Implementation Agreement
MEF 23.2

Carrier Ethernet Class of Service – Phase 3

August 2016

Disclaimer

The information in this publication is freely available for reproduction and use by any recipient and is believed to be accurate as of its publication date. Such information is subject to change without notice and the MEF Forum (MEF) is not responsible for any errors. The MEF does not assume responsibility to update or correct any information in this publication. No representation or warranty, expressed or implied, is made by the MEF concerning the completeness, accuracy, or applicability of any information contained herein and no liability of any kind shall be assumed by the MEF as a result of reliance upon such information.

The information contained herein is intended to be used without modification by the recipient or user of this document. The MEF is not responsible or liable for any modifications to this document made by any other party.

The receipt or any use of this document or its contents does not in any way create, by implication or otherwise:

- any express or implied license or right to or under any patent, copyright, trademark or trade secret rights held or claimed by any MEF member company which are or may be associated with the ideas, techniques, concepts or expressions contained herein; nor
- any warranty or representation that any MEF member companies will announce any product(s) and/or service(s) related thereto, or if such announcements are made, that such announced product(s) and/or service(s) embody any or all of the ideas, technologies, or concepts contained herein; nor
- any form of relationship between any MEF member companies and the recipient or user of this document.

Implementation or use of specific Metro Ethernet standards or recommendations and MEF specifications will be voluntary, and no company shall be obliged to implement them by virtue of participation in the MEF Forum. The MEF is a non-profit international organization accelerating industry cooperation on Metro Ethernet technology. The MEF does not, expressly or otherwise, endorse or promote any specific products or services.

Table of Contents

List of Figures

List of Tables

1. List of Contributing Members

The following members of the MEF Forum participated in the development of this document and have requested to be included in this list.

Albis Technologies

Bell Canada

Ciena Corporation

Cisco Systems

Colt Technology Services

Cox Communications

Ericsson AB

HFR, Inc.

Omnitron Systems Technology, Inc.

PLDT Corp. Business Solutions

Tata Communications

TDS Telecom

Telus

The Carrier Ethernet Academy

XO Communications

2. Abstract

In order to provide differentiated levels of service, it is necessary to classify incoming frames to a service level either based on context (e.g., which EVC or OVC) or content (i.e, the contents of a specific field within the frame).

MEF10.3 and MEF26.2 provide attributes for associating each ingress frame with a Class of Service Name (CoS Name) for this purpose. Those specifications also provide attributes for associating each ingress frame with a color.

This Implementation Agreement formalizes the CoS Name and defines three specific CoS Names called Class of Service Labels (CoS Labels).

For frames associated with a CoS Label, this IA provides:

- values for fields containing the CoS identifier
- values for fields containing the frame color
- definition of Performance Tiers. Performance Tiers provide a way to define sets of performance objectives based on inherent characteristics of the service (primarily geographic span).
- specific performance objectives. Required values for performance objectives are specified in this document for service with a Class of Service identified by one of the MEF CoS Labels.
- requirements associated with bandwidth profile applicability to frames associated with the CoS Labels.

This IA also provides guidelines for CoS Names, in general, in terms of how the performance objectives for OVCs are composed into performance objectives for EVCs.

3. Terminology

Terms defined in MEF 10.3 [1], MEF 6.2 [16], and MEF 26.2 [10] are included in this document by reference and, hence, are not repeated in the Terminology table.

Term	Definition	Reference
Carrier Ethernet Network	A network from a Service Provider or network Operator supporting the MEF service and architecture models.	MEF 12.2 [13]
CEN	Carrier Ethernet Network	MEF 12.2 [13]
Class of Service Identifier	The fields in a Service Frame or ENNI Frame, along with the values of those fields, that are used to identify the Class of Service Name that applies to the frame.	This document
Class of Service Frame Set	A set of Service or ENNI Frames that have a commitment from the Operator or Service Provider subject to a particular set of performance objectives.	This document
Class of Service Label	A CoS Name that is standardized in this document. Each CoS Label identifies several Performance Tiers where each Performance Tier contains a set of performance objectives and associated parameters.	This document
Class of Service Name	A designation given to one or more sets of performance objectives and associated parameters by the Service Provider or Operator.	This document
Class of Service Performance Objective	An objective for a given performance metric.	This document
Color ID	Color Identifier	This document
Color Identifier	The fields in a Service Frame or ENNI Frame, along with the values of those fields, that are used to identify the Color that applies to the frame.	This document
CoS	Class of Service	This document
CoS IA	Class of Service Implementation Agreement (this document)	This document
CoS FS	Class of Service Frame Set	This document
CPO	CoS Performance Objective	This document

Term	Definition	Reference
C-Tag	Subscriber VLAN Tag	IEEE 802.1Q [2]
DEI	Drop Eligible Indicator	IEEE 802.1Q [2]
DSCP	Differentiated Services Code Point	RFC 2474 [15]
EI	External Interface	MEF 4 [4]
ENNI	External Network Network Interface. An interface used to interconnect two CEN Operators	MEF 4 [4]
ENS	Ethernet Network Section	This document and Y.1540 [11]
Ethernet Network Section	A set of one or more CENs, each under a single or collaborative jurisdictional responsibility, for the purpose of managing CPOs.	This document and Y.1540 [11]
IA	Implementation Agreement	
Operator	Also Network Operator. The Administrative Entity of a CEN	Derived from MEF 4 [4]
N/A	Not Applicable	
N/S	Not Specified	
PCP	Priority Code Point	IEEE 802.1Q [2]
Performance Tier	A MEF CoS Performance Objectives (CPO) set	This document
PT	Performance Tier	This document
Service Level Specification	The technical specification of the service level being offered by either the Service Provider to the Subscriber in the case of an EVC or by an Operator to a Service Provider in the case of an OVC.	Adapted from MEF 10.3 [1] and MEF 26.2 [10]
S-Tag	Service VLAN Tag	IEEE 802.1Q [2]
VLAN	Virtual LAN	IEEE 802.1Q [2]

Table 1: Terminology and Definitions Table

4. Scope

CoS IA defines a set of three CoS Names called CoS Labels for EVCs and OVCs. This IA also defines values for CoS Performance Objectives (CPOs) grouped in Performance Tier sets, as well as Performance Parameters.

The CoS Name and Color requirements in this IA are applicable at UNIs and ENNIs (collectively External Interfaces or EIs), and the indicated CoS Performance Objectives are applicable to Qualified Frames[1] that arrive at those EIs for transport across an EVC or OVC. This IA also addresses Ethernet Network Sections associated with typical Operator domains that interconnect at ENNIs (e.g., concatenation of CPOs for OVCs to derive CPOs for EVCs).

The internal mechanisms for implementing the CoS IA are out of scope.

The three CoS Labels provide support for key applications. This document also sets requirements for the mapping of Class of Service IDs defined in [1] and [10] to CoS Labels. Operators can offer other proprietary CoS Names and map values of the CoS ID to these CoS Names. Consequently, an Operator or Service Provider can offer zero to all three of the CoS Labels in any combination simultaneously with zero or more proprietary CoS Names.

This document specifies values for Performance Parameters (e.g., Percentile (P), Time interval (T)) to allow determination of CPOs for One-way Frame Delay, One-way Mean Frame Delay, One-way Frame Delay Range, One-way Inter Frame Delay Variation, One-way Frame Loss Ratio, One-way Availability, One-way Resiliency (HLI and CLI) and One-way Group Availability. It does not include Performance Parameters for One-way Multiple EVC Group Availability or One-way Composite Performance[2].

Where possible this IA relies on CoS and performance-related service attributes already defined in other MEF specifications. To further define CoS, this IA identifies, and where necessary constrains or extends, current MEF specifications. The IA also builds upon previous work in IEEE, ITU and IETF for consistency and facilitation of end-to-end CoS. This previous standards work includes CoS definitions for the IP layer, thus facilitating synergies between Ethernet and IP services and networks.

Figure 1 provides examples of the scope and applicability of the CoS IA to both UNI and ENNI, to Point-to-point, and Multipoint-to-Multipoint EVCs and OVCs (Rooted-Multipoint are applicable but not shown), and to both single and multiple CENs.

---

[1] Consistent with the definition in MEF 26.2 [10], the term Qualified Frame is used to refer generically to Qualified Services Frames on a UNI and Qualified ENNI Frames on an ENNI.

[2] Parameters for One-way Multiple EVC Group Availability and One-way Composite Performance are not in scope for MEF23.2.

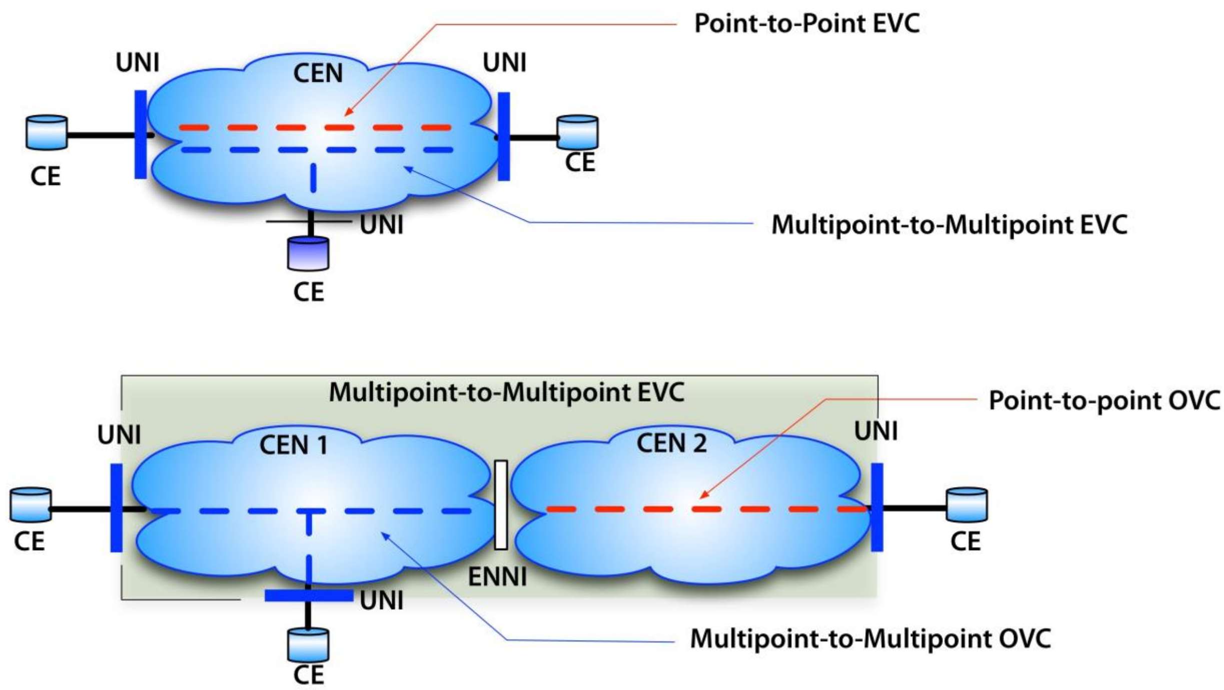


Figure 1: Examples of Scope and Applicability of CoS IA

With respect to the set of interfaces that are described as CEN External Interfaces in [4], the CoS IA uses the term External Interface (EI) to include UNI and ENNI.

4.1 New Material beyond MEF 23.1

New topics in this document include:

- A more formal definition of CoS Frame Set in section 8.1.3.
- A new Performance Tier ("PT0.3" or "City") with more stringent CPOs than Performance Tier 1, to support additional applications.
- CPOs for Multipoint Services in all Performance Tiers.
- Appendix F discussing the methodology for deriving CPOs for Multipoint Services and a description and recommended methods to remedy the focused overload condition.
- Appendix G on Burst Size, Shaper Considerations, and discussion of the interaction of TCP Congestion Control with the MEF policer.
- Appendix H providing guidance on the choice of CBS.

4.2 Revisions to Material in MEF 23.1

Revisions to previous material include:

- Terminology for service attributes is aligned with MEF 10.3.

- Section 8 has been reorganized to eliminate duplicate text and multiple definitions of the same terms.
- Table 3 and Table 4 that map CoS ID and Color ID to CoS Label and Color have been restructured.
- Requirements and recommendations regarding Bandwidth Profiles for CoS Labels have been updated to be consistent with MEF 10.3 (including token sharing model).
- L2CP recommendations have been updated to be consistent with MEF 10.3.
- Parameters for CPOs have been split into multiple tables. One for SLS-wide parameters, one for parameters for point-to-point xVCs and one for parameters for multipoint xVCs.
- Interpretation for N/S in the Performance Objective Parameter and CPO tables has been included.
- Older material from the Burst Size Appendix G (sections 8.7.1 in MEF 23.1) has been removed in favor of the new material in Appendix G and the new Appendix H.

This IA supersedes MEF 23.1 (CoS IA Phase 2).

5. Compliance Levels

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in RFC 2119 [3]. All key words use upper case, bold text.

Items that are **REQUIRED** (contain the words **MUST** or **MUST NOT**) are labeled as **[Rx]** for required. Items that are **RECOMMENDED** (contain the words **SHOULD** or **SHOULD NOT**) are labeled as **[Dx]** for desirable. Items that are **OPTIONAL** (contain the words **MAY** or **OPTIONAL**) are labeled as **[Ox]** for optional**.**

6. Numerical Prefix Conventions

This document uses the prefix notation to indicate multiplier values as shown in Table 2.

Decimal		Binary	
Symbol	**Value**	**Symbol**	**Value**
k	10^{3}	Ki	2^{10}
M	10^{6}	Mi	2^{20}
G	10^{9}	Gi	2^{30}
T	10^{12}	Ti	2^{40}
P	10^{15}	Pi	2^{50}
E	10^{18}	Ei	2^{60}
Z	10^{21}	Zi	2^{70}
Y	10^{24}	Yi	2^{80}

Table 2: Numerical Prefix Conventions

7. Introduction

With the introduction of Metro and Carrier Ethernet services, Service Providers and Operators started using this Ethernet "connectivity" technology to provide Ethernet "services". Various MEF specifications have added to IEEE 802 series standards in order to create a framework to define Ethernet services. This IA is motivated by the need to introduce and define specific "classes" or CoS Names called CoS Labels that will deliver a commitment for a particular level of performance for a set of Qualified Frames from the Service Provider or Operator. This is to further develop Carrier Ethernet services that are interoperable and predictably support Subscriber applications. For example, Operators and Service Providers that connect CENs will be able to do so with a set of commonly understood CoS Labels, CoS IDs, and CoS Performance Objectives (CPOs) in addition to any bilaterally-agreed CoS Names they want to support.

The requirements in this document are applicable to Subscribers, Service Providers and Operators who desire CoS Name interoperability across EIs. The requirements are developed based on the needs of Subscribers and their applications. Compliance with the CoS Labels in this IA does not limit an Operator from providing additional CoS Names using CoS Identifier values (e.g., PCP) that are left unused in this IA. Note that the CoS Performance Objective (CPO) and Parameter values are specified in this IA as maximums or minimums and thus do not limit Operators from providing conformant values that are less than the maximums or greater than the minimums. These other values could be described as more stringent, i.e., having more rigor or severity with respect to the standard or requirement value.

Figure 2 illustrates the need for a standard CoS Label model for mapping at an ENNI which is one key motivation for this IA. The problem addressed is that the Operators of CEN 1 and CEN 2 may have different CoS Names and different methods and values to indicate the CoS Names. The figure illustrates how the use of the CoS IA can provide a common set of CoS Labels that the Operators can map frames into, to facilitate interworking. For example:

- A frame with CoS Name HEART going from CEN 1 to CEN 2 maps to MEF CoS Label M on the ENNI which then maps to CoS Name PAPER in CEN 2.
- A frame with CoS Name SQUARE going from CEN 1 to CEN 2 also maps to MEF CoS Label M on the ENNI and then maps to CoS Name PAPER in CEN 2.
- A frame with CoS Name PAPER going from CEN 2 to CEN 1 maps to MEF CoS Label M on the ENNI which then maps to CoS Name SQUARE in CEN 1. (Note that frames from CEN 2 to CEN 1 can never result in CoS Name HEART in CEN 1.

MEF 26.2 defines the OVC End Point Egress Map Service Attribute which controls how the sending CEN maps its CoS Names and Color to egress ENNI Frame content in order to ensure that the receiving CEN processes the frame in a way that correctly maps CoS Names between CENs.

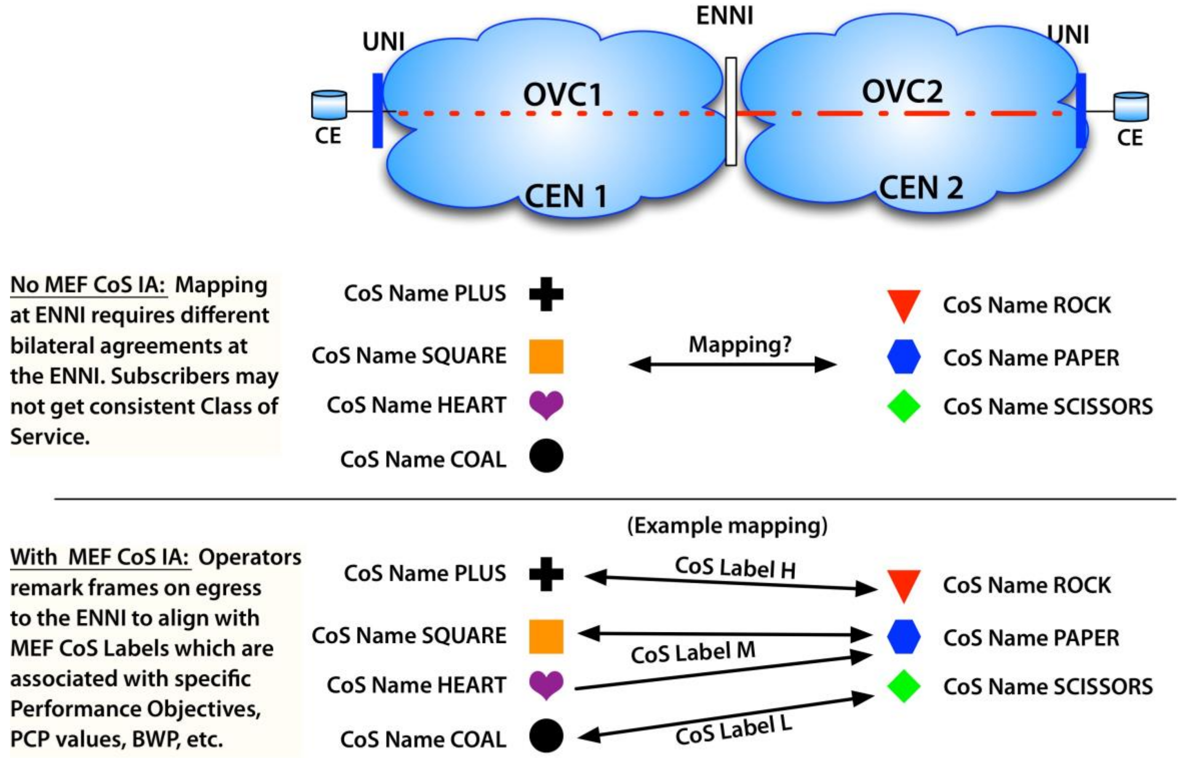


Figure 2: CoS IA Motivation Example – ENNI Mapping

Note that in the figure above the 3 CoS Names used by the Operator (ROCK, PAPER, SCISSORS) may align with the CoS IA Three CoS Label Model. A case could be constructed where neither CEN complies with the CoS IA Three CoS Label Model at the UNIs in their CEN, but both map to the CoS IA Model at the ENNI. A Three CoS Label Model is specified in order to satisfy the competing needs of a diversity of applications, finding common needs among Operators, limited CoS Identifier and Color Identifier field value space (e.g., 8 possible PCP values) and ensuring sufficiently simple interoperability. This CoS IA allows any subset of the 3 CoS Labels to be specified for a given EVC or OVC.

In addition, interconnection at the ENNI faces the challenge of providing UNI-to-UNI CoS with multiple Operators. Each Operator will provide a subset of the OVCs that make up the EVC. In addition to the need for CPOs associated with the UNI-to-UNI EVC, interworking and performance will be facilitated if each Operator has CPOs for their OVCs that are consistent with the EVC CPOs.

This document is organized as follows:

- Section 8 includes definitions of key terms followed by the requirements associated with the Class of Service Model.

- Section 9 includes a description of the Performance Metrics, the Parameters used for measuring Performance Objectives and the Performance Objectives for the three CoS Labels by Performance Tier.

8. Class of Service Model and Objectives (Normative)

This section includes definitions of key terms in section 8.1 followed by the details of the MEF Class of Service Model in sections 8.2 through 8.7.

8.1 Definitions of Key Terms

This section includes the definitions of key terms that are used throughout this document.

- **Class of Service Name (CoS Name)** in section 8.1.1
- **Class of Service Label (CoS Label)** in section 8.1.2 with additional detail in section 8.6
- **Ordered Endpoint Pairs (OEPP)** in section 8.1.3
- **Performance Tier (PT)** in section 8.1.4
- **Class of Service Frame Set (CoS FS)** in section 8.1.5
- **Class of Service Identifier (CoS ID)** in section 8.1.6 with additional detail in section 8.2
- **Color Identifier (Color ID)** in section 8.1.7 with additional detail in section 8.2

8.1.1 Class of Service Name (CoS Name)

A CoS Name is a designation given to one or more sets of performance objectives and associated parameters by the Service Provider or Operator. Examples of CoS Names are *Bronze, Gold, Silver,* and *Platinum.*

8.1.2 Class of Service Label (CoS Label)

A Service Provider or Operator can use many CoS Names, each with several different sets of performance objectives and associated parameters. A key goal of this document is to standardize three CoS Names and the values for the sets of performance objectives and associated parameters. These three CoS Names are called CoS Labels and are designated ***H***, ***M***, and ***L***. These informally refer to High, Medium and Low. The order of the CoS Labels is based on the traffic classes in [2] and their associated PCP values. Each CoS Label identifies five Performance Tiers where each Performance Tier contains a set of performance objectives and associated parameters.

CoS Labels do not imply any specific implementation of network priority mechanisms (e.g., strict priority queuing, weighted fair queuing, etc.) in handling a frame.

CoS Label is independent of all Service Provider, Operator and other standards' CoS Names. Users of this IA, such as Operators and Service Providers, can assign any brand or marketing names desired to the MEF compliant CoS Labels for their own services.

[R1] A CoS Label **MUST** be one of *H*, *M*, or *L*.

8.1.3 Ordered Endpoint Pairs (OEPP)

MEF 10.3 [1] and MEF 26.2 [10] specify a set of Performance Metrics on which this CoS IA is based. These performance metrics are all defined as one-way performance metrics so performance is defined on *ordered* UNI pairs for an EVC and *ordered* OVC End Point pairs for an OVC (in other words, a→b is distinct from b→a). Since many of the descriptions in this document apply to both EVCs and OVCs, the term Ordered Endpoint Pair[3] (OEPP) is used generically.

8.1.4 Performance Tier (PT)

Performance Objectives, with the exception of the One-way Availability Performance, apply to Qualified Frames in a EVC or OVC. Clearly, the objectives for a frame arriving at an External Interface (EI) depend on the EI that the frame will be delivered to. For example, the geographic distance between the EIs has a significant bearing on the Frame Delay. This Implementation Agreement provides guidance to Service Providers, Operators, and Subscribers by specifying five sets of CoS Performance Objectives (CPOs) called Performance Tiers (PTs). Each set includes objectives for seven performance metrics for point-to-point and multi-point CPOs.

The PTs are defined on the basis of geographic distance between the EIs, but the choice of a PT can depend on several considerations such as the number of switching hops or speed of links traversed, including access links. Note that the speed and technology used for links is a factor in delay that can be significant. For example, for a 1500 byte frame the serialization delay on a 2 Mb/s link can be about 6 ms and the delay for certain multiple physical link bonding technologies and associated fragmentation and de-fragmentation can add several additional milliseconds.

This Implementation Agreement requires, for a service that uses a CoS Name that is a MEF CoS Label, that CPOs that are specified in the SLS for frames with that CoS Label be consistent with the CPO ranges specified in an appropriate Performance Tier. This connection is made by associating a PT with a subset of OEPPs in the service. This is discussed in section 8.1.5 on CoS Frame Sets.

When an Operator (in agreement with the Service Provider) chooses a PT that is most applicable for a given set of frames for a given CoS Label, the Operator may base that choice on any criteria (e.g., distance, link speed). Setting the proper PT (i.e., CPO set) for OVCs requires a concept of CPOs for each OVC that composes an EVC that are consistent with the EVC CPOs. This is discussed in section 8.3.

In terms of the requirements of this IA, distance between EIs is not a performance-related parameter that must be measured and reported by an Operator. Distance is only used to derive CPOs in this IA. Therefore precise definitions regarding how to measure and report distances between EIs are not necessary. The CPOs for a given PT may be viewed as a set of CPOs for a particular 'field of use' or 'area of applicability' from the Operator point of view. *The Operator*

[3] The number of OEPPs in a service with N total endpoints and L leaf endpoints is $(N*(N-1))-(L*(L-1))$.

need not adhere to the distances used in the derivation of a PT in their use of a particular MEF PT.

In deriving PT CPOs for CoS IA, assumptions were made about mapping of applications to one or more CoS and PT. In CEN implementations, particular applications may be mapped differently. For example, a subset of the Mobile Backhaul traffic may have some of the smaller FD/MFD value requirements and these requirements may only be achievable in a particular PT set that is based on relatively low propagation (minimum) delay. CoS IA does not normatively make such application or service exclusions however.

This IA uses distance as the primary means of describing PTs and deriving minimum delays. The distances stated for each PT can be considered as approximate distance only if the assumptions stated in Appendix A are applicable. Below are the five PTs defined in this IA with the format: PT Number (PT Name) - Description (distance, derived propagation delay used in CPO constraints to establish a minimum per PT).

- PT0.3 (City PT) – derived from distances less than Metro in extent (<75 km, 0.6 ms),
- PT1 (Metro PT) – derived from typical Metro distances (<250 km, 2 ms),
- PT2 (Regional PT) - derived from typical Regional distances (<1200 km, 8 ms),
- PT3 (Continental PT) - derived from typical National/Continental distances (<7000 km, 44 ms),
- PT4 (Global PT) – derived from typical Global/Intercontinental distances (<27500 km, 172 ms)

Appendix A describes how PT sets were derived. Distances are not normative and are only used to provide per PT delay related CPO constraints. The intent is to provide a range of PT sets that address Carrier Ethernet Networks of different geographic coverage, design and scope. Thus a five PT model is adopted for MEF CoS Labels. CPO value sets are specified in a separate table per PT.

Note that in this document, the Parameters for the Performance Metrics (see section 9.2) have the same values across all Performance Tiers.

