



ATIS-0100020.2008(R2013)

Availability Metric for IP-Based Networks and Services

AMERICAN NATIONAL STANDARD FOR TELECOMMUNICATIONS



As a leading technology and solutions development organization, ATIS brings together the top global ICT companies to advance the industry's most-pressing business priorities. Through ATIS committees and forums, nearly 200 companies address cloud services, device solutions, emergency services, M2M communications, cyber security, ehealth, network evolution, quality of service, billing support, operations, and more. These priorities follow a fast-track development lifecycle – from design and innovation through solutions that include standards, specifications, requirements, business use cases, software toolkits, and interoperability testing.

ATIS is accredited by the American National Standards Institute (ANSI). ATIS is the North American Organizational Partner for the 3rd Generation Partnership Project (3GPP), a founding Partner of oneM2M, a member and major U.S. contributor to the International Telecommunication Union (ITU) Radio and Telecommunications sectors, and a member of the Inter-American Telecommunication Commission (CITEL). For more information, visit www.atis.org.

AMERICAN NATIONAL STANDARD

Approval of an American National Standard requires review by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer.

Consensus is established when, in the judgment of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that a concerted effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give an interpretation of any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

Notice of Disclaimer & Limitation of Liability

The information provided in this document is directed solely to professionals who have the appropriate degree of experience to understand and interpret its contents in accordance with generally accepted engineering or other professional standards and applicable regulations. No recommendation as to products or vendors is made or should be implied.

NO REPRESENTATION OR WARRANTY IS MADE THAT THE INFORMATION IS TECHNICALLY ACCURATE OR SUFFICIENT OR CONFORMS TO ANY STATUTE, GOVERNMENTAL RULE OR REGULATION, AND FURTHER, NO REPRESENTATION OR WARRANTY IS MADE OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR AGAINST INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. ATIS SHALL NOT BE LIABLE, BEYOND THE AMOUNT OF ANY SUM RECEIVED IN PAYMENT BY ATIS FOR THIS DOCUMENT, AND IN NO EVENT SHALL ATIS BE LIABLE FOR LOST PROFITS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES. ATIS EXPRESSLY ADVISES THAT ANY AND ALL USE OF OR RELIANCE UPON THE INFORMATION PROVIDED IN THIS DOCUMENT IS AT THE RISK OF THE USER.

NOTE - The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to whether use of an invention covered by patent rights will be required, and if any such use is required no position is taken regarding the validity of this claim or any patent rights in connection therewith. Please refer to [<http://www.atis.org/legal/patentinfo.asp>] to determine if any statement has been filed by a patent holder indicating a willingness to grant a license either without compensation or on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain a license.

ATIS-0100020.2008(R2013), *Availability Metric for IP-Based Networks and Services*

Is an American National Standard developed by the **ATIS Network Performance, Reliability and Quality of Service Committee (PRQC)**.

Published by

**Alliance for Telecommunications Industry Solutions
1200 G Street, NW, Suite 500
Washington, DC 20005**

Copyright © 2013 by Alliance for Telecommunications Industry Solutions

All rights reserved.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher. For information contact ATIS at 202.628.6380. ATIS is online at < <http://www.atis.org> >.

Printed in the United States of America.

American National Standard for Telecommunications

**QUANTIFYING THE IMPACT ON IP SERVICE AVAILABILITY FROM
NETWORK ELEMENT OUTAGES**

Alliance for Telecommunications Industry Solutions

Approved October 14, 2008

American National Standards Institute, Inc.

Abstract

This standard describes a metric that quantifies the impact on IP service availability due to an underlying network element outage. Currently, Network Management System (NMS) tools offer limited capabilities to collect necessary data for estimating this impact. The purpose of this metric is to encourage development of outage measurement capabilities/techniques for metric estimation by equipment vendors.

FOREWORD

The information contained in this Foreword is not part of this American National Standard (ANS) and has not been processed in accordance with ANSI's requirements for an ANS. As such, this Foreword may contain material that has not been subjected to public review or a consensus process. In addition, it does not contain requirements necessary for conformance to the Standard.

