International Telecommunication Union

# ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

# Series E.800
## Supplement 8
(11/2009)

SERIES E: OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS

# Guidelines for inter-provider quality of service

ITU-T E.800-series Recommendations – Supplement 8

ITU-T

ITU International Telecommunication Union

ITU-T E-SERIES RECOMMENDATIONS

**OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS**

| | |
|---|---|
| INTERNATIONAL OPERATION | |
| Definitions | E.100–E.103 |
| General provisions concerning Administrations | E.104–E.119 |
| General provisions concerning users | E.120–E.139 |
| Operation of international telephone services | E.140–E.159 |
| Numbering plan of the international telephone service | E.160–E.169 |
| International routing plan | E.170–E.179 |
| Tones in national signalling systems | E.180–E.189 |
| Numbering plan of the international telephone service | E.190–E.199 |
| Maritime mobile service and public land mobile service | E.200–E.229 |
| OPERATIONAL PROVISIONS RELATING TO CHARGING AND ACCOUNTING IN THE INTERNATIONAL TELEPHONE SERVICE | |
| Charging in the international telephone service | E.230–E.249 |
| Measuring and recording call durations for accounting purposes | E.260–E.269 |
| UTILIZATION OF THE INTERNATIONAL TELEPHONE NETWORK FOR NON-TELEPHONY APPLICATIONS | |
| General | E.300–E.319 |
| Phototelegraphy | E.320–E.329 |
| ISDN PROVISIONS CONCERNING USERS | E.330–E.349 |
| INTERNATIONAL ROUTING PLAN | E.350–E.399 |
| NETWORK MANAGEMENT | |
| International service statistics | E.400–E.404 |
| International network management | E.405–E.419 |
| Checking the quality of the international telephone service | E.420–E.489 |
| TRAFFIC ENGINEERING | |
| Measurement and recording of traffic | E.490–E.505 |
| Forecasting of traffic | E.506–E.509 |
| Determination of the number of circuits in manual operation | E.510–E.519 |
| Determination of the number of circuits in automatic and semi-automatic operation | E.520–E.539 |
| Grade of service | E.540–E.599 |
| Definitions | E.600–E.649 |
| Traffic engineering for IP-networks | E.650–E.699 |
| ISDN traffic engineering | E.700–E.749 |
| Mobile network traffic engineering | E.750–E.799 |
| QUALITY OF TELECOMMUNICATION SERVICES: CONCEPTS, MODELS, OBJECTIVES AND DEPENDABILITY PLANNING | |
| **Terms and definitions related to the quality of telecommunication services** | **E.800–E.809** |
| Models for telecommunication services | E.810–E.844 |
| Objectives for quality of service and related concepts of telecommunication services | E.845–E.859 |
| Use of quality of service objectives for planning of telecommunication networks | E.860–E.879 |
| Field data collection and evaluation on the performance of equipment, networks and services | E.880–E.899 |
| OTHER | E.900–E.999 |
| INTERNATIONAL OPERATION | |
| Numbering plan of the international telephone service | E.1100–E.1199 |

*For further details, please refer to the list of ITU-T Recommendations.*

# Supplement 8 to ITU-T E.800-series Recommendations

## Guidelines for inter-provider quality of service

### Summary

Supplement 8 to the ITU-T E.800-series of Recommendations was originally prepared as a white paper by the Interconnection Working Group of the Communications Futures Program (CFP) of the Massachusetts Institute of Technology (MIT). It presents an approach to the deployment of inter-provider quality of service (QoS) to enable further consideration of the topic. This Supplement discusses key issues that service providers need to agree upon if inter-provider QoS is to be readily deployable.

### Source

Supplement 8 to ITU-T E.800-series Recommendations was agreed on 12 November 2009 by ITU-T Study Group 12 (2009-2012).

## FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

## NOTE

In this publication, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this publication is voluntary. However, the publication may contain certain mandatory provisions (to ensure e.g., interoperability or applicability) and compliance with the publication is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the publication is required of any party.

## INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this publication may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the publication development process.

As of the date of approval of this publication, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this publication. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

## CONTENTS

# Supplement 8 to ITU-T E.800-series Recommendations

# Guidelines for inter-provider quality of service

## Executive summary

This Supplement presents an approach to the deployment of inter-provider quality of service (QoS). This Supplement begins from the observation that QoS, based on the differentiated services architecture defined in [IETF RFC 2475] is now widely deployed within the networks of single providers. This is especially the case for providers of network-based VPNs (see, for example [IETF RFC 2547], and [IETF RFC 4364]). Some providers are now beginning to interconnect with each other via "QoS-enabled peering" in an attempt to offer QoS that spans the networks of multiple providers. However, in the absence of appropriate standards and established procedures for management, trouble-shooting, monitoring, etc., such interconnections are likely to be challenging and labour-intensive. This Supplement discusses key issues that service providers need to agree upon if inter-provider QoS is to be readily deployable.

NOTE – As this Supplement is a vehicle for sharing research topics under consideration, there may be some conflicts between this Supplement and existing ITU-T Recommendations. In such conflicts, the reader is reminded that Supplements are only informative and are therefore not considered to be an integral part of any Recommendation; they do not imply any agreement on the part of the ITU-T.

This Supplement:

- Takes the approach of describing the simplest multi-class, multi-provider network scenario (i.e., a single end-to-end low-loss and low-latency class (Class 0) which is offered as a service to customers, in addition to the best-effort class). This does not limit the same approach being applied to multi-provider end-to-end services with IP QoS classes other than the two classes described.
- Assumes that most regional (intra-continental) end-to-end connections will span no more than 3 core provider networks.
- Identifies "common practices" that could ease the deployment of inter-provider QoS if adopted by a critical mass of providers.

While there is continued debate about how many service classes need to be supported across multiple providers, it is widely agreed that some moderate number of classes should be commonly supported and consistently defined among providers. This Supplement uses the two-class scenario as it is the simplest multi-class service offering. This is done as a way of exposing the issues that must be addressed. The additional service class that is defined is intended to be suitable for real-time voice applications, and is intended to be appropriate for use both in a provider-provisioned VPN context and in public networks. It is noted that, in many cases, providers may internally make use of an additional class of service that is restricted to network control traffic (such as routing protocol traffic and network management traffic).

The key issues that are addressed in this Supplement are:

– **Consistent definitions of metrics**: To support QoS meaningfully across multiple providers, it is essential that metrics such as delay, delay variation and loss are defined consistently.

– **Service class definition**: The "Class 0" service class is defined in terms of what the customer must do to receive the service (e.g., mark packets with a certain DSCP, conform to a certain token bucket) and what the provider in turn commits to deliver (e.g., statistical upper bounds on loss and delay). Although this Supplement only outlines detailed criteria for a single class of service beyond best effort, its goal is to remain flexible so that additional classes of service may be added. Furthermore, any individual service provider is free to offer additional service classes beyond those defined here.

- **Measurement, monitoring and reporting**: Because of the multiple parties involved in the delivery of QoS, it is necessary to have defined methods for measurement of QoS, ways to monitor the performance of different network segments, and ways to report performance consistently among providers. Such methods are defined in this Supplement.
- **Routing**: It may be necessary to route QoS-sensitive traffic to different providers or along different routes than those followed by best-effort traffic. Routing considerations are discussed.
- **Provider responsibilities**: There may need to be some agreed-upon responsibilities and "common practices" to which providers should agree. A set of such practices is proposed that simplifies deployment of inter-provider QoS among a large set of providers.

## 1 Introduction

Quality of service (QoS) technologies based on the differentiated services architecture outlined in [IETF RFC 2475] is now widely deployed within the networks of many service providers. This is especially the case for providers of network-based VPNs (see, for example [IETF RFC 2547], and [IETF RFC 4364]). Some providers are now beginning to interconnect with each other via "QoS-enabled peering" in an attempt to offer QoS that spans the networks of multiple providers. However, in the absence of appropriate standards and established procedures for management, trouble-shooting, monitoring, etc., such interconnections have proven to be challenging and labour-intensive. This Supplement seeks to identify the key issues that service providers need to agree upon if inter-provider QoS is to be readily deployable.

### 1.1 Scope

It is the intent of this Supplement to discuss guidelines that are applicable to the interconnection of VPNs spanning multiple service providers. Because QoS deployment is much better established in private networks than in public networks, VPN provider interconnection is the primary focus, but the intent is to outline solutions that may also be applicable to public network service interconnection as well. Specific guidelines for public network service interconnection are for further study.

This Supplement is primarily concerned with the network support for two service classes across multiple providers' networks. These two classes being:

- A service class suitable for latency and loss insensitive data (Class 5).
- An additional service class, considered more suitable for the transport of conversational voice over IP (VoIP) and other latency and loss-sensitive applications (Class 0).

A key distinguishing characteristic of private network services, compared to public network services, for services guaranteeing low packet loss and delay, is that the service provider(s) may not need to implement any means of avoiding congestion occurring for traffic in this class. This is discussed further in clause 3.10.

Within the VPN context, it is likely that many VPN providers will deliver a service based on [IETF RFC 4364] (BGP/MPLS VPNs, formerly known as 2547 VPNs[1]). This Supplement will not restrict itself to BGP/MPLS VPNs – any IP VPN service should be supported – but the specific QoS issues of interconnecting providers of BGP/MPLS VPNs are addressed, including the MPLS-based interconnection styles (referred to as options B and C in [IETF RFC 4364]).

[1] [IETF RFC 2547], which was the informational RFC that described MPLS/BGP VPNs, was superseded by the standards track [IETF RFC 4364].

## 1.2 Relationship to standards

This work draws heavily on the efforts of both the IETF (particularly the IPPM working group) and Recommendations ITU-T [ITU-T Y.1540], [ITU-T Y.1541] and [ITU-T Y.1543].

The focus of this Supplement is on practical operational methods for estimating and managing end-to-end metrics for services across multiple providers' networks. This focus has led to a pragmatic IP Delay Variation concatenation method that differs from that provisionally outlined in [ITU-T Y.1541].

# 2 Reference model and terminology

## 2.1 Definitions

This Supplement defines the following terms:

**2.1.1 access**: That part of an end-to-end IP path from the customer's side of the customer edge (CE) router to the customer's side of the first provider edge (PE) router.

**2.1.2 autonomous system border router (ASBR)**: The router at the edge of an autonomous system (AS), facing towards another AS. ASBRs will typically be located at inter-provider boundaries, and may also be at AS boundaries that are within a single provider when a provider has chosen to divide his network into several ASs.

**2.1.3 customer edge router (CE router)**: The router at the edge of a customer's network, usually facing towards a provider.

**2.1.4 Class 0**: An IP class of service for loss and delay sensitive applications as defined in [ITU-T Y.1541].

**2.1.5 core**: That part of a provider's network from the customer side of the provider edge (PE) router to the customer's side of the distant PE router, or the mid-point of the ASBR – ASBR network-to-network interface.

**2.1.6 customer to provider interface (CPI)**: The interface where a physical link interconnects a customer's network and a single provider's access network. This may also be referred to as a CE to PE interface or CE-PE link interface. [ITU-T Y.1541] refers to this interface as the UNI.

**2.1.7 customer**: The user of the services provided by a service provider. In the context of IPVPNs, a customer typically exists at multiple physical locations, all of which are under one administrative authority, with each site connecting to one or more VPN service providers. In the context of public networks, customers typically connect to a service provider at one or more locations.

**2.1.8 inter-provider link**: The transmission link between two providers. Such a link typically interconnects a pair of ASBRs.

**2.1.9 managed CE device**: A customer edge device that is configured and managed by the provider on behalf of the customer.

**2.1.10 measurement POP**: A service provider's point of presence (POP) that contains equipment capable of initiating and responding to measurement probes from another location.

**2.1.11 network-to-network interface**: The point on an inter-provider link that represents the agreed demarcation point for service performance responsibility between the two different interconnecting network providers.

**2.1.12 option A (B, C)**: Methods for interconnection of MPLS VPNs across service provider (and AS) boundaries, defined in [IETF RFC 4364].

**2.1.13 provider router** (**P router)**. A backbone router, within a public network or VPN service provider(s) network, that only attaches to PE routers of the same service provider.

**2.1.14 provider edge router** (**PE router)**. The router at the edge of a provider's network, usually facing towards a customer.

**2.1.15 provider**: A single public network and/or VPN service provider. In the context of this Supplement, more than one provider is required to deliver an end-to-end quality of service IP path for the service class(es) defined herein.

**2.1.16 trust boundary**: The line between two entities that do not fully trust each other. A CE-PE link is a typical example of a trust boundary because the provider does not trust the customer to configure his equipment correctly or to stay within his SLA parameters. Conversely, an internal link inside a single provider's network is usually not a trust boundary.