8.1.5 Class of Service Frame Set (CoS Frame Set)

Different pairs of end points can have different performance characteristics (driven heavily by their geographic span but also by other factors) so the performance objectives for different OEPPs can be different. Consider, for example, a service (EVC or OVC) that has an endpoint in San Francisco, an endpoint in Los Angeles, and an endpoint in New York. Clearly, the performance objectives for the various delay-related performance metrics will be different for frames flowing between LA and SF than for frames flowing between SF and NY or LA and NY.

The Service Level Specification[4] for an EVC or OVC contains a list, *PM*, of Performance Metrics. For each of the Performance Metrics defined in [1] and [10] (e.g., One-way Frame Delay), *PM* contains a list of Performance-Metric-Specific Parameters and a Performance Objective. *PM* can contain multiple entries with a given Performance Metric Name, but one or more of the parameter values associated with each objective for a given Performance Metric Name need to be different from each other (from [10]).

Two parameters that are common to all entries in the *PM* list are a Class of Service Name, ***CoS Name*** and a set of OEPPs, ***S***. Therefore, there can be different objectives for the same Performance Metric if the objectives apply to different sets of OEPPs or different CoS Names. For example, for an EVC with UNIs numbered 1 through 3:

Performance Metric	S	CoS Name	PM Objective
One-way Frame Loss Ratio,	{⟨1,2⟩,⟨ 2,1⟩},	*Bronze*,	0.02%
One-way Frame Loss Ratio,	{⟨ 3,1⟩,⟨ 1,3⟩,⟨ 3,2⟩,⟨ 2,3⟩},	*Bronze*,	0.025%
One-way Frame Loss Ratio,	{⟨ 1,2⟩,⟨ 2,1⟩},	*Gold*,	0.01%
One-way Frame Loss Ratio,	{⟨ 3,1⟩,⟨ 1,3⟩,⟨ 3,2⟩,⟨ 2,3⟩},	*Gold*,	0.015%

These PM entries for the same Performance Metric (One-way Frame Loss Ratio), indicate that for class of service *Bronze*, the objective is .02% on ordered UNI pairs ⟨1,2⟩ and ⟨ 2,1⟩, and it is .025% on ordered UNI pairs ⟨3,1⟩, ⟨1,3⟩, ⟨3,2⟩, and ⟨2,3⟩. A tighter requirement is listed for class of service *Gold* for the same sets of OEPPs. MEF 10.3 [1] and MEF 26.2 [10] do not include any constraints on how the ordered End Point pairs are organized, however the **CoS Frame Set** (defined below) along with [R1], and [R2] provide requirements for classes of service based on MEF CoS Labels.

Performance Objectives can be specified[5] on all of the OEPPs in a service, or on a subset. In this document we assume that for each CoS Name, ***CN***, there is a Metric Set, $\boldsymbol{MS_{CN}}$, that contains all OEPPs in a service for which performance objectives are specified (for that CoS Name). The cardinality of $\boldsymbol{MS_{CN}}$ can range from 0 to the total number of OEPPs in the service (except for leaf-to-leaf OEPPs) and it can be different for each CoS Name in the service.

It is, in theory, possible to specify different performance objectives for each OEPP in MS_{CN}, but this is usually not practical. It is more likely that OEPPs with similar performance characteristics are grouped together and have a common set of performance objectives applied to them.

Therefore subsets of MS_{CN}, *{S1…Sn}*, which we call *Endpoint Groups*, can be defined and performance objectives can be associated with each of these Endpoint Groups.

For example, in the PM list in the example above, MS_{Gold} and MS_{Bronze} have both been split into two Endpoint Groups, S_1 and S_2:

$$MS_{Bronze}=MS_{Gold}=\{\langle 1,2\rangle, \langle 2,1\rangle,\langle 3,1\rangle, \langle 1,3\rangle, \langle 3,2\rangle,\langle 2,3\rangle\}$$

[4] This reflects the definition in MEF 26.2 and assumes the same structure will be added to MEF 10.x in a future revision.

[5] The SLS can specify performance objectives for some sets of OEPPs even if there is no intention (capability) of continuously monitoring/measuring the performance.

S_1={⟨1,2⟩, ⟨2,1⟩} S_2={⟨3,1⟩,⟨1,3⟩, ⟨3,2⟩,⟨2,3⟩}

[D1] If, for CoS Name *CN*, the OEPP ⟨x,y⟩ is a member of a Endpoint Group $S_x \subseteq MS_{CN}$, then if the reverse OEPP ⟨y,x⟩ is in MS_{CN}, it **SHOULD** also be a member of S_x.

[D1] recommends that the forward and reverse paths for a given pair of Endpoints should be part of the same Endpoint Group.

A **CoS Frame Set** is a way to group the frames that are transported by an EVC or OVC into classes that are subject to common sets of performance objectives. A **CoS Frame Set** includes frames that are Qualified (i.e., subject to SLS performance objectives) and those that are not. For example, a frame declared yellow by a Bandwidth Profile is not subject to any performance objectives but is part of a **CoS Frame Set**.

A **CoS Frame Set** for a CoS Name can be described by the 2-tuple $FS=\{CN, S_x \subseteq MS_{CN}\}$ where *CN* is a CoS Name and S_x is an Endpoint Group. *FS* represents all of the frames with CoS Name *CN* that arrive at Endpoint *a* for delivery to Endpoint *b* for all OEPPs, ⟨*a*,*b*⟩, in S_x.

[R1] If *CL* is a CoS Name that is a MEF CoS Label, the Endpoint Groups defined for *CL* i.e., *{S1…Sn}*, **MUST** be a partition of MS_{CL}.

Requirement [R1] means that every OEPP ⟨x,y⟩ in MS_{CL} must reside in one and only one Endpoint Group, S_x.

[R2] If *CL* is a CoS Name that is an MEF CoS Label, each CoS Frame Set associated with *CL* **MUST** be associated with one of the Performance Tiers (PT) defined in this Implementation Agreement and all Performance Objectives defined in the SLS for the CoS Frame Set must be consistent with the ranges specified in section 9 for the PT.

8.1.6 Class of Service Identifier (CoS ID)

CoS ID is a Service Attribute that describes how the Service Frame or ENNI Frame indicates the CoS Name for the frame. CoS Identifiers for an EVC at a UNI are specified in [1] section 10.2. CoS Identifiers for an OVC End Point are specified in [10] section 16.6.

8.1.7 Color Identifier

Color Identifier (Color ID) is a Service Attribute that describes how the Service Frame or ENNI Frame indicates Color (e.g., Color Identifier can indicate a Yellow frame at an ENNI via the S-Tag PCP or DEI). Color Identifiers for an EVC at a UNI are specified in [1] section 10.3. Color Identifiers for an OVC End Point are specified in [10] section 16.7.

Note that the frame Color can be critical, even in the case where the receiving Operator has not applied an Ingress Bandwidth Profile. This is because frames with Color indicated as Yellow are not considered Qualified Frames, as described in [1] and [10], and hence any CPOs specified in the SLS do not apply to them. This in turn can affect how the receiving Operator chooses to queue and schedule the frame.

8.2 CoS Identifier and Color Identifier Requirements

At an ENNI (meaning an OVC End Point at an ENNI that is not in a VUNI) this IA does not impose any constraints on the selection of CoS ID and Color ID beyond those in [10].

At a UNI (meaning an EVC at a UNI or an OVC End Point at a UNI) or a VUNI (meaning an OVC End Point at an ENNI that is in a VUNI) this IA does not impose any constraints on the selection of CoS ID beyond those in [1] and [10]. This IA imposes the following constraints on the relationship of CoS ID to Color ID:

[R3] When CoS ID is based on C-Tag PCP at a UNI or VUNI, any Color ID used **MUST** be based on the C-Tag PCP or C-Tag DEI.

[R4] When CoS ID is based on Internet Protocol at a UNI or VUNI, any Color ID used **MUST** be based on the Internet Protocol.

Section 8.6.2 contains additional requirements on the mapping of CoS ID values to a CoS Name that apply when the CoS Name is a CoS Label.

8.3 Composing End-to-end CPOs

An EVC can be composed of multiple concatenated OVCs. When this is done, the per-OVC CPOs must be consistent with the end-to-end EVC CPOs.

ITU-T Recommendation Y.1541 [6] defines methods for concatenating performance objectives or measurements for Ethernet Network Sections (ENS) into end-to-end performance objectives (this is described in Appendix B). Per the ITU-T definition, an ENS aligns with a CEN.

For MEF services it is possible that a given EVC is supported by multiple OVCs (including the case where there are multiple OVCs in a single CEN). The ENS model in Appendix B can be applied to the relationships the various OVC CPOs have with the EVC CPOs and to other OVC CPOs that compose the EVC when there is one OVC in each CEN. Appendix B provides a concatenation method example and associated guidelines for a subset of Performance metrics based on the methods in [6]. Concatenation is sometimes described as accumulating or combining sections. Concatenation is part of composing the end-to-end (UNI-to-UNI) CPOs. Allocation is the inverse of concatenation. Appendix B provides no direct method of calculating allocation but does provide guidance for an indirect approach based on iteration. Allocation facilitates establishing CoS Frame Set performance budgets for each Operator or domain.

For the case where multiple OVCs in a single CEN support an EVC (e.g., with hairpin switching) the composition methods described for ENS can be applied by replacing “ENS” with “OVC”.

Note that the definition of delay in [1] and [10] includes the delay incurred in traversing each ENNI thus the calculated delay for the UNI-UNI using concatenated OVCs will be slightly overstated. See Appendix B for more information.

The ability to allocate EVC CPOs and concatenate OVC CPOs is motivated by several factors. These include:

- Existence of typical administrative and network boundaries that exist between CENs at ENNIs and within Operator networks between administrative and technology domains (e.g., between access networks and Ethernet networks).
- Establishment of clear responsibilities for an appropriate budgeted part of the UNI-to-UNI CPO for each OVC and its Operator (or domain within a CEN).
- Specification and reporting of CPO related SLS results (e.g., performance for each OVC) in an EVC that traverses multiple CENs.

Figure 3, Figure 4, and Figure 5 illustrate use cases for assignment of PTs to OEPP in OVCs . Figure 3 represents the simplest case, a point-point EVC in a single CEN. In this example, an EVC's CPOs utilize the PT3 set of CPOs for UNI-to-UNI SLS.

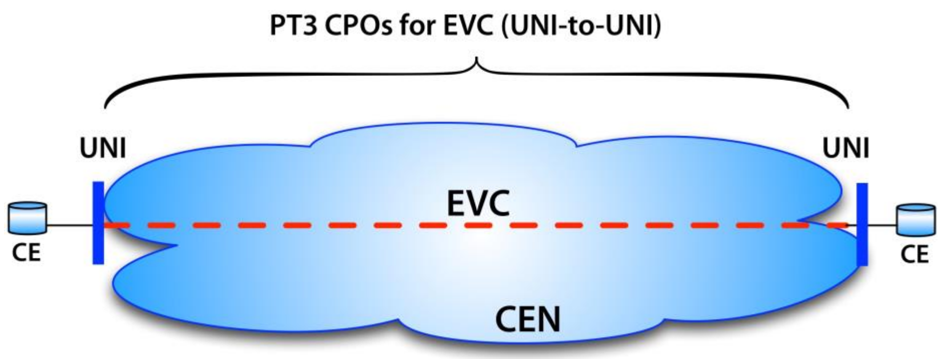


Figure 3: Example Performance Tier for a Single CEN EVC

In Figure 4 the EVC traverses an ENNI that connects two CENs.

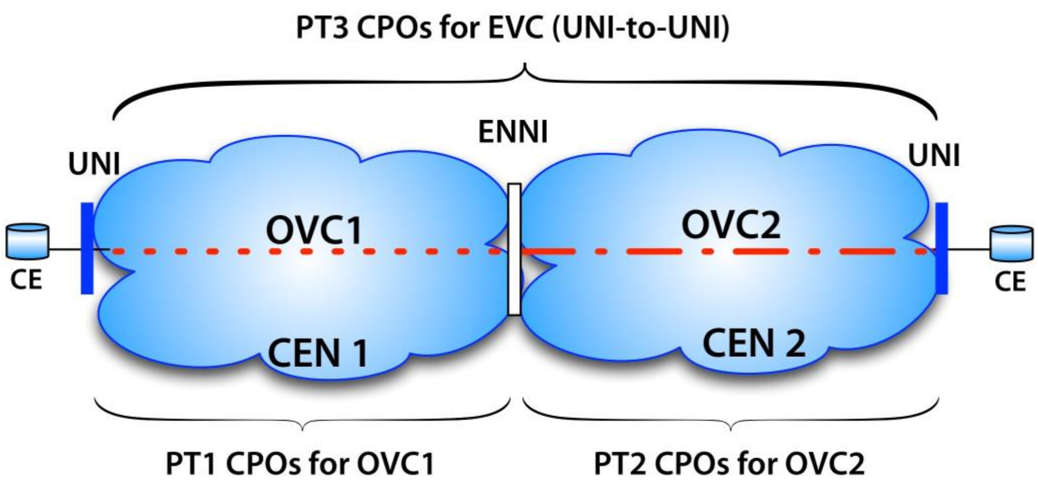


Figure 4: Example Performance Tiers for a Multiple CEN EVC and OVCs

The EVC will still have a UNI-to-UNI CPO set based on PT3 as represented by the bracket on top. The OVCs that compose the EVC may have CPOs as represented by the bottom brackets. In this example, the OVC in CEN1 (UNI-to-ENNI) and the OVC in CEN2 (ENNI-to-UNI) use the PT1 and PT2 set of CPOs, respectively. Note that the OVC CPO values in PT0.3–4 in this IA are not likely to concatenate precisely to the EVC CPO values in PT0.3–4 tables in this IA. How CEN Operators arrive at acceptable objectives is beyond the scope of this IA. As stated previously, the composition model includes both allocation and concatenation. While the example in Figure 4 is UNI-to-ENNI, a similar case can be constructed that includes ENNI-to-ENNI OVCs or the case of a multipoint EVC with a subset of ordered UNI pairs mapped to a PT.

Figure 5 represents the cases described above with a multipoint EVC that spans two CENs.

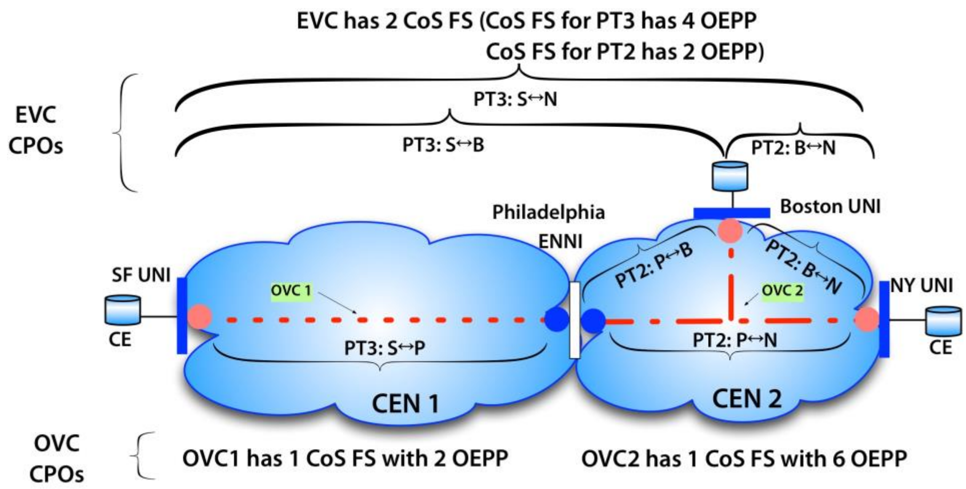


Figure 5: Example Performance Tiers for a Multi-CEN/Multi-Point EVCs and OVCs

The composition model can also be applied to scenarios in which a CEN that would appear from the outside as a single CEN is actually decomposed into multiple administrative based CENs. The CPOs for each of these component CENs can be composed into CPOs for the larger CEN. An example of this would be a Service Provider that has subsidiaries that provide access service CENs on each end and a CEN providing a transit service in the middle. These could be treated as three CENs for the purpose of setting CPOs. The EVC or OVC must meet the performance objectives agreed to with the Subscriber or Service Provider regardless of whether the EVC or OVC spans a single CEN or multiple CENs. The SP needs to carefully consider the performance objectives for each metric for each CEN in order to determine the end-to-end CPO. Appendix B provides guidance on this process.

See Appendix B.4 for guidelines on how to apply the concatenation methods.

8.4 Bandwidth Profile and Color

[1] and [10] provide no requirements or guidelines for how the various Bandwidth Profile models should be applied in the various CoS ID options. For example, at the UNI the choice of "per UNI", "per EVC" or "per CoS Name" Bandwidth Profile models are not constrained by the choice of CoS ID. For example, the choice of C-Tag PCP for CoS ID is very relevant when using a "per CoS Name" Bandwidth Profile, but the choice of C-Tag-PCP CoS ID does not preclude using a "per UNI" or "per EVC" Bandwidth profile model. The service specifications in [16] provide certain constraints for which Bandwidth Profile models are allowed for each MEF service. For example, [16] does not allow "per UNI" or "per EVC" Ingress or Egress Bandwidth Profiles for any EVC service, and does not allow any Egress Bandwidth Profile for Ethernet Private Line (EPL) service.

This IA complements those requirements by requiring that the Bandwidth Profile granularity matches CoS Name granularity. Only when a single CoS Name is present at an EVC will a "per EVC" Bandwidth Profile 'police' at the granularity of CoS Name. For example, if multiple CoS Names are mapped to an Ingress Bandwidth Profile "per EVC", the Bandwidth Profile will not be able to 'police' Service Frames per CoS Name.

[R5] When Ingress Bandwidth Profiles and an SLS with at least one Performance Objective are present, Ingress Bandwidth Profiles **MUST** utilize the "per CoS Name" model in [1] and [10].

[R5] means that the Ingress Bandwidth Profile model needs to have a Bandwidth Profile Flow for each CoS Name to provide the best chance of delivering on the CPOs. [R5] applies at a UNI, at a VUNI, and at an ENNI.

8.4.1 Bandwidth Profile Compliance

In this IA, use of the terms CoS ID, Bandwidth Profile and Color is consistent with [1] and [10] for the UNI and [10] for the ENNI. Indication of Color can be used to indicate which frames are deemed to be within or outside of the SLS according to the Ingress Bandwidth Profile and the definition of Qualified Frames from [1] and [10]. As stated in [1] and [10] all performance metrics (except One-way Availability Performance) are defined such that they only apply to Qualified Frames.

Levels of Ingress Bandwidth Profile compliance are Green when fully compliant (compliant with CIR, CBS), Yellow when there is sufficient Ingress Bandwidth Profile compliance for transmission but without SLS Performance Objectives (compliant with EIR, EBS) and Red when not Ingress Bandwidth Profile compliant with either. Green and Yellow frames are identified as such in this IA. Red frames are discarded. Note that the ITU terminology in [6] for Green is Discard Ineligible frames and for Yellow/Red it is Discard Eligible frames.

When there is no Ingress Bandwidth Profile, implicit rate limiting is provided by the bandwidth limits of the EI Ethernet links. The requirements in this CoS IA for the case of no Ingress Bandwidth Profile apply. In particular, any frame successfully transmitted across an EI is considered a Green frame unless the Color Identifier indicates that it is Yellow in which case it is considered a Yellow frame.

8.5 Service Type Applicability

Any of the MEF CoS Labels can be used with services based on any of the EVC types defined in MEF 10.3 [1] and any type of the OVC types defined in MEF 26.2. In particular, Point-to-Point EVCs/OVCs can use the same CoS Labels as Multipoint-to-Multipoint EVCs/OVCs. At an External Interface a specific implementation might serve these different service types using separate treatment (e.g., queues). MEF CoS IA is intended to be applicable to Point-to-Point, Multipoint-to-Multipoint and Rooted-Multipoint EVCs and OVCs including the case where some or all are present simultaneously on a given EI.

For example, serving an EVP-LAN might be more complex than an EVPL. A given OEPP on a Multipoint-to-Multipoint EVC may communicate Service Frames using different paths within a CEN and among different Operators' CENs compared to the paths and network traversed by Service Frames from another OEPP in the same EVC. This and the variability of traffic between UNI pairs within a given Endpoint Group (with >2 OEPPs) within compliance of the Ingress Bandwidth Profile can complicate meeting CoS Performance Objectives for Multipoint EVCs and OVCs. Careful use of multiple CoS Frame Sets can help to better characterize the traffic in a multipoint EVC or OVC.

Consistent with [1] and [10], the MEF CPOs apply to frames in CoS Frame Sets associated with CoS Labels. When the CoS Frame Set is associated with an Endpoint Group containing two or more OEPPs, the performance is based on the worst pair's performance.

8.6 Three CoS Label Model

The Three CoS Label Model provides normative information for the CoS Labels defined in this IA: H, M, and L. The normative information includes Ingress Bandwidth Profile constraints, CoS Identifier and Color Identifier values, and CPOs. The requirements in this model apply to EVCs and OVCs. All CPO requirements refer to UNI-to-UNI, UNI-to-ENNI, ENNI-to-UNI, and ENNI-to-ENNI performance.

CoS Labels H, M and L informally refer to High, Medium and Low, and are differentiated by their performance requirements.

- H – intended for applications that are very sensitive to loss, delay and delay variation such as VoIP and mobile backhaul control.
- M – intended for applications that are sensitive to loss but more tolerant of delay and delay variation such as near-real-time or critical data applications.
- L – intended for applications that are more tolerant of loss as well as delay and delay variation such as non-critical data applications.

Since this CoS IA supports the use of all or of any subset of the three CoS Labels, there is a need for interworking or mapping when different operators use different subsets. For example, Operator of CEN 1 adopts all CoS Labels in the Three CoS Label Model and Operator of CEN 2 adopts a subset with 2 CoS Labels including CoS Labels H and L. If CEN 1 and CEN 2 are connected via an ENNI there is a need for mapping between the two models.

8.6.1 Ingress Bandwidth Profiles for CoS Labels

This IA does not mandate that an EVC or OVC with CoS Labels have ingress bandwidth profiles, however in order to support the intended applications for those CoS Labels this IA does impose constraints when ingress bandwidth profiles are present.

[R6] When an ingress Bandwidth Profile Flow is present for a CoS Label at an EI, the value of the CBS parameter for that flow **MUST** be either equal to zero or greater than or equal to:

- For an EVC, the EVC Maximum Service Frame Size as defined in MEF 10.3 [1]
- For an OVC, the lower bound specified in Table 47 of MEF 26.2 [10]

[R7] When an ingress Bandwidth Profile Flow is present for a CoS Label at an EI, the value of the EBS parameter for that flow **MUST** be either equal to zero or greater than or equal to:

- For an EVC, the EVC Maximum Service Frame Size as defined in MEF10.3 [1]
- For an OVC, the lower bound specified in Table 47 of MEF26.2 [10]

[R9] and [R10] mean that setting CBS or EBS to a value greater than zero is necessary and sufficient to ensure that the BWP is capable of declaring ingress frames of any allowable size to be green or yellow, respectively. This is not meaningful in a practical sense, however, unless the bandwidth profile is also configured to allow tokens to be replenished as they are consumed.

[R8] When an ingress Bandwidth Profile Flow is present for a CoS Label at an EI, and when the value of CBS is greater than zero, the other BWP parameters for that flow **MUST** be configured in a way that allows tokens to be added to the committed token bucket over time.

[R9] When an ingress Bandwidth Profile Flow is present for a CoS Label at an EI, and when the value of EBS is greater than zero, the other BWP parameters for that flow **MUST** be configured in a way that allows tokens to be added to the excess token bucket over time.

The configurations for a Bandwidth Profile Flow that allow tokens to be added to a committed token bucket include:

1) $CIR^i_{max} > 0$ and $CIR^i > 0$, or
2) $CIR^i_{max} > 0$ and the Bandwidth Profile Flow is in an envelope where the next higher rank flow has $CF^i = 0$ and has a configuration that allows tokens to be added to its committed token bucket over time.

The bandwidth profile configurations that allow tokens to be added to an excess token bucket include:

1) $EIR^i_{max} > 0$ and $EIR^i > 0$, or
2) $EIR^i_{max} > 0$ and $CF^i = 1$ and the configuration allows tokens to be added to the committed bucket over time, or
3) $EIR^i_{max} > 0$ and the Bandwidth Profile Flow is in an envelope where the next higher rank flow has a configuration that allows tokens to be added to its excess token bucket over time.

CoS Label H is typically considered a "green-only" service, however it is allowed to be "green-yellow". CoS Label M is typically considered a "green-yellow" service, however it is allowed to be "green-only". CoS Label L is typically considered a "green-yellow" or "yellow-only" service, however it is allowed to be "green-only". This is formalized in the following requirements:

[R10] When an ingress Bandwidth Profile Flow is present for CoS Label H, it **MUST** have CBS > 0.

[R11] When an ingress Bandwidth Profile Flow is present for CoS Label M, it **MUST** have CBS > 0.

[R12] When an ingress Bandwidth Profile Flow is present for CoS Label L, it **MUST** have CBS + EBS > 0.

When a bandwidth profile specifies CBS = 0 for CoS Label L, the CoS Frame Set associated with the bandwidth profile will contain only yellow frames. Since there are no green frames in such a CoS Frame Set the Performance Objectives for CoS Label L in section 9.2 would not apply to any frames in the CoS Frame Set.

8.6.2 Mapping CoS ID and Color ID to CoS Label and Frame Color

To promote consistency in the way CoS Identifier values map to CoS Labels and frame color, this IA recommends mappings at a UNI or VUNI and requires mappings at an ENNI.

For convenience, the phrase "at a UNI" means an EVC at a UNI or an OVC End point that is located at a UNI. Likewise "at a VUNI" means an OVC End Point that is located at an ENNI and is in a VUNI, and "at an ENNI" means an OVC End Point that is located at an ENNI and is not in a VUNI.

Table 3 provides the Color ID to ingress Service Frame Color mapping when the CoS Identifier is based on EVC or OVC only at a UNI or VUNI.

[D2] At a UNI or VUNI with CoS Identifier based on EVC or OVC End Point only and a Color Identifier based on Internet Protocol, if the CoS Identifier maps to a CoS Label, then an ingress frame with a DSCP value matching an entry in the first column of Table 3 **SHOULD** map to the corresponding color in Table 3.

[D3] At a UNI or VUNI with CoS Identifier based on EVC or OVC End Point only and a Color Identifier based on C-tag PCP, if the CoS Identifier maps to a CoS Label, then an ingress frame with a C-tag PCP value matching an entry in the second column of Table 3 **SHOULD** map to the corresponding color in Table 3.

[D4] At a UNI or VUNI with CoS Identifier based on EVC or OVC End Point only and a Color Identifier based on C-tag DEI, if the CoS Identifier maps to a CoS Label, then an ingress frame with a C-tag DEI value matching an entry in the third column of Table 3 **SHOULD** map to the corresponding color in Table 3.