The Alliance for Telecommunication Industry Solutions (ATIS) serves the public through improved understanding between carriers, customers, and manufacturers. The Network Performance, Reliability, and Quality of Service Committee (PRQC) -- formerly T1A1 -- develops and recommends standards, requirements, and technical reports related to the performance, reliability, and associated security aspects of communications networks, as well as the processing of voice, audio, data, image, and video signals, and their multimedia integration. PRQC also develops and recommends positions on, and foster consistency with, standards and related subjects under consideration in other North American and international standards bodies.

ANSI guidelines specify two categories of requirements: mandatory and recommendation. The mandatory requirements are designated by the word *shall* and recommendations by the word *should*. Where both a mandatory requirement and a recommendation are specified for the same criterion, the recommendation represents a goal currently identifiable as having distinct compatibility or performance advantages.

Suggestions for improvement of this document are welcome. They should be sent to the Alliance for Telecommunications Industry Solutions, PRQC Secretariat, 1200 G Street NW, Suite 500, Washington, DC 20005.

At the time it approved this document, PRQC, which is responsible for the development of this Standard, had the following members:

M. Neibert, PRQC Chair
 N. Seitz, PRQC Vice-Chair
 C. Underkoffler, ATIS Chief Editor
 P. Tarapore, PRQC Technical Editor

Organization Represented	Name of Representative
Alcatel-Lucent	Ken Biholar Tim Pantalis (Alt)
AT&T	Percy Tarapore Charles A. Dvorak (Alt)
Department of Defense	Chris Fitzgerald
Embarq Corporation	Carl M. Coppage John M. Heinz (Alt)
Ericsson Incorporated	Mustafa Kocaturk Susana Sabater-Maroto (Alt)
ETRI	Tae-Soo Chung Sung-Soo Kang (Alt)
Intelsat	Jose Albuquerque
National Communications System	An Nguyen Carol-Lyn Taylor (Alt)

Organization Represented	Name of Representative
NTIA	Neal B. Seitz Arthur Webster (Alt)
Nokia Siemens Networks	Nagaraja Rao David E. Francisco (Alt)
Nortel	Joseph A. Zebarth
Qwest	Steve Showell Michael Fargano (Alt)
Sprint	Steve Oliva
Telcordia Technologies	Spilios Makris Cliff Halevi (Alt)
Verizon Communications	John Colombo Wendy Pugh (Alt)

The PRQC QoS Task Force, which was responsible for the development of this document, had the following members:

N. Seitz, QoS Chair
 P. Tarapore, QoS Technical Editor
 J. Bennett, QoS Technical Editor

Active Participants:

J. Bennett
 G. Choudhury
 C. Dvorak
 R. Holley
 Y. Kogan

S. Makris
 A. McCain
 A. Morton
 M. Neibert
 A. Nguyen

E. Rojek
 N. Seitz
 P. Tarapore
 A. Webster

TABLE OF CONTENTS

1 SCOPE & PURPOSE	1
2 NORMATIVE REFERENCES	2
3 DEFINITIONS	2
4 ACRONYMS & ABBREVIATIONS	2
5 INTRODUCTION & RATIONALE	3
6 IP NETWORK ELEMENTS	3
7 METRIC FOR SERVICE AVAILABILITY IMPACT FROM NETWORK ELEMENT OUTAGES	5
7.1 DISCUSSION ON SERVICE AFFECTING ELEMENT OUTAGE DURATION.....	6
7.2 DISCUSSION ON ELEMENT FAILURE TYPES	6
7.3 DISCUSSION ON SERVICE PRIORITY	7
A HIGH LEVEL DISCUSSION ON ESTIMATING AVAILABILITY IMPACT ON IP-BASED SERVICES FROM NETWORK ELEMENT OUTAGES	8
A.1 ELEMENT OUTAGE DOWNTIME.....	8
A.2 FRACTION OF SERVICE LOST	9