**2.1.17 unmanaged CE router**: A customer edge router that is managed by the customer, rather than by a provider.

## Abbreviations

This Supplement uses the following abbreviations:

| | |
|---|---|
| AFI | Address Family Identifier |
| AS | Autonomous System |
| ASBR | Autonomous System Border Router |
| ATM PVC | Asynchronous Transfer Mode Permanent Virtual Circuit |
| BGP | Border Gateway Protocol |
| CDF | Cumulative Distribution Function |
| CE | Customer Edge |
| CPI | Customer to Provider Interface |
| Diffserv | Differentiated services |
| DSCP | Differentiated Services Code Point |
| E2E | End-to-End |
| EF | Expedited Forwarding |
| EXP | Experimental (field in MPLS header that carries Class of Service information) |
| FCAPS | Fault-management, Configuration, Accounting, Performance, and Security |
| FR DLCI | Frame Relay Data Link Connection Identifier |
| GigE | Gigabit Ethernet |
| IP | Internet Protocol |
| IPDV | IP Delay Variation |
| IPLR | IP Loss Ratio |
| IPPM | IP Performance Metrics |
| IP-QoS | IP Quality of Service |
| IPTD | IP Transfer Delay |
| IPVPN | Internet Protocol Virtual Private Network |
| LL | Low Latency |

| | |
|---|---|
| LSP | Label-Switched Path |
| MPLS | MultiProtocol Label Switching |
| MPoP | Measurement Point of Presence |
| NLRI | Network Layer Reachability Information |
| NMS | Network Management System |
| NNI | Network-to-Network Interface |
| OAM | Operationals and Management |
| OPSEC | Operational Security Capabilities |
| OWAMP | One-Way Active Measurement Protocol |
| PDU | Protocol Data Unit |
| PE | Provider Edge |
| PHB | Per-Hop Behaviour |
| PWE3 | Pseudo-Wire Emulation Edge to Edge |
| RPSEC | Routing Protocol Security Requirements |
| SP | Service Provider |
| SAFI | Subsequence Address Family Identifier |
| SLA | Service Level Agreement |
| ToS | Type of Service |
| TWAMP | Two-Way Active Measurement Protocol |
| TWPD | Two-Way Packet Delay |
| TWPL | Two-Way Packet Loss |
| UTC | Coordinated Universal Time |
| VC | Virtual Circuit |
| VoIP | Voice over IP |
| VPN | Virtual Private Network |

## 2.2 Reference approach

The key practice underpinning this Supplement is that DSCP in the IP header is the default definitive indicator to all providers in the end-to-end path of the underlying QoS treatment an IP packet is expected to receive. Providers may use parameters in other protocol headers to convey QoS treatment required (e.g., where encapsulation of the IP packet occurs) but in the event that these markings differ from the QoS class indicated by the DSCP parameter, the latter will be the definitive indication of the treatment expected by the customer.

## 2.3 Basic reference case

For simplicity, the single provider case depicted in Figure 1 is initially considered. In this model, customer sites connect to the provider via a CE (customer edge) device, and the provider's routers that connect to customer sites are PE (provider edge) devices.

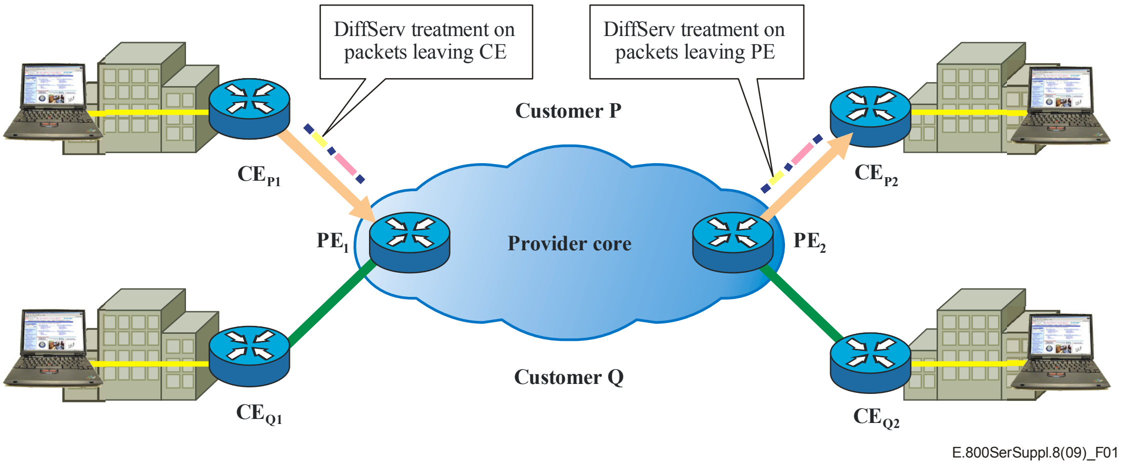

**Figure 1 – Basic Diffserv model for a single provider**

The customer interface is the interface between the customer equipment and the CE router. The CE router can therefore prioritize and police traffic put onto the CE to PE link to treat differently various classes of traffic. The basic single provider reference model places no restrictions on the mechanisms that are deployed by the provider within his access and core networks. Services are defined in terms of externally measurable performance parameters (e.g., loss, delay), with the mechanisms for achieving those performance targets left to the single provider.

Two CE models typically occur in the marketplace and these impact how and where a provider may manage the performance of its service to its customers even in the single provider case.

### 2.3.1 Customer-managed CE device model

In the customer-managed CE device model, it is the responsibility of the customer to ensure that the traffic that traverses the CE to PE link is "correctly" marked before it reaches the PE. "Correct" in this context simply means that the customer needs to ensure that packets are marked in a way that ensures they receive the service desired. For example, if the customer has subscribed to a Class 0 service and the provider/customer contract for this service dictates that packets must be marked "EF" to receive the service, then the customer must decide which of his packets are to receive the Class 0 service and mark them before they arrive at the PE. The selection of packets to receive the Class 0 service is thus entirely up to the policies of the customer.

The PE router, and/or other devices in the access transport network, may enforce various aspects of the service performance contract, such as policing the amount of Class 0 traffic received from, or sent to, a given customer. The details of such policing will be an aspect of a provider's service definition that enables the provider's commitments for Class 0 to be managed and therefore met.

### 2.3.2 Provider-managed CE device model

When the service provider manages the CE devices on behalf of the customer, it is possible to move the trust boundary to the CE. This means that the CE router, rather than the PE router, can be responsible for ensuring that the amount of traffic sent on an access IP path for any service class does not exceed the contract for that service class. This may be achieved by policing or shaping of the customer traffic before sending it to the PE.

## 2.4 Inter-provider reference case

Figure 2 illustrates the Diffserv model extended to a simple inter-provider scenario. Its main difference from Figure 1 is that there are now two providers in the path between the two sites of each customer. The connection between the two providers is referred to as the inter-provider link.

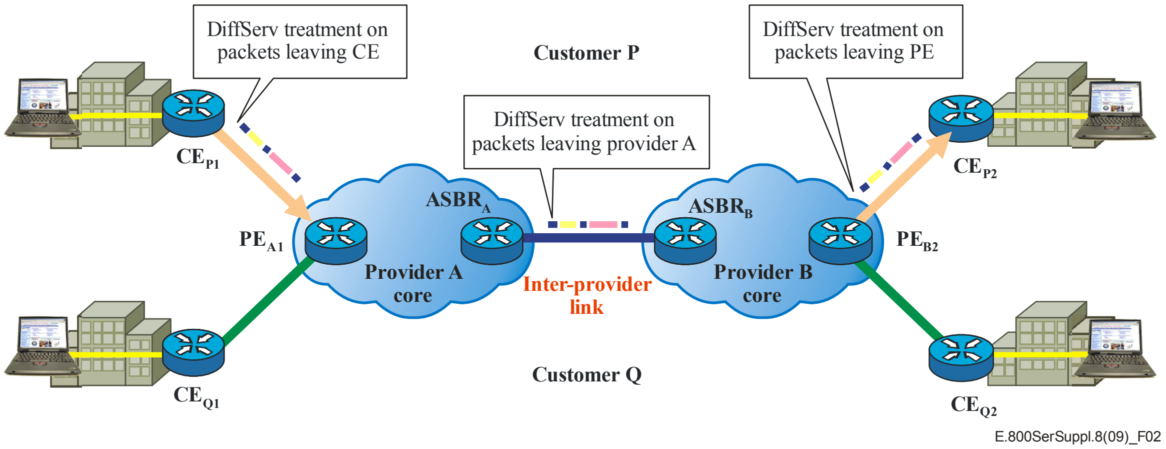

**Figure 2 – Basic inter-provider model**

When the problem of delivering a particular service (e.g., the "Class 0" service) to customer P is considered, several issues that were not present in the single provider case must be addressed, including:

- Packets should be "correctly" marked on the inter-provider link to obtain the desired onward service, and the providers may have different markings for a given service.
- It may be desirable to carry that marking in a manner that avoids modification of the customer's data packets, e.g., in an extra header.
- Providers A and B should each offer a service that, when concatenated with the service of the other provider, provides a useful end-to-end service to the customer (e.g., for a service with a fixed maximum delay, the allowable delay may need to be agreed between both providers).
- Monitoring the end-to-end performance experienced by the customer is now likely to involve both providers in the path.

Marking on inter-provider links is the subject of clause 2.5. Service definitions are discussed in clause 3. QoS measurement issues are discussed in clause 4.

It is desirable to support a wide range of interconnection methods. It should be possible to support a simple IP interconnect (which would include option A interconnection of [IETF RFC 4364] VPNs) as well as an MPLS interconnect of both option B and option C styles defined in [IETF RFC 4364]. Interconnection using MPLS traffic engineered LSPs should also be possible. It should also be possible to support any sort of layer 2 interconnect (e.g., ATM, Ethernet, etc.). The encapsulation and style of interconnection used at the inter-provider boundary has consequences on the marking and policing requirements, as discussed below.

### 2.4.1 Provider-managed CE device model with multiple providers

The managed CE device case is more complicated when there are multiple providers as in Figure 2. For example, if customer P purchases a managed service from provider A, who manages all of that customer's CEs, then the link between customer P and provider B still represents a trust boundary, while the link between customer P and provider A does not. In summary, the management of CEs by providers may or may not cause trust boundaries to be different than in the unmanaged CE case.

## 2.5 DSCP marking

It is recommended that customer packets egress a provider's network with the same DSCP marking they had on ingress. If a network provider modifies any DSCP values within their network, then the network provider should implement a means to accurately restore the original value of the packet's marking at the egress of their network.

When a packet ingresses an access provider's network from an unmanaged CE, the packet marking on egress from that provider's network at any NNI should match the class that the packet should be carried in by the subsequent network provider(s). Whether the ingress access provider re-marks the packet in their network (if the packet is non-conforming to the service contract and the provider elects to remark) is the access provider's choice.

It is often necessary for the provider to impose a QoS treatment on customer packets that differs from that which might be indicated by the customer's DSCP. For example, a customer may have an SLA that allows him to send traffic with DSCP=X up to a rate *r*, with excess packets being downgraded to best effort. However, even if the packets are treated as best effort by the provider, the customer wishes to retain the DSCP marking of X for his own use when the packets arrive at his remote site. In the single provider environment, this capability is readily provided by encapsulating the customer's data with a header that exists only in the service provider network, e.g., an MPLS label header. This header is used to carry the service provider's desired marking for the traffic, while leaving the customer's headers intact.

When there are multiple providers in the path, as in Figure 2, the marking issue is slightly more complex. Packets need to be marked appropriately to receive the desired service from the provider on the receiving side of the link. (That is, packets from provider A need to arrive at the edge of provider B with an appropriate marking for the desired service.) In options B and C, or when MPLS-TE is used across the inter-provider boundary, the MPLS EXP header may be able to be used to carry the marking, thus leaving the customer header unchanged. In option A or pure IP interconnection, it may be possible to encode the marking in a layer-2 dependent way to again leave the customer header unchanged. For example, an Ethernet 802.1q tag may be used to carry the marking across the boundary, or multiple ATM VCs may be used, one per service, with provider A placing the packets on the appropriate VC to receive the desired service from provider B, and vice versa. Such border arrangements can be established though bilateral agreement between two interconnecting providers; however, when three or more providers are in the end-to-end path, such practices become very challenging to be reliably supported and therefore should be avoided.

## 2.6 Routing

In a network as simple as that shown in Figure 2, there are no real routing issues since there is only one path between any two customer sites. However, it is clear that in a true multi-provider environment there may be many alternate paths between customer sites. The preferred path among providers is typically determined by BGP policies. However, when multiple classes of service exist, it may be desired to route some traffic preferentially via providers who support the enhanced QoS class(es) while best-effort traffic takes the conventional route. This issue is discussed further in clause 5.

## 2.7 Measurement and management

Performance measurement and management is both important and challenging in the inter-provider QoS context because of the number of potential providers in the path between two customer sites. In the single provider case, a customer can conduct performance measurement between CEs; if the performance targets are not met, it can be assumed that the problem lies with the provider (unless of course the customer has overbooked and thus congested the access links). Even in a network as simple as the one shown in Figure 2, there are now five possible locations of performance problems for a given site-site pair: the two access links, the core networks of each of the two providers, and the inter-provider link.

In order to deal with troubleshooting and performance monitoring issues, QoS measurement needs to be addressed as part of an inter-provider QoS solution. This topic is addressed in detail in clause 4 and draws on [ITU-T Y.1543].

# 3 Service class

## 3.1 Network service classes

A best-effort service is the default service class that is assumed to be available everywhere. This Supplement is primarily concerned with the support for an additional service class.

Service classes are treated here in terms of "black-box" behaviours – that is, the externally visible attributes of the service are considered (e.g., loss, delay) rather than internal network implementation mechanisms (e.g., Diffserv PHBs). In this respect, service classes are similar to the Diffserv concept of a per-domain behaviour (PDB) defined in [IETF RFC 3086], but the scope of a service is not limited to a single domain. To be precise, a service class is defined by the externally visible treatment that the packets in that class receive as they traverse a network (in terms of loss, delay, and delay variation, and potential transfer capacity). There may be additional aspects of a service class definition such as a default marking for packets in that service class.

The two network classes considered here are two of the classes defined in [ITU-T Y.1541]. These are:

- Class 5 (a "default best effort" class);
- Class 0.

### 3.1.1 Class 5

Traffic that has not been explicitly identified and associated with another service class will receive Class 5 ("best effort") treatment. Class 5 has no explicit guarantee with respect to latency, delay variation, or packet loss; however, providers typically do endeavour to provide for satisfactory delivery of packets in this service class, and provider service commitments for Class 5 are not uncommon.

### 3.1.2 Class 0

The additional service class in this Supplement is Class 0, which is intended to be used for the transmission of services that require low delay, low delay variation and low loss between customers of different service providers, and where the traffic between the customers traverses two or more disparate carriers. The class is intended to be suitable for real-time applications such as VoIP, but there is no restriction on which applications may actually use the service. Mapping of applications to service classes is left to the customers.