Table 4 provides the CoS and Color Identifier to ingress EI Frame CoS Label and Color mapping when the CoS and Color Identifiers are based on Internet Protocol, PCP or PCP and DEI. [1] and [10] require that when the CoS Identifier is based on Internet Protocol or PCP, all DSCP or PCP values map to a CoS Name. Ingress frames with CoS Identifier values that are not constrained to map to a specific CoS Label by Table 4 can be mapped to any CoS Name, including a CoS Name with a performance characteristic of discarding all ingress frames. For a multi-CoS EVC that supports only the standard MEF CoS Labels as defined in this document, tables providing examples of full PCP and DSCP mapping at a UNI are located in Appendix D. Providing the same CoS Label mapping on all UNIs for a given EVC will minimize subscriber confusion.

Since a given EVC or OVC is not required to have all CoS Labels, an ingress frame could contain a CoS Identifier value that Table 4 indicates is mapped to a CoS Label that is not in the EVC or OVC. In this case the CoS Identifier value can be used to map the ingress frame to any CoS Name.

When the CoS and Color Identifier are both based on S-tag or C-tag PCP, there are only two PCP values that are not included Table 4 and are always available to be used to map ingress frames to a different CoS Name. As noted above, when a PCP value indicates a CoS Label that is not in the EVC or OVC at the receiving CEN, the PCP value can be used to map ingress frames to a different CoS Name. Furthermore, it is possible that Table 4 indicates a CoS Label and Color combination for a particular PCP value where the indicated Color is not relevant to the indicated CoS Label due to ingress Bandwidth Profile configuration. Specifically, the CoS Label and Color combination is not relevant to the CoS Label in the receiving CEN when:

1. When the indicated CoS Label has a Color-Blind ingress Bandwidth Profile and has PCP Preservation disabled, the PCP value indicating Yellow is not relevant to this CoS Label. (Ingress frames with this PCP value would be treated exactly like ingress frames with the PCP value indicating Green.)
2. When the indicated CoS Label has a Color-Aware ingress Bandwidth Profile with EBS = 0, the PCP value indicating Yellow is not relevant to this CoS Label. (Ingress frames with this PCP value would be declared Red and discarded by the ingress Bandwidth Profile.)
3. When the indicated CoS Label has a Color-Aware ingress Bandwidth Profile with CBS = 0 and has PCP Preservation disabled, the PCP value indicating Green is not relevant to

this CoS Label. (Ingress frames with this PCP value would be treated exactly like ingress frames with the PCP value indicating Yellow.)

In these cases the Service Provider and Operator(s) can agree that all frames on the EI that map to this CoS Label will use a single PCP value, which allows the PCP value that would indicate the Color that is not relevant to the CoS Label to be used for any other CoS Name.

[D5] At a UNI or VUNI with CoS and Color Identifiers based on Internet Protocol, an ingress frame with a DSCP value matching an entry in the first column of Table 4 **SHOULD** map to the corresponding CoS Label and color in Table 4 if the EVC/OVC has that CoS Label.

[D6] At a UNI or VUNI with CoS and Color Identifiers based on C-tag PCP, an ingress frame with C-tag PCP value matching an entry in the second column of Table 4 **SHOULD** map to the corresponding CoS Label and color in Table 4 if the EVC/OVC has that CoS Label and the corresponding color is relevant to that CoS Label in the receiving CEN.

[D7] At a UNI or VUNI with CoS Identifier based on C-tag PCP and Color Identifier based on C-tag DEI, an ingress frame with C-tag PCP and DEI values matching an entry in the third column of Table 4 **SHOULD** map to the corresponding CoS Label and color in Table 4 if the EVC/OVC has that CoS Label.

[R13] At an ENNI with CoS and Color Identifiers based on S-tag PCP, an ingress ENNI frame with and S-tag PCP value matching an entry in the second column of Table 4 **MUST** map to the corresponding CoS Label and color in Table 4 if the OVC has that CoS Label and the corresponding color is relevant to that CoS Label in the receiving CEN .

[R14] At an ENNI with CoS Identifier based on S-tag PCP and Color Identifier based on S-tag DEI, an ingress ENNI frame with S-tag PCP and DEI values matching an entry in the third column of Table 4 **MUST** map to the corresponding CoS Label and color in Table 4 if the OVC has that CoS Label.

CoS ID = EVC or OVC EP			Color[1]
Color ID = IP	Color ID = PCP	Color ID = DEI	
DSCP (PHB)	PCP	DEI	
28 (AF32), 30 (AF33), 12 (AF12), 14 (AF13), 0 (Default)	4, 2, 0	1	Yellow
All other values	7, 6, 5, 3, 1	0	Green

[1] Note that [R103] of [1] requires that a Service Frame without a Color ID (e.g., an untagged Service Frame when Color ID is based on PCP or DEI, or a non-IP Service Frame when Color ID is based on Internet Protocol) to be Green.

Table 3: Color when the CoS ID is based on Only EVC or OVC EP at a UNI or VUNI

CoS and Color Identifiers [1]			Color [2]	CoS Label
CoS ID = IP Color ID = IP	CoS ID = PCP Color ID = PCP	CoS ID = PCP Color ID = DEI		
DSCP (PHB)	PCP	PCP ; DEI		
46 (EF), 44 (VA)	5	5 ; 0	Green	H
	4	5 ; 1	Yellow	
26 (AF31)	3	3 ; 0	Green	M
28 (AF32), 30 (AF33)	2	3 ; 1	Yellow	
10 (AF11)	1	1 ; 0	Green	L
12 (AF12), 14 (AF13), 0 (Default)	0	1 ; 1	Yellow	

[1] Full CoS Identifier includes the EVC or OVC End Point Identifier. This table specifies only the PCP or DSCP values to be used with EVC or OVC End Point Identifier to form a CoS ID.
[2] Note that [R103] of [1] requires that a Service Frame without a Color ID (e.g., an untagged Service Frame when Color ID is based on PCP or DEI, or a non-IP Service Frame when Color ID is based on Internet Protocol) to be Green.

Table 4: CoS Label and Color when the CoS Identifier is based on Internet Protocol or PCP at a ENNI, UNI, or VUNI

The specific values for PCP in Table 4 were derived from [2] using Tables 6-4 and G-5 Priority Code Point Decoding. The table row used is “5P3D” scheme (5 traffic classes of which 3 also have drop eligibility PCP values). See Section Appendix E for table excerpts.

In [2] (Table 6-4 "5P3D" row) there is a traffic class called "Best Effort" which is associated with PCP=1 when not drop eligible and PCP=0 when drop eligible. In this IA CoS Label L is aligned with this traffic class in [2]. In terms of Bandwidth Profile, note that CoS Label L allows CBS or EBS = 0. The special case of CBS = 0 effectively results in no CPOs for the Performance Attributes in this IA while the case of CBS > 0 requires conformance with CPOs. From a DSCP perspective CoS Label L is a combination of AF1 (for CBS>0) and Default (for CBS=0) classes.

Note that the tables contain both the DSCP value and associated Per Hop Behavior (PHB), however it is the DSCP value that is actually contained in EI Frames containing IP datagrams. A service supporting IPv4 must map the DSCP values per the table for each frame carrying an IPv4 datagram. A service supporting IPv6 must map the DSCP values per the table for each frame carrying an IPv6 datagram. A service supporting both IPv4 and IPv6 must map the DSCP values per the table for each frame carrying either an IPv4 or an IPv6 datagram. Consistent with MEF 10.3 [1] and MEF 26.2 [10], if CoS Name or color are based on DSCP, the CoS Name or color for a non-IP frame is determined by agreement of the parties. If a service does not support one of the IP versions, EI Frames containing that version are treated as non-IP frames. Note that MEF 26.2 [10] allows the mappings to be specified independently for IPv4 and IPv6.

Per [1] and [10], the CoS Identifier value in an ingress EI Frame determines the CoS Name assigned to the EI frame by the receiving CEN. In an egress EI Frame, the value of the PCP and DEI fields is specified by the OVC End Point Egress Map Service Attribute in [10] whose value is agreed upon by the Service Provider and the sending Operator. At an ENNI, the Service Provider knows the CoS Names and corresponding S-Tag PCP values agreed to by the receiving Operator. This knowledge allows the Service Provider to agree to a value of the OVC End Point Map Service Attribute with the transmitting Operator such that the egress ENNI Frame is given the desired CoS Name by the receiving Operator CEN. When the receiving Operator CEN supports CoS Labels, the value of the OVC End Point Egress Map Service Attribute needs to use the PCP values in Table 4. See the appendix in [10] for an example of the use of the OVC End Point Egress Map Service Attribute. At a UNI, the values of the C-Tag PCP and DEI fields in an egress Service Frame are not constrained by this IA.

8.6.3 L2CP to CoS Label Mapping

The CoS Identifier for L2CP frames at a UNI or VUNI can be independent of the CoS Identifier for data frames. [1] and [10] provide for a list of Layer 2 Control Protocols and the CoS Names to which those L2CP frames will be mapped. When the CoS Name for L2CP frames is a CoS Label, CoS Label M is recommended, based on its superior loss performance over CoS Label L, and a desire to keep it separate from real-time applications.

When L2CP frames are mapped to a CoS Label (e.g. CoS Label M), they will typically share that CoS Label with Data frames. One means of accomplishing this is to use a single CoS Name, with a single Bandwidth Profile Flow, for the combination of Data and L2CP frames that map to the CoS Label. Alternatively a Service Provider might prefer to have separate Bandwidth Profile Flows for data and L2CP frames. Mapping ingress L2CP and Data Frames at a UNI or VUNI to distinct CoS Names, both of which correspond to the same CoS Label (i.e. both conform to all requirements of that CoS Label) allows them to have distinct ingress Bandwidth Profiles.

[D8] At a UNI or VUNI that lists a specific L2CP to CoS Name mapping:

- If the indicated CoS Name is a MEF CoS Label, it **SHOULD** be a CoS Label M or another CoS Label whose CoS Frame Sets have objectives for One-way Frame Loss Ratio that meet the constraints for CoS Label M for the associated Performance Tiers as specified in Table 8 through Table 12.
- If the indicated CoS Name is not a MEF CoS Label, it **SHOULD** be associated with CoS Frame Sets that have objectives for One-way Frame Loss Ratio that meet the constraints for CoS Label M for Performance Tiers that best align with the OEPPs that the L2CPs are transported between (as specified in Table 8 through Table 12).

At an ENNI, L2CP frames are not distinguished from data frames and the CoS Identifier is based on S-tag PCP.

[1] and [10] do not distinguish between ingress L2CP frames and ingress data frames for the purposes of determining color. Therefore the color of an ingress L2CP frame is determined by the Color Identifier according to Table 3 and Table 4 and the associated requirements and recommendations.

9. Performance Metrics, Parameters, and Objectives

Definitions of Performance attributes, metrics, and associated parameters are found in [1] and [10].

This section has three parts:

1. Requirements and recommendations on the use of the various Performance Metrics defined in MEF 10.3 [1] and MEF 26.2 [10] (section 9.1),
2. Required values for the SLS Parameters for the Performance Metrics when applied to CoS Frame Sets associated with a MEF CoS Label (section 9.2), and
3. Required Performance Objectives (CPOs) for each Performance Tier for CoS Frame Sets associated with a MEF CoS Label (section 9.3)

The requirements stated in bullets 2 and 3 above are established in [R15].

9.1 Performance Metrics

One-way Frame Delay (FD) and One-way Mean Frame Delay (MFD) form a pair for which this IA requires support for at least one. Either one or both of these two can apply to a given SLS. Similarly, One-way Inter-Frame Delay Variation (IFDV) and One-way Frame Delay Range (FDR) Performance form a pair for which this IA requires support for at least one. Either one or both of these two can apply to a given SLS. The combination where a given SLS includes MFD and IFDV but not FD or FDR is not recommended because this does not allow an estimate of an upper bound one way delay. Requirements below formalize this normatively. However, it should be noted that to support EVCs end-to-end with OVCs it is recommended in [17] that operators support all four frame delay metrics for OVCs. Furthermore, in this case there are issues of allocation and concatenation to consider for Performance Objectives. See sections 8.3 and Appendix B.

All Performance metrics are one-way. As noted in section 8.1.5, it is not a requirement that the SLS include Performance Objectives for both directions of an endpoint pair. If both directions of an endpoint pair have Performance Objectives it is not a requirement that the two OEPPs are associated with the same Endpoint Group (and hence the same performance tier) although [D1] recommends that they are.

The following requirement applies to all of the Performance Metrics supported by this IA.

[R15] In an EVC or OVC that uses a MEF CoS Label, an SLS entry for a given performance metric and a given CoS Frame Set associated with that CoS Label **MUST** be specified per:

(1) The parameter values for that performance metric defined in Table 5, Table 6 and Table 7, as appropriate for the EVC/OVC type, and;

(2) The objective for that performance metric for the associated CoS Label and EVC/OVC Type in Table 8, Table 9, Table 10, Table 11, or Table 12, where table selection is dependent on the PT chosen for that CoS Frame Set.

9.1.1 Delay-related Performance Metrics

There are four Performance Metrics associated with frame delay: One-way Frame Delay (FD), One-way Mean Frame Delay (MFD), One-way Frame Delay Range (FDR), and One-way Inter-Frame Delay Variation (IFDV). These Performance Metrics are defined in MEF 10.3 [1] and in MEF 26.2 [10] for Service Frames and ENNI frames, respectively. The following requirements and recommendations apply to the use of these Performance Metrics in an SLS for a Class of Service based on a MEF CoS Label.

[R16] An SLS **MUST** include at least one Performance Objective for either One-way FD or One-way MFD for each CoS Frame Set associated with a MEF CoS Label.

[O1] An SLS **MAY** include Performance Objectives for both One-way FD and One-way MFD for any CoS Frame Set.

[R17] An SLS **MUST** include at least one Performance Objective for either One-way FDR or One-way IFDV for each CoS Frame Set associated with a MEF CoS Label.

[O2] An SLS **MAY** include Performance Objectives for both One-way FDR and One-way IFDV for any CoS Frame Set.

[D9] If an SLS includes Performance Objectives for a CoS Frame Set that includes objectives for One-way MFD but not One-way FD, the objectives for that CoS Frame Set **SHOULD** include objectives for One-way FDR.

CEN changes that alter delay such that delay is still within the SLS performance objectives for FD and MFD may lead to increases in FDR that cause the FDR SLS objectives to be missed. For example, a topology change during the interval T can significantly change the delay characteristics with the result being that the difference between the percentile delay and the minimum delay over the interval become large. If this event is isolated in time, however, the actual impact of the event at the application layer will be transient and may be insignificant. In such cases, the Service Provider and Subscriber or Service Provider and Operator may agree to ignore the FDR violation, especially if it can be shown that the impact of the topology change is the source of the miss or a One-way IFDV objective, if one is specified, is met. Procedures and/or criteria for reaching such an agreement are beyond the scope of this document.

9.1.2 Loss-related Performance Metrics

One-way Frame Loss Ratio (FLR) Performance, One-way Availability Performance, One-way Resiliency Performance (HLI and CHLI), and One-way Group Availability Performance are defined in MEF 10.3 [1] and in MEF 26.2 [10] for Service Frames and ENNI frames, respectively.

Figure 6 illustrates how the two One-way Resiliency Performance attributes defined in [1] and [10], counts of High Loss Intervals and counts of Consecutive High Loss Intervals, fit into the hierarchy of time and other attributes.

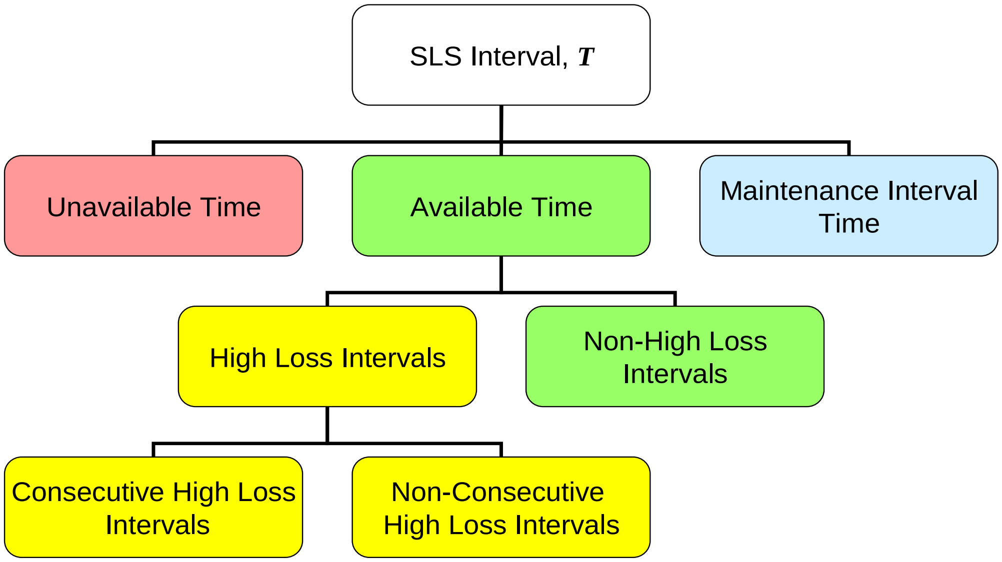


Figure 6: Hierarchy of Time Showing the Resiliency Attributes[6]

9.2 Performance Parameters

Parameters associated with Performance Metrics are specified in two groups. [1] and [10] require that certain parameters have a single value for all Performance Metrics specified in an SLS. These parameters are specified in Table 5. Parameters for the remaining Performance Metrics are listed in Table 6 and Table 7. Table 6 lists the parameters for point-to-point services and Table 7 lists the parameters that are different for multipoint services. They are stated separately for each CoS Label but the values are uniform across PTs. (In future phases they may be stated per PT.)

Table 5, Table 6, and Table 7 specify Performance Parameters required to derive and specify the CPOs in Table 8, Table 9, Table 10, Table 11, and Table 12.

MEF 10.3 [1] and MEF 26.2 [10] include parameters S (a set of OEPPs) and c (a CoS Name) for every Performance Metric. These are not listed in the Parameters tables since they are implicit in the CoS Frame Set to which the CPOs apply. A CoS Frame Set associates an Endpoint Group and a CoS Label and a Performance Tier (PT).

The Parameters are stated as inequalities, therefore for each Parameter an Operator may agree to a value less than the maximum or more than the minimum Parameter values.

In Appendix B.4 there is a non-normative guidance for parameter uniformity across particular OVCs that compose an EVC.

Consistent with the requirements in section 9.1, if the SLS includes a performance metric for a CoS Frame Set that is associated with a CoS Label, the parameter values need to meet the constraints in Table 5 – Table 7 and the CPO value needs to meet the constraints in Table 8 –

[6] Figure 6 is from MEF 10.3 [1]. See [1] or [10] for definitions of the terms inside the figure (e.g., Available Time).

Table 12. The entries in the tables are either a numerical limit on the parameter or CPO value, or "N/S", or both. The interpretation of these entries is as follows:

1. If the SLS includes a CPO for a performance metric whose parameter value constraints and CPO value constraints are listed in the table with a numerical limit, the SLS is required to use parameter values and CPO values consistent with the tables.
2. If the SLS includes a CPO for a performance metric whose parameter value constraints and/or CPO value constraint are listed with "N/S", then the N/S value is determined by agreement of the parties and not constrained by this document.
3. If the SLS includes a CPO for a performance metric with a parameter value constraint that is specified with both a numerical limit and "N/S" (e.g., "≥ 1sec or N/S"), the recommended constraint for the parameter value is as stated in the table, but the SLS can include a different value that does not meet the constraint and that is agreed on by the parties.

Parameter Symbol (Description)	Used in Performance Metric	Parameter Value
t_s (SLS Start Time)	ALL	Any time and date
T (Time Interval)	ALL	≤ 1 Month

Table 5: SLS Common Parameters

Parameter Symbol (Description)	Used in Performance Metric	Parameter Values for CoS Label H	Parameter Values for CoS Label M	Parameter Values for CoS Label L
P_d (Percentile)	FD	$\geq 99.9^{th}$	$\geq 99^{th}$	$\geq 95^{th}$
P_v (Percentile)	IFDV	$\geq 99.9^{th}$	$\geq 99^{th}$ or N/S	N/S
Δτ (Pair Interval)		≥ 1sec	≥ 1sec or N/S	N/S
P_r (Percentile)	FDR	$\geq 99.9^{th}$	$\geq 99^{th}$ or N/S	N/S
C (Loss Threshold)	ALL	≤ 0.1	≤ 0.1	≤ 0.5
Δt (Loss Interval)	ALL	≤ 10 sec	≤ 10 sec	≤ 10 sec
n (Consecutive Δt)	ALL	≤ 10	≤ 10	≤ 10
p (Consecutive Δt)	HLI, CHLI	≤ 5	≤ 5	≤ 5
K (Number of OEPPs)	Group Availability (A_g)	N⍰⍰	N/S	N/S

Table 6: CoS Label H, M and L Parameter Values for Point-to-Point Services

Parameter Symbol (Description)	Used in Performance Metric	Parameter Values for CoS Label H	Parameter Values for CoS Label M	Parameter Values for CoS Label L
P_d (Percentile)	FD	≥ 98.5th	≥ 98th	≥ 94th
P_v (Percentile)	IFDV	≥ 98.5th	≥ 98th or N/S	N/S
P_r (Percentile)	FDR	≥ 98.5th	≥ 98th or N/S	N/S
Note that the Parameters Below have the same values as the Parameters for Point-to-Point Services				
Δτ (Pair Interval)	IFDV	≥ 1sec	≥ 1sec or N/S	N/S
C (Loss Threshold)	ALL	≤ 0.1	≤ 0.1	≤ 0.5
Δt (Loss Interval)	ALL	≤ 10 sec	≤ 10 sec	≤ 10 sec
n (Consecutive Δt)	ALL	≤ 10	≤ 10	≤ 10
p (Consecutive Δt)	HLI, CHLI	≤ 5	≤ 5	≤ 5
K (Number of OEPPs)	Group Availability (A_g)	N/S	N/S	N/S

Table 7: CoS Label H, M and L Parameter Values for Multipoint Services

9.3 CoS Performance Objectives Per Performance Tier

Table 8, Table 9, Table 10, Table 11 and Table 12 provide CPOs for each Performance metric per each CoS Label. Each Table provides CPOs for one of the PTs. These are normative as per the requirements that refer to them. Note: Multipoint also includes Rooted Multipoint as per [1] and [10].

Derivation of the CPOs is found in Appendix C. The remainder of this section describes the Performance metrics and requirements for CPOs. Appendix A provides information on the derivation of the Performance Tiers.

The CPOs are stated as inequalities, therefore for each CPO an Operator may agree to a value less than the maximum or more than the minimum.

The tables of CPOs define objectives for multipoint services that differ from point-to-point services. These performance values apply to multipoint services with 100 or fewer external interfaces. This document does not specify objectives and parameters for multipoint services larger than 100 external interfaces.

These multipoint objectives also do not apply in time periods where the focused overload condition is present. The focused overload condition is described in Appendix F.1.2. A definition of the method an operator should use to measure this condition is out-of-scope of this version of the document.

In order to meet CPOs, in the case of an EVC that is composed of multiple OVCs, alignment of CBS between Operators and/or shaping at the ENNI is recommended. Otherwise, the EVC CPOs in Table 8 – Table 12 may not be met even if CoS Label mapping is aligned. In other words, the EVC performance may be impacted enough to cause performance results that miss some CPOs for the EVC or create the need to utilize a less stringent PT. For informative guidance on these issues see Burst Size and Shaper Considerations, Appendix G. In addition, Appendix H includes guidance (informative) on the choice of value for Burst Size (CBS).

Performance Metric	CoS Label H		CoS Label M		CoS Label L[1]	
	Pt-Pt	Multipt	Pt-Pt	Multipt	Pt-Pt	Multipt
FD (ms)	≤ 3	≤ 3	≤ 6	≤ 6	≤ 11	≤ 11
MFD (ms)	≤ 2	≤ 2	≤ 4	≤ 5	≤ 9	≤ 10
IFDV (ms)	≤ 1	≤ 1	≤ 2.5 or N/S	≤ 2.5 or N/S	N/S	N/S
FDR (ms)	≤ 1.25	≤ 1.25	≤ 3 or N/S	≤ 3 or N/S	N/S	N/S
FLR (percent)	≤ .001% i.e., 10^{-5}	≤ .001% i.e., 10^{-5}	≤ .001% i.e., 10^{-5}	≤ .001% i.e., 10^{-5}	≤ .1% i.e., 10^{-3}	≤ .1% i.e., 10^{-3}
Availability High Loss Interval (HLI) Consecutive HLI (CHLI) One Way Group Availability	N/S	N/S	N/S	N/S	N/S	N/S

[1] Ingress Bandwidth Profile parameters may be chosen such that no frames are subject to SLS.

Table 8: PT0.3 CPOs

Performance Metric	CoS Label H		CoS Label M		CoS Label L[1]	
	Pt-Pt	Multipt	Pt-Pt	Multipt	Pt-Pt	Multipt
FD (ms)	≤ 10	≤ 10	≤ 20	≤ 20	≤ 37	≤ 37
MFD (ms)	≤ 7	≤ ▯	≤ 13	≤ 15	≤ 28	≤ 30
One-way IFDV (ms)	≤ 3	≤ 3	≤ 8 or N/S	≤ 8 or N/S	N/S	N/S
FDR (ms)	≤ 5	≤ 5	≤ 10 or N/S	≤ 10 or N/S	N/S	N/S
FLR (percent)	≤ .01% i.e. 10^{-4}	≤ .01% i.e. 10^{-4}	≤ .01% i.e. 10^{-4}	≤ .01% i.e. 10^{-4}	≤ .1% i.e. 10^{-3}	≤ .1% i.e. 10^{-3}
Availability High Loss Interval (HLI) Consecutive HLI (CHLI) One Way Group Availability	N/S	N/S	N/S	N/S	N/S	N/S

[1] Ingress Bandwidth Profile parameters may be chosen such that no frames are subject to SLS.

Table 9: PT1 CPOs

Performance Metric	CoS Label H		CoS Label M		CoS Label L[1]	
	Pt-Pt	Multipt	Pt-Pt	Multipt	Pt-Pt	Multipt
FD (ms)	≤ 25	≤ 25	≤ 75	≤ 75	≤ 125	≤ 125
MFD (ms)	≤ 18	≤ 20	≤ 30	≤ 32	≤ 50	≤ 52
One-way IFDV (ms)	▯▯8	▯▯8	≤ 40 or N/S	≤ 40 or N/S	N/S	N/S
FDR (ms)	≤ 10	≤ 10	≤ 50 or N/S	≤ 50 or N/S	N/S	N/S
FLR (percent)	≤ .01% i.e., 10^{-4}	≤ .01% i.e., 10^{-4}	≤ .01% i.e., 10^{-4}	≤ .01% i.e., 10^{-4}	≤ .1% i.e., 10^{-3}	≤ .1% i.e., 10^{-3}
Availability High Loss Interval (HLI) Consecutive HLI (CHLI) One Way Group Availability	N/S	N/S	N/S	N/S	N/S	N/S

[1] Ingress Bandwidth Profile parameters may be chosen such that no frames are subject to SLS.