TABLE OF FIGURES

FIGURE 1: GENERIC IP BACKBONE NETWORK INTERCONNECTION ARCHITECTURE.....	4
---	---

Quantifying the Impact on IP Service Availability from Network Element Outages

1 SCOPE & PURPOSE

Service availability is a critical component in designing IP-based networks. It is also a key measure in Service Level Agreements (SLA) between service providers and customers. Hence, it is important that the concept of availability is well-defined and understood, and that the capabilities for accurate estimation are readily available and implemented by service providers. This document provides a definition of a commonly accepted availability metric for IP-based networks and services that can be correlated to network element outages. Future documents will discuss this metric within the context of MPLS-based transport networks.

Current network trends indicate a significant increase in traffic of different types over IP networks. Traffic types include real-time Voice over IP (VoIP) and video services, data services, “Enterprise Customer” services over “Virtual Private Networks” (VPN), IPTV services, as well as traditional Internet services such as e-mail and Web browsing. In designing SLAs with critical customers for key services such as VPN and video, service providers typically agree to a specified level of availability within the bounds of a pre-determined traffic volume for the service in question. Under such conditions, the predominant and overwhelming contribution to the unavailability of the specified service is a failure or outage in the network. A hardware or software outage may result in downtime for one or more network elements, thus resulting in service disruption with negative impacts on service availability.

This standard describes a metric that quantifies the impact on service availability of an underlying network element outage. Currently, Network Management System (NMS) tools offer limited capabilities to collect necessary data for estimating this metric. The purpose of this metric is to encourage development of outage measurement capabilities/techniques for metric estimation by equipment vendors.

The scope of this document is restricted to estimating service availability impacts within a single IP network domain. It is also restricted to service availability impacts caused by network element outages; it does not include service unavailability periods outside of such conditions (e.g., service degradation as a result of packet loss or jitter). It is recognized that communications services can span several network domains, possibly of differing technologies (e.g., 3GPP Access-IP Backbone-PSTN Termination). Once agreements have been reached on service availability metrics and estimation methodologies, discussions on extending availability on an end-to-end basis over all domains and technologies can commence. Finally, it is recognized that evolving technologies such as MPLS are further transforming IP-based networks. This document serves as the first in a series of documents that will also examine availability estimation within the context of MPLS-based networks and services.

2 NORMATIVE REFERENCES

The following standards contain provisions which, through reference in this text, constitute provisions of this American National Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this American National Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

[Y.1540] ITU-T Recommendation Y.1540, *IP packet transfer and availability performance parameters*.¹

[ATIS-Avail-TR] – “End to End Service Availability”, ATIS Draft Technical Report, ATIS Contribution PRQC-2007-017R4, July 2007.²

[ATIS-Router-Avail-TR] ATIS-0100016, *End to End Service Availability*.²

[Y.2171] ITU-T Recommendation Y.2171, *Admission Control Priority Levels in Next Generation Networks*.¹

[Y.2172] ITU-T Recommendation Y.2172, *Service Restoration Priority Levels in Next Generation Networks*.¹

3 DEFINITIONS

3.1 Service Affecting Element Outage Duration: Total amount of time that a network element is unavailable to transport IP packets. Note that if network restoration successfully re-routes traffic around the failed element, then the latter is no longer considered to be part of the transporting path; hence, outage duration is the time taken to re-route traffic -- see 7.1 for further discussion.

3.2 Element Outage: Failure of a network element or elements that cause a disruption in the transport of an IP service, thus impacting service availability. Such failures can generally be considered as hardware (e.g., power supplies, line cards, routers, links, etc.) failures or software (e.g., software protocols) failures -- see 7.2 for further discussion.