For many applications, Class 0 will carry traffic bidirectionally (e.g., media packets in both directions of an individual VoIP call). However, there is no requirement to provide a symmetric path for the bidirectional traffic flow between any given source and destination.

### 3.1.3 Other service classes

Service providers are at liberty to offer any number of service classes above and beyond those defined in this Supplement. Indeed, it is typical to offer four or more service classes to customers and also to use one or more internal classes for network control (e.g., routing protocol) traffic. It is expected that more classes will probably need to be agreed upon for inter-provider use at some point in the future, but the discussion of additional classes is for further study. As noted above, even agreeing on how to support one additional class (i.e., Class 0), beyond the default Class 5, raises many, if not most, of the hard problems that need to be addressed.

Even if there were a larger number of "standard" service classes that could be offered in an inter-provider context, it is likely that providers would continue to offer some additional classes beyond the standard set as a means of competitive differentiation. An inter-provider QoS model should allow for such flexibility.

## 3.2 Customer to provider interface (CPI) behaviour

### 3.2.1 Marking of customer traffic

At the CPI, the customer should appropriately mark packets that are to receive Class 0 service. It is recommended that the DSCP value for Expedited Forwarding (101110) recommended in [IETF RFC 3246] should be used for packets that the customer intends to receive Class 0 service.

If packets at the CPI are MPLS encapsulated (e.g., because a carrier's carrier service is being offered to the customer), then the top MPLS header should contain an EXP value of 5.

If packets at the CPI are encapsulated in Ethernet frames (e.g., because an Ethernet access service is being used by the provider to reach the customer), then the priority information in the Ethernet tag should contain a value of 5, or provisionally 6, in the case where the access network service solely determines QoS treatment of packets based on Ethernet priority information values. The recommendation of a single Ethernet priority information value for this class is for further study. A default value of 5 should be used to indicate this class by providers when a preference for use of a value of 5 or 6 is not specified by or agreed with the customer.

For traffic that is to receive Class 5 service, the customer should mark the packets with a DSCP value, MPLS EXP value, or Ethernet Priority information value where appropriate, of 0.

Providers are free to specify the use of other DSCP, MPLS EXP or Ethernet Priority values for other service classes beyond Class 5 and Class 0. It is recommended that, wherever practical, the DSCP values used by providers for other classes should be consistent with those defined in [IETF RFC 4594]. This will reduce the risk that markings used will conflict with any future end-to-end multi-provider class values adopted.

There is no restriction as to what type of traffic the customer may place in any service class. For example, if the customer chooses to place bulk data traffic with long packets in the Class 0 service, it may degrade the performance of that customer's voice traffic experiences, but that is up to that customer to decide, provided it complies with any service requirements defined by the customer's service provider (e.g., MTU for the class, class capacity).

If traffic from the customer is marked with a value that does not match any of the acceptable values that have been agreed with the customer's service provider, the ingress provider may take any action that the provider considers appropriate (such as dropping or remarking). Note that this issue also arises at the NNI and is discussed below.

### 3.2.2 Policing and re-marking

Policing to manage the network transfer capacity of Class 0 is performed by the access network provider at, or as close as practically possible to, the CPI and it is recommended that a single rate policer with burst size control is utilized. That is, the SLA includes a token bucket rate and burst size; traffic sent by the customer that exceeds this token bucket at the CPI will be dropped. Such policing should be performed at the PE or in the access network in the case of unmanaged CE devices. It should be performed at the CE if, and only if, the CE device is managed by the provider.

Re-marking of excess traffic may be appropriate for future service classes, but it is not recommended for Class 0.

The configuration of egress queuing (from egress SP's PE device to ingress CE device) is a local matter for the service provider.

## 3.3 Network-to-network interface (NNI) behaviour

### 3.3.1 Marking of traffic at the NNI

At the NNI, packets that are to receive Class 0 service should be appropriately marked. In keeping with the recommendation that DSCP values are preserved or restored on egress from a provider's network, it is recommended that the DSCP value for Expedited Forwarding (101110) should also be used for packets at this interface that are to receive Class 0 service. If packets at the NNI are MPLS encapsulated (e.g., because options B or C are in use at the NNI) or encapsulated in an Ethernet VLAN (because option A is in use at the NNI), then the top MPLS header or outer Ethernet tag should contain an EXP or priority value of 5.

For traffic that is to receive Class 5 service, packets should retain their default marking with a DSCP value of 0 and/or, particularly where other DSCP values are in the packet header, the packets encapsulated with an outer tag EXP or Priority value of 0.

Providers are free to negotiate with their peers the use of other DSCP or MPLS EXP values for other service classes beyond Class 5 and Class 0.

If traffic from one provider to another does not match one of the agreed-upon values for that interface, then the behaviour is unspecified – that is, traffic may be dropped, re-marked, or transmitted unmodified with any QoS treatment the receiving provider chooses.

### 3.3.2 Policing and re-marking at the NNI

Where any policing required to manage the network transfer capacity of Class 0 is performed at the NNI, it is recommended that a single rate policer with burst size control is utilized. That is, the service agreement between the peering providers includes a token bucket rate and burst size; traffic sent by a provider that exceeds this token bucket at the NNI will be dropped. Such policing should be performed at the ASBR of the receiving provider. It is expected that the token bucket parameters will be statically configured as a result of offline configuration in the absence of any other admission management mechanisms. Admission management is discussed further in clause 3.10.

Re-marking of excess traffic may be appropriate for future service classes, but is not recommended for Class 0.

On receipt of packets from the NNI, an SP may encapsulate packets, (e.g., using IP or MPLS), and mark the encapsulating header with a 'local-use' DSCP or, EXP values within that provider's backbone, as long as the underlying IP packet header marking is not modified or can be reliably restored to its value on egress.

It is generally considered desirable to avoid re-marking of customer traffic in a way that the customer can detect, i.e., by modifying the customer's DSCP values. This means that if re-marking is required for some reason (e.g., to deal with unknown or unexpected DSCP values), it is desirable to encapsulate the customers traffic with a header that can carry the desired marking, rather than modifying the customer's DSCP. The implication of this policy at the NNI is that it is preferable to carry traffic in an encapsulation that supports some sort of marking other than the customer's DSCP. Option B and Option C meet this requirement, since there is an MPLS header to carry the marking. It is also possible to apply an MPLS (or IP) header at the NNI purely for the purposes of carrying an EXP (or DSCP) value – this is feasible even with Option A or a pure IP interconnect.

The configuration of egress queuing (from one provider's ASBR as he transmits onto the link) is a local matter for the service provider.

### 3.4 Definitions of metrics

#### 3.4.1 Initial considerations

Class 0 is characterized by three network performance metrics: mean one-way delay, one-way packet loss, and one-way delay variation. The metrics for this class are those defined in [ITU-T Y.1540]. The main challenge in the context of managing and measuring end-to-end performance across multiple provider networks is to restrict the options available. The IPPM RFCs allow a great deal of latitude in metric definitions. Since the desire in this Supplement is to produce service classes with metrics that can be meaningfully concatenated, it is important to have commonality of metrics across providers. The problem of practically and meaningfully concatenating metrics across multiple providers is considered, and this has motivated recommending an operationally simplified method for estimating compliance with end-to-end one-way delay variation targets, rather than the more precise estimation method provisionally recommended in [ITU-T Y.1541].

Additional metrics can also be defined for Class 0 traffic but their use is not considered in this Supplement and is for further study. These include: availability, connectivity, transfer capacity, severely errored seconds and packet reordering.

#### 3.4.2 Measurement and reporting considerations

The primary measurement method recommended for inter-provider performance metrics is active measurement as described in [ITU-T Y.1543]. The metrics discussed in this clause relate both to customer traffic and active probes.

Passive measurement may also assist a provider in supplementing active measurements for service management and also for a provider's operational requirements. However, passive measurement is not recommended as an alternative to active measurement as the sole basis for inter-provider measurement and reporting.

There is a widespread practice of reporting two-way metrics or one-way metrics derived from two-way measurements. However, it is preferable one-way metrics are measured and reported, as they reflect most accurately the performance of the network. One-way measurements do, however, require accurately synchronized clocks. It is recommended that one-way metrics should be reported whenever possible; one-way values derived from two-way measurements may be reported only if one-way measurement is not practically possible, and the fact that they are not true, one-way metrics should be reported.

All performance guarantees are only for conforming packets/traffic – packets sent outside the service capacity (token bucket) parameters for a given interface are not counted in performance measurements of their senders' service class, even if they are delivered.

Metrics are always defined by the relevant single instance of the metric measurement and the reported statistics of the metric. Single measurements are rarely reported and rarely stored in the network-wide, operational performance measurement systems. Single measurements may be used and reported during the debugging or calibration process.

For Class 0, all metrics should be defined for packets that are representative of the traffic that will use that class. Thus they should use IP/UDP/RTP packets with a payload size of 160 bytes (representative of common voice codecs today). Test probe packets should be marked with the appropriate DSCP or MPLS EXP values as defined above for Class 0.

For all the metrics defined here, a number of measurements should be taken over a defined time interval (rollup period). When reporting these measurements, the time of the start of the rollup period shall be reported, referenced to UTC.

For all the metrics defined here, the recommended mean probe transmission period (PTP) and rollup period (RP) are defined in the following clauses. Choosing the sampling frequency is clearly a complex trade-off between accuracy and load on the network itself and on the measurement devices. It is believed a 200 ms mean sampling interval is a reasonable compromise, and it is noted that providers may probe more frequently if they wish (perhaps on an exceptional basis, e.g., for troubleshooting).

Consideration of the pattern of inter-probe timing is important. Where probing with equal inter-probe intervals is utilized, probing should occur in blocks with randomized start times of each probe block. Test inter-probe timing may alternatively be continuous and generated by an interval randomizing process to avoid any periodic effects. In either case, the test probe pattern used should be clearly stated with any metrics reported.

### 3.4.3 One-way packet transfer delay

The one-way IP packet transfer delay (IPTD) metric is defined in [ITU-T Y.1540].

The single instance of the IP packet transfer delay is defined for all successful and errored packet outcomes traversing network segment(s) between two reference points. The metric is defined as the time from the time the first bit of the packet is put on the wire at the source reference point to the time the last bit of the packet is received at the receiver reference point.

The IPTD metric is reported as the arithmetic mean of several (specified) single measurements over the rollup period. Errored packets are included in the calculation but, for obvious reasons, lost packets are excluded.

### 3.4.4 One-way IP packet delay variation (IPDV)

A number of 2 point definitions of IP packet delay variation (IPDV) have been described in [IETF RFC 3393] and also in [ITU-T Y.1540].

Use of the minimum delay as the 2-point reference is preferred, but it cannot be certain that the lowest delay observed be any single instance of IPTD in any rollup period. This Supplement proposes a practical definition for IPDV, which closely follows the definition adopted in [ITU-T Y.1541].

A definition of the IP packet delay variation (IPDV) can be given for packets inside a stream of packets. The singleton measure of IPDV for a pair of packets within a stream of packets is defined for a selected pair of packets in the stream going from measurement point MP1 to measurement point MP2.

In this Supplement, the IPDV$_{(n)}$ is the difference between the one-way-delay of the selected packet and the packet with the lowest IPTD measured in the rollup period.

$$IPDV_{(n)} = IPTD_{(n)} - IPTD_{(\mathrm{min})}$$

When reporting IPDV, it is more practical and useful to report at least one point on the IPDV distribution in an evaluation interval rather than the entire distribution of singleton measures.

The selected point of the distribution should follow the quantile-based limits on IP packet delay variation recommended in [ITU-T Y.1540]/[ITU-T Y.1541]. Specifically, at least the 99th percentile of the IPDV distribution over the rollup period is to be reported. Furthermore, in the rest of this Supplement and unless noted otherwise, the term "IPDV" is used indistinguishably with its 99th percentile IPDV, the subscript in IPDV$_{(99\text{th P})}$ is omitted where it causes no confusion. Recommendation of the 99th percentile here differs from the 1 x $10^{-3}$ percentile recommended in [ITU-T Y.1541]. This is because of the very large measurement sample required to adequately determine a 1 × $10^{-3}$ percentile. The 99th percentile is considered an option for pragmatic operational service management purposes.

### 3.4.5 IP packet loss ratio (IPLR)

The IP packet loss ratio (IPLR) metric is defined in [ITU-T Y.1540].

A single instance of packet loss measurement is defined as a record of the packet sent by the sender reference point at the destination reference point. The record is 0 if the packet was received or 1 if the packet was not received. A packet is deemed to be lost if its one-way delay exceeds an agreed time $T_{max}$. [ITU-T Y.1540] recommends a value of 3 seconds for $T_{max,}$ for general use when a packet is deemed to be lost. A packet is also counted as not received if it is corrupted in transit.

Packet loss ratio is defined as a metric measured for packets traversing the network segment between the source reference point and the destination reference point. The IPLR metric is reported as the number of packets not received at the destination reference point after $T_{max}$, or received corrupted, divided by the number of packets sent at the sender reference point to that destination.

### 3.5 Class 0 service definition

Using the metrics defined in clause 3.4, it is possible to determine the end-to-end performance characteristics of a Class 0 service. Details of the measurement approach to be taken are presented in clause 4. In this clause, performance characteristics are recommended for IP or MPLS traffic that may traverse the networks of multiple providers. The issue of how the total impairment budget is allocated among multiple providers is discussed below.

[ITU-T Y.1541] defines two classes that are potentially suitable for VoIP (Classes 0 and 1). Class 0 is the more stringent and is the class considered here. Wherever possible, providers should aim to achieve the end-to-end targets for [ITU-T Y.1541] Class 0.

The parameters specified in [ITU-T Y.1541] for Class 0 are as follows:

**Table 1 – ITU-T Y.1541 Class 0 targets**

| Parameter | Recommended upper bound |
|---|---|
| IPTD | 100 ms |
| IPDV | 50 ms |
| IPLR | 0.1% |

Where geographic distance prevents the delivery of Class 0 IPTD performance, the upper bound on IPTD should be as close a practically possible to the ITU-T Y.1541 Class 0 IPTD upper bound. This is discussed further below.