Table 10: PT2 CPOs

Performance Metric	CoS Label H		CoS Label M		CoS Label L[1]	
	Pt-Pt	Multipt	Pt-Pt	Multipt	Pt-Pt	Multipt
FD (ms)	≤ 77	≤ 77	≤ 115	≤ 115	≤ 230	≤ 230
MFD (ms)	≤ 70	≤ 72	≤ 80	≤ 82	≤ 125	≤ 127
One-way IFDV (ms)	≤ 10	≤ 10	≤ 40 or N/S	≤ 40 or N/S	N/S	N/S
FDR (ms)	≤ 12	≤ 12	≤ 50 or N/S	≤ 50 or N/S	N/S	N/S
FLR (percent)	≤ .025% i.e., 2.5×10^{-4}	≤ .025% i.e., 2.5×10^{-4}	≤ .025% i.e., 2.5×10^{-4}	≤ .025% i.e., 2.5×10^{-4}	≤ .1% i.e., 10^{-3}	≤ .1% i.e., 10^{-3}
Availability High Loss Interval (HLI) Consecutive HLI (CHLI) One Way Group Availability	N/S	N/S	N/S	N/S	N/S	N/S

[1] Ingress Bandwidth Profile parameters may be chosen such that no frames are subject to SLS.

Table 11: PT3 CPOs

Performance Metric	CoS Label H		CoS Label M		CoS Label L[1]	
	Pt-Pt	Multipt	Pt-Pt	Multipt	Pt-Pt	Multipt
FD (ms)	≤ 230	≤ 230	≤ 250	≤ 250	≤ 390	≤ 390
MFD (ms)	≤ 200	≤ 202	≤ 220	≤ 222	≤ 240	≤ 242
One-way IFDV (ms)	≤ 32	≤ 32	≤ 40 or N/S	≤ 40 or N/S	N/S	N/S
FDR (ms)	≤ 40	≤ 40	≤ 50 or N/S	≤ 50 or N/S	N/S	N/S
FLR (percent)	≤ .05% i.e., $5x10^{-4}$	≤ .05% i.e., $5x10^{-4}$	≤ .05% i.e., $5x10^{-4}$	≤ .05% i.e., $5x10^{-4}$	≤ .1% i.e., 10^{-3}	≤ .1% i.e., 10^{-3}
Availability High Loss Interval (HLI) Consecutive HLI (CHLI) One Way Group Availability	N/S	N/S	N/S	N/S	N/S	N/S

[1] Ingress Bandwidth Profile parameters may be chosen such that no frames are subject to SLS.

Table 12: PT4 CPOs

10. References

[1] MEF 10.3, "Ethernet Services Attributes Phase 3"

[2] IEEE 802.1Q – 2014, "IEEE Std 802.1Q™– 2014, IEEE Standards for Local and metropolitan area networks— Bridges and Bridged Networks, 3 November 2014"

[3] RFC 2119, "Key words for use in RFCs to Indicate Requirement Levels", S. Bradner

[4] MEF 4, "Metro Ethernet Network Architecture Framework - Part 1: Generic Framework"

[5] Inter-provider Quality of Service, MIT Communications Futures Program, 2006

[6] ITU-T Recommendation Y.1541, "Network performance objectives for IP-based services", December 2011

[7] MEF 3, "Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks"

[8] RFC 2597, "Assured Forwarding PHB Group", Heinanen

[9] ITU-T Recommendation I.356, "B-ISDN ATM layer cell transfer performance," March 2000

[10] MEF Technical Specification MEF 26.2, "External Network Network Interface (ENNI) and Operator Service Attributes"

[11] ITU-T Recommendation Y.1540, "Internet protocol data communication service – IP packet transfer and availability performance parameters", March 2011

[12] MEF 45, "Multi-CEN L2CP"

[13] MEF 12.2, "Carrier Ethernet Network Architecture Framework Part 2: Ethernet Services Layer"

[14] ITU-T Recommendation Y.1563, "Ethernet frame transfer and availability performance", January 2009

[15] Internet Engineering Task Force RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," December 1998

[16] MEF 6.2, "EVC Ethernet Services Definitions Phase 3"

[17] MEF 51, "OVC Services Definitions"

[18] MEF 13, "User Network Interface (UNI) Type 1 Implementation Agreement"

[19] Guido Appenzeller, Isaac Keslassy, and Nick McKeown. Sizing Router Buffers. Proceedings of SIGCOMM 2004, Portland OR, USA, August 30-September 3, 2004

Appendix A Performance Tier Model Derivation (Informative)

Assumptions for PTs:

- PT distances represent the path a frame would traverse and thus drive associated propagation delay minimums for FD/MFD/FDR
- Though number of switch hops generally increases with longer distance PTs, hops will not be quantified
- For simplicity, PT CPOs are expressed as constants based on the maximum distance for the PT rather than formulas with distance variables
- PTs are derived with certain distance and application assignments
- PTs can be arbitrarily assigned to given services by Operators based on factors in or outside the scope of this IA
- All links, including access links, will have a link speed of at least 10 Mb/s, with the notion that a given service may utilize a "higher" PT for slower links based on Operator discretion. For PT0.3, the minimum link speed is 1 Gbps.

A five PT model is chosen to allow for sufficient granularity and cover range from small area networks and applications to global. This IA uses distance as the primary means of describing PTs. Below are the five PTs defined in this IA with the format: PT Number (PT Name) - Description (distance, derived propagation delay used in CPO constraints to establish a minimum per PT).

- PT0.3 (City PT) – derived from sub-Metro distances (<75 km, 0.5ms*)
- PT1 (Metro PT) – derived from Metro distances (<250 km, 2 ms*)
- PT2 (Regional PT) - derived from Regional distances (<1200 km, 8 ms*)
- PT3 (Continental PT) - derived from National/Continental distances (<7000 km, 44 ms*)
- PT4 (Global PT) – derived from Global/Intercontinental distances (<27500 km, 172 ms*)
    - Based on I.356 [9].

*Minimum Frame Delay based on distance * .005 ms/km * 1.25 where distance is in kilometers (km), .005 ms/km propagation delay and 1.25 is route/airline distance ratio. Distance is difficult to ascertain in real-networks as path (i.e., circuit) distance is unknown or may vary due to routing or other path changes (e.g., dynamic control protocols). In real CENs there may be additional delays (e.g., switch hops, buffering, shaping, serialization for low speed links).

An Operator's Ethernet service compliance with this IA does not depend on adherence to PT distances. As stated in the normative sections, a given service may utilize a particular PT for reasons other than EI to EI distance of the service.

A.1 Low Speed Link Considerations

Low speed access and internal links in a CEN can have a significant impact on frame delay. In CoS IA this is accounted for by the choice of PT for a service or UNI pair within a service. This is simpler than a Low Speed Factor that is applied to the CPO per CoS Label. For example, if a service would otherwise utilize PT1 CPOs it could utilize PT2 due to its use of sub-10Mb/s low speed links in the access between the NID and the core of the CEN. Additional low speed performance considerations are contained in [6] and [5].

Appendix B Ethernet Network Section Model – Composing UNI-UNI Values (Informative)

ITU-T Recommendation Y.1541 [6] defines methods for concatenating performance objectives or measurements associated with network sections, thus combining their performance to estimate the complete path (i.e., composing). This Appendix reproduces the equations using MEF variables where possible and uses MEF terminology whereby ENS replaces the term network sections used in [6].

While these methods are applicable to both objective setting and measurements, the methods are often not needed for measurements if ENS (e.g., UNI-ENNI for the OVCs) and end-to-end (e.g., UNI-UNI for the EVC) measurements are available.

When combining the metrics based on percentiles, it is a gross over-estimate to simply add the performance values for each ENS. However, there may be circumstances when even this over-estimate will suffice. For example, consider two ENSs, each of which has FDR of 2 ms. If the Subscriber is satisfied with 4 ms, simple addition could suffice. If the Subscriber requires 3 ms, then simple addition is not sufficient.

This IA provides no direct method of calculating allocation but the concatenation methods can be used to evaluate proposed OVC ENS CPOs against an EVC CPOs and through iteration adjust EVC or OVC objectives to guide the determination of OVC CPOs. Iteration is practical based on a small range/set of potential CPOs for the OVCs under consideration and a small number of ENS (i.e., usually 2-4).

The following table illustrates the mapping used, to the extent possible. Note that many ITU-T variables do not have a counterpart in MEF and that [6] does not address a metric equivalent to the MEF One-way IFDV.

Metric/Parameter	MEF 23.2/26.2	Y.1541	Notes
UNI-UNI One-way Delay Distribution		T	No MEF equivalent
SLS Interval	T		No ITU-T equivalent
Subset of ordered UNI pairs	S		No ITU-T equivalent
k^{th} Network Section		k	No MEF equivalent
Mean One-way Delay	$\bar{\mu}_{TS}$	μ_k	
Variance of One-way Delay		σ_k^2	No MEF equivalent
Probability or Percentile of interest	P_d or P_r	p	P_d for Frame Delay and or P_r for Frame Delay Range
Delay at Percentile	$\bar{d}_{TdS}$, $\bar{d}_{TrS}$ or $\bar{d}_{TRS}$	t_k, t	Frame Delay (d), or Frame Delay (r),

Metric/Parameter	MEF 23.2/26.2	Y.1541	Notes
			Frame Delay Range (R); t_k & t are values of delay used in the Steps below
Skewness		γ_k	No MEF equivalent
Third moment		ω_k	No MEF equivalent
Value of the standard normal distrib. at *p*		x_p	No MEF equivalent
Loss Ratio	$FLR_{T,S}$	$IPLR_k$	

Table 13: MEF – ITU Variable Mapping

B.1 Mean delay

For the Mean Frame Delay (MFD), or $\bar{\mu}_{TS}$ performance parameter (grouped with performance metrics in this IA), the UNI-UNI performance is the sum of the means contributed by Ethernet Network Sections.

$$\bar{\mu}_{TS} = \bar{\mu}_{TS1} + \bar{\mu}_{TS2} + \bar{\mu}_{TS3} + ... + \bar{\mu}_{TSn}$$

The units of $\bar{\mu}_{TS}$ values are seconds.

Note that the definition of delay in MEF per [1] and [10] includes the delay incurred in traversing the External Interface thus the calculated delay for the UNI-UNI using this concatenation method will be overstated. The sum of per OVC delays will be greater than the UNI-UNI delay (see the formula at the bottom of page 31, section 7.2.16.1, in [10]). In general this overstatement is likely to be small in terms of modeling objectives and in terms of measurements may not be feasible to capture precisely as defined. This is not addressed in this phase of CoS IA.

B.2 Loss ratio

For the Frame Loss Ratio ($FLR_{T,E}$) performance metric, the UNI-UNI performance may be estimated by inverting the probability of successful frame transfer across *n* Ethernet Network Sections (*En*), as follows:

$$FLR_{T,E} = 1 - \{(1 - FLR_{T,E1}) \times (1 - FLR_{T,E2}) \times (1 - FLR_{T,E3}) \times ... \times (1 - FLR_{T,En})\}$$

This relationship does not have limits on the parameter values, so it is preferred over other approximations, such as the simple sum of loss ratios.

The units of $FLR_{T,E}$ values are lost Qualified Frames per total Qualified Frames sent. This is equivalent to MEF FLR except that it is not expressed as a percentage.

B.3 Relationship for delay and delay range

The relationship for estimating the UNI-UNI Frame Delay ($\bar{d}_{TdS}$) or the Frame Delay Range ($\bar{d}_{TRS}$) performance from the Ethernet Network Section values must recognize their sub-additive nature and is difficult to estimate accurately without considerable information about the individual delay distributions. If, for example, characterizations of independent delay distributions are known or measured, they may be convolved to estimate the combined distribution. This detailed information will seldom be shared among Operators, and may not be available in the form of a continuous distribution. As a result, the UNI-UNI delay estimation may have accuracy limitations.

The relationship for combining Frame Delay at P_d, or Frame Delay Range (i.e., delay at P_r less minimum delay) values is given below. Note that P_d parameter value is equal to P_r parameter value for this IA for a given CoS Label and PT.

The problem under consideration can be stated as follows: estimate the quantile $\bar{d}_{TRS}$ of the UNI-UNI Frame Delay Range T as defined by the condition:

$\Pr(T < \bar{d}_{TRS}) = p$ where $p = P_r/100$ for UNI-UNI Frame Delay Range.

A similar relation for UNI-UNI Frame Delay would be based on $\bar{d}_{TdS}$ and $p=P_d/100$.

When using the methods below to calculate Frame Delay Range, the calculations are based on using the difference between the delay and the minimum delay. In other words, all delay values are normalized by removing the minimum delay observed over T.

Step 1

Measure the mean and variance for the delay for each of n Ethernet Network Sections. Estimate the mean and variance of the UNI-UNI delay by summing the means and variances of the component distributions.

$$\bar{\mu}_{TS} = \sum_{k=1}^{n} \bar{\mu}_{TSk}$$

$$\sigma^2 = \sum_{k=1}^{n} \sigma_k^2$$

Step 2

Measure the quantiles for each delay component at the probability of interest, e.g., $P_d = 99.9$ and p = 0.999. Estimate the corresponding skewness and third moment using the formula shown below, where $x_{0.999} = 3.090$ is the value satisfying $\Phi(x_{0.999}) = 0.999$ where Φ denotes the standard normal (mean 0, variance 1) distribution function. Note that $\chi_{0.999}$ is an example based on 99.9th percentile. This IA also recommends use of other percentiles including 95th and 99th which yield $x_{0.95}$= 1.645 and $x_{0.99}$= 2.33.

$$\gamma_k = 6 \cdot \frac{x_P - \frac{t_k - \bar{\mu}_{TSk}}{\sigma_k}}{1 - x_P^2}$$

where t_k represents delay at x_P based on $P_d/100$ for Frame Delay or where t_k represents delay less minimum delay at x_P based on $P_r/100$ for Frame Delay Range.

$$\omega_k = \gamma_k \cdot \sigma_k^3$$

Assuming independence of the delay distributions, the third moment of the UNI-UNI delay is just the sum of the Ethernet Network Section third moments.

$$\omega = \omega_1 + \omega_2 + \omega_3 + \ldots = \sum_{k=1}^{n} \omega_k$$

The UNI-UNI skewness is computed by dividing by σ^3 as shown below.

$$\gamma = \omega / \sigma^3$$

Step 3

The estimate of the 99.9-th percentile ($p = 0.999$) of UNI-UNI delay, $\bar{d}_{TdS}$ or $\bar{d}_{TRS}$ is represented by t as follows:.

$$t = \bar{\mu}_{TS} + \sigma \cdot \left\{ x_P - \frac{\gamma}{6}\left(1 - x_P^{\ 2}\right) \right\}$$

where t represents $\bar{d}_{TdS}$ at x_P based on $P_d/100$ for Frame Delay or where t represents $\bar{d}_{TRS}$ at x_P based on $P_r/100$ for Frame Delay Range.

B.4 Ethernet Network Section Recommendations

Below are recommendations for how to apply the concatenation methods in section Appendix B.

- Suggest that the choice of MFD and/or FD metrics be the same for each OVC that is a component of the EVC and the same for the EVC CPOs.
- Suggest that the FDR Performance be used for each OVC that is a component of the EVC and for the EVC CPOs.
- Suggest that the choice of Parameter values for the Performance metrics from Table 6 and Table 7 (as appropriate) be the same for each OVC that is a component of the EVC and the same for the EVC.

- Suggest that the SLS time interval, T, be the same for each OVC that is a component of the EVC.
- Suggest that the SLS time interval, T, be aligned for each OVC that is a component of the EVC. This implies that for all pairs of OVCs, (A,B), $t_{sA} = t_{sB} \pm (n{*}T)$ for some value n.

Appendix C Key Applications to Derive Performance Requirements (Informative)

The intent of the CoS IA is to provide sufficient CoS Labels and Performance Objectives to efficiently support the vast majority of well-known applications. Identification of the applications supported, quantification of CPOs, specification of associated parameters (e.g., *P*, *T*, etc) and mapping to CoS Labels is described in this section.

Application mapping is for the purpose of determining the quantitative Performance Objectives for each CoS Label. It is not intended to mandate how an Operator, Service Provider or Subscriber maps a particular instance of an application. For example, a Subscriber could map some VoIP for certain types of communication to CoS Label L and other VoIP to CoS Label H if desired. This IA is constructed such that VoIP (of the high-quality type defined in this appendix) will be supported in the CoS Label it is mapped to if the Operator conforms with this IA for that CoS Label. The proposed mapping shows how the CoS Performance Objectives are derived and not meant to imply a requirement for application mapping in actual implementations.

Similar to Application mapping, L2CP needs to be mapped to CoS Labels. There may be different CoS Labels for different L2CP types. At a minimum, there is a need to specify a CoS Label that meets the L2CP application requirements.

The applications considered in the process of generating CPOs and mapping requirements to CoS Labels are shown in Table 14. The applications fall into three general user segments: Consumer, Business, and Mobile. The user segments are not mutually exclusive, and many applications are aligned with more than one segment.

Application	Consumer	Business	Mobile
VoIP Data	X	X	X
Interactive Video (Video Conferencing)	X	X	?
VoIP and Video Signaling	X	X	X
Web Browsing	X	X	X
IPTV Data Plane	X	X	?
IPTV Control Plane	X	X	?
Streaming Media	X	X	X
Interactive Gaming	X		X
Best Effort	X	X	X
Circuit Emulation		X	X
Telepresence		X	
Remote Surgery (Video)		X	
Remote Surgery (Control)		X	
Telehealth (Hi-res image file transfer)		X	

Application	Consumer	Business	Mobile
Email	X	X	X
Broadcast Engineering (Pro Video over IP)		X	
CCTV	X	X	X
Financial/Trading		X	
Database		X	
Real Time Fax over IP	X	X	
Store and Forward Fax over IP	X	X	
SANs (Synchronous Replication)		X	
SANs (Asynchronous Replication)		X	
Wide Area File Services		X	
Network Attached Storage	X	X	
Text Terminals (telnet, ssh)		X	
Graphics Terminals (Thin Clients)		X	
Point of Sale Transactions		X	
E-Commerce (Secure transactions)	X	X	X
Mobile Backhaul System Requirements			X

Table 14: Application list

C.1 Application-specific Performance Objectives

Each of the applications listed in Table 14 was researched to determine the performance requirements associated with the application and the corresponding application-specific Performance Objectives associated with CEN Performance metrics. The requirements for application performance are usually specified from end-to-end. Since the CEN of interest may only be a portion of the end-to-end network which can also include customer network segments and endpoint devices, allocation or budgeting of the objective is generally required as the application-specific Performance Objectives are quantified. In addition, application level requirements for zero loss frequently assume the use of a loss recovery mechanism such as TCP operating above the CEN.

Table 15 through Table 35 show the requirements compiled for each application. Each table comprises two or three general sections. The top section provides application-level requirements and supporting measurement parameters compiled directly from the available sources. The second section maps the application level requirements to application-specific Performance Objective values for each CEN Performance metric and applies the appropriate parameters to each metric. The third section (if present) provides supplementary information about the application.

Application requirements were compiled from a variety of public sources. The first and most desirable category for source references is standards-based. Where standards-based references are unavailable, industry-based Best Practices are used, as well as vendor-specific and product-specific information. The sources for all application requirements are provided in their respective tables.

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One-way delay**	< 150 ms preferred < 400 ms limit	G.1010, TS 22.105	Total mouth-to-ear, includes encoding, decoding, and all buffering in addition to network delays.
		< 150 ms	TR-126	
	Delay variation	< 1 ms	G.1010, TS 22.105	Total mouth-to-ear, achieved using de-jitter buffer in receiver.
	Meas. Params.	$T \approx 1$ minute $P = 0.999$	Y.1541 Y.1541	Suggested value (section 5.3.2) Table 1/Y.1541
Appl. Perform. Objectives	**FLR**	< 3e-2	G.1010, TS 22.105	Assumes use of a packet loss concealment algorithm to minimize effect of packet loss.
	FD	< 125 ms preferred < 375 ms limit	See text	$P_d = 0.999$
	FDR	< 50 ms	Y.1541	$P_r = 0.999$
	MFD	< 100 ms preferred < 350 ms limit	See text	
	IFDV	< 40 ms		$P_v = 0.999$
Info	**Bit rates**	4 to 64 kbps	G.1010	
	Frame sizes	≤ 200 bytes		200 bytes based on G.711 with 20 ms frames. Most other codecs result in equal or smaller frame sizes.
	Availability	≥ 99.99%	TR-NWT-000418, TA-NWT-000909	Bellcore standard for the PSTN (quoted from TR-126).

Table 15: VoIP Parameters

The values in Table 15 provide an example of how application level requirements are mapped to application-specific Performance Objectives. The preferred value for one way delay for VoIP is 150 ms. The scope of this parameter includes everything between the talker's mouth and the listener's ear – the microphone, analog-to-digital conversion, speech encoding, buffering and framing, network delays, dejitter buffering, decoding, digital-to-analog conversion, and the speaker which converts the decoded analog signal to sound waves. Of all these elements, only network delays are within the scope of the CEN.

Typical non-network delays are identified and summed with guidance from ITU-T Recommendations G.114 and Y.1541. Per G.114, the buffering and framing delays associated with a G.711 encoder with 20 ms voice frames is 20.125 ms. Using Table VII.2/Y.1541 in Appendix VII of Y.1541 for guidance, a dejitter buffer of 50 ms is assumed and half of that value (25 ms) is allocated as its contribution to mean delay. A total of 5 ms is used for the contributions of other processes and equipment, for a total non-network contribution of approximately 50 ms to mean delay. The resulting Mean Frame Delay that can be allocated to the CEN as a Performance metric is 100 ms.

Frame Delay is mapped using a similar process. In this case, all non-network sources of delay except for the dejitter buffer are subtracted from the application parameter. The dejitter buffer acts to "smooth out" the variation in received voice frames resulting from network jitter. As a result, frames that arrive at the receiver with minimum delay are held in the dejitter buffer for its maximum duration, and frames arriving at the receiver at the maximum end of the jitter range are forwarded immediately, with no added delay in the dejitter buffer. Since the non-network delays (not including the dejitter buffer) total approximately 25 ms, the preferred value of 150 ms for one way application delay maps to a Frame Delay (at P_d = 0.999, close to the maximum value) of approximately 125 ms.

Application level parameters are mapped to Performance Objectives in Table 16 through Table 35 using the process described in the above example. Where source data is available, recommended measurement parameter values are also provided.

Real-time and streaming applications typically make use of a dejitter buffer such as that described above in the VoIP example. For those applications, frames which do not arrive at the dejitter buffer within a delay window corresponding to the length of the buffer are likely to be discarded. As a result, there is an implicit relationship between the percentile valued parameters used to define maximum delay or jitter (P_d for Frame Delay, P_v for Inter Frame Delay Variation and P_r for Frame Delay Range) and the Frame Loss Ratio for those types of applications, since frames which arrive too late to be accepted into the dejitter buffer are effectively lost to the application. The relationship is:

$$P_r \text{ (or } P_v \text{ or } P_d) = 1 - \text{FLR}.$$

For real-time and streaming applications in the tables below, the above relationship has been used to derive P_r or P_v if recommended values for the parameters are not directly available from the source documentation.

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One-way delay**	< 150 ms preferred < 400 ms limit	G.1010, TS 22.105	Total user-to-user, includes encoding, decoding, and all buffering in addition to network delays.
	Delay variation	< 1 ms	G.1010	Total user-to-user, achieved using de-jitter buffer in receiver.
	Meas. Params.	$T \approx$ 1 minute P = 0.999	Y.1541 Y.1541	Suggested value (section 5.3.2) Table 1/Y.1541
Appl. Perform. Objectives	**FLR**	< 1e-2	G.1010, TS 22.105	Assumes use of a packet loss concealment algorithm to minimize effect of packet loss.
	FD	< 125 ms preferred < 325 ms limit		P_d = 0.999
	MFD	< 100 ms preferred < 350 ms limit		Network and de-jitter delays similar to VoIP case H.264 supports sub-frame encoding/decoding delays (20 ms used for conversion)
	FDR	< 50 ms	Y.1541	P_r = 0.999
	IFDV	< 40 ms		P_v = 0.999
Info	**A/V synch**	< 80 ms	G.1010	
		< 100 ms	TS 22.105	
	Bit rates	16 to 384 kbps	G.1010	
		32 to 384 kbps	TS 22.105	
		Up to ≈ 2 Mbps	H.264	Configurable to 2 Mbps in current applications
	Frame sizes	≤ 1500 bytes		
	Availability			Not specified

Table 16: Interactive Video Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**Latency**	< 200 ms	TR-126	Further detail unspecified in source, interpreted as upper bound on network delay.
	Jitter	< 50 ms	TR-126	
	Packet Loss Rate	< 5.26E-6	TR-126	End-to-end application layer objective. Minimum value from TR-126 Tables 12 and 13. Assumes no or minimal loss concealment (tolerable loss rates may be higher depending on degree and quality of STB loss concealment).
Appl. Perform. Objectives	**FLR**	< 1E-3	Y.1541 Amd. 3	Network objective assuming Application Layer Forward Error Correction (AL-FEC) sufficient to provide application layer packet loss rate objective.
	FDR	< 50 ms	Y.1541	Assumes AL-FEC sufficient to provide application layer packet loss rate objective. P_r = 0.999*
	MFD	< 100 ms	See Notes	Encoding delay not included. Allow 100 ms for de-jitter buffer, decoding and AL-FEC delays.
	FD	< 125 ms		P_d = 0.999*
	IFDV	< 40 ms		P_v = 0.999*
Info	**Bit rates (MPEG-2)**	3 to 5 Mbps	TR-126	From TR-126 Table 12
	Bit rates (MPEG-4)	1.75 to 3 Mbps	TR-126	From TR-126 Table 13
	Frame sizes	≤ 1500 bytes		
	Availability	≥ 99.99%	TR-126	

*No direct reference for percentiles, but dejitter buffering is required

Table 17: Standard Definition Video Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**Latency**	< 200 ms	TR-126	Further detail unspecified in source, interpreted as upper bound on network delay.
	Jitter	< 50 ms	TR-126	
	Packet Loss Rate	< 1.16E-6	TR-126	End-to-end application layer objective. Minimum value from TR-126 Tables 14 and 15. Assumes some loss concealment.
Appl. Perform. Objectives	**FLR**	< 1E-3	Y.1541 Amd 3	Network objective assuming AL-FEC sufficient to provide application layer packet loss rate objective.
	FDR	< 50 ms	Y.1541	Assumes AL-FEC sufficient to provide application layer packet loss rate objective. P_r = 0.999*
	MFD	< 100 ms	See Notes	Encoding delay not included. Allow 100 ms for de-jitter buffer, decoding and AL-FEC delays.
	FD	< 125 ms		P_d = 0.999*
	IFDV	< 40 ms		P_v = 0.999*
Info	**Bit rates (MPEG-2)**	15 to 18.1 Mbps	TR-126	From TR-126 Table 14
	Bit rates (MPEG-4)	8 to 12 Mbps	TR-126	From TR-126 Table 15
	Frame sizes	≤ 1500 bytes		
	Availability	≥ 99.99%	TR-126	

*No direct reference for percentiles, but dejitter buffering is required

Table 18: High Definition Video Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**Delay**	< 10 s	G.1010, TS 22.105	Further detail unspecified in source, interpreted as time from request to initiation of playout.
	Delay Variation	<< 1 ms	G.1010	Value specified in G.1010 for audio as parameter at ear (post de-jitter buffer). Unspecified for video.
Appl. Perform. Objectives	**FDR**	< 2 s	TS 22.105	Transport path, implies a 2 s de-jitter buffer. P_r values unspecified in source.
	FLR	< 1%	G.1010	
	MFD			Not specified
	FD			Not specified
	IFDV	< 1.5 s		$P_v = 0.99$*
Info	**Bit rates (audio)**	16 to 128 kbps	G.1010	
		5 to 128 kbps	TS 22.105	
	Bit rates (video)	16 to 384 kbps	G.1010	
		20 to 384 kbps	TS 22.105	
		Up to 2+ Mbps		Measured video playout rates
	Frame sizes	≤ 1500 bytes		
	Availability			Not specified

*No direct reference for percentiles, but dejitter buffering is required

Table 19: Internet Streaming Audio/Video Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One way delay**	< 250 ms	G.1010, TS 22.105	Telemetry/two-way control/command and control category.
	IPTV control plane response	< 200 ms	TR-126	Set-top box (STB) command processing - time interval between the remote control action (button push) and GUI update. May include middleware server processing time for some functions.
	Channel change response	< 2 s	TR-126	Remote button to stable video on new channel.
	Delay Variation	N.A.	G.1010, TS 22.105	
	Loss	0	G.1010, TS 22.105	
Appl. Perform. Objectives	**FDR**	N.A.	G.1010, TS 22.105	
	FLR	1e-3	G.1010, TS 22.105	Assumes TCP or other loss recovery
	MFD	< 75 ms		Uses STB command processing with middleware server processing as worst case. Allocates 50 ms to combined STB/middleware server processing, 150 ms to round trip delay.
	FD	N.A.		
	IFDV	N.A.		
Info	**Bit rates**	< 1 kbps	G.1010	
		< 28.8 kbps	TS 22.105	
	Frame sizes	≤ 1500 bytes		
	Availability	≥ 99.99%	TR-126	Same as VoIP and SD/HD Video data plane requirements.