4 ACRONYMS & ABBREVIATIONS

IP	Internet Protocol
IPLR	IP Packet Loss Ratio
ITU-T	International Telecommunications Union-Telecommunication
MPLS	Multi-Protocol Label Switching
NNI	Network-Network Interface
PRQC	Performance Reliability and Quality of Service Committee
PSTN	Public Switched Telephone Network
SLA	Service Level Agreement
UNI	User-Network Interface
VoIP	Voice over IP
VPN	Virtual Private Network

¹ This document is available from the International Telecommunications Union. < <http://www.itu.int/ITU-T/> >

² This document is available from the Alliance for Telecommunications Industry Solutions (ATIS), 1200 G Street N.W., Suite 500, Washington, DC 20005. < <https://www.atis.org/docstore/default.aspx> >

5 INTRODUCTION & RATIONALE

The concept of availability has received significant attention in various standards bodies. A general, high-level definition of availability has been recently defined by ATIS [ATIS-Avail-TR]. According to this definition, end-to-end service availability for a communications network is the fraction of time that service is available between an arbitrarily-specified ingress point and arbitrarily-specified egress point to the communications network. The time fractions are appropriately weighted to account for different capacities of the various access and egress points. ITU-T Recommendation Y.1540 [Y.1540] describes service availability of the IP network in terms of the ability to transport IP packets between any pair of ingress and egress points in an IP network. It defines IP service availability in terms of an availability function that classifies total scheduled service time into available and unavailable periods. This Recommendation also presents a detailed rationale for the availability function, which is expressed in terms of the *IP Loss Ratio (IPLR)*. IP service is considered to be available if IPLR is less than a threshold c_1 :

$$IPLR \leq c_1$$

The value of the threshold c_1 (currently set to 0.75) is considered to be provisional.

The metric defined in this standard differs from prior work as follows. In the development of SLA agreements with key customers for critical services (e.g., VPN, multimedia, video, VoIP), service providers typically agree to availability levels based upon pre-determined traffic volumes for the given set of services. Under such conditions, the predominant and overwhelming contribution to service unavailability is loss of service resulting from a failure or outage of network elements. Quantification of the impact on service availability from such outages can facilitate the administration of the SLA. It can also enable the design and deployment of necessary spare resources such that negative service impacts can be mitigated. This standard provides the necessary metric definition. It should be noted that to date, limited Network Management System capabilities exist for collecting the necessary network status information for estimating the metric. The purpose of defining this metric is to provide encouragement to vendors for developing such capabilities.

6 IP NETWORK ELEMENTS

Figure 1 shows a generic IP backbone network architecture for an Internet Service Provider (ISP). The main elements of the network are *Customer equipment* (triangles), *Access routers* (squares), *Backbone routers* (circles), *Internet Gateway routers* (rhombuses), *Access facilities* (links) connecting customer equipment to access routers and access routers to backbone routers (represented by segments of straight line), and *Backbone facilities* interconnecting backbone routers (not shown because there exist different backbone topologies with respective impact on the backbone availability). In addition, the IP backbone may contain *Route reflectors*, which are not shown here. All nodes (A, B, C, D, & E) have access routers with customer equipment connected to them. For simplicity, Figure 1 does not show all the elements at nodes B, C, D, and E.

This network architecture has the following redundant elements:

1. Customer equipment can be connected to *two different access routers* (double-homing). These access routers may be located in one node or in different nodes. Double homing reduces the likelihood of customer equipment being decoupled from service by a single network element outage.

2. Access routers are connected by *two diverse uplinks* to *two backbone routers*. These backbone routers may be located in one or different nodes.
3. Internet gateway (peering) routers provide a bridge to other service providers. They are connected by *two diverse uplinks* to *two backbone routers*. These backbone routers are usually located in the same node.
4. Backbone routers are interconnected by a set of facilities with pre-designed redundancy that prevents congestion and excessive packet delay for any single backbone router or facility failure. Multiple backbone failures may increase packet delay and cause packet loss, which may impact the service severely enough to be considered service outages.