[ITU-T Y.1541] assumes that the above values are calculated on a 24 hours/7 days-per-week basis, unless specified otherwise. This Supplement proposes that the above metrics are determined 24/7 excluding periods of unavailability and planned outages. See clause 7 for further discussion of maintenance windows.

### 3.6 Impairment budgets

#### 3.6.1 Introduction

To support real-time traffic in multi-provider VPNs with the desired quality of service, the end-to-end impairment objectives for Class 0, as defined above, should be met. The topic of this clause is the impairment allocation among multiple providers in order to meet those end-to-end objectives.

The guidance provided here is intended to accelerate the planning, deployment and management of networks and systems that can interoperate with a clear goal of supporting the end-to-end performance objectives detailed in [ITU-T Y.1541].

### 3.6.2 Division of end-to-end Class 0 budget

Any algorithm for impairment budget division among providers should be evaluated along with its probable implementation(s), which the following objectives reflect:

1) The algorithm should be:
   a) Scalable – It should be able to support paths between the many edges of every network provider.
   b) Robust – It should be widely applicable to the majority of situations including unusual topologies and distances, and recognize that capabilities of access and core networks are different (core networks typically have high capacity and more resilient paths between points, whereas access networks may not).
   c) Low overhead – The amount of extra traffic and extra infrastructure should be considered.
   d) Timing appropriate to path selection needs – Business needs may dictate the need for frequent usage of allocations on multi-second, multi-month or indefinite sessions, starting immediately or at some time in the future.
   e) As simple as possible.
   f) Secure – considering:
      i) Access control.
      ii) Authentication.
      iii) Non-repudiation.
      iv) Data confidentiality.
      v) Communication security.
      vi) Data integrity.
      vii) Availability.
      viii) Privacy.
   g) Resistant to gaming – Providers which do not meet expected objectives must be detectable.
2) Time sensitivity of solution:
   a) The evolving nature of requirements and technology are recognized. Consideration of solutions should target particular deployment timeframes and evolving technology trends.
3) Consideration of how service providers handle cases where the aggregated impairments exceed those specified for a network QoS class.

Some algorithms will, by their very nature, support additional capabilities that are not seen as current requirements. For example, a provider may offer a menu of impairment capabilities between edges based upon offered financial cost. It is recognized that the evaluation of solutions may change if requirements change.

### 3.6.3 Challenge of budget division and subsequent concatenation

Compared to networks and systems that are circuit-based, those based on IP pose distinctly different challenges for planning and achieving the end-to-end performance levels necessary to adequately support the wide array of customer applications (voice, data, fax, video, etc.). The fundamental quality requirements for these applications are well understood and have not changed as perceived by the customer; what has changed is the technology (and associated impairments) in the layers below these applications. The very nature of statistically multiplexed IP-based networks makes balancing capital efficiency with end-to-end performance across multiple network operators a major challenge for applications with stringent performance requirements.

Clause 3.5 outlines end-to-end targets for the classes being considered in this Supplement. These end-to-end targets are valuable to aid the definition of the service experience of an end customer, but are of a lesser value in themselves to network planners.

Where an end-to-end service is provided across a single provider's network, the service planner can singly apportion or allocate impairments in various parts of the provider's network for each service IP path offered. The network planner has the full visibility of all network components contributing to the overall outcome and can plan the resultant service in the manner that best fits the desired technical and commercial outcome.

Where an end-to-end IP path spans multiple provider's networks, this end-to-end visibility is no longer so readily available. Network planners need to understand the boundaries to work within that will result in a high probability that an acceptable outcome can be achieved across any reasonable combination of providers that may be required to collectively provide an end-to-end service.

The approaches that could be taken in allocating total impairment targets among network segments can be characterized by the amount of information shared among segments and at what point in the design process and subsequent operation of the network that information should be gathered and assessed. Each approach has its pros and cons. [ITU-T Y.1542] outlines a framework for achieving end-to-end performance objectives.

IP network architectures, currently being deployed to build IP network segments, do not readily lend themselves to be able to change the impairments consumed on a per flow and/or per destination host basis. Network segments therefore tend to be designed and constructed using relatively static architectures and consequential impairment consumption remains within relatively static bounds by design. For example, a particular access network segment to a customer will consume impairments in much the same way for the same service Class regardless of whether the packet flow destination is in the same geographical area or is very distant.

Underlying path delay does not rapidly change and can be numerically added to accurately derive the likely end-to-end outcome. However, rapidly varying metrics such as delay variation cannot be so easily numerically concatenated. As part of the impairment budget allocation process, therefore, an operationally pragmatic approach to setting budgets must factor in the statistical nature of these metrics while still, as much as reasonably possible, resulting in a high probability that the concatenation of any reasonable number of network sections will still meet the desired end-to-end outcomes with an acceptable level of confidence.

A key consideration for the approach adopted in this Supplement is the requirement that any provider should not be reliant on *a priori* knowledge of the performance of other providers networks before being able to design their own network or prepare a commercial proposal for an end-to-end service. Guidelines are required that enable providers to design their networks in isolation, but at the same time have a high degree of confidence that end-to-end targets will be met when any reasonable concatenation of provider network segments are subsequently required to achieve an end-to-end service.

## 3.7 Proposed allocation approach

This Supplement proposes a static weighted segment allocation of impairment budgets to each provider rather than apportionment or accumulation on a path by path basis. Weighting for reference segments is discussed below.

## 3.8 Assessment reference model

The reference model used is as follows:

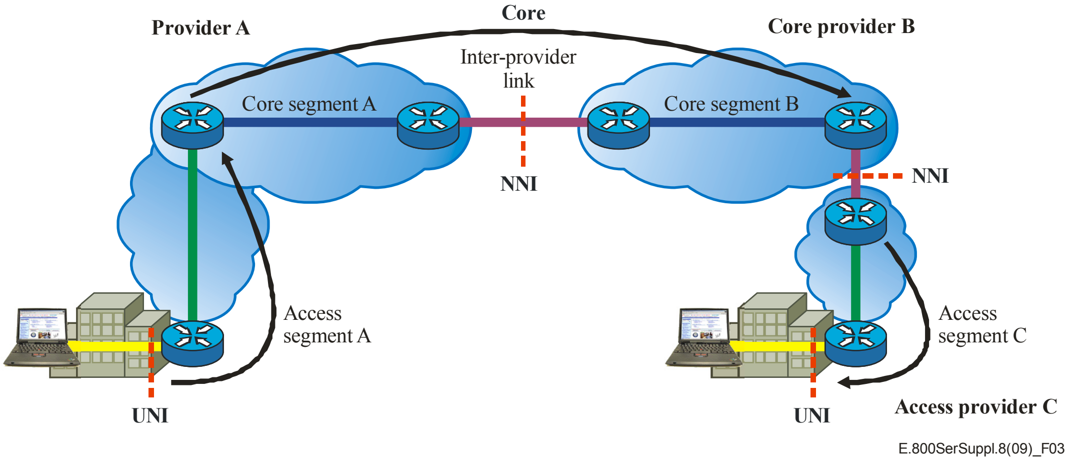

**Figure 3 – Impairment budget reference model**

For the purposes of impairment allocation, the edge of the core providers' segments is the midpoint between their ASBR's, unless otherwise agreed by the interconnecting service providers.

The inter-provider link may need to be dimensioned with greater capacity than might otherwise be required, to ensure its contribution to the consumption of the interconnecting provider's respective budget allocations is not excessive.

### 3.8.1 Access segment

The access segment is that part of an end-to-end IP path from the customer's side of the CE router to the customer's side of the first PE router.

Where the CE router is not managed by the access service provider, the allocation of impairment budgets for the access segment between the CE and the access service shall be mutually agreed between the access provider and the CE router provider (or the customer, if the customer provides its own CE router) and is for further study.

For many networks, the access segment is the network domain where bandwidth per customer is the most cost sensitive. Total bandwidth is therefore limited, and in many cases the peak transmission data rate may be 2 Mbit or less in one or both directions of transmission.

The ability for any access provider to dynamically change the size of the access link to compensate for higher consumption of the impairment budget in a core segment for a given IP path destination is very limited, and, in most cases, not operationally or commercially practical.

Long, low-capacity links using copper- or radio-based technologies are typically subject to higher interference than high speed optical core links. As a consequence, a significant proportion of interference-related impairments such as packet loss needs to be allocated to this part of an IP path to optimize the price performance of the overall outcome.

### 3.8.2 Number of providers

The number of core segment providers in any end-to-end IP path will vary based on both technical and commercial considerations.

It is assumed in this Supplement that use of no more than three concatenated core segments would be a reasonable maximum for any intra continental or national connection. In addition, it is assumed there is one access segment per end, making a total of five concatenated segments. It should be noted, however, that in many practical cases, one access segment and one core segment will be provided by the same service provider, meaning the five concatenated path segments will typically be comprised of no more than three interconnecting service providers.

Metro networks are not treated differently from core networks as far as budget allocation is concerned. However, if the metro network is provided by the same service provider as the adjacent core segment, then they are considered as a single core segment for budget allocation purposes. In the remainder of this clause, "Core" is used to describe all "non-access" provider segments.

For intercontinental connections, an additional two core service providers, for a maximum of seven concatenated segments, is assumed to be the typical maximum, that is, two access segments and up to five "non-access" segments are considered to be the upper bound for most practical cases on intercontinental paths.

For any national (e.g., trans United States or Australia) or regional (e.g., Western Europe or Eastern Asia) service, the end-to-end target is ITU-T Y.1541 Class 0. For intercontinental connections or those between major global regions of North America, Asia-Pacific, Central Asia and Europe, it is assumed the end-to-end impairment targets will be those of ITU-T Y.1541 Class 1, but should be as close to Class 0 targets as practical. In either case, the end-to-end budgets should have a high probability of being met if:

1) all providers consume their maximum impairment budget allocation for normal base-line performance; and

2) the probability of more than one operator simultaneously operating in the upper range of the performance budget for varying impairments is very low, i.e., less than 0.1%.

In practice, some negotiation or signalling of impairments between operators may be employed to ensure the end-to-end budget will be met or exceeded for any individual connection, but a single provider's network planners or sales representatives cannot rely on this when dimensioning or offering a network service. Clause 3.9 proposes an approach to dimensioning and allocating impairment budgets among providers.

### 3.9 Budget allocation

This clause outlines a simple, pragmatic solution for the apportionment of impairments for network planning and operational purposes.

The most complex impairment metric to allocate is delay variation, discussed in clause 3.9.2. Impairment budget allocation methods for IPTD and IPLR are discussed in clause 3.9.3.

Appendix I gives examples of the end-to-end outcomes resulting from this allocation of impairment budgets.

### 3.9.1 Budget allocation for providers offering services between a UNI and NNI

For budget allocation and reporting purposes, where a provider offers services between a UNI and core segment NNI, the allocations for the access segment and one core segment can be combined as a single budget for that provider's segment of the IP path. There should be no requirement for that provider to separately measure and report the performance of the separate access and core segments of the UNI to NNI IP path.

### 3.9.2 Budget allocation for IPDV and IPDV concatenation

The simple arithmetic division of delay variation budget across multiple providers would ordinarily result in a more stringent requirement than is actually required to achieve the end-to-end targets.

The method of determining budget allocation described in this clause is less demanding on individual providers as it acknowledges the statistical nature of delay variation. The solution can also be used for the advanced signalled or accumulation approaches, or it can be further refined to obtain tighter end-to-end performances values. The solution can be extended for more network segments, if needed.

The approach presented here allocates a significant proportion of the IPDV impairment budget to each access segment, with each core segment having a lesser fixed budget. The approach also allocates a fixed IPDV budget for core network segments, irrespective of the number of core network segments in any resulting services.

### 3.9.2.1 The "2-point promise"

For each reporting period, each provider will collate its own distribution of measured IPDV values for each rollup period in that reporting period. In practice, this distribution of rollup period IPDV values can typically be expected to be heavily skewed towards lower "background" IPDV values with a small tail of "above normal" values.

To estimate the resultant IPDV distribution for any end-to-end IP path, with a high degree confidence, each provider would need to determine and report the best-fit distribution function for all IP path segments for the reporting period. This is not considered practical or desirable.

In practice, it is considered adequate for the purpose of operationally managing end-to-end IPDV if each provider committed to a frequency probability of measured IPDV values for the reporting period for three bands, defined by two thresholds. This is referred to as the "2-point promise".

More precisely, each provider commits to deliver a service with the IPDV rollup period values within specified probability bounds for each reporting period. The reporting period is defined in [ITU-T Y.1543] and the recommended default is one calendar month. Note that rollup period IPDV here is understood to mean the 99th percentile of the singleton measurements as discussed in clause 3.4.4.

This approach gives an indication to all participants in the IP path of the shape of the upper tail of the probability distribution curve of IPDV values for each provider's path segment for that reporting period.

### 3.9.2.2 IPDV performance bands

The three IPDV performance bands for a reporting period are indicated in Table 2.

**Table 2 – IPDV performance bands**

| Band name | Band description | Operational impact |
|---|---|---|
| Normal IPDV | Measured IPDV value for rollup period is below lower threshold | Normal conditions – high confidence end-to-end IP path performance can be met |
| Above normal IPDV | Measured IPDV value for rollup period is between lower and upper threshold | Unusual conditions handled by moderate buffers but end-to-end performance is likely to be met, unless several network segments are in this region during the same rollup period |
| High IPDV | Measured IPDV value for rollup period is above upper threshold | Conditions where there is a high risk that the end-to-end IP path performance will not be met |

### 3.9.2.3 IPDV core budget thresholds and probabilities

Table 3 gives an example of possible allocated IPDV budget thresholds and probabilities for core network segments.