Table 20: Interactive Transaction Data Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One way delay**	< 200 ms	G.1010	TR-126 refers to this value as "likely too high."
		< 75 ms preferred	TS 22.105	
		< 50 ms objective	TR-126	Includes application layer (game server and game client) and network layer delays.
	Delay Variation	N.A.	G.1010, TS 22.105	
		< 10 ms objective	TR-126	
	Loss	0	G.1010	
Appl. Perform. Objectives	**FDR**	< 10 ms objective	TR-126	
	MFD	< 40 ms objective		TR-126 does not provide typical client/server delays. 10 ms used as a strawman value for the combination.
	FLR	1e-3	G.1010	Assumes TCP or other loss recovery
	FD	< 50 ms objective		
	IFDV	< 8 ms objective		
Info	**Data**	< 1 KB	G.1010, TR-126	Data per transaction.
	Bit rates	< 60 kbps	TS 22.105	
	Frame sizes	≤ 1500 bytes		
	Availability			Not specified

Table 21: Interactive Gaming Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**Web browsing response time**	< 2 s/page preferred < 4 s/page acceptable	G.1010, TR-126	Multiple round trip delays for most web pages imply requirement for MFD of less than 100 ms to meet 4 s response time. Typical page size of ≈10 kbytes specified. Current page sizes range from ≈20 kbytes to >1 Mbyte.
		< 4 s/page	TS 22.105	Multiple round trip delays for most web pages imply requirement for MFD of less than 100 ms to meet 4 s response time.
	Transaction services (e.g., e-commerce)	< 2 s preferred < 4 s acceptable	G.1010	Multiple round trip delays for most web pages imply requirement for MFD of less than 100 ms to meet 4 s response time.
		< 4 s	TS 22.105	Multiple round trip delays for most web pages imply requirement for MFD of less than 100 ms to meet 4 s response time.
Appl. Perform. Objectives	**FDR**	N.A.	G.1010, TR-126, TS 22.105	
	FLR	N.A.	G.1010, TS 22.105	
	MFD	N.A.		Not specified
	FD	N.A.		
	IFDV	N.A.		
Info	**Frame sizes**	≤ 1500 bytes		
	Availability			Not specified

Table 22: Best Effort Parameters

Category	Param.	Value	Source	Notes
Appl. Req's.	**FD**	25 ms	MEF 3	P_d = 99.9999%
	Packet loss	1e-5 to 1e-7	MEF 3	Dependent on TDM service
	Jitter	10 ms	MEF 3	P = 99.9999%
Appl. Perform. Objectives	**FLR**	1E-6		
	FDR	15 ms	Inferred from IFDV	P_r = 99.9%
	MFD	20 ms	Inferred from FD, IFDV	
	IFDV	10 ms	MEF 8	P_v = 99.9%, Δt = 900s, T = 3600s
	FD	25 ms	MEF 3	P_d = 99.9999%

Table 23: Circuit Emulation Parameters

Circuit Emulation is further defined in [7].

Category	Param.	Value	Source	Notes
Appl. Req's.	**Delay**	< 2 s preferred < 4 s acceptable	G.1010	Transaction services
	Packet loss	0	G.1010	Transaction services Application level requirement
	Jitter	N.A.	G.1010	Transaction services
Appl. Perform. Objectives	**FLR**	1e-3	Y.1541 Class 3	
	FDR	Not specified		
	MFD	1 s		
	IFDV	Not specified		
	FD	2 s		

Table 24: Point of Sale Transaction Parameters

Category	Param.	Value	Source	Notes
Appl. Req's.	**RTT**	10 ms	IBM/Cisco SAN Multiprotocol Routing IBM Redbook SG24-7543-01	Round trip Includes jitter
		5 ms	EMC SRDF Connectivity Guide	Best practice
		15 ms	IBM/Brocade SAN Multiprotocol Routing IBM Redbook SG24-7544-01	Referring to iSCSI implementation
	Packet loss	0.1% limit 0.01% rec.	EMC SRDF Connectivity Guide	Network requirement
		0.01% rec.	IBM SAN Multiprotocol Routing IBM Redbook SG24-7321-00	Network requirement
	Jitter	25% of latency or 25 ms	EMC SRDF Connectivity Guide	Use the lower value
Appl. Perform. Objectives	**FLR**	≤ 1e-4		
	FDR	≤ 1.25 ms		25% of 5 ms (one way)
	MFD	≤ 3.75 ms		75% of 5 ms (one way)
	IFDV	≤ 1 ms		
	FD	≤ 5 ms		

Table 25: Synchronous Replication Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	80 ms	IBM SAN Volume Controller Configuration Guide IBM Redbook SC23-6628-02	Round trip, Includes jitter SVC version 4.1.1 or higher
		200 ms	EMC SRDF Connectivity Guide	Round trip
	Packet loss	1% limit 0.01% rec.	EMC SRDF Connectivity Guide	Network requirement
		0.01% rec.	IBM SAN Multiprotocol Routing IBM Redbook SG24-7321-00	Network requirement
	Jitter	25% of latency or 25 ms	EMC SRDF Connectivity Guide	Use the lower value
Appl. Perform. Objectives	**FLR**	≤ 1e-4		
	FD	≤ 40 ms		
	MFD	≤ 30 ms		75% of 40 ms (one way)
	FDR	≤ 10 ms		25% of 40 ms (one way)
	IFDV	≤ 8 ms		

Table 26: Asynchronous Replication Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**Delay**	15 s preferred 60 s acceptable	G.1010 bulk data	Time for entire file to transfer
	Packet loss	0	G.1010 bulk data	Application level requirement
	Jitter	N.A.	G.1010 bulk data	
Appl. Perform. Objectives	**FLR**	≤ 1e-3	Y.1541 Class 4	Assumes reliable delivery protocol (e.g., TCP)
	FDR	Unspecified		
	MFD	≤ 1 s	Y.1541 Class 4	
	IFDV	Unspecified		
	FD	Unspecified		

Table 27: Network Attached Storage Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One way delay**	< 200 ms	G.1010	
	Packet loss	0	G.1010	At application layer
	Jitter	N.A.	G.1010	
Appl. Perform. Objectives	**FLR**	1e-3	Y.1541 Class 3	Assumes TCP
	FDR	Unspecified		
	MFD	< 200 ms		
	IFDV	Unspecified		
	FD	Unspecified		

Table 28: Text and Graphics Terminal Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**One-way delay**	< 400 ms	G.1010	VoIP "acceptable" value
	Delay variation	< 1 ms	G.1010, TS 22.105	Achieved using de-jitter buffer in T.38 gateway
Appl. Perform. Objectives	**FLR**	< 3e-2	G.1010, TS 22.105	RTP, UDPTL, TCP all provide protection against frame loss
	FDR	< 50 ms	Y.1541	P_r = 0.999
	MFD	< 350 ms		From VoIP "acceptable" value
	IFDV	< 40 ms	Y.1541	P_v = 0.999
	FD	< 400 ms	Y.1541	From VoIP "acceptable" value P_d = 0.999

Table 29: T.38 Fax Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	< 3 ms	IBM System Storage Business Continuity Planning Guide	Synchronous copy / replication
		≤ 10 ms	Oracle Configuration Best Practices	Synchronous multiple log writer (LGWR) process
		≤ 12 ms	Oracle9i Data Guard Best Practice	Physical standby database distance
		≤ 100 ms	Active/Active clusters in SQL Server	Server Clustering
	Packet loss	0	G.1010	Transaction Service
	Jitter	N.A	G.1010 Transaction services	
Appl. Perform. Objectives	**FLR**	1e-5	Y.1541 TCP Performance	
	FD	≤ 5 ms		
	MFD	N/S		
	FDR	N/S		
	IFDV	N/S		

Table 30: Database Parameters – Hot Standby

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	≤ 100 ms	Oracle9i Data Guard: Primary Site and Network Cfg BP	Asynchronous LGWR process
		100 ms	Active/Active clusters in SQL Server	Server Clustering
	Packet loss	1e-5	Y.1541 TCP Performance	
	Jitter	N.A	G.1010 Transaction services	
Appl. Perform. Objectives	**FLR**	10^{-5}	Y.1541 TCP Performance	
	FD	≤ 50 ms		
	MFD	N/S		
	FDR	N/S		
	IFDV	N/S		

Table 31: Database Parameters – WAN Replication

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	≤ 300 ms	Oracle On Demand Reference Guide	End user to Oracle hosted servers
		≤ 2 s	G.1010 Transaction services	Preferred < 2 s Acceptable < 4 s
		≤ 7 s	Zona Research	eCommerce threshold (abandon rate)
	Packet loss	≤ 0.1%	Oracle On Demand Reference Guide	End user to Oracle hosted servers
		zero	G.1010 Transaction services	
	Jitter	N.A.	G.1010 Transaction services	
Appl. Perform. Objectives	**FLR**	1e-3	Y.1541 Class 3 (Transaction data, interactive)	Assumes TCP
	FD	N/S		
	MFD	≤ 1 s	G.1010 Transaction services	
	FDR	N/S		
	IFDV	N/S		

Table 32: Database Parameters – Client/Server

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	≤ 1 s	SEC Regulation NMS Self-Help	
		< 1 s	SEC Regulation NMS Intermarket Sweep Order Workflow	
	Packet loss	Extremely low	Cisco Trading Floor Architecture	
	Jitter	N/S		
Appl. Perform. Objectives	**FLR**	1e-5		
	FD	N/S		
	MFD	≤ 2 ms		
	FDR	N/S		
	IFDV	N/S		
Info	**Availability**	99.999%		Various sources

Table 33: Financial Trading Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	≤ 500 ms	Various, use cases	Based on 250ms (one way) PTZ requirement
		≤ 80 ms	Cisco Video Surveillance Best Practice	between client viewing station and VSOM
	Packet loss	≤ 0.01%	MPEGIF	Based on MPEG-4 with Simple Profile
	Jitter	< 1 ms	G.1010	Total user-to-user, achieved using de-jitter buffer in receiver.
Appl. Perform. Objectives	**FLR**	< 1e-2	G.1010, TS 22.105	Assumes use of a packet loss concealment algorithm to minimize effect of packet loss.
	FD	≤ 150 ms (MPEG-4) ≤ 200 ms (MJPEG)		Based on 250ms for PTZ, leaves 100ms for MPEG-4 encoding / decoding, 50ms for MJPEG encoding / decoding
	MFD	N/S		
	FDR	50 ms	Y.1541	$P_r = 0.999$
	IFDV	N/S		
Info	**Availability**			

Table 34: CCTV Parameters

Category	Parameter	Value	Source	Notes
Appl. Req's.	**RTT**	≤ 300 ms	Cisco TelePresence (1)	240 ms Service Provider budget
		≤ 300 ms	Polycom (2)	Video endpoints and multipoint server delay is in addition
	Packet loss	≤ 0.05%	Cisco TelePresence (1)	0.025% Service Provider budget
		≤ 0.1%	Polycom (2)	Average over 5-minute interval
	Jitter	≤ 10 ms	Cisco TelePresence (1)	
		≤ 40 ms	Polycom (2)	
Appl. Perform. Objectives	**FD**	≤ 120 ms	Cisco TelePresence (1)	P_d = 0.999
	MFD	≤ 110 ms	Cisco TelePresence (1) Polycom (2)	= 120 – 10 ms = 150 – 40 ms
	FLR	≤ 0.025%	Cisco TelePresence (1) Service Provider budget	
	FDR	≤ 40 ms	Polycom jitter	P_r = 0.999
	IFDV	≤ 10 ms	Cisco TelePresence (1)	P_v = 0.9999
	Bandwidth	15 Mbps	Cisco TelePresence (1)	

Table 35: Telepresence Parameters

The CPO ranges proposed relative to Mobile Backhaul are listed separately in Table 36. These CPO ranges map values associated with H, M, and L required classes as developed jointly between the CoS and Mobile Backhaul projects. Note that the driver for the requirements in the CoS Label H are often based on MBH for the older Mobile technologies (2G and 3G). For example, due to the tight control/signaling requirements when Ethernet MBH is inserted in the 3G UMTS RAN between the NodeB and the RNC (e.g., soft handover).

CoS Label	Example CoS Performance Objectives for each Metric#					
	MFD*	FD*	FDR	IFDV	FLR	Availability^
H	7 ms	10 ms	5 ms	3 ms	10^{-4}	TBD
M	13 ms	20 ms	10 ms	8 ms	10^{-4}	TBD
L	28 ms	37 ms	N/S	N/S	10^{-3}	TBD

Table 36: Mobile Backhaul Proposed CPOs

Notes:

Values are not recommendations for or reflections of actual services from contributing companies but rather represent reasonable industry values based on a wide range of MBH requirement sources, wide variety of applications, on any possible 2G-4G technologies. Less stringent values could be used for certain technologies or under certain mix of services/applications or network assumptions. Values will evolve (to more or less stringent values) as technologies mature and relational constraints between attributes are better understood and applied, and when SP field experiences will be available**.** SPs are free to provide CPOs that are more stringent for their specific services based on their field experience**.**

* MFD and FD Objectives assume geographic area/scope of limited size/distance (i.e., a Metro Performance Tier)

^ Availability metric is added as a Placeholder for MBH Phase 3 and CoS IA Phase 3. Values are TBD in future phase.

All of the applications and their respective Performance Objectives are summarized in Table 37. Not all applications from the list in Table 14 are represented in Table 37. The remote control aspects of remote surgery and the IP-based transport of professional video were applications for which no clear guidance was found.

Application	FD	MFD	FLR	FDR	IFDV
VoIP Data	125 ms pref 375 ms limit P_d = 0.999	100 ms pref 350 ms limit	3e-2	50 ms P_r = 0.999	40 ms P_v = 0.999
Video Conferencing Data	125 ms pref 375 ms limit P_d = 0.999	100 ms pref 350 ms limit	1e-2	50 ms P_r = 0.999	40 ms P_v = 0.999
VoIP and Videoconf Signaling	Not specified	250 ms pref	1e-3	Not specified	Not specified
IPTV Data Plane	125 ms P_d = 0.999	100 ms	1e-3	50 ms P_r = 0.999	40 ms P_v = 0.999
IPTV Control Plane	Not specified	75 ms	1e-3	Not specified	Not specified
Streaming Media	Not specified	Not specified	1e-2	2 s	1.5 s P_v = 0.99
Interactive Gaming	50 ms	40 ms	1e-3	10 ms	8 ms
Circuit Emulation	25 ms P_d = .999999	20 ms	1e-6	15 ms P_r = .999	10 ms P_v = .999, Δt = 900s, T = 3600s
Telepresence, includes: Remote Surgery (Video)	120 ms P_d = 0.999	110 ms	2.5e-4	40 ms P_r = 0.999	10 ms
Financial/Trading	Unknown	2 ms	1e-5	Unknown	Unknown
CCTV	150 ms (MPEG-4) 200 ms (MJPEG) P_d=0.999	Not specified	1e-2	50 ms P_r = 0.999	Not specified
Database (Hot Standby)	5 ms	Not specified	1e-5	Unknown	Unknown
Database (WAN Replication)	50 ms	Not specified	1e-5	Unknown	Unknown
Database (Client/Server)	Not specified	1 s	1e-3	Not specified	Not specified

Application	FD	MFD	FLR	FDR	IFDV
T.38 Fax	400 ms P_d = 0.999	350 ms	3e-2	50 ms P_r = 0.999	40 ms P_v = 0.999
SANs (Synchronous Replication)	5 ms	3.75 ms	1e-4	1.25 ms	1 ms
SANs (Asynchronous Replication)*	40 ms	30 ms	1e-4	10 ms	8 ms
Network Attached Storage	Not specified	1 s	1e-3	Not specified	Not specified
Text and Graphics Terminals	Not specified	200 ms	1e-3	Not specified	Not specified
Point of Sale Transactions	2 s	1 s	1e-3	Not specified	Not specified
Best Effort, includes: Email Store/Forward Fax WAFS Web Browsing File Transfer (including hi-res image file transfer) E-Commerce	Not specified	Not specified	Not specified	Not specified	Not specified
Mobile Backhaul H	10 ms	7 ms	1e-4	5 ms	3 ms
Mobile Backhaul M	20 ms	13 ms	1e-4	10 ms	8 ms
Mobile Backhaul L	37 ms	28 ms	1e-3	Not specified	Not specified

Table 37: Summarized CPOs

C.2 Derivation of CPOs from Application Performance Requirements

The values for CoS Performance Objectives (CPOs) are derived using multiple criteria. First, the set of applications described in section C.1 is mapped into CoS Labels and Performance Tiers to determine the set of application-specific Performance Objectives applicable for each case. Candidate CPO values are determined which meet the Performance Objectives for most or all of the applications mapped into a CoS Label/Performance Tier combination. Ideally, all of the application-specific Performance Objectives will be satisfied for each application mapped into a specific CoS Label/Performance Tier combination – however, given the limited number of CoS Labels in the 3-CoS Label model and the breadth of the applications considered, this is not always possible.

Second, a set of statistical and other constraints are applied to the candidate CPO values to make sure that they maintain the correct relationships to each other across CoS Labels, across Performance Tiers, and between the CPOs within a single CoS Label/Performance Tier. The candidate CPO values are modified as necessary to meet the constraints while still satisfying the application-specific Performance Objectives.

C.2.1 Mapping Applications to CoS Labels and Performance Tiers

Table 38 below is a table representing the explicit mapping of the applications in the tables in Section C.1 above to the MEF 3 CoS Label Model. This mapping is informative for the purpose of derivation of CPOs, and does not constrain any mapping of actual applications to CoS Labels or Performance Tiers by Subscribers or Operators.

CoS Label	**H**					**M**					**L**				
Performance Tier	**0.3**	**1**	**2**	**3**	**4**	**0.3**	**1**	**2**	**3**	**4**	**0.3**	**1**	**2**	**3**	**4**
VoIP		X	X	X	X										
VoIP & videoconf signaling							X	X	X	X					
Videoconf data							X	X	X	X					
IPTV data							X	X	X						
IPTV control							X	X	X						
Streaming media												X	X	X	X
Interactive gaming		X	X				X	X							
SANs synch replication	X						X								
SANs asynch replication							X								
Network attached storage												X	X	X	X
Text & graphics terminals												X	X	X	X
T.38 fax over IP							X	X	X	X					
Database hot standby	X						X								
Database WAN replication						X	X								
Database client/server												X	X	X	X
Financial/Trading	X	X													
CCTV							X	X	X	X					
Telepresence		X	X	X											
Circuit Emulation	X	X													
Mobile BH H		X													
Mobile BH M							X								
Mobile BH L												X			

Table 38: Explicit Application Mapping for Derivation of CPOs[7]

C.2.2 Constraints on CPO Values

The set of CPOs for each class in a given tier is derived initially from the objectives of one or more applications, subject to minimum FD/MFD values implied by the distance range of that tier.

The following constraints on CPOs are necessary in order to avoid a statistical contradiction:

- FDR > FD – MFD
- MFD < FD
- IFDV < FDR

Also, assuming that the distribution of delays has a long tail to the right:

- FD – MFD >> .5 FDR (.5 represents a symmetric distribution)

We also apply two constraints to ensure consistency between the values for FD and FDR and the estimated maximum Propagation Delay PD associated with each performance tier, calculated as described in Section Appendix A. When the percentile parameter $P_d = P_r$, then the Minimum Delay (MinD) associated with a given CoS Label/Performance Tier can be calculated as MinD = FD – FDR. This value MinD should be no less than PD. MinD should also not be significantly higher than PD. The first constraint is satisfied by:

- FD – FDR ≥ PD.

The second constraint is satisfied if the CPO values meet either of two tests. The first test scales PD by a ratio and then compares it to MinD. The second test, which prevents the constraint from becoming too severe for very low propagation delays, adds a fixed offset to PD before comparing it to MinD. Therefore, the second constraint is expressed as:

- (FD – FDR ≤ PD * 1.5) OR (FD – FDR ≤ PD + 20ms)

Finally, for PT constraints we assume that CPOs should never improve as tier number increases and that the MFD for each PT must exceed the estimated maximum propagation delay for the PT.

Below is a tabular summary of the various constraints that are applied to the Application driven performance objectives in order to derive CPOs.

[7] The red "X" marks indicate new mappings based on the inclusion of PT0.3 and in each case the grey "X" depicts where this application was mapped in MEF 23.1.

Statistical and Inter-CoS Label Constraints	Notes
H CoS Label CPOs ≤ all other CoS Label CPOs, except H FLR ≥ M FLR	For all in-scope metrics CPO (assumes Parameters are consistent across CoS Labels)
FD – MFD >> .5 FDR *	Where .5 represents a symmetric distribution
MFD < FD	
FDR > FD – MFD *	
IFDV < FDR	
FD – FDR ≥ PD	PD = estimated max Propagation Delay for a given PT
(FD – FDR ≤ PD * 1.5) OR (FD – FDR ≤ PD + 20ms)	PD = estimated max Propagation Delay for a given PT

- *Note: can be combined into various forms, e.g., MFD + .5 FDR << FD < MFD + FDR.

PT Constraints	Notes
PTm CPO ≤ PTn CPO	Where m<n (assumes Parameters are consistent across PTs. Includes all in-scope CPOs.)
PT0.3 MFD > 0.5 ms	Estimated max Propagation Delay for PT0.3
PT1 MFD > 2 ms	Estimated max Propagation Delay for PT1
PT2 MFD > 8 ms	Estimated max Propagation Delay for PT2
PT3 MFD > 44 ms	Estimated max Propagation Delay for PT3
PT4 MFD > 172 ms	Estimated max Propagation Delay for PT4

Standards and Other Constraints	Notes
MEF CPOs ≤ Y.1541 IP QoS Class Objectives CoS Label H PT1-3 for ITU QoS Class 0, 2 CoS Label H PT4 for ITU QoS Class 1 CoS Label M PT1-4 for ITU QoS Class 3 CoS Label L PT1-4 for ITU QoS Class 4	Includes MFD (IPTD) and FLR (IPLR). Where PT1, PT2, PT3 comparable to National and PT4 comparable to Global
PT1 (Metro) ≤ CPOs for MBH	Not including any synchronization-only driven objectives that could be developed. These are for future phase
CPOs and Parameters will be expressed as maximum or minimum values (not ranges)	

Table 39: CPO Derivation Constraints

C.2.3 The CoS Performance Objective Compliance Tool

The CoS Performance Objective Compliance Tool is a Microsoft Excel spreadsheet used to test candidate CPO values against the application-specific Performance Objectives and the constraints identified above. The tool comprises a worksheet for each Performance Tier as well as two summary worksheets. The first worksheet summarizes all CPO values in one table and displays whether they meet the constraint tests. The second summary worksheet shows how the CPO values compare to the mapped application-specific Performance Objectives.

Performance Tier worksheets

There are a total of five Performance Tier worksheets, one for each PT. At the bottom left of the table for each tier is a set of proposed CPO values (MFD, FDR, FLR, FD, and IFDV) for each class (H, M, L) in the 3-CoS Label model. The tool checks the compliance of each set of class objectives against the Application Performance Metrics objectives contained in the upper part of the table; the result of the compliance checks is displayed to the right of the application objective values.

In its current form, the definition of compliance used in the tool is as follows.

1. Each CPO value is compared to the corresponding Application Objective (AO) value. If the CPO value is less stringent than the AO value, it is considered Not Compliant; otherwise, the CPO value is considered Compliant. Two types of compliance are defined: Loose and Tight. If the AO value is within a (configurable) range of the CPO value, it is considered Tight compliance; otherwise it is Loose compliance. As an example, if an AO for MFD is 50% higher (less stringent) than the corresponding CPO, it is considered Loose compliance. An Unspecified or Unknown application objective also results in Loose compliance.