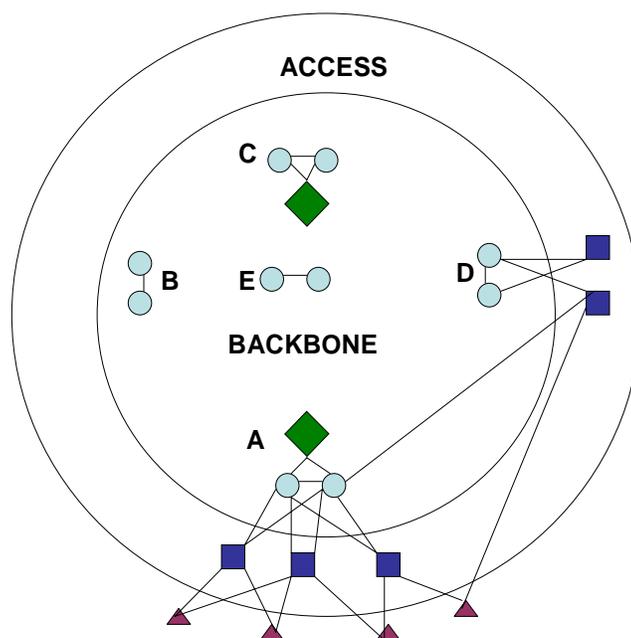


Figure 1: Generic IP Backbone Network Interconnection Architecture

In this illustration, “end-to-end” within the network domain can be determined as follows:

- ◆ For customer traffic originating and terminating at customer equipment, the IP network domain is delimited by the ingress port on the originating Access Router and the egress port on the terminating Access Router in the network. If the access links between customer equipment and Access Routers are owned by the service provider, then the IP network domain is delimited by the originating User-Network Interface (UNI) and the terminating UNI.
- ◆ For customer traffic originating at customer equipment and transiting onto other networks, the IP network domain is delimited by the ingress port of the originating Access Router and the egress port of the NNI Gateway.
- ◆ For Internet Service Provider (ISP) traffic that transits the IP network, the IP network domain is delimited by the ingress port on the originating Gateway Router and the egress port on the terminating Gateway Router.

7 METRIC FOR SERVICE AVAILABILITY IMPACT FROM NETWORK ELEMENT OUTAGES

IP network operations systems need to do the following during a failure:

- ◆ *Determine the downtime of the failed network element(s).* Often, the downtime is of short duration (< 1 minute). Depending on the type of service, the customer may experience a negative impact from a short-duration failure. It is critical that IP network operations capabilities can correctly detect short failure times.
- ◆ *Determine the extent of lost services.* Due to the “connectionless” property of IP networks, the task of correctly determining the degree of service outages resulting from a network failure is daunting. Several assumptions, simplifications, and traffic routing models are necessary in order to provide a reasonable estimate of the degree and extent of lost services.

The definition of an availability metric A_o quantifying the impact on service availability from network element outages for a specified period of time, T , depends on two quantities:

1. Fraction of service lost for each outage that occurs in time period T .
2. Duration of each network/element outage that occurs in T .

Thus for an IP network comprising several components, the availability for any given service over the network that experiences N failures during T can be written as (see 5.1 in [ATIS-Router-Avail-TR]):

$$A_o = 1 - \frac{\sum_{i=1}^N f_i D_i}{T}$$

Where:

- ◆ f_i = Fraction of service lost during outage i . The units of service need to be predetermined. Some examples of service units are:

- Packets.
- Transactions/sessions/calls (e.g., VoIP calls).
- Label Switched paths (LSP) weighted by bandwidth in the case of MPLS networks.

Thus, some examples of the fraction of service lost are:

- Lost Packets/Total Packets.
- Lost VoIP Calls/Total VoIP Calls.
- Lost LSPs/Total LSPs.
- ◆ D_i = Duration of outage i in units of time (e.g., hours).
- ◆ T = Total observation period (e.g., hours in one month).

Additional clarifications are provided in the following subclauses.