**Table 3 – An example of IPDV budget allocations and thresholds for core network segments**

| Budget region | IPDV range | Probability of any rollup period in the reporting period being in this region |
|---|---|---|
| Normal IPDV | 2 ms or less | > 0.99 |
| Above Normal IPDV | Greater than 2 ms but no more than 6 ms | < 0.0099 |
| High IPDV | Greater than 6 ms | $\leq 1 \times 10^{-4}$ (Note) |
| NOTE – For each rollup period, this should be calculated and reported for 12 rolling rollup periods, consisting of the current rollup period and the 11 preceding rollup periods. | | |

### 3.9.2.4 IPDV access budget thresholds and probabilities

Table 4 gives an example of allocated IPDV budget thresholds and probabilities for access network segments.

**Table 4 – An example of IPDV budget allocations and thresholds for access network segments**

| Budget region | IPDV range | Probability of any rollup period with the reporting period being in this region |
|---|---|---|
| Normal IPDV | 16 ms or less | 0.99 |
| Above normal IPDV | Greater than 16 ms but no more than 20 ms | 0.0099 |
| High IPDV | Greater than 20 ms | $\leq 1 \times 10^{-4}$ (Note) |
| NOTE – For each rollup period, this should be calculated and reported for 12 rolling rollup periods, consisting of the current rollup period and the 11 preceding rollup periods. | | |

For access segments with a peak data rate of under 2 Mbit/s, which are also used to carry best-effort traffic on the same access link as the Class 0 traffic, packet fragmentation techniques may need to be employed to enable the delay variation target to be achieved. This is to avoid a Class 0 packet getting "stuck" behind a large Class 5 packet.

### 3.9.2.5 Estimation of IP path IPDV

Once the 2-point promise is thus specified, the probability of end-to-end IPDV being less than the specified target is approximated as the probability of seeing a combination of "normal" and "above normal" intervals such that the sum of the maximum IPDV thresholds specified in the corresponding 2-point promise is less than the end-to-end target. The intuition behind this is that if end-to-end traffic encounters three "normal IPDV" core segments (with the IPDV threshold of no more than 2 ms each) and two "above normal IPDV" core segments (with the IPDV threshold of no more than 6 ms each), then the end-to-end IPDV across the maximum expected five concatenated core network segments will have a high probability of being no more than 3 × 2 + 2 × 6 = 18 ms.

For example, if one is interested in approximating the probability of end-to-end IPDV exceeding the target across five core segments, each one of those declaring the 2-point promise as specified above for the core (or metro) segments, then the following computation can be performed:

$$\mathrm{IPDV}_{\mathrm{bound}}$$

$$= (0.99)^5 + \binom{5}{4}(0.99)^4 * 0.0099 + \binom{5}{3}(0.99)^3 * (0.0099)^2 + \binom{5}{2}(0.99)^2 * (0.0099)^3 = 0.9995$$

If any network has a rollup period value of high IPDV, there can be no assurance that any end-to-end IPDV bound is met for that period.

The meaning of this computation is that if each of the five core network segments declare the "2-point promise" with the thresholds as specified above, then the probability of end-to-end IPDV across the concatenation of these five core networks exceeding the desired target is very small (1-0.9995 is < 0.1%).

Note that the above computation does not yield a reliable "theoretical" bound on the end-to-end probability of IPDV. However, in practice it is a very good (and typically conservative) approximation of this probability and is considered operationally pragmatic and adequate for practical management of end-to-end service quality outcomes.

[ITU-T Y.1541] outlines a different method of concatenating delay variation values but acknowledges possible practical operational limitations in the methodology described.

### 3.9.3 Impairment allocation budget for IPTD and IPLR

For allocation of IPLR and IPTD, more straightforward methodologies can be used, since these metrics can be considered to be additive for practical operational purposes, provided IPLR is small. These allocations would allow an end-to-end IPLR of no more than $8.5 \times 10^{-4}$ for connections with five concatenated core networks, which is within the limits for Class 0 and Class 1. (See clause 4.7.2 for discussion of the issues related to reporting IPLR.)

For example, each network section may be allocated IPTD as given in Table 5 below. Core networks less than 1200 km edge to edge are allocated of the same IPTD budget to accommodate practical physical paths in a local geographic area. An additional allowance for propagation delay for long network segments is also provided. Core network segments only need to have knowledge of the distance between their edges when the total distance between the edges of any core network segment (ingress NNI to egress NNI) exceeds an air path (great circle) distance of 1200 km.

**Table 5 – An example of possible impairment budgets for IPTD and IPLR**

| Network section | | IPTD budget | IPLR budget |
|---|---|---|---|
| Access network | | 25 ms | $4 \times 10^{-4}$ |
| Core network | ≤ 1200 km air path | 12 ms | $1 \times 10^{-5}$ |
| | > 1200 km air path | *CoreIPTD* (Note) | $1 \times 10^{-5}$ |

NOTE – For core segments with an air path (great circle) edge to edge distance greater than 1200 km, the following formula would apply:

*CoreIPTD* = max [12,10 + (total segment g distance in km –1200) × $D_f$ × 0.005] ms, where $D_f$ is provisionally set to 1.4

The IPTD budget should be rounded up to the nearest integer number of milliseconds.

This approach requires lowest latency services (ITU-T Y.1541 Class 0) to have no more than three core network segment providers. Typically, no more than this would be expected to be used in most "national" or regional IP paths to achieve lowest latency performance. Intercontinental services may only meet ITU-T Y.1541 Class 1 performance (IPTD relaxed to 400 ms end-to-end) under this approach, unless network segment providers negotiated lower budgets for a service. For these longer path length services, the number of core network segment operators can be greater than three.

## 3.10 Admission management

For Class 5 (best effort) traffic, if congestion occurs at any point in the IP path, network providers will be able to queue and/or discard excess packets without breaching the performance guarantees, because Class 5 traffic impairment budgets are unspecified and typical applications allow for this to occur. No special admission management mechanism is required to enable performance budgets for this class to be met.

However, for Class 0 traffic, if IP packets are discarded or otherwise delayed through queuing at a point of network congestion, then the resulting performance impairments incurred are likely to result in a failure to meet the IP path budget performance for Class 0 traffic. Customers and/or service providers, therefore, need to include different response mechanisms to avoid congestion of Class 0 traffic from occurring.

The total Class 0 traffic that can be permitted to compete for any onward transmission resource anywhere in a network must be limited to an aggregate volume that is less than the maximum resource transmission capacity available on that path. How much less is a matter for individual network designs and is outside the scope of this Supplement; however, to reliably meet the IP path performance targets, service providers must take steps to ensure that available transmission resources are not exceeded for any possible combination of ingress and egress points if customers are unable or unwilling to do this themselves.

If congestion information is communicated back to customers (for example, using Explicit Congestion Notification outlined in [IETF RFC 3168]), then the customers themselves may be able to take suitable action to avoid congestion. However, providers cannot rely on this occurring particularly when service providers choose to seek higher network utilization by not dedicating full contracted Class 0 capacity to every VPN.

To manage congestion, some form of provider managed admission control is therefore likely to be required. DSCP-based per-hop behaviours alone are not adequate to ensure the required end-to-end Class 0 performance will always be achieved. Before Class 0 traffic can be carried by a network provider, the receiving network provider should therefore:

- implement a mechanism within the network, which is aware of the current utilization and capacity to carry Class 0 traffic,
- provide a mechanism for the sending party at the ingress UNI or NNI to request the required Class 0 capacity to the network egress point (NNI or UNI) of interest,
- make a decision based on this request to admit or refuse carriage of the requested Class 0 traffic at the network interface ingress point, and
- communicate this decision back to the sending party.

This, in its simplest form, may be achieved by prior agreement between the service provider and the sending party on ingress Class 0 volume limits, and by the service provider permanently dedicating adequate capacity to the sender at any concentration points in its network in a way that Class 0 congestion is very unlikely to be experienced by the sender. However, to achieve higher network efficiency, providers may need to implement more dynamic forms of admission control.

Specific methods of dynamic admission control for IP traffic is for further study.

## 4 QoS measurement

The monitoring and troubleshooting of inter-provider SLAs require measurement of QoS-related information along the path between customer sites. Some agreement among co-operating providers on common approaches to measurement will simplify the tasks of service monitoring and troubleshooting. This clause lays out the requirements for QoS measurement in the inter-provider context and proposes some practices.

### 4.1 QoS measurement requirements

The measurement methodology, protocol and reporting should be capable of estimating at least the set of QoS metrics defined in clause 3 (one-way delay, one-way loss, one-way delay variation) of packets transmitted between specified measurement points. It should be possible to perform measurements on-demand or on a periodic, ongoing basis.

In this Supplement, all metrics are defined to be one-way. Thus, measurements should also be made one-way. Because this raises some practical challenges (e.g., clock synchronization), there may be occasions where two-way measurements will be made (and one-way metrics may be estimated from the two-way measurements). If this is the case, it should be noted and reported.

Measurement probe packets should traverse as much as possible the same path as customer packets having the same QoS service class. They should also be subject to the same QoS mechanisms in routers along the path, implying that the DSCP value of probe packets should be appropriately set for the QoS class to be measured.

The measurement approach should not significantly impact production traffic, either through excessive link load from measurement probes or as the result of load placed on routers by the measurement processes such as generating and responding to probes.

Measurements are generally made between two points in the network. Any of the points mentioned in clause 2.3 (PE, CE, ASBR) and their associate measurement reference points in clause 3.8 may be useful points for one end of a measurement. The concept of a measurement point of presence (MPoP) is introduced. This is an MPoP which is specifically designated as a suitable endpoint for certain measurements. This concept is discussed in more detail below.

The measurement methodology should not require that providers provide access to measurement points nor exchange measurement data. However, the protocols should support access to measurement points or measurement data between consenting providers for authorized requestors. It should ideally be possible to make PE-PE or CE-CE measurements, even when the PEs or CEs are contained in, or attached to, the networks of different providers. (Note however that large amounts of PE-PE or CE-CE probing raises scalability issues.)

The measurement methodology should specify how the errors in measurements are treated, and how results are processed in terms of any statistical treatment of data.

Finally, the measurement methods and protocol must provide means to limit and detect attempts to tamper with or alter the QoS metric estimates.

### 4.1.1 Service provider measurement agreements

One of the major challenges of inter-provider measurement is that there are so many valid options. This Supplement narrows the options so that measurements made across the networks of multiple providers could be compared and combined to create meaningful and reasonably accurate end-to-end measurements. To that end, the set of things that service providers would need to agree upon in the measurement area are listed.

Service providers should agree upon the metrics defined in clause 3.4. The methodology for measurement of these metrics should define the size of measurement packets, the measurement protocol (e.g., OWAMP), the frequency of tests, and the distribution of probe packets (e.g., uniform or random) in test series. Note that this Supplement proposes values for all these parameters.

It should also be possible to make measurements from within the network of one provider to the ASBR of a neighbouring provider. A provider may also designate an MPoP as a location that has specific capabilities for measurement. In these cases, service providers should agree on the volume of the test traffic that they will generate into each others' networks.

Service providers should publish enough information about the location of measurement devices that are available for customers and/or other service providers to enable customers or other service providers to make rational choices of where to direct their measurement traffic.

Co-operating providers should agree on the clock accuracy they will support. A maximum error of 100 μs for measurement devices in MPoPs, and a maximum error of 1 ms for other measurement devices (e.g., CEs, PEs, or devices co-located with them) is provisionally recommended.

In order to support diagnostics and service conformance tracking, each provider should retain QoS measurement data for some agreed-upon period.

## 4.2 QoS measurement methodologies

Ideally, the measurement methodology would be common among providers; however, this may not be practical in the near to mid-term since a number of measurement methodologies are already in use. In this clause some of the options that exist within the realm of active (i.e., probe-based) measurement are described, as distinct from passive measurement in which the actual data traffic is monitored to gather performance data.

The sources and sinks of probes may be either dedicated measurement devices, routers that are dedicated to measurement tasks, or routers that support both data traffic and measurement probes. The location of measurement points may include:

- each CE or a subset of CEs;
- each PE router or a set of PE routers;
- each P router or a set of P routers.

The measurements may be reported as point-to-point measurements between two measurement points or a matrix of such measurements among various points. It is also possible to report average measurements or other statistics computed over a number of different point-to-point measurements – such statistics clearly become less useful if the measurement points span widely different geographic areas.

When selecting measurement points, the goal is to capture the properties of the paths traversed by real customer traffic as much as possible. In general, it will only be possible to approximate the path of customer traffic with a bounded number of measurement devices. See clause 4.4 for further discussion of this issue.

To enable measurement of QoS parameters across multiple provider networks, one of the following methods could be used:

- Each provider agrees to use a common measurement protocol and to make probe points available to other providers, enabling measurements to be made along the end-to-end path.
- Each provider network uses its own methods and probe devices to collect measurements on a per-provider basis, with these measurements being combined to estimate the concatenated end-to-end performance.

Note that even the latter requires co-operation among interconnected service providers in terms of the protocol and availability of probe points to measure the QoS parameters of the inter-provider links.

## 4.3 QoS measurement protocols

ICMP-based PING measurement of TWPD, TWPL, and instantaneous bidirectional connectivity have historically been used by a number of providers when monitoring networks to deliver QoS-oriented SLAs. Vendor-proprietary measurement protocols have also been developed and used by some providers and end customers. In general, for inter-provider performance testing, open testing protocols should be used.

The IPPM protocol OWAMP [IETF RFC 4656] (or a protocol compatible with it) should be used for one-way measurements, with TWAMP [IETF RFC 5357] as an alternative if two-way measurements are to be used. (Note all measurements should be one-way but measurements may be two-way as long as the distinction is reported.) In addition, there is an ongoing work that would allow the use of a lightweight version of TWAMP for one-way measurements. With this approach, TWAMP and its simple version TWAMP-lite can provide simple but reliable one-way and two-way performance measurements. One of the possible avenues that require further study is the use of OAM multi-hop protocols for inter-provider performance testing. Such an approach could reduce significantly the operational burden of network performance monitoring.