2. The compliance results for the set of CPO values for a class as compared to an application's requirements are then combined as follows:

    a. If any CPO value is Not Compliant, the overall compliance of the class to that application's requirements is considered "Bad."

    b. If any CPO value for the class yields Loose compliance, the overall compliance of the CPOs to that application's requirements is considered "OK" (which may be interpreted as "overkill," i.e., the stringency of the CPO is greater than required by the application).

    c. Otherwise, the overall compliance of the CPOs for the class to that application's requirements is considered "Good."

The spreadsheet based tables below illustrate the derivation of CPOs per PT. The derivation was based on a visual basic macro incorporated in the spreadsheet to provide a best fit for the application objectives into the 3 CoS Labels. In addition the constraints above were applied.

(Note that the figures below are illustrative of the process used to derive the CPOs, and that the specific values may not reflect the normative CPO values in this document.)

The CPOs for PT0.3 and PT1 are primarily driven by the MBH application.

First, the derivation of the PT0.3 objectives:

(yellow)	=Unspecified application objective
(brown)	=Unknown application objective

Application Performance Attributes / **MEF CPOs Compliance**

Application Attributes	Application	Context	CIR-only?	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	H	M	L
Consumer Applications	VoIP	PE-PE*	FALSE	100	50	3.E-02	125	40	OK	OK	OK
	VoIP and Videoconf Signaling	PE-PE*	FALSE	250		1.E-03	250		OK	OK	OK
	Video Conferencing Data	PE-PE*	FALSE	100	50	1.E-02	125	40	OK	OK	OK
	IPTV data plane	PE-PE*	FALSE	100	50	1.E-03	125	40	OK	OK	OK
	IPTV control plane	PE-PE*	FALSE	75		1.E-03			OK	OK	OK
	Streaming media	PE-PE*	FALSE		2000	1.E-02		1500	OK	OK	OK
	Interactive gaming	PE-PE*	FALSE	40	10	1.E-03	50	8	OK	OK	OK
Business Applications	SANs (Synchronous Replication)	PE-PE*	FALSE	3.75	1.25	1.E-04	5	1	OK	Bad	Bad
	SANs (Asynchronous Replication)	PE-PE*	FALSE	30	10	1.E-04	40	8	OK	OK	Bad
	Network Attached Storage	PE-PE*	FALSE	1000		1.E-03			OK	OK	OK
	Text and Graphics Terminals	PE-PE*	FALSE	200		1.E-03			OK	OK	OK
	T.38 Real-time Fax over IP	PE-PE*	FALSE	350	50	3.E-02	400	40	OK	OK	OK
	Database (Hot Standby)	PE-PE*	FALSE			1.E-05	5		OK	Bad	Bad
	Database (WAN Replication)	PE-PE*	FALSE			1.E-05	50		OK	OK	Bad
	Database (Client-Server)	PE-PE*	FALSE	1000		1.E-03			OK	OK	OK
	Financial/Trading	PE-PE*	FALSE	2		1.E-05			OK	Bad	Bad
	CCTV	PE-PE*	FALSE		50	1.E-02	150		OK	OK	OK
	Telepresence (includes Remote Surgery video)	PE-PE*	FALSE	110	18	3.E-04	120	10	OK	OK	Bad
	Circuit Emulation	PE-PE*	FALSE	20	15	1.E-06	25	10	Bad	Bad	Bad
HetNet / Small Cell Applications [Tight Coordination]	H/S H	PE-PE*	FALSE	0.7	0.5	1.E-04	1	0.3	Bad	Bad	Bad
	H/S M	PE-PE*	FALSE	3	1	1.E-04	5	1	Bad	Bad	Bad
	H/S L	PE-PE*	FALSE	8		1.E-03	10		OK	OK	Bad

MEF CPOs (PT0.3)

MEF CoS Parameter Objectives (CPOs)	Description (MEF Example Suggested Applications)	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
	Sync, Voice, Near-RT	H	FALSE	2	1.25	1.E-05	3	0.9
	Control/Signaling, Data	M	FALSE	4	3	1.E-05	6	2.4
	Data, Background	L	FALSE	8	4.8	1.E-03	11	4.2

Statistical Constraints	CoS	MFD<FD	IFDV<FDR	FD<MFD+FDR	FD>MFD+FDR/2		(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset)	
	H	Good	Good	Good	Good		Good	[Minimum Delay Test (aka "Bob
	M	Good	Good	Good	Good		Good	
	L	Good	Good	Good	Good		Good	

Non-Statistical Constraints	CoS					
As stringent as Y.1541		MFD/IPTD		FLR/IPLR		
	H	Good		Good		
	M	Good		Good		
	L	Good		Good		
As stringent as higher tiers		MFD	FDR	FLR	FD	IFDV
	H	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good
		MFD	FDR	FLR	FD	IFDV
H<=M (FLR:H>=M)		Good	Good	Good	Good	Good
H<=L		Good	Good	Good	Good	Good
		MFD	Air D	CRD km	CRD ms	
MFD > Calculated route delay	H	Good	75	93.75	0.46875	
	M	Good				
	L	Good				

The derivation of the PT1 objectives:

Yellow = Unspecified application objective
Brown = Unknown application objective

Application Attributes	Application	Context	CIR-only?	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	MEF CPOs Compliance H	M	L
				Application Performance Attributes							
Consumer Applications	VoIP	PE-PE*	FALSE	100	50	3.E-02	125	40	OK	OK	OK
	VoIP and Videoconf Signaling	PE-PE*	FALSE	250		1.E-03	250		OK	OK	OK
	Video Conferencing Data	PE-PE*	FALSE	100	50	1.E-02	125	40	OK	OK	OK
	IPTV data plane	PE-PE*	FALSE	100	50	1.E-03	125	40	OK	OK	OK
	IPTV control plane	PE-PE*	FALSE	75		1.E-03			OK	OK	OK
	Streaming media	PE-PE*	FALSE		2000	1.E-02		1500	OK	OK	OK
	Interactive gaming	PE-PE*	FALSE	40	10	1.E-03	50	8	OK	OK	Bad
Business Applications	SANs (Synchronous Replication)	PE-PE*	FALSE	3.75	1.25	1.E-04	5	1	Bad	Bad	Bad
	SANs (Asynchronous Replication)	PE-PE*	FALSE	30	10	1.E-04	40	8	OK	OK	Bad
	Network Attached Storage	PE-PE*	FALSE	1000		1.E-03			OK	OK	OK
	Text and Graphics Terminals	PE-PE*	FALSE	200		1.E-03			OK	OK	OK
	T.38 Real-time Fax over IP	PE-PE*	FALSE	350	50	3.E-02	400	40	OK	OK	OK
	Database (Hot Standby)	PE-PE*	FALSE			1.E-05	5		Bad	Bad	Bad
	Database (WAN Replication)	PE-PE*	FALSE			1.E-05	50		Bad	Bad	Bad
	Database (Client-Server)	PE-PE*	FALSE	1000		1.E-03			OK	OK	OK
	Financial/Trading	PE-PE*	FALSE	2		1.E-05			Bad	Bad	Bad
	CCTV	PE-PE*	FALSE		50	1.E-02	150		OK	OK	OK
	Telepresence (includes Remote Surgery video)	PE-PE*	FALSE	110	18	3.E-04	120	10	OK	OK	Bad
	Circuit Emulation	PE-PE*	FALSE	20	15	1.E-06	25	10	Bad	Bad	Bad
MBH Applications	MBH H	PE-PE*	FALSE	7	5	1.E-04	10	3	Good	Bad	Bad
	MBH M	PE-PE*	FALSE	13	10	1.E-04	20	8	OK	Good	Bad
	MBH L	PE-PE*	FALSE	28	16	1.E-03	37	14	OK	OK	Good

MEF CoS Parameter Objectives (CPOs)
(PT1, e.g., Metro)

MEF CPOs (PT1)

Description (MEF Example Suggested Applications)	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
Sync, Voice, Near-RT	H	FALSE	7	5	1.E-04	10	3
Control/Signaling, Data	M	FALSE	13	10	1.E-04	20	8
Data, Background	L	FALSE	28	16	1.E-03	37	14

Statistical Constraints

CoS	MFD<FD	IFDV<FDR	FD<MFD+FDR	FD>MFD+FDR/2	(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset) OR [Minimum Delay Test (aka "Bob Test")
H	Good	Good	Good	Good	Good
M	Good	Good	Good	Good	Good
L	Good	Good	Good	Good	Good

Non-Statistical Constraints

As stringent as Y.1541

CoS	MFD/IPTD	FLR/IPLR
H	Good	Good
M	Good	Good
L	Good	Good

As stringent as higher tiers

CoS	MFD	FDR	FLR	FD	IFDV
H	Good	Good	Good	Good	Good
M	Good	Good	Good	Good	Good
L	Good	Good	Good	Good	Good

	MFD	FDR	FLR	FD	IFDV
H<=M (FLR:H>=M)	Good	Good	Good	Good	Good
H<=L	Good	Good	Good	Good	Good

MFD > Calculated route distance

CoS	MFD	Air D	CRD km	CRD ms
H	Good	250	312.5	1.5625
M	Good			
L	Good			

As stringent as MBH (PT1 only)

CoS	MFD	FDR	FLR	FD	IFDV
H	Good	Good	Good	Good	Good
M	Good	Good	Good	Good	Good
L	Good	Good	Good	Good	Good

The following chart illustrates derivation of PT2 objectives:

Legend	
(yellow)	=Unspecified application objective
(brown)	=Unknown application objective

Application Attributes	Application	Context	CIR-only?	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	MEF CPOs Compliance H	M	L
Consumer Applications	VoIP	PE-PE*	FALSE	1.E+02	5.E+01	3.E-02	1.E+02	4.E+01	OK	Good	Bad
	VoIP and Videoconf Signaling	PE-PE*	FALSE	3.E+02		1.E-03	3.E+02		OK	OK	OK
	Video Conferencing Data	PE-PE*	FALSE	1.E+02	5.E+01	1.E-02	1.E+02	4.E+01	OK	Good	Bad
	IPTV data plane	PE-PE*	FALSE	1.E+02	5.E+01	1.E-03	1.E+02	4.E+01	OK	Good	Bad
	IPTV control plane	PE-PE*	FALSE	8.E+01		1.E-03			OK	OK	Good
	Streaming media	PE-PE*	FALSE		2.E+03	1.E-02		2.E+03	OK	OK	OK
	Interactive gaming	PE-PE*	FALSE	4.E+01	1.E+01	1.E-03	5.E+01	8.E+00	OK	Bad	Bad
Business Applications	SANs (Synchronous Replication)	PE-PE*	FALSE	4.E+00	1.E+00	1.E-04	5.E+00	1.E+00	Bad	Bad	Bad
	SANs (Asynchronous Replication)	PE-PE*	FALSE	3.E+01	1.E+01	1.E-04	4.E+01	8.E+00	OK	Bad	Bad
	Network Attached Storage	PE-PE*	FALSE	1.E+03		1.E-03			OK	OK	OK
	Text and Graphics Terminals	PE-PE*	FALSE	2.E+02		1.E-03			OK	OK	OK
	T.38 Real-time Fax over IP	PE-PE*	FALSE	4.E+02	5.E+01	3.E-02	4.E+02	4.E+01	OK	Good	Bad
	Database (Hot Standby)	PE-PE*	FALSE			1.E-05	5.E+00		Bad	Bad	Bad
	Database (WAN Replication)	PE-PE*	FALSE			1.E-05	5.E+01		Bad	Bad	Bad
	Database (Client-Server)	PE-PE*	FALSE	1.E+03		1.E-03			OK	OK	OK
	Financial/Trading	PE-PE*	FALSE	2.E+00		1.E-05			Bad	Bad	Bad
	CCTV	PE-PE*	FALSE		5.E+01	1.E-02	2.E+02		OK	OK	Bad
	Telepresence (includes Remote Surgery video)	PE-PE*	FALSE	1.E+02	2.E+01	3.E-04	1.E+02	1.E+01	OK	Bad	Bad
	Circuit Emulation	PE-PE*	FALSE	2.E+01	2.E+01	1.E-06	3.E+01	1.E+01	Bad	Bad	Bad
MBH Applications	MBH H	PE-PE*	FALSE	6.E+00	3.E+00	1.E-05	8.E+00	2.E+00	Bad	Bad	Bad
	MBH M	PE-PE*	FALSE	1.E+01	1.E+01	1.E-05	2.E+01	8.E+00	Bad	Bad	Bad
	MBH L	PE-PE*	FALSE	3.E+01	2.E+01	1.E-03	4.E+01	1.E+01	OK	Bad	Bad

Application Performance Attributes

MEF CPOs (PT2)

MEF CoS Parameter Objectives (CPOs) (PT2, e.g., Regional)	Description (MEF Example Suggested Applications)	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
	Sync, Voice, Near-RT	H	FALSE	18	10	1.E-04	25	8
	Control/Signaling, Data	M	FALSE	30	50	1.E-04	75	40
	Data, Background	L	FALSE	50	100	1.E-03	125	80

	CoS							
Statistical Constraints		MFD<FD	IFDV<FDR	FD<MFD+FDR	FD>MFD+FDR/2		(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset) OR	
	H	Good	Good	Good	Good		Good	[Minimum Delay Test (aka "Bob Test"
	M	Good	Good	Good	Good		Good	
	L	Good	Good	Good	Good		Good	
Non-Statistical Constraints			**MFD/IPTD**		**FLR/IPLR**			
As stringent as Y.1541	H		Good		Good			
	M		Good		Good			
	L		Good		Good			
			MFD	**FDR**	**FLR**	**FD**	**IFDV**	
As stringent as higher tiers	H		Good	Good	Good	Good	Good	
	M		Good	Good	Good	Good	Good	
	L		Good	Good	Good	Good	Good	
			MFD	**FDR**	**FLR**	**FD**	**IFDV**	
H<=M (FLR:H>=M)			Good	Good	Good	Good	Good	
H<=L			Good	Good	Good	Good	Good	
			MFD	**Air D**	**CRD km**	**CRD ms**		
MFD > Calculated route distance	H		Good	1200	1500	7.5		
	M		Good					
	L		Good					

Likewise, the following chart illustrates derivation of PT3 objectives:

(yellow) =Unspecified application objective
(brown) =Unknown application objective

Application Attributes	Application	Context	CIR-only?	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	MEF CPOs Compliance H	MEF CPOs Compliance M	MEF CPOs Compliance L
			Application Performance Attributes								
Consumer Applications	VoIP	PE-PE*	FALSE	1.E+02	5.E+01	3.E-02	1.E+02	4.E+01	OK	OK	Bad
	VoIP and Videoconf Signaling	PE-PE*	FALSE	3.E+02		1.E-03	3.E+02		OK	OK	OK
	Video Conferencing Data	PE-PE*	FALSE	1.E+02	5.E+01	1.E-02	1.E+02	4.E+01	OK	OK	Bad
	IPTV data plane	PE-PE*	FALSE	1.E+02	5.E+01	1.E-03	1.E+02	4.E+01	OK	OK	Bad
	IPTV control plane	PE-PE*	FALSE	8.E+01		1.E-03			OK	Bad	Bad
	Streaming media	PE-PE*	FALSE		2.E+03	1.E-02		2.E+03	OK	OK	OK
	Interactive gaming	PE-PE*	FALSE	4.E+01	1.E+01	1.E-03	5.E+01	8.E+00	Bad	Bad	Bad
Business Applications	SANs (Synchronous Replication)	PE-PE*	FALSE	4.E+00	1.E+00	1.E-04	5.E+00	1.E+00	Bad	Bad	Bad
	SANs (Asynchronous Replication)	PE-PE*	FALSE	3.E+01	1.E+01	1.E-04	4.E+01	8.E+00	Bad	Bad	Bad
	Network Attached Storage	PE-PE*	FALSE	1.E+03		1.E-03			OK	OK	OK
	Text and Graphics Terminals	PE-PE*	FALSE	2.E+02		1.E-03			OK	OK	OK
	T.38 Real-time Fax over IP	PE-PE*	FALSE	4.E+02	5.E+01	3.E-02	4.E+02	4.E+01	OK	OK	Bad
	Database (Hot Standby)	PE-PE*	FALSE			1.E-05	5.E+00		Bad	Bad	Bad
	Database (WAN Replication)	PE-PE*	FALSE			1.E-05	5.E+01		Bad	Bad	Bad
	Database (Client-Server)	PE-PE*	FALSE	1.E+03		1.E-03			OK	OK	OK
	Financial/Trading	PE-PE*	FALSE	2.E+00		1.E-05			Bad	Bad	Bad
	CCTV	PE-PE*	FALSE		5.E+01	1.E-02	2.E+02		OK	OK	Bad
	Telepresence (includes Remote Surgery video)	PE-PE*	FALSE	1.E+02	2.E+01	3.E-04	1.E+02	1.E+01	OK	Bad	Bad
	Circuit Emulation	PE-PE*	FALSE	2.E+01	2.E+01	1.E-06	3.E+01	1.E+01	Bad	Bad	Bad
MBH Applications	MBH H	PE-PE*	FALSE	6.E+00	3.E+00	1.E-05	8.E+00	2.E+00	Bad	Bad	Bad
	MBH M	PE-PE*	FALSE	1.E+01	1.E+01	1.E-05	2.E+01	8.E+00	Bad	Bad	Bad
	MBH L	PE-PE*	FALSE	3.E+01	2.E+01	1.E-03	4.E+01	1.E+01	Bad	Bad	Bad

MEF CPOs (PT3)

MEF CoS Parameter Objectives (CPOs) (PT3, e.g., National)	Description (MEF Example Suggested Applications)	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
	Sync, Voice, Near-RT	H	FALSE	70	12	2.5E-04	77	10
	Control/Signaling, Data	M	FALSE	80	50	2.5E-04	115	40
	Data, Background	L	FALSE	125	165	1.0E-03	230	130

Statistical Constraints		MFD<FD	IFDV<FDR	FD<MFD+FDR	FD>MFD+FDR/2		(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset) OR	
	H	Good	Good	Good	Good		Good	[Minimum Delay Test (aka "Bob Test"
	M	Good	Good	Good	Good		Good	
	L	Good	Good	Good	Good		Good	
Non-Statistical Constraints			MFD/IPTD		FLR/IPLR			
As stringent as Y.1541	H		Good		Good			
	M		Good		Good			
	L		Good		Good			
			MFD	FDR	FLR	FD	IFDV	
As stringent as higher tiers	H		Good	Good	Good	Good	Good	
	M		Good	Good	Good	Good	Good	
	L		Good	Good	Good	Good	Good	
			MFD	FDR	FLR	FD	IFDV	
H<=M (FLR:H>=M)			Good	Good	Good	Good	Good	
H<=L			Good	Good	Good	Good	Good	
			MFD	Air D	CRD km	CRD ms		
MFD > Calculated route distance	H		Good	7000	8750	43.75		
	M		Good					

Finally, the following chart illustrates the derivation of PT4 objectives:

(yellow) =Unspecified application objective
(brown) =Unknown application objective

Application Attributes	Application	Context	CIR-only?	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	H	M	L
				Application Performance Attributes					MEF CPOs Compliance		
Consumer Applications	VoIP	PE-PE*	FALSE	4.E+02	5.E+01	3.E-02	4.E+02	4.E+01	OK	OK	Bad
	VoIP and Videoconf Signaling	PE-PE*	FALSE	3.E+02	(unspecified)	1.E-03	3.E+02	(unspecified)	OK	OK	Bad
	Video Conferencing Data	PE-PE*	FALSE	3.E+02	5.E+01	1.E-02	4.E+02	4.E+01	OK	OK	Bad
	IPTV data plane	PE-PE*	FALSE	1.E+02	5.E+01	1.E-03	1.E+02	4.E+01	Bad	Bad	Bad
	IPTV control plane	PE-PE*	FALSE	8.E+01	(unspecified)	1.E-03	(unspecified)	(unspecified)	Bad	Bad	Bad
	Streaming media	PE-PE*	FALSE	(unspecified)	2.E+03	1.E-02	(unspecified)	2.E+03	OK	OK	OK
	Interactive gaming	PE-PE*	FALSE	4.E+01	1.E+01	1.E-03	5.E+01	8.E+00	Bad	Bad	Bad
Business Applications	SANs (Synchronous Replication)	PE-PE*	FALSE	4.E+00	1.E+00	1.E-04	5.E+00	1.E+00	Bad	Bad	Bad
	SANs (Asynchronous Replication)	PE-PE*	FALSE	3.E+01	1.E+01	1.E-04	4.E+01	8.E+00	Bad	Bad	Bad
	Network Attached Storage	PE-PE*	FALSE	1.E+03	(unspecified)	1.E-03	(unspecified)	(unspecified)	OK	OK	OK
	Text and Graphics Terminals	PE-PE*	FALSE	2.E+02	(unspecified)	1.E-03	(unspecified)	(unspecified)	OK	Bad	Bad
	T.38 Real-time Fax over IP	PE-PE*	FALSE	4.E+02	5.E+01	3.E-02	4.E+02	4.E+01	OK	OK	Bad
	Database (Hot Standby)	PE-PE*	FALSE	(unspecified)	(unknown)	1.E-05	5.E+00	(unknown)	Bad	Bad	Bad
	Database (WAN Replication)	PE-PE*	FALSE	(unspecified)	(unknown)	1.E-05	5.E+01	(unknown)	Bad	Bad	Bad
	Database (Client-Server)	PE-PE*	FALSE	1.E+03	(unspecified)	1.E-03	(unspecified)	(unspecified)	OK	OK	OK
	Financial/Trading	PE-PE*	FALSE	2.E+00	(unknown)	1.E-05	(unknown)	(unknown)	Bad	Bad	Bad
	CCTV	PE-PE*	FALSE	(unspecified)	5.E+01	1.E-02	2.E+02	(unspecified)	Bad	Bad	Bad
	Telepresence (includes Remote Surgery video)	PE-PE*	FALSE	1.E+02	2.E+01	3.E-04	1.E+02	1.E+01	Bad	Bad	Bad
	Circuit Emulation	PE-PE*	FALSE	2.E+01	2.E+01	1.E-06	3.E+01	1.E+01	Bad	Bad	Bad
MBH Applications	MBH H	PE-PE*	FALSE	6.E+00	3.E+00	1.E-05	8.E+00	2.E+00	Bad	Bad	Bad
	MBH M	PE-PE*	FALSE	1.E+01	1.E+01	1.E-05	2.E+01	8.E+00	Bad	Bad	Bad
	MBH L	PE-PE*	FALSE	3.E+01	2.E+01	1.E-03	4.E+01	1.E+01	Bad	Bad	Bad

MEF CPOs (PT4)

MEF CoS Parameter Objectives (CPOs) (PT4, e.g., Global)	Description (MEF Example Suggested Applications)	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
	Sync, Voice, Near-RT	H	FALSE	200	40	5.E-04	230	32
	Control/Signaling, Data	M	FALSE	220	50	5.E-04	250	40
	Data, Background	L	FALSE	240	200	1.E-03	390	160

Statistical Constraints

	MFD<FD	IFDV<FDR	FD<MFD+FDR	FD>MFD+FDR/2	(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset) OR
H	Good	Good	Good	Good	Good [Minimum Delay Test (aka "Bob Test"
M	Good	Good	Good	Good	Good
L	Good	Good	Good	Good	Good

Non-Statistical Constraints

As stringent as Y.1541

	MFD/IPTD	FLR/IPLR
H	Good	Good
M	Good	Good
L	Good	Good

As stringent as higher tiers

	MFD	FDR	FLR	FD	IFDV
H	NA	NA	NA	NA	NA
M	NA	NA	NA	NA	NA
L	NA	NA	NA	NA	NA

	MFD	FDR	FLR	FD	IFDV
H<=M (FLR:H>=M)	Good	Good	Good	Good	Good
H<=L	Good	Good	Good	Good	Good

MFD > Calculated route distance

	MFD	Air D	CRD km	CRD ms
H	Good	27500	34375	171.875
M	Good			
L	Good			

Minimum Delay Test (aka "Bob test")

	(FD-FDR > CRD) AND ((FD-FDR < CRD+Offset) OR (FD-FDR < CRD*Ratio))
H	Good
M	Good
L	Good

Table 40: PT0.3 – 4 CPO Derivation and Evaluation Spreadsheets

CPO Summary worksheet

The CPO Summary worksheet displays numerical values for all CPOs (even for those CPOs defined as "Not Specified" in Table 8 through Table 12) and shows the results of the constraint tests applied to those CPO values (note that PT 1–4 assume 10Mbps Ethernet for serialization delay, but PT 0.3 assumes 100Mbps in order to meet the FD requirements). Figure 7 shows the summary displays.

See MEF10.3 for definitions

	MEF CoS	CIR-only	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)
PT 0.3	H	FALSE	2	1.25	1.E-05	3	0.9
	M	FALSE	4	3	1.E-05	6	2.4
	L	FALSE	8	4.8	1.E-03	11	4.2
PT1	H	FALSE	7	5	1.E-04	10	3
	M	FALSE	13	10	1.E-04	20	8
	L	FALSE	28	16	1.E-03	37	14
PT2	H	FALSE	18	10	1.E-04	25	8
	M	FALSE	30	50	1.E-04	75	40
	L	FALSE	50	100	1.E-03	125	80
PT3	H	FALSE	70	12	2.5E-04	77	10
	M	FALSE	80	50	2.5E-04	115	40
	L	FALSE	125	165	1.0E-03	230	130
PT4	H	FALSE	200	40	5.E-04	230	32
	M	FALSE	220	50	5.E-04	250	40
	L	FALSE	240	200	1.E-03	390	160

	MEF CoS	MFD (ms)	FDR (ms)	FLR (ratio)	FD (ms)	IFDV (ms)	Minimum Delay Test
PT 0.3	H	Good	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good	Good
PT1	H	Good	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good	Good
PT2	H	Good	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good	Good
PT3	H	Good	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good	Good
PT4	H	Good	Good	Good	Good	Good	Good
	M	Good	Good	Good	Good	Good	Good
	L	Good	Good	Good	Good	Good	Good

				Implied Values			
	MEF CoS	MinD = FD-FDR	Prop Delay (ms)	Shaping Delay budget factor	Serialization Delay (ms)	Queueing Delay + Shaping Delay Budget (ms)	Shaping Delay from budget (ms)
PT 0.3	H	1.8	0.5	0.50	0.32	2.18	1.09
	M	3.0	0.5	0.50	0.32	5.18	2.59
	L	6.2	0.5	0.50	0.32	10.18	5.09
PT1	H	5.0	2	0.50	3.2	4.8	2.4
	M	10.0	2	0.50	3.2	14.8	7.4
	L	21.0	2	0.50	3.2	31.8	15.9
PT2	H	15.0	8	0.50	3.2	13.8	6.9
	M	25.0	8	0.50	3.2	63.8	31.9
	L	25.0	8	0.50	3.2	113.8	56.9
PT3	H	65.0	44	0.50	3.2	29.8	14.9
	M	65.0	44	0.50	3.2	67.8	33.9
	L	65.0	44	0.50	3.2	182.8	91.4
PT4	H	190.0	172	0.50	3.2	54.8	27.4
	M	200.0	172	0.50	3.2	74.8	37.4
	L	190.0	172	0.50	3.2	214.8	107.4

Figure 7: CPO Summary worksheet

Application Mapping Summary Worksheet

The Application Mapping summary worksheet contains several tables. The lower table defines the explicit mapping of applications to CoS Label/Performance Tier combinations used to test the CPO values. An 'X' in a cell maps the application in the cell's row to the CoS Label/Performance Tier in the cell's column. The right side of the table includes a summary of the application-specific Performance Objectives for each application. The upper left table shows how well the mapped application-specific Performance Objectives match the CPO values, using the criteria described for the Performance Tier worksheets in Section C.2.3 above. The upper right table provides a summary of how well the application-specific Performance Objectives match the CPO values for all applications, CoS Labels and Performance Tiers, both mapped and unmapped. Figure 8 shows the application mapping tables.

