- Data protection within the same site
- Disaster recovery and avoidance between two sites
- Data center migration

vSphere Replication can be the first step toward a full, multisite disaster recovery solution.



vSphere Replication is an alternative to array-based replication. vSphere Replication protects VMs from partial or complete site failures by replicating the VMs between a source site and a target site:

- Within a single site from one cluster to another
- From multiple source sites to a shared remote target site

vSphere Replication provides the following benefits as compared to storage-based replication:

- Data protection at lower cost per VM
- Flexibility in storage vendor selection at the source and target sites
- Lower overall cost per replication

For more information about vSphere Replication, see vSphere Replication Documentation at https://docs.vmware.com/en/vSphere-Replication/index.html.

11-16 Design Decision: Disaster Recovery with vSphere Replication (2)

Review the RPO and RTO requirements defined in SLAs and determine whether vSphere Replication should be implemented:

- The RTO depends on the configuration.
- The RPO can be set at 5 minutes to 24 hours, and you can configure the RPO for each VM:
  - With vSphere Replication, you can set the RPO to 5 minutes.
  - Target and source site use one of the following storage types:
    - VMFS 5.x and VMFS 6
    - NFS 3 and NFS 4.1
    - vSphere Virtual Volumes
    - vSAN 6.2 U3 storage and later
- Multiple recovery point-in-time instances can be retained.

To avoid RPO violations, provide sufficient storage and network bandwidth for vSphere Replication.

After you set up the replication infrastructure, you can choose the VMs to be replicated, and you can give different RPOs to each VM. RPOs are defined in the replication schedules. The minimum RPO that you can set is 5 minutes using the appropriate storage types.

Retention of multiple recovery points can be useful when an issue is discovered several hours, or even a few days, after it occurred. For example, a replicated VM with a 4-hour RPO contracts a virus, but the virus is not discovered until 6 hours later. As a result, the virus replicates in the target location. With multiple recovery points, the VM can be recovered and then reverted to a recovery point retained before the virus issue occurred.

Storage and network bandwidth requirements can increase when you use vSphere Replication. Network bandwidth requirements increase if all storage is network-based because data operations between the host and the storage also use the network. If slow bandwidth exists between the source hosts and the vSphere Replication servers, an RPO violation can occur. To resolve the RPO violation, more bandwidth should be provided to vSphere Replication (which should shorten replication time) or the RPO policy should be increased.

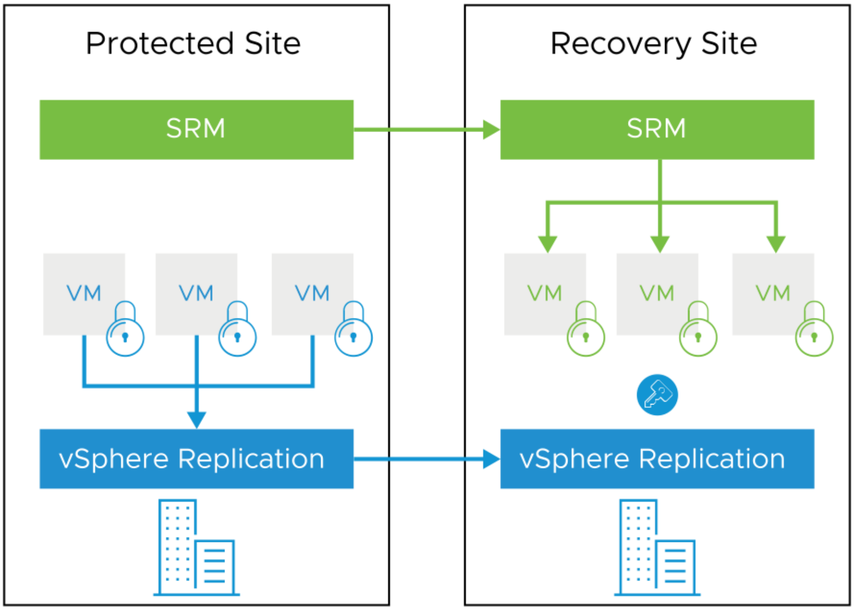
For vSphere Replication bandwidth requirements, see VMware vSphere Replication Documentation at https://docs.vmware.com/en/vSphere-Replication/index.html. Search for bandwidth requirements.

11-17 Design Decision: Disaster Recovery with Site Recovery Manager

What disaster recovery strategies, if any, should you consider?

Use Site Recovery Manager to benefit from the following features:

- Centralized disaster recovery management
- Automated orchestration and testing of centralized recovery plans
- Integrates with vSphere Replication
- Supports several array-based replication products and asynchronous and synchronous replication
- Ensures fast and predicable RTOs to maintain business continuity:
  - RTO: Depends on the configuration
  - RPO: Zero, if the synchronous replication is used



Site Recovery Manager complements vSphere and integrates tightly with vCenter Server to automate disaster protection and planned site migrations for all applications.

Site Recovery Manager provides disaster recovery for an unplanned failover. This use case is the most critical but least frequently used for Site Recovery Manager. Unexpected site failures do not happen often but when they do, fast recovery is critical to business. Site Recovery

Manager can help in this situation by automating and orchestrating the recovery of critical business systems for partial or full site failures, ensuring the fastest RTO.