### 4.3.1 OAM-based active measurement

The use of ICMP-based PING as a measurement protocol is not recommended as a reliable protocol for measurement of customer IP path performance. IP network elements often treat ICMP messages quite differently to end-customer traffic, particularly under higher network traffic conditions.

Other lower layer OAM protocols, such as Ethernet OAM, with its performance measurement parameters defined in [ITU-T Y.1731], may be suitable for deducing IP delay and loss performance where Ethernet maintenance entities exist at or near representative IP network segment measurement points. Care needs to be taken to ensure that the measured performance of Ethernet OAM frames is truly representative of IP customer traffic performance over the same network path. Use of Ethernet OAM may be a valid protocol across Ethernet access segments where achieving a large enough sample of paths makes deployment of dedicated IP measurement probes otherwise uneconomic.

## 4.4 Measurement considerations for VPN services

When a VPN service spans the networks of multiple providers, there are additional challenges in providing accurate end-to-end measurements for a given VPN customer. For example, it may be difficult for any one provider to determine the path that is taken by a particular VPN customer's traffic. And even if the path is known, it may be difficult to conduct measurements along that exact path, e.g., due to a lack of devices to respond to measurement probes at various points on the path.

### 4.4.1 Measurement of each service provider's network performance

The goals of the measurement techniques described above, therefore, are more modest than the delivery of precise performance data to a particular VPN customer. Instead, the primary goal is to allow a provider to make certain QoS assurances to a customer, knowing that the impairments that can be expected from other providers in the path, as described in clause 3.9, will enable those assurances to be met if all providers meet their impairment targets. The reported measurements of each provider should indicate when a provider has potentially jeopardized the end-to-end targets.

### 4.4.2 Measurement of individual customer VPNs

Practical measurement methodologies for individual customer VPNs will vary depending on the nature of the specific VPN service configuration and the service commitments offered to the customer. The precision of any measurement will be constrained by the transmission capacity of the VPN service offered and will often need to be conducted end to end (i.e., UNI to UNI). Measurement of individual customer VPN service instances is therefore a matter of negotiation between services provider's and their customers and is outside the scope of this Supplement.

It is expected that the test VPN performance measurements will be able to supplement any less accurate measurement of the specific service VPN instances to enable service providers to demonstrate, with a high confidence, that the performance commitments for that VPN instance is likely to have been achieved.

## 4.5 Recommended measurement approach

It is recommended that, as a minimum, providers establish test VPNs between the reference MPoPs in their network that:

a) are provisioned, as much as practically possible, using the same processes as production customer service VPN instances across their network;

b) traverse the same network elements and use the same (or typical) forwarding policies as used for customer VPN service instances wherever practical;

c) are not in-service VPN instances actually carrying a customer's traffic.

This approach will allow intensive active measurement to be undertaken to achieve the resolution required to reliably observe the network service class performance thresholds without impacting any individual customer's VPN service, whilst maintaining a high probability that the performance observed is likely to be representative of actual customer service VPN instances between the reference network interfaces.

## 4.6 Performance measurement metrics

The following measurement metrics are recommended to be used for network performance measurement (i.e., active measurement between MPoPs over a test VPN).

### 4.6.1 IP delay and delay variation measurement metrics

**Table 6 – IPTD and IPDV recommended measurement metrics**

| Metric | Value |
|---|---|
| Maximum rollup period | 5 minutes |
| Mean probe transmission period | 200 ms |
| IPTD reported unit | milliseconds |
| Minimum reported value | 0 ms |
| Metric accuracy | 1 ms, rounded up |

NOTE 1 – Providers are at liberty to measure more often than every 200 ms and to report that fact.

NOTE 2 – The 99th percentile value, i.e., $IPDV_{(99)}$, is chosen so that a stable value is achievable for the 1500 singleton IPDV values (5x60x5) obtained over the rollup period.

NOTE 3 – At least one IPDV value is recorded for each test rollup period. Other percentiles or singleton values may optionally be recorded for each rollup period where a better estimate of the distribution of delay variation is required by the provider.

NOTE 4 – Where the same test packet stream is used to measure both delay and loss metrics, then the probe frequency for delay will need to meet the minimum recommended loss probe interval

### 4.6.2 IP loss ratio measurement metrics

**Table 7 – IPLR recommended metrics**

| Parameter | Value<br>(access segment) | Value<br>(core segment) |
|---|---|---|
| Maximum rollup period | 5 minutes | 5 minutes |
| Maximum mean probe transmission period | 200 ms | 20 ms |
| IPLR reported unit | Percentage | Percentage |
| Minimum reported IPLR | 0% | 0% |
| Minimum IPLR metric accuracy | 0.1%, rounded up | 0.01% rounded up |

NOTE – The recommended default value of $T_{max}$ = 3 seconds. Providers are at liberty to use values of $T_{max}$ other than 3 seconds, but shall report the value used.

### 4.7 Reporting of measurement results

Service providers need to agree on the reporting methods. At least there should be agreed processes for the exchange of hard copies of the performance results, including the content and format of such reports. [ITU-T Y.2173] provides guidance on this.

It is highly desirable that service providers agree on methods for the electronic exchange of measurement reports. Such an agreement would include both the content of the reports and a protocol for exchange of the reports.

The frequency of reports should be agreed upon. It would also need to be agreed whether reporting of QoS information among providers is a normal, ongoing activity or whether it is only triggered by requests (e.g., to troubleshoot a particular customer problem). Daily exception reporting (as outlined in the following clause) is provisionally recommended between providers to meet network operations and service management requirements.

Reports should contain only aggregated data. Aggregated data should be available at different aggregated levels (by the fraction of an hour, by hour, daily, monthly – depending on the report) and statistics of the aggregates (mean, median, quantiles, number of measurements) should be reported.

The report should include at least:

- start date and time (UTC);
- location of end points;
- measurement/report period (duration and/or finish time);
- measurement type;
- measurement statistics.

#### 4.7.1 Proposal for reporting of measurement results

This clause defines one method to report the set of QoS metrics defined in clause 3 (IPLR, IPTD, IPDV) of packets transmitted between specified measurement points. As for the rest of this Supplement, VPN provider interconnection is the primary focus, but the intent is that the reporting is applicable in the broader public network context as well.

As noted above, statically dividing impairment budgets among the participating networks is recommended so that the budget per network segment can be better designed and managed. Three steps below are recommended to ensure that each provider can formulate an attractive end-to-end SLA and also have the information necessary to troubleshoot for a VPN customer across multiple providers.

Each operator should measure the following metrics as defined in clause 3.4 for each rollup period:

- loss (IPLR);
- mean delay (IPTD);
- delay variation (IPDV).

As noted above, individual measurements will not be reported, but the appropriate rollup period statistics (loss ratio, mean for IPTD, and 99th percentile for IPDV) may be reported, as described below.

#### 4.7.2 Monitoring and comparison to threshold

For each metric, the rollup period measurements are monitored and compared with the threshold values suggested in Tables 3 and 4. From a practical point of view, it is an advantage to report (or act upon) as little as possible. Thus, not all quantities need to be reported (or acted upon) for all classes.

**Table 8 – Thresholds for reporting between providers**

| Result of rollup period measurement | Class 0 report |
|---|---|
| IPDV > 2 ms (core) | Report value |
| IPDV > 6 ms (core) | Report value |
| IPDV > 16 ms (access) | Report value |
| IPDV > 20 ms (access) | Report value |
| IPLR > $10^{-5}$ (core) | Report value |
| IPLR > 4 × $10^{-4}$ (access) | Report value |
| PLR > $10^{-2}$ | Report unavailable |
| IPTD > 25 ms (access) | Report value |
| IPTD > 12 ms or *coreIPTD* (core) | Report value |

Providers are only required to report (or act) when during a certain rollup period a measured value is above the threshold. The suggested threshold values are set with such a margin that they would not normally need to be reported (or acted upon), but still not so large enough that the E2E budget is in danger when measured values are below this threshold. In practice, reporting only occurs when a link or a cluster is very highly loaded and thus has problems with the achieved QoS levels.

This reporting philosophy is for further discussion.

### 4.7.3 Reporting the measurement results

With the threshold above, providers need only to report to each other when something unusual occurs. The assumption is that providers do not try to cheat each other on purpose, but rather providers report events that might endanger the end-to-end performance objectives, so that the owner of any end-to-end service contract can troubleshoot and constructively resolve problems.

The provider should specify the QoS problems for a relevant part of its domain in each case and not just report "problems anywhere in the domain" quite frequently. The provider should not report more to any provider than what is relevant to him. The reports could potentially therefore specify the VPNs which a certain measurement with a bad value will affect.

## 4.8 QoS measurement security considerations

Security is discussed in some detail in clause 6. Some specific security issues related to measurement involve the authentication of access and protecting the integrity of data. In particular:

- Integrity of measurement reports needs to be protected by standard cryptographic techniques.
- Authentication and access control mechanisms should be used to ensure that measurement reports are only made available to authorized parties.
- Access to measurement probe devices, especially when access is permitted by other providers or customers, needs to be controlled by standard access control mechanisms.

# 5 Routing

While existing routing may be sufficient in some inter-provider QoS deployment scenarios, it may also be desirable to select among multiple inter-domain paths based on the QoS requirements of different classes of traffic. That is, there may be cases in which the current route selection capabilities of BGP, which yield only a single best path for a given prefix, may not be sufficient.

Extending BGP to support QoS-aware routing inherently implies increasing the amount of information carried in BGP. This could have some implications for the convergence and scaling of BGP, at least in principle. Moreover, in order to maximize the stability of inter-domain routing in very large VPNs (and public networks), it is highly desirable that the QoS-related information that is to be advertised into BGP be stable (in terms of not changing rapidly over time). These issues should be taken into consideration if BGP is extended to carry QoS-aware information.

## 5.1 Current BGP capabilities

BGP is good at passing end-to-end routing reachability between two peers. There are no additional semantics, of which the protocol is aware, that are carried in the update messages. All additional semantics attached to a prefix are opaque to the protocol (e.g., extended communities) and have local semantics.

BGP is not a suitable protocol for passing rapidly changing path characteristics (delay, delay variation, etc.) as the protocol is based on a distance vector architecture and not one that floods data or has full network topology awareness.

BGP is capable of carrying multiple classes of routing information through its AFI/SAFI hierarchy. QoS class or service context could be considered as a class of routes and BGP could simply announce reachability and service/QoS classes would be passed along in an opaque manner. If, as this Supplement proposes, there is a very small, bounded number of service classes that are infrequently changing, use of BGP in this way should be tractable.

## 5.2 BGP considerations

There are a few issues that need to be resolved with respect to the BGP protocol architecture before use of BGP for multi-class interconnection across inter-provider boundaries would work perfectly.

BGP has no way to carry multiple routes to the same destination. The protocol is based on "implicit withdraw" semantics. This means that every new announcement of a prefix causes any other announcement of the same prefix to be "withdrawn" or be no longer reachable. Thus, announcing a prefix multiple times (e.g., once per QoS class) may not work well.

BGP in most current implementations is based upon multiplexing all AFI/SAFI onto one BGP peering session, which implies shared fate in the state of the peering session. An error in one AFI/SAFI update message causes all prefixes in all AFI/SAFIs to be purged. Due to this multiplexing, it is also impossible to prioritize the convergence of the prefixes associated with one service, AFI or SAFI upon reception of a new update. All are treated equally in a "first in, first converged" manner.

BGP is limited in its ability to distinguish NLRI ("prefixes") associated with different services (e.g., different QoS classes). A possible future enhancement to BGP to address this for interconnection may be to provide BGP with the means to mark address families (AFIs, SAFIs) and prefixes via a simple, opaque (to BGP) marking, to associate them with a "service context" (e.g., QoS class).

There are other additional features that are needed to build an inter-domain system for service separation that can enable revenue generating service level agreements. They include:

- BGP peering session separation;
- passing of redundant or backup routes;
- fast failure notification propagation;
- the ability to have 'service topologies' or network overlays and pass 'context' information within the new hierarchy.

### 5.3 Routing recommendation

Specific solutions to the routing issues outlined are for further study. A potentially useful solution is an enhancement to BGP under consideration by the IETF referred to as "multi-session BGP".

## 6 Securing QoS

### 6.1 Motivation

In order to provide high quality service to specific customers, it is necessary to secure the network infrastructure as well as the use and provisioning of the service. What to secure and how to secure it depends on what is done and how it is done (i.e., how the network is operated and what services are offered). For example, if all signalling and provisioning is done via manual configuration, then securing the network may be limited to securing the protocols used for configuration, as well as maintaining an audit trail of operator actions (e.g., to protect against insider attacks). Thus, this clause is more a set of considerations to be taken into account.

### 6.2 Areas which need to be secure

There are multiple areas that need to be secured, including:

1) Securing the network infrastructure to ensure high availability of the network.

2) Securing the customer site.

3) Securing the use of preferential services.

The first two of these are critical to ensure that services are available and operate correctly, but are outside of the scope of this Supplement. Methods for securing the network infrastructure are, for example, being worked on in the IETF OPSEC working group (Operational Security Capabilities for IP Network Infrastructure; see http://www.ietf.org/html.charters/opsec-charter.html) and the Secure Inter-Domain Routing (SIDR) working group (see http://www.ietf.org/dyn/wg/charter/sidr-charter.html). Methods for securing a customer site are not currently the subject of standards efforts, but are the purpose of a variety of products such as firewalls and intrusion detection and/or prevention devices. A survey of current practices for securing service provider networks can be found in [IETF RFC 4778]. A survey of standards efforts related to network security can be found in [SecurityEfforts]. A set of best practices for cyber security and physical security can be found at www.nric.org, by clicking on "NRIC Best Practices", and then searching on the keyword "Cyber Security" or "Physical Security", respectively.