7.1 Discussion on Service Affecting Element Outage Duration

Traditionally, *outage duration* is defined as the total amount of time that a network element is in a failed state. From the perspective of the availability metric, the element downtime ends when one of the following “repair” options is completed first:

1. *Network Restoration* – Complete re-routing of IP packets around the failed element. In this case, the downtime is the network re-convergence time, regardless of the actual element repair or replacement time. The assumption is that there is redundant capacity and resources that are engineered in the IP network in order to carry out significant re-routes of IP packet flows. For example:
 - a. *If the failed element is a router Line Card and a spare Line Card exists in the router*, then the downtime is the switchover time. If switchover is instantaneous, then the downtime is zero regardless of the actual time to replace or repair the Line Card.
 - b. *If the failed element is a backbone facility*, then the downtime is the time taken to re-route all IP flows over a spare facility that is engineered over an alternate route. If a large optical trunk group with several facilities fails and there is inadequate capacity to re-route all IP flows, then services with higher re-route priorities have a better chance for successful restoration over others. In such cases, downtime for the re-routed service flows is the restoration/re-convergence time, whereas downtime for all other service flows is the repair time for the optical trunk group.
2. *Element Repair* – In the case where sufficient spare capacity is unavailable for complete restoration/re-route, then the downtime for service flows that are not re-routed is the time to repair the failed element.

Note that maintenance windows are not considered for contribution to service availability. That is, maintenance window periods are subtracted from T and outage durations during those periods are not included in the calculation of A_0 .

7.2 Discussion on Element Failure Types

Network failures/outages can be generally considered as either hardware failures or software failures. Hardware failures relate to elements such as power supplies, line cards, routers, links, etc. Such failures could be captured in terms of system logs (Syslog) and NMS scripts.

Software failures are more complicated. For example, a software protocol failure could impact several routers, causing them to be in a failed state. As such, software failures could potentially be more disruptive. The manifestation of software failures could be captured as the simultaneous failure of several hardware elements such as routers and line cards. Presumably, Syslog and NMS scripts may be able to highlight these simultaneous failures; further analysis would be required to trace the cause back to a software failure.

Other forms of software failures could be even more complex. For example, control plane software errors may result in “message storms” causing a significant strain on network resources. Software failures and their impacts on network operations are for further study. Detailed development of NMS capabilities for tracking downtimes arising from hardware and software failures is also for further study.

7.3 Discussion on Service Priority

When there is network failure, services having higher priority of admission [Y.2171] will be selectively admitted into the network over other services depending on the availability of adequate resources. Similarly, service flows that have been established will get re-routed in the case of failed elements depending on the service restoration priority [Y.2172]. Hence, for example, critical services such as Emergency Telecommunications (ETS) have a very high probability that service availability will suffer minimum impact even under conditions of a regional or national failure. While the availability metric does not include priority as a parameter explicitly, the effect of priority on the availability metric is implicit in the definitions of fraction of service lost and outage duration. A higher admission priority results in a lower fraction of service lost while a higher restoration priority leads to a decrease in outage duration. Hence, the proposed metric captures the effect of service priority on service availability.

Annex A (informative)

A HIGH LEVEL DISCUSSION ON ESTIMATING AVAILABILITY IMPACT ON IP-BASED SERVICES FROM NETWORK ELEMENT OUTAGES

This annex provides a high level discussion on how the two factors – *downtime* and *fraction of service lost* – can be estimated. The discussions offered are generic in nature. Detailed discussions require specific operational and network management structures and are for further study.

A.1 Element Outage Downtime

Estimating element *downtime* could potentially be quantified via historical data. This method looks at all possible element failures in an IP network and tracks failure durations D_i for given elements over their life cycle. It also tracks network re-convergence times (time taken to re-route packet flows over alternate paths) for given element failures. This assumes that there is sufficient spare bandwidth to route all new packet flows after network re-convergence – see Figure 1 for redundancies in the backbone. For example, a complete router failure in the backbone typically results in an Open Shortest Path First (OSPF) protocol re-convergence time of the order of a few minutes. Thus, incoming packet flows may experience an inability to get proper routing over a period typically less than three minutes - - hence, they would be lost. However, subsequent flows would be successfully routed around the failed element. The result is that downtime in this case equals the re-convergence time, even though the actual element downtime may be longer.