Merge of actual and desired states

Category	Application	PT1 CoS H	PT1 CoS M	PT1 CoS L	PT2 CoS H	PT2 CoS M	PT2 CoS L	PT3 CoS H	PT3 CoS M	PT3 CoS L	PT4 CoS H	PT4 CoS M	PT4 CoS L
Real time	VoIP	OK			OK			OK			OK		
Interactive	VoIP and videoconf signaling		OK			OK			OK			OK	
Real time	Videoconf data		OK			Good			OK			OK	
Near real time	IPTV data		OK			Good			OK				
Interactive	IPTV control		OK			OK			Bad				
Streaming	Streaming media			OK			OK			OK			OK
Low delay	Interactive gaming	OK	OK		OK	Bad							
Very low delay	SANs synchronous replication		Bad										
Low delay	SANs asynchronous replication		OK										
Best effort	Network attached storage			OK			OK			OK			OK
Best effort	Text and graphics terminals			OK			OK			OK			Bad
Near real time	T.38 fax over IP		OK			Good			OK			OK	
Very low delay	Database hot standby		Bad										
Low delay	Database WAN replication		OK										
Best effort	Database client/server			OK			OK			OK			OK
Very low delay	Financial/Trading	Bad											
Near real time	CCTV		OK			OK			OK			Bad	
Real time	Telepresence	OK			OK			OK					
Real time	Circuit Emulation	Bad											
Very low delay	MBH H	Good											
Very low delay	MBH M		Good										
Low delay	MBH L			Good									

Applications to CoS Levels (Current state)

PT1 CoS H	PT1 CoS M	PT1 CoS L	PT2 CoS H	PT2 CoS M	PT2 CoS L	PT3 CoS H	PT3 CoS M	PT3 CoS L	PT4 CoS H	PT4 CoS M	PT4 CoS L
OK	OK	OK	OK	Good	Bad	OK	OK	Bad	OK	OK	Bad
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	Bad
OK	OK	OK	OK	Good	Bad	OK	OK	Bad	OK	OK	Bad
OK	OK	OK	OK	Good	Bad	OK	OK	Bad	Bad	Bad	Bad
OK	OK	OK	OK	OK	Good	OK	Bad	Bad	Bad	Bad	Bad
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	OK	Bad	OK	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	Bad	OK	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	Bad	Bad
OK	OK	OK	OK	Good	Bad	OK	OK	Bad	OK	OK	Bad
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	Bad	OK	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	OK	OK	OK	Bad	OK	OK	Bad	Bad	Bad	Bad
OK	OK	Bad	OK	Bad	Bad	OK	Bad	Bad	Bad	Bad	Bad
Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
Good	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	Good	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad
OK	OK	Good	OK	Bad	Bad	Bad	Bad	Bad	Bad	Bad	Bad

Applications to CoS Levels (Desired state)

Category	Application	Very low jitter (<< 50 ms) CoS H PT1	PT2	PT3	PT4	Low jitter (50 ms) CoS M PT1	PT2	PT3	PT4	Non-critical data CoS L PT1	PT2	PT3	PT4
Real time	VoIP	X	X	X	X								
Interactive	VoIP and videoconf signaling					X	X	X	X				
Real time	Videoconf data					X	X	X	X				
Near real time	IPTV data					X	X	X					
Interactive	IPTV control					X	X	X					
Streaming	Streaming media									X	X	X	X
Low delay	Interactive gaming	X	X			X	X						
Very low delay	SANs synchronous replication					X							
Low delay	SANs asynchronous replication					X							
Best effort	Network attached storage									X	X	X	X
Best effort	Text and graphics terminals									X	X	X	X
Near real time	T.38 fax over IP					X	X	X	X				
Very low delay	Database hot standby					X							
Low delay	Database WAN replication					X							
Best effort	Database client/server									X	X	X	X
Very low delay	Financial/Trading	X											
Near real time	CCTV					X	X	X	X				
Real time	Telepresence	X	X	X									
Real time	Circuit Emulation	X											
Very low delay	MBH H	X											
Very low delay	MBH M					X							
Low delay	MBH L									X			

Performance Attributes

MFD	FDR	FLR	FD	IFDV
100	50	0.03	125	40
250		0.001	250	
100	50	0.01	125	40
100	50	0.001	125	40
75		0.001		
	2000	0.01		1500
40	10	0.001	50	8
3.75	1.25	0.0001	5	1
30	10	0.0001	40	8
1000		0.001		
200		0.001		
350	50	0.03	400	40
		1E-05	5	
		1E-05	50	
1000		0.001		
2		1E-05		
	50	0.01	150	
110	18	0.0003	120	10
20	15	1E-06	25	10
6	3	1E-05	8	2
13	10	1E-05	20	8
28	16	0.001	37	14

=Unspecified application objective

=Unknown application objective

Figure 8: Application Mapping summary worksheet

Appendix D Example PCP and DSCP Mapping at UNI for Multi-CoS EVCs Supporting Only Standard MEF Classes of Service (Informative)

The CoS IA requires that all PCP (or DSCP) values that may occur in any service deployment are to be supported in some way by the service. Several alternatives exist. For example, any specific CEN service may support additional CoS Names beyond those defined in this IA, and PCP (or DSCP) values not specified as CoS Identifiers in the CoS IA may be mapped to a CoS Name provided as an addition to the CoS IA defined CoS Labels. Alternatively, a service may include at least one additional CoS Name intended specifically to handle frames not associated (by PCP/DSCP value) with a defined CoS Identifier. If a specific CEN service supports *only* the CoS Labels defined by this IA, there needs to be a mapping of all possible PCP (or DSCP) values to one of the CoS Labels defined in the CoS IA or to a CoS defined in [1] called "Discard" which simply discards all frames that are classified as such.

This section provides example mappings for this case assuming no "Discard" CoS Name. Note that in some cases the use of a "Discard" CoS with only the PCP and DSCP values specified in Table 4 may be the simplest way to negotiate markings. In this case all PCP and DSCP values not shown in Table 4 would be discarded at the EI.

D.1 Example PCP Mappings

The following tables provide examples of full mappings of PCP at a UNI for multi-CoS Label EVCs that support only standard MEF CoS Labels.

Table 41 shows an example mapping in which PCP value 5 is assumed to be handled by CE routers as "EF" traffic. This may be a common approach in handling low latency traffic based on a PCP marking – particularly when using (for instance) IP Routers.

MEF CoS Label Combination Supported on EVC	PCP Mapping per Class of Service Label – Color-Blind Mode		
	H	M	L
{H + M + L}	5	2-4, 6, 7	0, 1
{H + M}	5	0-4, 6, 7	N/A
{H + L}	5	N/A	0-4, 6, 7
{M + L}	N/A	2-7	0, 1

Table 41: Example PCP Mapping for Multi-CoS Label EVC Supporting Only Standard CoS Labels at UNI – "Router-Application-Friendly" mapping

Table 42 shows a similar mapping that may apply in an application that bases choices of PCP values on the assumption of Ethernet CE bridges forwarding based on strict priority. In this case, higher PCP values would be handled at a higher priority. This mapping works in an application where very-high priority traffic is (by nature) very low volume (possibly less than 1 percent of the total traffic volume). This mapping is needed, for instance, if the application is not necessarily able to distinguish traffic that is carried natively in Ethernet over the local LAN from traffic that may be carried by a CEN service.

MEF CoS Label Combination Supported on EVC	PCP Mapping per Class of Service Label – Color-Blind Mode		
	H	M	L
{H + M + L}	4-7	2,3	0, 1
{H + M}	4-7	0-3	N/A
{H + L}	4-7	N/A	0-3
{M + L}	N/A	2-7	0, 1

Table 42: Example PCP Mapping for Multi-CoS Label EVC Supporting Only Standard CoS Labels at UNI – "Bridging-Application-Friendly" mapping

D.2 Example DSCP Mappings

The following table provides an example of a full mapping of DSCP values at a UNI for multi-CoS Label EVCs that support only standard MEF CoS Labels.

MEF CoS Combination Supported on EVC	DSCP Mapping per Class of Service – Color-Blind Mode		
	H	M	L
{H + M + L}	40-47	16-39, 48-63	0-15
{H + M}	40-47	0-39, 48-63	N/A
{H + L}	40-47	N/A	0-39, 48-63
{M + L}	N/A	16-63	0-15

Table 43: Example DSCP Mapping for Multi-CoS EVC Supporting Only Standard Classes of Service at UNI

Appendix E Other Relevant Standards and Industry models (Informative)

This section excerpts information from relevant standards that may be helpful in reading this document.

Below are excerpted tables from Section 6 and Annex G (informative) of [2]. Specifically this IA used the 5P3D row PCP values (bottom row on the excerpt below) from Table 6-4 for the CoS Identifier PCP values in Table 4 because 5 Priorities (i.e., classes) is the closest match to the 3 CoS Label Model. There is no row in the table for a smaller number of Priorities than 5P3D. Note that in Table G-2 of [2] the VO (voice) class specifies 10ms latency and jitter.

PCP Allocation		PCP Values and Traffic Classes							
# PCP Priorities	# PCP Drop Eligible	PCP = 7	6	5	4	3	2	1	0
8	0	IEEE Traffic Class = 7	6	5	4	3	2	1	0
5	3	IEEE Traffic Class = 7	6	4	4 DE	2	2 DE	0	0 DE

Table 44: PCP Decoding (Adapted from [2])

Appendix F Guidelines for Multipoint Services (Informative)

F.1 Guidelines for Multipoint Services

The following section outlines the rationale for defining separate CPOs for point-to-point and multipoint services, and explains the focused overload condition introduced with this service type.

F.1.1 Multipoint CoS Performance Objectives

Tables 7-11 define less stringent CPOs for multipoint services in comparison to point-to-point services. The origin of these objectives is in relation to the additional processing required to achieve one-to-many connectivity. Multipoint services require two types of additional processing not commonly experienced in point-to-point services: frame replication and address table lookup.

A variety of methods are available to operators for transporting Ethernet frames across the CEN. Some methods are more sensitive to multipoint service impairment than others. VPLS for example must replicate flooded frames at ingress, placing the entire processing burden on a single node. Provider bridging distributes the replication task throughout the CEN. Design enhancements, including Hierarchical VPLS (H-VPLS) or point-to-multipoint LSPs, should be considered by operators using this method to reduce the effect of replication processing.

Frame replication involves the duplication of the same frame within the CEN network to reach all external interface members of an EVC or OVC (subsequent references to EVC in this Appendix include OVC). Only frames that require flooding are replicated. Replication is needed for broadcast and multicast frames, as well as unicast frames for which the MAC forwarding table does not have a matching destination address entry (termed unknown unicast). Unknown unicast traffic is not a common occurrence, but can be experienced in various circumstances:

- when end stations first begin transmitting on an EVC
- unidirectional traffic streams
- after a network topology change that flushes the MAC forwarding table.
- MAC forwarding table exhaustion

The additional processing required to replicate frames can introduce a consistent and measurable delay that is represented in the performance objectives for multipoint services. This delay was quantified in operator lab environments for common vendor equipment used to deliver these services at the time of this document revision. EVCs of sizes 2, 4, 10, 24, 50, and 100 UNIs were tested. A significant increase in delay occurred under the following conditions:

- EVCs comprising 100 UNIs
- Flooding traffic reaching 80% of link speed (1 Gbps links tested)

The relaxed delay-related performance allocated to multipoint services is defined in two areas depending on the metric type:

- delay-related metrics that support a percentile (FD, IFDV, FDR) have the point-to-point parameter value reduced by one percent
- delay-related metrics that do not have a percentile (MFD) have the point-to-point value increased by 1–2 milliseconds

The relaxed objectives defined in Table 8 though Table 12 are recommended for EVCs comprising 100 or fewer UNIs. As with all CoS IA performance objectives, operators can always define more stringent objectives. If an operator constrains multipoint service design (e.g.: modest maximum EVC size, ingress rate-limiting of flooding traffic), CPOs equal to that of point-to-point services can be achieved. Operators are encouraged to test their equipment for performance impairment under flooding conditions.

The test participants did not observe a measureable amount of delay impairment due to MAC address table lookup. Even excessive MAC table sizes did not produce a sufficient degree of delay to justify adjustment of CPOs. These tests were performed on higher-end distribution routers/switches. The same result may not be observed in lower-end devices. Operators are encouraged to measure the effect of MAC address table lookup as well on their equipment and services.

F.1.2 Focused Overload

The focused overload condition occurs when the sum of network traffic from ingress external interfaces that are members of one or more multipoint EVCs exceeds the available capacity of an egress external interface or a CEN internal link. Point-to-point services can introduce similar conditions, but only in a hub-and-spoke architecture.

When frames are discarded due to focused overload of egress traffic at a UNI for a Multipoint-to-Multipoint or a Rooted-Multipoint EVC, MEF 10.3 [1] provides the option to exclude those discarded frames from the Availability and Frame Loss Ratio performance.

With multipoint services, any one of possibly many egress external interfaces can be the destination of input traffic. Also, multipoint services support the transmission of flooding frames, which each need to be delivered to a subset or all of the remaining UNIs in an EVC. While the service definition of multipoint services has provisions for operators to control the rate of these specific flooding frames, the exercise of capacity planning and controlled provisioning remains difficult for this service type. Figure 9 shows an example multipoint EVC, where the combined traffic of UNI-1, UNI-2, and UNI-3 focused onto UNI-4 will overload its 10 Mbps access with as high as 12 Mbps of CIR traffic.

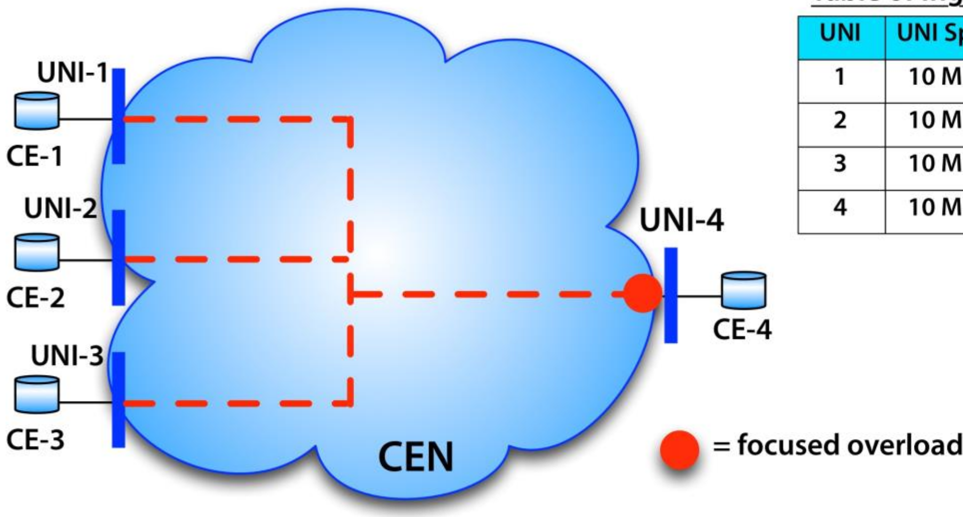


Table of Ingress BWP parameters

UNI	UNI Speed	EVC CIR	EVC EIR
1	10 Mbps	4 Mbps	6 Mbps
2	10 Mbps	4 Mbps	6 Mbps
3	10 Mbps	4 Mbps	6 Mbps
4	10 Mbps	4 Mbps	6 Mbps

Figure 9: Example of focused overload

Figure 10 shows an example of focused overload within an operator CEN, where the combined traffic of multiple UNIs in a multipoint EVC can exhaust the capacity of intra-CEN network elements. The combined traffic of UNI-1, UNI-2, and UNI-3 focused onto link-1 will overload its 1 Gbps capacity with as high as 1.2 Gbps of CIR traffic.

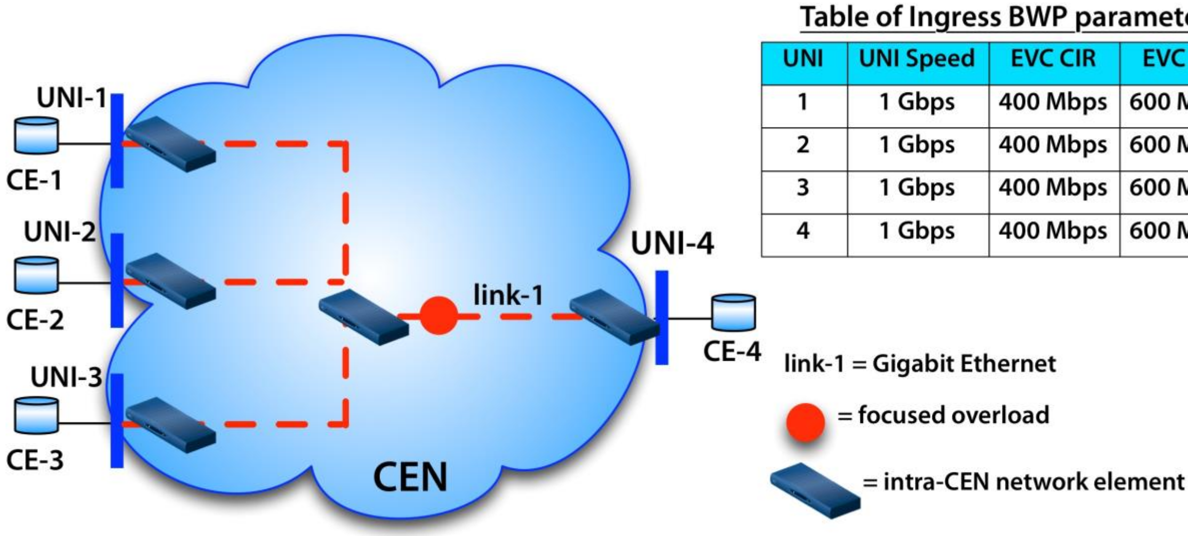


Table of Ingress BWP parameters

UNI	UNI Speed	EVC CIR	EVC EIR
1	1 Gbps	400 Mbps	600 Mbps
2	1 Gbps	400 Mbps	600 Mbps
3	1 Gbps	400 Mbps	600 Mbps
4	1 Gbps	400 Mbps	600 Mbps

Figure 10: Example of focused overload of an intra-CEN link

While the focused overload condition cannot be entirely avoided in multipoint service deployments, the following methods should be considered by operators to improve the situation.

Separate Multipoint and Point-to-point queues

When one or more EVCs produce congestion on internal CEN links, all services that share the associated network queues are impacted. To avoid the situation where all services in a common traffic class are affected, an operator might choose to place multipoint services in different queues than point-to-point services. Depending on the number of queues supported by the network element provider, this can be achieved for all or a subset of the traffic classes offered by the operator. An operator might also use this approach to differentiate between multiple multipoint service categories based on either relative priority/importance of the service or the risk of a specific service type introducing unpredictable high volume flooding traffic. An operator might consider, at a minimum, separating queues for the H CoS Label, ensuring the strict performance requirements for point-to-point services are not impacted by this condition. This separation does not require different CoS IDs for the same CoS label across point-to-point and multipoint services, but instead can be achieved with unique intra-CEN transport header CoS markings.

Appendix G Burst Size and Shaper Considerations (Informative)

G.1 Shaping Considerations for Burst Alignment

This section extends the example shaper from Section A.3 of MEF 10.3 [1], which is specific to a single color shaper at a CE, to a color aware shaper applicable at either a UNI or ENNI.

Section A.3 of MEF 10.3 [1] ("Traffic Shaping") describes a pair of algorithms that together describe a shaper implementation at the CE. The algorithm in MEF 10.3 Figure 37 ("Periodic Algorithm") is run every Δt seconds where Δt is the period between updating the token bucket values C(t) and E(t) (i.e. the token bucket refresh rate). The algorithm in MEF 10.3 Figure 38 ("New Frame Algorithm") is run every time a new frame arrives at the shaper.

Similarly, we define a pair of example algorithms that together describe a shaper implementation at the egress of CEN-1 at the ENNI. The algorithms are updated to be Color Aware, so that they handle Yellow frames coming from CEN-1. In the example New Frame Algorithm, a limited number of Yellow frames can be placed in the shaper buffer[8] (and subsequent Yellow frames will be dropped if required). This controls the delay that may be experienced by Green frames due to the presence of Yellow frames. There may be Yellow frames ahead of a Green frame in the transmission buffer, but that is no different from current practice.

The following parameters are used in the example algorithms (using the notation from [1]):

- *CIR* = the shaping rate of Green frames (average output rate of the shaper);
- *CBS* = the shaping burst size of Green frames (maximum output burst of the shaper);
- *EIR* = the shaping rate of Yellow frames (average output rate of the shaper);
- *EBS* = the shaping burst of Yellow frames (maximum output burst of the shaper).

The following notation is used in the example algorithms (following the definitions in [1]):

- *B(t)* = the instantaneous shaper buffer occupancy in bytes,
- *C(t)* = the instantaneous value of the tokens in the Committed token bucket with C(0) = *CBS*,
- *E(t)* = the instantaneous value of the tokens in the Excess token bucket with E(0) = *EBS*,
- *L* = the length of the frame at the head of the shaper buffer, and
- *LNF* = the length of the newly-arrived frame.

Note that for *B(t)*, *C(t),* and *E(t), t* represents the time that the algorithm is run.

[8] Note that we differentiate between the shaper buffer, and the transmission buffer (outgoing link queue). Frames taken from the head of the shaper buffer are enqueued on the transmission buffer and transmitted at line rate. We assume that the transmission buffer remains unchanged from the current situation.

The following parameters are new or modified from [1]:

- *CF* = the "Coupling Flag" that controls whether tokens that overflow the Committed token bucket are added to the Excess token bucket,
- *CM* = the "Color Mode" can be either Color-blind or Color-aware and controls whether the color of a frame is considered in determining when to transfer it from the shaper buffer to the transmission buffer,
- *SBL* = the "Shaper Buffer Limit" is the depth of the shaper buffer (in bytes) above which no new frames will be placed in the shaper buffer,
- *YBL* = the "Yellow Buffer Limit" is the depth of the shaper buffer (in bytes) above which no new yellow frames will be placed in the shaper buffer.

Note that the *CM* parameter only affects whether the color of a frame is considered when removing that frame from the shaper buffer. Whether the color of a frame is considered when placing that frame into the shaper buffer is controlled by *YBL* and *SBL*. If *YBL* = *SBL* then yellow and green frames receive identical treatment when determining whether to place the frame in the shaper buffer (i.e. color blind behavior), and otherwise the color affects whether to place the frame in the shaper buffer (i.e. color aware behavior).

Note that *YBL* is the maximum accepted burst of yellow frames and *SBL* is the maximum accepted burst of all frames (green and yellow). This means that the maximum burst of green frames that is guaranteed to be accepted is *SBL* – *YBL*, however up to *SBL* green frames may be accepted in the absence of yellow frames. Typically the shaper buffer limits are configured such that $SBL - YBL \geq CBS$, which means the shaper accepts larger bursts of green frames at its input and generates smaller bursts of green frames at its output.

The maximum delay that can be experienced by a green frame is *SBL* divided by *CIR*, provided that *CIR* is the minimum rate at which frame are transmitted from the shaper buffer. This will be the case as long as either *CF* = 1 (so tokens overflowing the Committed token bucket are added to the Excess token bucket) or $EIR \geq CIR$.

Yellow frames arriving at the shaper will only be transmitted using yellow tokens. Green frames arriving at the shaper will be transmitted using green tokens when they are available, or yellow tokens when no green tokens are available.

```
C(t) = min(CBS, (C(t) + (CIR/8)*Δt));
O(t) = max(0, (C(t) + (CIR/8)*Δt - CBS));
E(t) = min(EBS, (E(t) + (EIR/8)*Δt + CF*O(t)));
while((L <= C(t)) && (B(t) > 0) && (frame at head of shaper buffer is green || CM =
Color_Blind)) ||
      ((L <= E(t)) && (B(t) > 0)))
{
   if((L <= C(t)) && (B(t) > 0) && (frame at head of shaper buffer is green || CM =
Color_Blind))
   {
      C(t) = C(t) - L;
      B(t) = B(t) - L;
      send the frame at the head of the shaper buffer to the transmission buffer;
//Transmit using green tokens
   }
   Else
   {
      E(t) = E(t) - L;
      B(t) = B(t) - L;
      send the frame at the head of the shaper buffer to the transmission buffer;
//Transmit using yellow tokens
   }
}
```

Figure 11: Periodic Algorithm

The revision of the New Frame Algorithm from [1] to handle transmission of Yellow frames is shown below.

```
if(B(t) == 0) // If shaper buffer is empty
{
    if(new frame color is Yellow && CM = Color_Aware)
    {
        if(LNF <= E(t))
        {
            E(t) = E(t) - LNF;
            send new frame to transmission buffer; //Transmit using yellow tokens
        }
        else if (B(t) + LNF <= YBL)
        {

            place new frame in shaper buffer;
            B(t) = B(t) + LNF;
        else
        {
            discard new frame;
        }
        }
    }
    else // new frame is Green or shaper is configured to be Color_Blind
    {
        if(LNF <= C(t))
        {
            C(t) = C(t) - LNF;
            send new frame to transmission buffer; //Transmit using green tokens
        }
        else if(LNF <= E(t))
        {
            E(t) = E(t) - LNF;
            send new frame to transmission buffer; //Transmit using yellow tokens
        }
        else if (new frame color is Green &&(B(t) + LNF <= SBL))
        {
            place new frame in shaper buffer;
            B(t) = B(t) + LNF;
        }
        else
        {
            discard new frame;
        }
    }
}
else  // shaper buffer is not empty
{


    if(((new frame color is Green) && (B(t) + LNF <= SBL)) || ((new frame color is
Yellow) && (B(t) + LNF <= YBL) && (EBS >= max frame size)))
    {
        place new frame in shaper buffer;
        B(t) = B(t) + L;
    }
    else // no room in shaper buffer for another frame
    {
       discard new frame;
    }
}
```

Figure 12: New Frame Algorithm

G.2 Shaping and Bandwidth Profile Considerations with TCP Traffic

Transport Control Protocol (TCP) is a rate-adaptive protocol. In theory this means that a TCP source will dynamically adjust its data transmission rate to match the bandwidth available at the point between the source and destination where the bandwidth is most constrained (the "bottleneck bandwidth"). This leads to the expectation that if a TCP flow goes through an EVC where the bandwidth is restricted by an ingress bandwidth profile configured with a given Committed Information Rate (CIR), TCP will adjust its transmission rate such that the average throughput equals the CIR. That is not what happens. At least, that is not what happens unless the ingress bandwidth profile is also configured with an unexpectedly large value for the Committed Burst Size (CBS). The following section examines the interaction of TCP traffic with functions that constrain that traffic to meet the bandwidth profile specifications of an EVC or OVC, including ingress bandwidth profile policers and shapers either before or after the ingress bandwidth profile policers. The conclusion is that without shaping it is possible to configure the ingress bandwidth profile such that TCP throughput matches the CIR. In many cases doing so requires very large values for CBS, and the analysis in this section provides some guidelines for estimating what the CBS values need to be. On the other hand, shaping the traffic to CIR allows significantly smaller values for CBS at the ingress bandwidth profile and also provides much better predictability of TCP throughput when large values of CBS cannot be accommodated. The analysis in this section also provides guidelines for the configuring and positioning the shaping.