The set of practices and guidelines for network security is constantly changing and evolving. Network operators should constantly be reviewing them and altering their procedures and practices accordingly.

Another general security issue is the design of protocols and the implementation of the protocols in software and hardware. This issue is also beyond the scope of this Supplement.

There are two broad areas of security that apply to IP-QoS: i) Provisioning security; and ii) Service security. Provisioning is the mechanism by which services are created and managed. Provisioning security is how those mechanisms are protected against attack. A service is some kind of ToS which is available to a subset of customers (and their packets) in a network. Service security protects that service.

### 6.3 Provisioning security

The goal of "Provisioning Security" is to secure the protocol aspects of the provisioning system, that is, the transfer of provisioning information between network elements. Provisioning information includes, but is not limited to:

– QoS parameters such as bandwidth and latency; and

– traffic signatures, such as DSCP.

Routers, switches, network management stations, and end nodes all comprise network elements.

A service provider should also secure its network management elements and provisioning data (configuration files, audit trails, logs, and so on). If an NMS or configuration data are compromised, then the attacker can alter the ToS provisioning. If audit trails and logs are compromised, usage and billing data could be lost. Securing these elements is the same as general end-system and data-file security and, as such, is beyond the scope of this Supplement.

There are also manual activities with regard to provisioning (business development people negotiating to create an IP-QoS, operators cooperating to implement and debug it, and so on). These activities can be vulnerable to attack and therefore must be secured, but discussion of these attacks and security mechanisms is beyond the scope of this Supplement.

Details of security (e.g., protocols and algorithms) are dependent on the exact protocols, algorithms, and procedures that provisioning uses. As such, these details are beyond the scope of this Supplement. Instead, the requirements of security are considered, outlining possible vulnerabilities, threats and attacks.

### 6.3.1 Goals

There are three goals of provisioning security:

**Protection against unauthorized or inappropriate provisioning** – Attackers and other unauthorized parties must not be allowed to install services in a provider's network. They must also be prevented from altering, deleting, or otherwise reconfiguring existing services. A primary technique is to use cryptographically strong authentication.

**Protection against DoS attack –** Attackers and other unauthorized parties must be prevented from attacking the provisioning protocols in ways that prevent legitimate provisioning protocol operations from being performed.

**Non-repudiation of provisioning requests –** Insofar as provisioning represents a business relationship between two providers, with concomitant financial considerations, it is necessary that provisioning operations cannot be repudiated. That is, if Bob sends a valid provisioning protocol operation to Alice, Bob must not be able to deny that he sent the operation.

### 6.3.2 Attacks

There are a number of attacks to which protocols in general are susceptible [IETF RFC 3552]:

- eavesdropping;
- replay;
- message insertion;
- deletion;
- modification;
- denial of service.

It may be argued that a particular attack is not of concern because the protocols in question will be used only in a way that obviates that attack, or the underlying network technology is such that the attack cannot happen. This view is not supported. Protocol use and network topology have consistently evolved in ways that were quite unforeseen by the original designers.

The following subclauses contain comments on each of the attacks.

#### 6.3.2.1 Eavesdropping

Protection against eavesdropping is not necessary for safe operation of IP-QoS. It may be necessary or desired in order to prevent commercially sensitive information from being disclosed to a third party.

This non-requirement presumes that the provisioning protocols do not do things like carry clear-text passwords.

#### 6.3.2.2 Replay

A replay attack is one where the attacker makes a copy of packets on the network and then retransmits them. Provisioning protocols must be safe from this attack.

#### 6.3.2.3 Message insertion

A message insertion attack occurs when an attacker creates a new message (or messages) and transmits it to the target. The provisioning system should protect against this as it could be used to send messages that alter or destroy existing services, or create new (unauthorized) ones.

#### 6.3.2.4 Deletion

Message deletion attacks occur when the attacker prevents the proper reception of a message. Most good protocols are not very susceptible to this attack as the deleted message would appear as if the network lost the packet for other ("good") reasons. Well designed protocols will detect lost messages and retransmit them. If subsequent packets continue to be lost, then a failure of the communication channel will be detected and brought to the attention of network operators.

#### 6.3.2.5 Modification

If an attacker can intercept, alter, and retransmit a message, then it is a modification attack. These attacks can be used to alter a provisioning request. Provisioning protocols should protect against this form of attack.

#### 6.3.2.6 Denial of service

Denial of service attacks in this context refers to attacks against the provisioning system that prevent the provisioning system from working. These attacks can take a couple of forms:

**Flooding**: Flooding DoS attacks work by simply sending so much traffic to the target that it spends so much time, memory, and so on, receiving, queuing, processing, and discarding the traffic that it has no resources left to process good traffic.

**Algorithmic**: These attacks utilize a weakness or vulnerability in the provisioning protocols (such as the TCP Timestamp vulnerability [CERT637934]).

A particularly insidious DoS attack can occur if the protocol uses cryptographic techniques to secure the packets. Cryptographic algorithms typically require significant amounts of resources. Thus, an attacker could overload a router's processor by sending a relatively moderate number of packets, each of which consumes a fairly large amount of resources to discard. The target could spend all of its time evaluating and discarding these packets. All other services provided by that target would then be effectively disabled. This attack can even occur indirectly. If some other protocol is attacked in this manner (e.g., BGP with MD5 authentication), in some cases there might not be enough resources available to process provisioning protocol messages.

Some provisioning protocols make use of Soft State that needs to be periodically refreshed. If the refresh does not happen, the state is discarded (and thereby, the IP-QoS). An attacker can prevent that refresh. It could overload queues or the processor in the target. It could also prevent the refresh packets from reaching the target (e.g., by corrupting them in the network).

### 6.3.3 Security of provider-provisioned CE devices

Where the service provider manages CE-based devices, the service provider cannot ensure the physical security of the CE device. This leads to the possibility that a physical breach of security could occur at the customer site, leading to a possible misconfiguration of the CE device (for example, if a hacker were to obtain access to the console port of a CE router). The CE device therefore cannot be trusted.

### 6.3.4 Carrier of carriers issues

In some cases, a service provider may make use of services provided by a different service provider in order to interconnect its network. This is common in at least two situations: i) where the carrier of carriers service is used to interconnect backbone routers in a service provider; ii) where the carrier of carriers service is used to interconnect a customer site with a service provider's network. In this case, the data plane and control plane may both be extended across the carrier of carrier's service.

In many cases, the carrier of carrier's service may be provided through use of virtual private network services (for example see [IETF RFC 4364]). Security issues with VPN approaches are discussed in [IETF RFC 4111],VPN Security Framework.

## 6.4 Service security

"Service Security" means protecting the service itself from attack, abuse, and misuse. It is essential to protect the network from unauthorized use of premium services. For example, unauthorized use has the potential of defeating the provisioning efforts that are necessary for ensuring premium services.

As discussed in clauses 3.2 and 3.3, packets should be marked correctly when crossing trust boundaries (CPI or NNI) in order to receive the appropriate service. Routers must therefore be able to examine packets and determine whether they are requesting a particular service or not (and if so, which one) without significant performance degradation. If they cannot do so, then the service is subject to attack by simply flooding a router with too much traffic for it to examine.

Policing is also discussed in clauses 3.2 and 3.3. The policing tests should be low-cost. If policing is too expensive (i.e., causes significant performance degradation), then it is possible to attack the policer by flooding it with packets.

A service provider cannot trust that a peer service provider has adequate security. Thus, service security measures must be provided on inter-provider links.

## 6.5 Security guidelines

This clause provides a brief list of procedures and practices that network operators should follow:

- Be in contact with, understand, and constantly review all available security practices, guidelines, alerts and other pertinent information. The nature of security threats and the methods for dealing with them is constantly changing. Network operators should constantly adapt their own security procedures.

  Good sources of security information include CERT, NRIC, the IETF and NANOG.

  Operators should also review all security-related announcements and information available from their equipment vendors. Security patches should be installed as soon as practical.

- Do not rely on clear-text passwords and the like. Assume that all network traffic is subject to sniffing and analysis. Cryptographically strong algorithms should be enabled and used. This is critical for network management protocols and service provisioning protocols.

Whenever packets/messages/operations fail, the failures should be counted and logged. Security personnel should be notified and take appropriate actions. One should never ignore a small violation as "one of those things". Large attacks start as small probes.

- Do not trust customer networks. You cannot assume that the customer's security practices are good. The customer could easily generate excessive traffic for a particular service, even if the customer's CE device is provisioned and/or managed by the provider. Since the device is not under the physical control of the provider, it can be reconfigured or be otherwise compromised.
- Do not trust peer networks. Just as a customer's net can be compromised, so too a peer provider's network can be compromised. Security practices which are deployed on links facing customers must also be deployed on links facing other providers.
- Filter and drop traffic that comes from a place where it should not. If a peer or customer is not supposed to be sending you traffic for a particular service, do not accept packets from that peer or customer that requested the service. This might just be a routing or configuration issue on the part of the peer or customer, but it could also be an attack.

  This is especially critical for management and provisioning protocol traffic.
- Filter and rate-limit ingress traffic. The best mechanism to ensure that a service is not attacked is to detect all packets that are to get that service and rate-limit them at the point they enter the network. Packets which are in violation of this limit may either be dropped or re-marked as nonconforming or "not to receive the service". Which mechanism to use depends on the business agreements and the service being requested.

  Selecting the rate at which the traffic is limited is complex. Factors include contractual obligations and available network resources. From a security perspective, it is assumed that the network resources are available to meet the contractual obligations. Therefore, the rate limit should be no higher than the contractual obligation. This prevents someone from using "more than they should".

  Traffic that is not to receive the service also should be rate-limited. If the non-QoS traffic is "too much", it could constitute a denial of service attack.
- Read, understand, and apply the practices in [IETF RFC 4778]. If you do not apply one of these practices, you should understand the practice, understand the vulnerabilities (if any) that you will create by not applying the practice, and have a good reason for doing so.

  Keep up to date with this document as it is revised.
- Read, understand, and apply the practices in [SecurityEfforts].

  Keep up to date with this document as it is revised.
- Read, understand, and apply the practices in [IETF RFC 3871]. This document spells out a number of practices and requirements for operators and network equipment. You should understand the extent to which any device you have deployed either meets the requirements or why it does not (understanding that there is no perfect device and that tradeoffs are needed).

## 7 Operational issues

The advent of interconnections where undertakings are made to deliver traffic with a specified quality brings new operational challenges. These are related to the operation of the differentiated services enabled interconnections, to QoS-related capabilities such as timely re-routing of traffic across domain borders, or to functions supporting the business relationship of the interconnecting parties such as accounting functions.

This clause is structured according to the FCAPS model. Some of the FCAPS topics central to inter-provider QoS have been covered already in other clauses:

- Performance monitoring has been given extensive coverage in the measurement clauses.
- Policing, scheduling and dimensioning have been covered in the service class definition clause.
- Fast re-routing is covered as part of the routing clause.
- Security issues are covered in the "Securing QoS" clause.

### 7.1 Fault

Fault management is not specific to inter-provider QoS but the requirements on timely fault detection and service restoration are more stringent as a consequence of the QoS guarantees. This means that fault detection and notification mechanisms and performance used between interconnected parties both in the control plane level and network management level should be agreed on as part of the service commitment. This is valid for both the NNI and the CPI.

Fault isolation and troubleshooting may require a coordinated effort by the providers involved. To make the process efficient, some prior agreement on the responsibilities of the providers regarding notification, troubleshooting and sharing trouble shooting information should be made.

The basic assumption is that each provider is responsible for troubleshooting his own domain. Therefore, it should not be a requirement for a provider to react to active probes (e.g., traceroute and ping) other than on the PE and ASBR nodes (although, as noted above, this capability may be made available selectively with appropriate authorization).

In the event of lost connectivity, service availability will depend on the efficiency of re-routing traffic. Each provider is responsible for re-routing the traffic within his domain and slow convergence will impact the service contract. This means that there is a direct connection between the requirement on fast re-routing of traffic and the formulation of the service performance metrics. Note that there is no need for any exchange of information on internal routing protocol re-routing performance.

In the case of service-affecting faults, it is considered good practice to notify customers of the expected duration of problems. This should be done via the same channels as notification of service windows.

### 7.2 Configuration and maintenance

Due to the higher demands on performance (or specified availability), there will be a need for correlating configuration events that might affect service performance.

Regarding configuration work on inter-provider interfaces (NNI and sometimes CPI), there should be a common change process that minimizes the affect on customer traffic due to bad correlation. This process includes approval, planning and scheduling the work to be done while still allowing for urgent corrective action to be performed.

To allow for service affecting management activities to be performed on networks with a minimum of customer impact, it is customary to define service windows when degradation or loss of service is accepted as being within the limits of the SLA. Due to the global scope and the number of different administrations that may be involved in the inter-provider QoS case, it is not possible to schedule a regular service window that is suitable to everyone. As a consequence of this, a provider that wishes to utilize a service window should notify all partners and customers ahead to give forewarning and to make sure that the intent of the definition of a service window is not abused.

Other providers may wish to take action as a consequence of the activation of a service window. This could be to notify their customers, rescheduling of some activities or to take precautionary action. To allow for efficient processes to be implemented, the length of the notification period and other constraints such as the frequency and length of the service windows allowed need to be generally agreed upon between providers.

If a provider needs to perform urgent service affecting management, it is considered good practice to give notification as early as possible, even though this does not validate a service window.

Providers of the real-time network class are expected to need similar maintenance periods as other providers. That is, every provider will have both planned and unplanned maintenance periods. Since industry practice does not consider planned maintenance outage as unavailability, planned maintenance periods should be considered separately. Unplanned maintenance should be considered as a component of unavailability.

In the case of a single provider, network performance objectives need not be met during planned maintenance. The service contract should make the hours clear, and whether notification of a customer affecting activity is required, how much notice, etc. Providers may try to plan maintenance for local low usage periods, say 2 am – 4 am local time.