Examining and tracking network re-convergence behaviors over different types of failed elements permit the ability to provide reasonable estimates for downtimes related to these failed elements. This method requires network management tools capable of tracking element status, the state of the network (e.g., fully re-converged), as well trouble ticket systems that accurately track customer service status. Some examples of element failures are as follows:

- ◆ *Complete Router Failure*: Typical re-convergence times ~ tens of seconds.
- ◆ *Partial Router Failure* (e.g., Failed Router Card): Typical re-convergence time ~ few seconds.
- ◆ *Complete Link Failure*: Typical re-convergence time ~ few seconds.

To complete the discussion, it is important to consider the case where spare bandwidth is insufficient for re-routing all new incoming flows around the failed element. In such a case, a percentage of new incoming packet flows will be lost until the failed element is repaired. The extent of the loss will have to be monitored by operations systems as well as customer trouble tickets. One example of this possibility is a complete failure of an access router that is linked to a network gateway through which incoming and outgoing traffic flows originate and terminate. If the gateway is dual-homed to another access router, and there is sufficient link bandwidth to accommodate peak traffic flows across both access routers, then a complete access router failure results in a downtime = Border Gateway Protocol (BGP) re-convergence time. If not, then some percentage of incoming/outgoing traffic flows will be lost and the extent will have to be tracked. Repair-time distributions for such cases can be derived by careful tracking.

A.2 Fraction of Service Lost

The *fraction of service lost*, f_i , depends on the service under consideration. Specific services such as VoIP and VPN services require careful tracking. Availability over all services can also be of interest to the service provider. Regardless of the specific service (or total traffic over all services) under consideration, the following data is essential for estimating f_i :

- ◆ *Network and Router Topology* – Connectivity between routers in the IP network.
- ◆ *Network Traffic Matrix* – Point-to-point traffic patterns between pairs of routers depending on time-of-day and day-of-week.
- ◆ *Routing Model* – Traffic routing patterns based on OSPF, Label Distribution Protocol (LDP), and BGP.

Typically, this level of data can be derived for total network traffic. Service specific data can be derived if service specific details are available.

Sampling methods can be used to estimate fraction of service lost for specific services such as VPN services. A service provider may have several VPN services for specific “Enterprise Customers”. For each VPN service, the customer facing ports on access routers are known. To estimate fraction of service lost for any given VPN, some assumptions need to be made:

- ◆ A VPN service in an IP network is modeled as a set of available paths or “connections” between pairs of customer facing ports on access routers that link the customer to the IP network. In other words, if VPN service traffic can ingress the network on port a on access router A and egress at port z on access router Z , then $(A,a)-(Z,z)$ is assumed to represent an available “connection” for the VPN service.
- ◆ A full mesh of bi-directional “connections” is assumed to be available between all relevant pairs of ingress/egress ports related to the specified VPN service.
- ◆ The traffic distribution on these “connections” is assumed to be symmetric and equal³. In other words, VPN service traffic on any given “connection” is likely to be the same as that on any other “connection” regardless of time-of-day and day-of-week.

Detection of VPN service packet streams (incoming and outgoing) at any given port could be done via sampling probes. In the event of a network element outage, the sampling probes would not detect the appropriate VPN streams. The lack of appropriate data for a subset of VPN customer facing ports therefore is an indication that the “connections” comprising these ports have temporarily failed. If j is the number of lost connections and J is the total number of “connections” associated with the VPN service, then the fraction of VPN service lost is the ratio j/J . The assumption here is that the sampling techniques are sufficiently optimized⁴.

In the case of MPLS networks, logical LSPs or Traffic Engineering (TE) tunnels are setup between the end points and their state (“up” or “down”) can be tracked with Network Management Systems. Such LSPs/tunnels can be weighted according to their bandwidth and the fraction of service lost can then be determined.

³ It is understood that port sizes may not be the same. Appropriate weighting and/or normalization factors will need to be implemented in such cases.

⁴ Discussion on sampling techniques is beyond the scope of this document.

For services such as VoIP transactions, it may be possible to track the number of calls lost due to the element outage and compare that with the time-of-day, day-of-week traffic forecasts to determine the fraction of service lost.