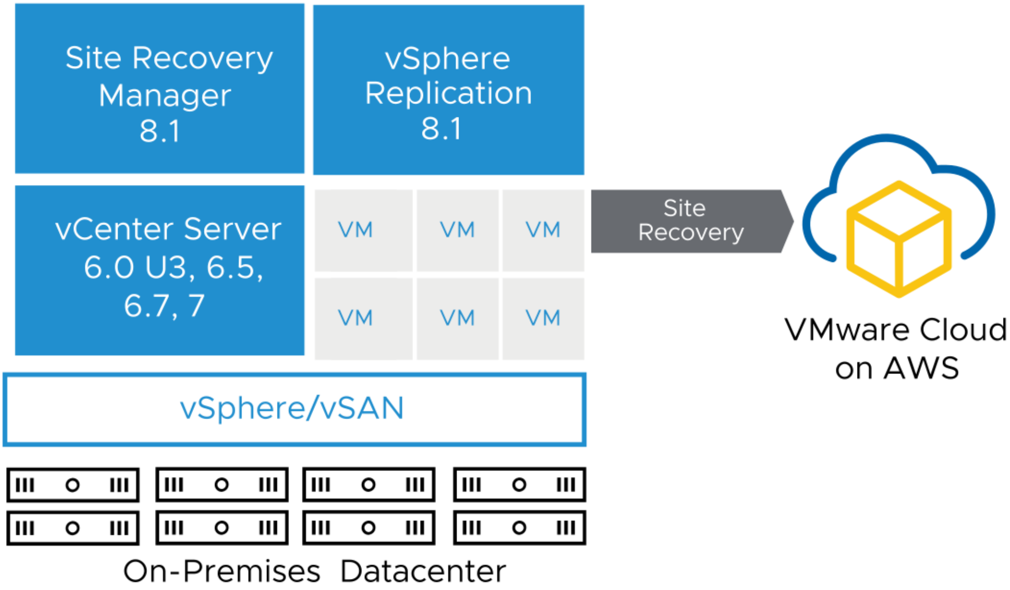
Site Recovery Manager simplifies and automates the key elements of disaster recovery. Using Site Recovery Manager, you can create and test disaster recovery plans, execute fail overs or planned migrations, and fail back to the primary protected site.

For information about Site Recovery Manager and stretched clusters, see https://storagehub.vmware.com/t/site-recovery-manager-3/.

11-18 Disaster Recovery with VMware Cloud on AWS

Simplify DR protection by pairing Site Recovery sites with software-defined data center stacks in VMware Cloud on AWS.

Offload infrastructure maintenance tasks to an environment managed by VMware.



Disaster Recovery as a Service (DRaaS) is compatible with multiple versions, and all editions of vCenter Server. You can protect sites running versions vCenter Server 7.0, 6.7, 6.5, and 6.0 U3. Workloads are protected using the same Site Recovery Manager interface and workflows as used in a regular site-to-site protection scenario.

The same vSphere environment is on-premises and on Cloud, managed through a single interface with no application replatforming required.

11-19 Summary of RPOs and RTOs by VMware Product

Method	RTO	RPO	Use Cases
vCenter Server Appliance high availability feature	5 minutes or less	Zero	- Recoverability
vSphere HA	Depends on the number of VMs to restart	N/A	- Recoverability
vSAN stretched cluster	Depends on the number of VMs to restart	Zero	- Recoverability
vSphere Replication	Depends on the configuration	5 minutes to 24 hours	- Recoverability
Site Recovery Manager	Depends on the configuration	Zero, with synchronous replication	- Asynchronous replication - Synchronous replication - Planned migrations - Patch and upgrade testing

vSphere Replication uses an internal scheduling mechanism that considers factors such as time required for previous replications to complete, number of concurrent replications, and so on. The schedule is constantly adjusted to avoid violating RPO policies and to support balancing the replication workload.

11-20 Lab 15: Designing Infrastructure Recoverability

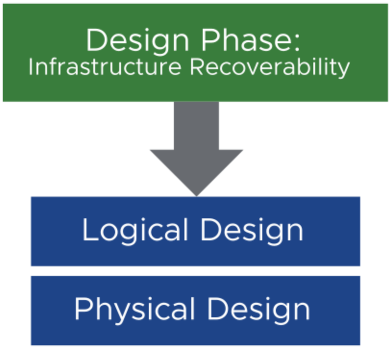
Create an infrastructure recoverability design:

1. Review the Conceptual Design

2. Evaluate Recoverability Design Options

11-21 Review of Lab

The instructor facilitates a class discussion of the design decisions of one or more teams.



11-22 Review of Learner Objectives

After completing this module, you should be able to meet the following objectives:

- Make infrastructure recoverability design decisions that adhere to business requirements
- Design an infrastructure recoverability strategy that meets the needs of the vSphere environment and follows VMware best practices

11-23 Key Points

- Use RTOs and RPOs to determine which data protection tools and technologies to use.
- Use host profiles to recover an ESXi host configuration.
- Back up vCenter Server Appliance with its native backup and recovery tool.
- Back up VMs according to business policies and SLAs.
- Use vSAN stretched clusters for disaster and downtime avoidance.
- With vSphere Storage API – Data Protection, backup solutions can perform full VM image backups, individual disk backups, image-level restores, and file-level recoveries.
- vSphere Replication and Site Recovery Manager can be used as disaster recovery solutions.
- Consider Disaster Recovery as a Service when using VMware Cloud on AWS.

Questions?