Extending SLAs to multiple providers is more complex. How can a customer-facing provider inform a customer of maintenance periods for traffic having a multitude of destinations which wind their way through multiple providers – each of which have their own planned maintenance periods? For global traffic, what is the likelihood of traffic crossing a provider who means well by doing planned maintenance during the "graveyard shift" when that traffic impacted may be for a customer's "busy hour"? It would be beneficial if planned maintenance notification could be extended to network partners, as well as customers, but how much value real or perceived is there for future or long lived sessions to have this foresight?

Inter-provider maintenance windows could be defined per path as the super set of all individual windows, providing that the result is acceptable to the customer. How windows match could be a key criterion to decide over which providers a path is routed. If end-to-end maintenance through a particular set of providers is unacceptable, an alternate set might be found.

A non-signalled static approach could only be statistical, possibly based upon heuristics, though this seems unlikely to satisfy customers.

Global agreement concerning a specific absolute time for when planned maintenance occurs is clearly impractical. However, there may be practical methods to coordinate within constraints. A notification scheme that is communicated to all potential affected parties seems to be the most practical and satisfactory. Providing the notification period and procedure were complied with, the planned maintenance could proceed. Any provider that has customers that were likely to be unreasonably impacted by another provider's planned outage would have the right to negotiate changes to the requested window. In this case, any changes agreed should still be communicated in accordance with the notification period and procedures to all other affected providers. This regime would require all notification requests to be cascaded through providers as one provider may not know what is used beyond the adjacent provider's network.

Current industry practice is for the communication of "planned maintenance" via electronic text (email). The format of the notice and its contents need to be well defined to avoid any misunderstanding. No current industry standards have been identified for this. The format is a matter for possible future study. Planned maintenance periods could be signalled during session set-up, during sessions, and/or indicated along with measurement exchanges, via a database using a standardized message structure.

A minimum notice of 15 days is recommended for service affecting planned maintenance, unless it is otherwise agreed upon by all affected parties. For urgent work, it is good practice, and the practice is encouraged, to give as much advanced notice as practically possible that a service impact is about to occur. Where outage notification is less than the recommended 15 days, then it is at the discretion of the affected parties as to whether the outage is accounted as unavailability or "planned maintenance" for service reporting purposes. Provided the notice period (15 days) is adhered to, the notification would be accepted by all parties unless there were exceptional circumstances. During both planned maintenance periods and periods of unavailability, the predicted resumption of service should be indicated to partners using the same communication channels.

Issues, practices and potential solutions to this maintenance window aspect of inter-provider QoS is an area for further study.

## 7.3 Accounting

New settlement models are expected to emerge in NGN networks where traffic with service quality guarantees is carried end to end across two or more providers. What provider settlement models might be appropriate for inter-provider QoS traffic is outside the scope of this Supplement.

It is noted that most NNI links carry aggregated traffic from many end customers and do not readily allow traffic from specific customers to be identified. (Option A VPN interconnects may be the exception to this). Thus it seems unlikely that accounting on a per-end-customer basis can be practically achieved in most cases.

In the context of this Supplement, it is a reasonable basic assumption that a receiving provider, promising to deliver IP traffic with a defined quality of service, performs a service to the sending provider. In this context, the sending provider and receiving provider are the two providers directly interconnecting via a NNI link, irrespective or what other providers may be participating in any overall end to end customer service.

Based on the above, it is the receiver of the traffic that will be responsible for measuring the IP traffic volume received from each interconnecting provider. The default unit of measure is GigaBytes but other volume metrics may be agreed between interconnecting parties. This measurement may be used for service agreement management purposes, and may also be used directly or indirectly to determine settlement compensation under any agreed commercial interconnection arrangement. In order for the sending provider to verify the measurements (if needed), this should be done using a well known and well specified method, e.g., standard interface counters that may be applied both on the outgoing interface and the incoming interface on the NNI link.

It is recommended that the received volume count on each NNI link should be reported by the receiving provider to the sending provider when requested by that sending provider, unless otherwise agreed. The time duration over which counts are made in any volume report is for further study, as is the frequency of exchanging usage volume reports between providers.

It is recommended that the volume measurement period is defined by stating the time at the commencement, the time at the end and the duration of the measurement period using UTC as the default time standard. Either the end of the period or the duration may be omitted but the beginning of the period shall always be reported. Where the time is not reported using UTC, the time standard used for the report should be clearly stated in the report.

If received packets are discarded by the receiving provider, as may occur as a result of a breach in the service agreement by the sending provider or as otherwise permitted under the interconnection service agreement, then whether this discarded traffic is included or excluded in the volume count for any class is for further study.

## 7.4 Performance

On the NNI links, there will be a need to agree upon how utilization is to be measured and the upgrading rules and process to use. In some cases, there will be a clear customer-provider relationship where the customer will have the responsibility to upgrade. In other cases (when there is a peer relationship), the need for upgrading might not coincide completely and must therefore be otherwise agreed.

It will be common practice to set up a number of interconnection points between two providers. These will be used as backup paths for each other. A provider might also wish to utilize several downstream providers in order to ensure high availability. A provider might choose to try to spread the utilization over the different paths or may prefer a certain path due to, e.g., delay or cost reasons. This means that the network split of the load in case of failures cannot be assumed to be known. To ensure that dimensioning of the networks (both inter-provider links and the networks in general) is based on the correct information, the backup requirements (and possibly re-routing policy?) should be agreed upon between interfacing providers.

## 8 References

[ITU-T Y.1540] Recommendation ITU-T Y.1540 (2007), *Internet protocol data communication service – IP packet transfer and availability performance parameters*.

[ITU-T Y.1541] Recommendation ITU-T Y.1541 (2006), *Network performance objectives for IP-based services*.

[ITU-T Y.1542] Recommendation ITU-T Y.1542 (2006), *Framework for achieving end-to-end IP performance objectives*.

[ITU-T Y.1543] Recommendation ITU-T Y.1543 (2007), *Measurements in IP networks for inter-domain performance assessment*.

[ITU-T Y.1731] Recommendation ITU-T Y.1731 (2008), *OAM functions and mechanisms for Ethernet based networks*.

[ITU-T Y.2173] Recommendation ITU-T Y.2173 (2008), *Management of performance measurement for NGN*.

[IETF RFC 4778] IETF RFC 4778 (2007), *Current Operational Security Practices in Internet Service Provider Environments*.

[IETF RFC 2330] IETF RFC 2330 (1998), *Framework for IP Performance Metrics*.

[IETF RFC 2385] IETF RFC 2385 (1998), *Protection of BGP Sessions via the TCP MD5 Signature Option*.

[IETF RFC 2475] IETF RFC 2475 (1998), *An Architecture for Differentiated Services*.

[IETF RFC 2547] IETF RFC 2547 (1999), *BGP/MPLS VPNs*.

[IETF RFC 2678] IETF RFC 2678 (1999), *IPPM Metrics for Measuring Connectivity*.

[IETF RFC 2679] IETF RFC 2679 (1999), *A One-way Delay Metric for IPPM*.

[IETF RFC 2680] IETF RFC 2680 (1999), *A One-way Packet Loss Metric for IPPM*.

[IETF RFC 2681] IETF RFC 2681 (1999), *A Round-trip Delay Metric for IPPM*.

[IETF RFC 3086] IETF RFC 3086 (2001), *Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification*.

[IETF RFC 3168] IETF RFC 3168 (2001), *The Addition of Explicit Congestion Notification (ECN) to IP*.

[IETF RFC 3246] IETF RFC 3246 (2002), *An Expedited Forwarding PHB (Per-Hop Behavior)*.

[IETF RFC 3393] IETF RFC 3393 (2002), *IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)*.

[IETF RFC 3552] IETF RFC 3552 (2003), *Guidelines for Writing RFC Text on Security Considerations*.

[IETF RFC 3871] IETF RFC 3871 (2004), *Operational Security Requirements for Large Internet Service Provider (ISP) IP Network Infrastructure*.

[IETF RFC 4111] IETF RFC 4111 (2005), *Security Framework for Provider-Provisioned Virtual Private Networks (PPVPNs)*.

[IETF RFC 4271] IETF RFC 4271 (2006), *A Border Gateway Protocol 4 (BGP-4)*.

[IETF RFC 4364] IETF RFC 4364 (2006), *BGP/MPLS IP Virtual Private Networks (VPNs)*.

[IETF RFC 4594] IETF RFC 4594 (2006), *Configuration Guidelines for DiffServ Service Classes*.

[IETF RFC 4656] IETF RFC 4656 (2006), *A One-way Active Measurement Protocol (OWAMP)*.

[IETF RFC 4778] IETF RFC 4778 (2007), *Current Operational Security Practices in Internet Service Provider Environments*.

[IETF RFC 5357] IETF RFC 5357 (2008), *A Two-way Active Measurement Protocol (TWAMP)*.

[AGG] Chan, K., Babiarz, J., and Baker, F. *Aggregation of DiffServ Service Classes,* Work in Progress, draft-chan-tsvwg-diffserv-class-aggr-03.txt, January 2006.

[CERT637934] CERT Vulnerability Note VU#637934, *TCP does not adequately validate segments before updating timestamp value*, <http://www.kb.cert.org/vuls/id/637934>.

[Multisession BGP] draft-ietf-idr-bgp-multisession-04.txt, Chandra Appanna, John G. Scudder, July 2009.

[SecurityEfforts] Lonvick C., and Spak, D. *Security Best Practices Efforts and Documents*, draft-ietf-opsec-efforts-10.txt, April 2009.

# Appendix I

## Examples of the application of budget allocations

(This appendix does not form an integral part of this Recommendation)

In this appendix, the worst case scenarios that may result are considered. These occur when all participants in an end-to-end IP path use their maximum impairment allocations. This situation will be rare in actual networks and real network elements cannot be that precisely configured.

Note that the allocation of IPDV in these examples uses the "low IPDV" thresholds from clause 3.9, and the arithmetic sum of those thresholds is shown just for illustrative purposes. Refer to clause 3.9 for more complete details of IPDV allocation.

### I.1 Case 1: Three core providers

For this example, it is assumed the total air path distance is 4000 km (e.g., Trans USA), and there are three core segment operators involved in the end-to-end connection. Additionally, it is assumed Provider A offers an integrated access and first core segment service for the connection, so a single aggregate budget for Provider A applies

| | **Core segment link air path distance** | **IPTD budget (base)** | **IPTD for long sections** | **Total IPTD** | **IPDV (low threshold)** | **IPLR** |
|---|---|---|---|---|---|---|
| **Provider A** Access plus first core | 300 km | 37 ms | 0 | 37 ms | 18 ms | $4.1 \times 10^{-4}$ |
| **Provider B** core | 3000 km | 0 ms | 22.6 ms | 23 ms | 2 | $1 \times 10^{-5}$ |
| **Provider C** core | 700 km | 12 ms | 0 | 12 ms | 2 | $1 \times 10^{-5}$ |
| Access provider 2 | | 25 ms | | 25 ms | 16 ms | $4 \times 10^{-4}$ |
| **Total CE to CE** | **4000 km** | | | **97 ms** | **< 38 ms** | $\mathbf{8.3 \times 10^{-4}}$ |

Note that this meets the UNI to UNI targets for [ITU-T Y.1541] Class 0.

### I.2 Case 2: Transcontinental service, five core providers

| | **Link air path distance** | **IPTD budget (base)** | **IPTD for long sections** | **Total IPTD** | **IPDV (low threshold)** | **IPLR** |
|---|---|---|---|---|---|---|
| Access provider 1 | | 25 ms | | 25 ms | 16 ms | $4 \times 10^{-4}$ |
| Core provider A | 300 km | 12 ms | 0 ms | 12 ms | 2 ms | $1 \times 10^{-5}$ |
| Core provider B | 3000 km | 0 ms | 22.6 ms | 23 ms | 2 ms | $1 \times 10^{-5}$ |
| Core provider C | 10.000 km | 0 ms | 71.6 ms | 72 ms | 2 ms | $1 \times 10^{-5}$ |
| Core provider D | 2.000 km | 0 ms | 15.6 ms | 16 ms | 2 ms | $1 \times 10^{-5}$ |
| Core provider E | 400 km | 12 ms | 0 ms | 12 ms | 2 ms | $1 \times 10^{-5}$ |
| Access provider 2 | | 25 ms | | 25 ms | 16 ms | $4 \times 10^{-4}$ |
| **Total CE to CE** | **15.700 km** | | | **185 ms** | **< 42 ms** | $\mathbf{8.5 \times 10^{-4}}$ |

Note that this meets the targets for [ITU-T Y.1541] Class 1. Core providers A and E might be considered "metro" providers in this example.

## SERIES OF ITU-T RECOMMENDATIONS

| | |
|---|---|
| Series A | Organization of the work of ITU-T |
| Series D | General tariff principles |
| **Series E** | **Overall network operation, telephone service, service operation and human factors** |
| Series F | Non-telephone telecommunication services |
| Series G | Transmission systems and media, digital systems and networks |
| Series H | Audiovisual and multimedia systems |
| Series I | Integrated services digital network |
| Series J | Cable networks and transmission of television, sound programme and other multimedia signals |
| Series K | Protection against interference |
| Series L | Construction, installation and protection of cables and other elements of outside plant |
| Series M | Telecommunication management, including TMN and network maintenance |
| Series N | Maintenance: international sound programme and television transmission circuits |
| Series O | Specifications of measuring equipment |
| Series P | Terminals and subjective and objective assessment methods |
| Series Q | Switching and signalling |
| Series R | Telegraph transmission |
| Series S | Telegraph services terminal equipment |
| Series T | Terminals for telematic services |
| Series U | Telegraph switching |
| Series V | Data communication over the telephone network |
| Series X | Data networks, open system communications and security |
| Series Y | Global information infrastructure, Internet protocol aspects and next-generation networks |
| Series Z | Languages and general software aspects for telecommunication systems |

Printed in Switzerland
Geneva, 2010