G.2.1 TCP Bottleneck at Ingress Bandwidth Profile Policer

When the bottleneck bandwidth of a TCP flow is enforced by an ingress bandwidth profile policer (with no shaping), the TCP throughput is a function of the CIR and CBS parameters of the bandwidth profile. Figure 13 shows simulation results of TCP source transmitting to a TCP receiver through an EVC with an ingress bandwidth policer (and no traffic shapers).[9] The figure plots TCP throughput versus CBS for various values of CIR. As can be seen, the TCP throughput is typically much smaller than the CIR until large values of CBS are reached. The goal of this section is to explain these results. It will be helpful to begin with a quick summary of TCP congestion control and its interaction with a bandwidth profile policer.

[9] In the simulation the TCP source has an unlimited amount of data to send, has a minimum retransmission timeout (RTO) value of 250 ms, and uses Selective Acknowledgment (SACK) information for retransmitting lost segments. The receiver window size is 60 KB and the round trip time (RTT) of the TCP connection is 5 ms. These parameters are the same for all of the simulation results in this section.

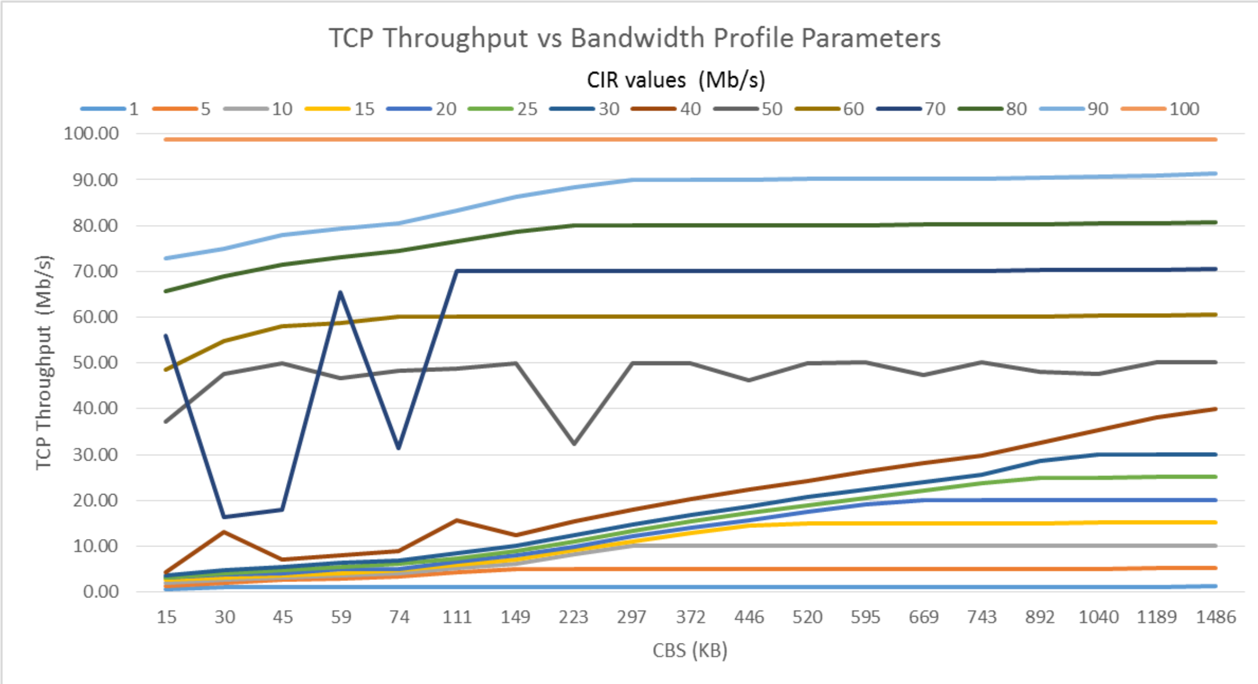


Figure 13: TCP Throughput through a Bandwidth Profile Policer

TCP throughput is limited by the amount of data it can send to the receiver before receiving an acknowledgment that some data has reached the receiver. This is the amount of data "in flight". The throughput is the amount of data in flight divided by the Round Trip Time (RTT) of the flow. TCP will attempt to maximize utilization of the available bandwidth by ramping up the amount of data in flight on the flow until the flow capacity is reached. The flow capacity can be calculated as the bottleneck bandwidth times the average Round Trip Time (RTT) of the flow. Putting all this together means the utilization of the available bandwidth is maximized when the throughput equals the bottleneck bandwidth, which is also when the amount of data in flight equals the flow capacity. TCP probes for changes in the flow capacity by gradually increasing the amount of data in flight until it detects that the capacity has been exceeded. TCP detects that the flow capacity has been reached when there is evidence of congestion, typically in the form of lost packets. When TCP detects that one, or a few, packets have been lost, it reacts by reducing the allowable amount of data in flight by a factor of two, retransmitting the lost packet(s), and then beginning to ramp up the allowable data in flight again. When multiple packets are lost within a short time interval, however, TCP reacts by waiting for a retransmission timeout (RTO) period before restarting transmission.

The characteristic of a Bandwidth Profile with a bucket size CBS and rate CIR is that it will allow traffic to pass at whatever rate it is offered while there are tokens available in the bucket, but when the bucket is depleted the amount of traffic allowed to pass is limited to CIR. To TCP this appears that the full line rate is available while there are tokens in the bucket, and TCP will ramp up transmissions to utilize the full line rate. When the bucket is depleted and the available rate suddenly drops to CIR, there are typically multiple TCP packets lost and TCP reacts with a timeout. During the timeout interval the token bucket gets a chance to refill, and the process repeats when TCP resumes transmission. The result is that a first approximation of the overall throughput of a single TCP flow is roughly one CBS worth of data every RTO interval, or CIR, whichever is less.

$$\text{IndividualTCPFlowThroughput} \sim \min\left\{\text{CIR}, \frac{\text{CBS}}{\text{RTO}}\right\} \quad (1)$$

TCP dynamically adjusts the RTO based on a weighted moving average of the measured flow round trip time, however there is a minimum value for the RTO. Although RFC 6298 recommends a minimum RTO value of one second, most common implementations reduce to this value to between 200 and 250 milliseconds. The simulation results described in this section all have the minimum RTO set to 250ms. Using an assumed value for the minimum RTO equal to 250ms it is possible to compute the minimum value of CBS necessary for the TCP throughput to equal CIR.

$$\text{CBS} \geq \text{CIR} \times \text{RTO} = \text{CIR} \times 250\text{ms} \quad (2)$$

This looks very much like the flow bandwidth delay product which is equal to CIR times RTT. The fact that the size of CBS needed to achieve throughput equal to CIR is dependent upon the RTO, not the RTT, is what leads to the required CBS being unexpectedly large on flows where RTT is significantly smaller than the minimum RTO.

Figure 14 shows simulation results of a TCP flow (with unlimited data to send) exhibiting the behavior described above when restricted by an ingress bandwidth profile policer on a 100 Mb/s UNI. The figure plots the aggregate amount of data acknowledged by the receiver over time, with a policer configured with a CIR of 10 Mb/s and four different values of CBS. The data received at the receiver is measured in units of TCP segments, and for simplicity CBS is also measured in units of segments.[10] The average TCP throughput for each CBS value is the slope of a line fit to the data set corresponding to that value. A gray line with a slope of CIR is shown as a reference for the target TCP throughput. The periodic transmit-then-timeout pattern is clearly visible. The TCP throughput is well below CIR for CBS equal to 10 segments. At a CBS of 100 segments the throughput is closer to CIR because the increased CBS allows more data to be transmitted during each cycle. At CBS equal 200 segments the throughput matches CIR and at any given time T the cumulative data received is between $CIR * T$ and $CBS + CIR * T$. Increasing CBS still further does not increase the TCP throughput because the average throughput cannot exceed CIR, so the plot for CBS equal 700 segments has a slope equal to CIR. The plot is elevated above the CIR line because the initial condition of a full token bucket causes

[10] Using “segments” as a data unit because it remains constant when applied to TCP data (where one segment equals 1460 bytes), or data in a policer or shaper (where one segment equals 1522 bytes due to the addition of TCP/IP and MAC headers, VLAN tag and FCS), or data “on the wire” (where one segment equals 1542 bytes due to the addition of preamble and accounting for a minimum inter-packet gap).

a large initial burst. This does not affect a long term average because the token bucket never completely fills again so the initial burst size is never repeated.

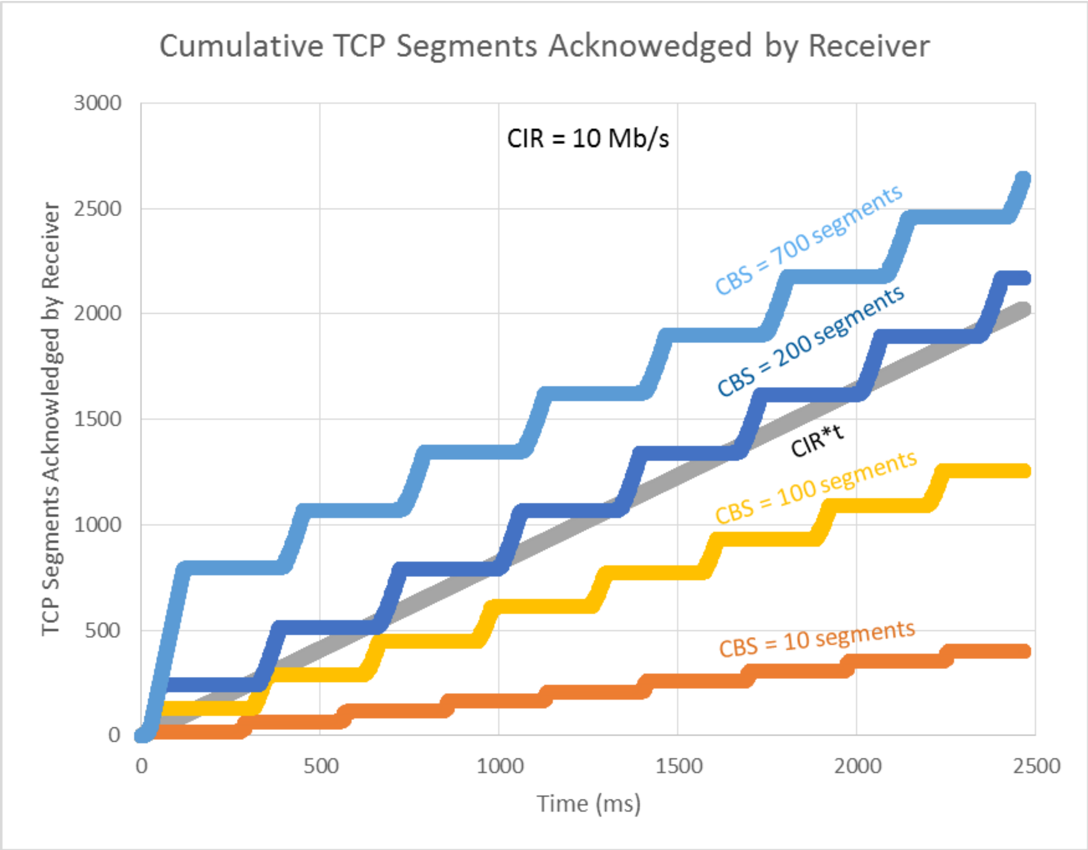


Figure 14: TCP segments received over time through a BWP Policer

Using equation (2) to calculate the minimum value of CBS that results in TCP throughput equal to a CIR of 10 Mb/s with RTO of 250 ms results in a CBS equal to 205 segments. This correlates well with the plot for CBS equal to 200 segments in Figure 14. Using equation (1) to calculate the expected TCP throughput for CBS equal to 10 segments results in a throughput of 0.49 Mb/s. This makes sense since this equation predicts a linear relationship between CBS and throughput for a given CIR. However the slope of the line in the simulation results for CBS equal to 10 segments is 2.1 Mb/s, which is significantly higher than predicted. The source of the discrepancy is that the derivation of this equation implicitly assumes the TCP transmitter sends CBS of data instantly at the start of each cycle, and no more for a timeout interval. Actually it takes time for the transmitter to send the data, during which time the bandwidth profile is also adding tokens to the token bucket at a rate of CIR. Furthermore the transmitter doesn't stop transmitting when the first packet is discarded at the bandwidth profile, but continues sending up to a TCP window size worth of additional data, some of which is received. Both of these factors are significant when CBS is small, but diminish as CBS approaches the value where TCP throughput equals CIR. Nonetheless the predictions of equation (1) are sufficiently close to the simulation results to justify a conclusion that the actual TCP throughput will be well below CIR for small values of CBS, and the predictions of equation (2) correlate very well with simulation results in Figure 13 for the value of CBS when the TCP throughput reaches CIR, at least for values of CIR up to 40 Mb/s.

Looking at Figure 13 it is clear that something different is happening for CIR values greater than 40 Mb/s than for CIR values less than 40 Mb/s. A closer look at what is happening when CIR equals 70 Mb/s is particularly revealing. Figure 15 plots TCP segments received versus time for four values of CBS when CIR equals 70 Mb/s. The plots for CBS = 30 segments (45 KB) and CBS = 50 segments (74 KB) clearly show the periodic transmit-then-timeout pattern seen in Figure 14, but the plots for CBS = 40 segments (59 KB) and CBS = 60 segments (89 KB) do not show any timeouts. Checking the simulation results for other points in Figure 13 support the conclusion that the transmit-then-timeout pattern is far less likely to be seen for CIR values greater than 40 Mb/s. This is not surprising in view of what needs to occur to cause a TCP timeout. The details of this vary with TCP implementations, but there are two sequences of events that pretty consistently lead to timeouts. The first is if no more than two packets get acknowledged after the first packet loss in a cycle. In this scenario the transmitter never receives the three duplicate acknowledgements (acknowledgements when there is a gap in the received packet sequence) which would trigger a fast recovery process that preempts a timeout. This is very unlikely at high CIR values because the token bucket is replenished quickly enough that some packets will get through to trigger duplicate acknowledgements. The second sequence of events leading to a timeout is when the fast recovery process is triggered causing the retransmission of lost packets, but one of these retransmissions is lost. This also gets less likely at higher CIR both because fewer packets need to be retransmitted and because more tokens are being added to the token bucket during the interval when those packets are retransmitted. This leads to two conclusions. First, for large values of CIR the TCP throughput will reach CIR with CBS values far smaller than predicted by equation (1). Second, for medium values of CIR the TCP throughput will appear chaotic as small variations in CBS can effect whether or not timeouts occur and thus cause a large variation in the TCP throughput.

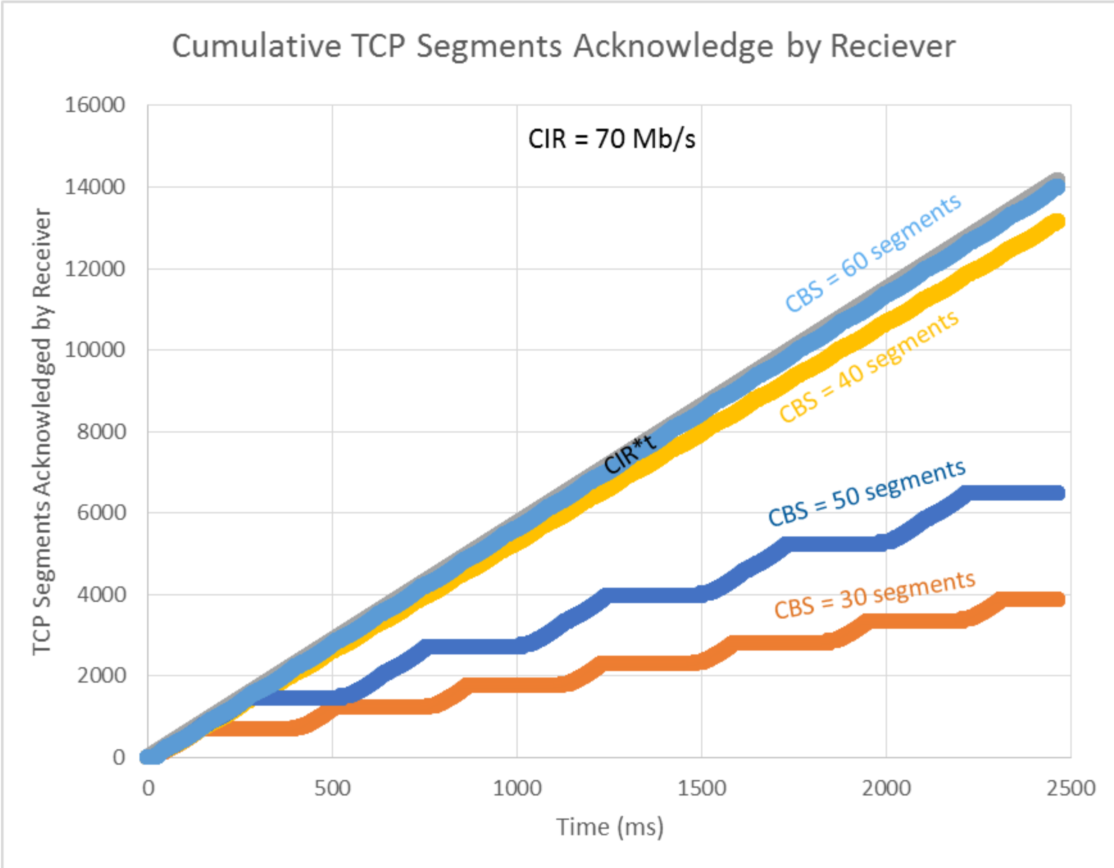


Figure 15: TCP segments received over time through a BWP Policer

A final observation on Figure 13 is that even though the value of CBS where TCP throughput reaches CIR is much lower than predicted by equation (1), it is still much larger than might be expected assuming expectations are on the order of the TCP window size or the bandwidth delay product (CIR * RTT). In this simulation the TCP window size is 42 segments (62 KB) and the bandwidth delay product for a CIR of 90 Mb/s is 38 segments, but the CBS required to have the TCP throughput reach CIR is 200 segments (297 KB). The reason for this is that even though the transmit pattern does not involve a timeout, it still involves a cycle of ramping up the allowable number of packets in flight until the token bucket empties and packets are lost, then cutting the allowable number of packets in flight in half and ramping it up again. When the allowable number of packets in flight is small the transmission rate is less than CIR and the number of tokens in the bucket increases. As the allowable number of packets in flight increases the transmission rate will increase above CIR. Whether the average throughput is equal to CIR or less than CIR depends upon the CBS value. To maintain an average rate equal to CIR the CBS has to be large enough to hold all the tokens that accumulate during the period when the transmission rate is below CIR so that this much extra data can be sent when the transmission rate is greater than CIR. With a small CBS the token bucket will overflow, with each token that overflows representing a transmit opportunity that cannot be made up, resulting in a throughput less than CIR.

Figure 16 shows TCP simulation results with a CIR of 90 Mb/s and CBS of 10 segments for one full cycle of the TCP transmitter ramping up the number of packets in flight until the token bucket empties and packets are lost. The plot on the left shows the packets transmitted at the TCP source, discarded by the bandwidth profile policer, and acknowledged by the TCP receiver. The plot on the right shows the token count in the bandwidth profile token bucket and the tokens discarded over the same time interval. Packets are lost when the token count reaches zero. The token bucket refills during the TCP recovery period, and stays near full as the TCP source ramps up its transmission rate. When the TCP transmission rate exceeds CIR the number of tokens in the bucket drops until it reaches zero and the cycle repeats. In this case the CBS is not large enough to hold all of the tokens that could accumulate during the ramp-up portion of the cycle (indicated by the token overflow), and therefore the average throughput is less than CIR (indicated by the average slope of the TCP segments delivered curve being less than the slope of the CIR*t line). Figure 17 shows simulation results with CBS increased to 200 segments. The TCP source still goes through a cycle of ramping up the transmission rate until packets are lost and then backing off, but in this case the token bucket never completely fills and no tokens overflow. Since tokens are generated at a rate of CIR, and no tokens are lost, the average TCP throughput equals CIR.

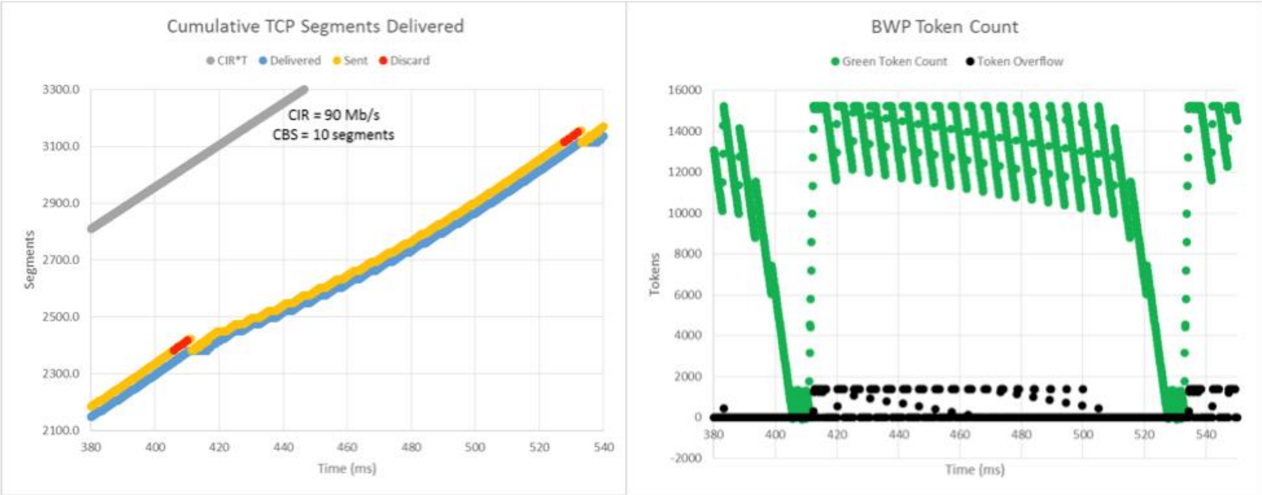


Figure 16: TCP cycle with CBS = 10

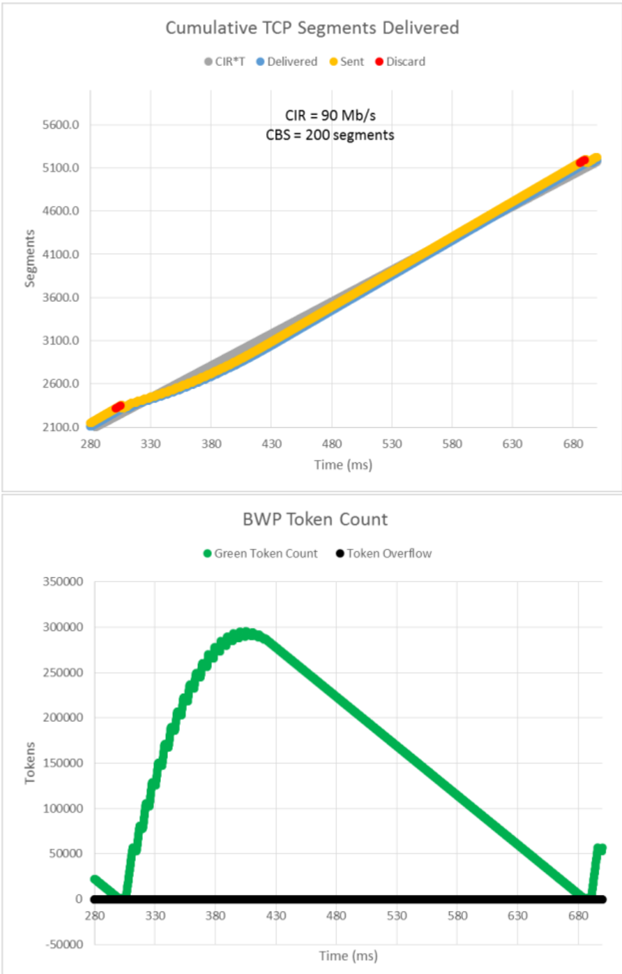


Figure 17: TCP cycle with CBS = 200

It can be argued that the single TCP flow case is artificial and only likely to be seen in test scenarios (though this may be reason enough to design for the single TCP flow case). If there are multiple TCP flows sharing the EVC then it is likely that the aggregate TCP throughput will be higher. As long as the timeout events of each TCP flow are evenly distributed, then the aggregate throughput will increase with the number of TCP flows (n) up to the limit set by CIR.

$$(3) \quad \text{AggregateTCPThroughput} \leq \min\left\{\text{CIR}, \sum_{i=1}^{n} \frac{\text{CBS}}{\text{RTO}_i}\right\}$$

How closely the aggregate throughput approaches CIR depends upon the validity of the “evenly distributed timeouts” assumption. Assume the token bucket is full and one TCP source starts transmitting. It will drain the token bucket, lose packets, and start a timeout. At this point the token bucket starts to refill. Once it refills another TCP source can start transmitting with the same results. As long as only one source is transmitting at a time, and is consuming tokens that otherwise would have been discarded, the sources will not interfere and the aggregate throughput will be the sum of each independent source. If the transmission interval of the sources overlaps, however, they will both experience packet loss at the same time when the bucket empties. If they are using the same value for a minimum RTO then they will begin transmitting again at about the same time. The result could be that once sources with the same minimum RTO value overlap transmission they will lock into a cycle where they compete for available tokens. As more sources get locked into the same cycle the aggregate throughput will be limited to the same throughput achievable by a single source. Whether this tendency to synchronize actually occurs is for future study.

This section has focused on the effect of ingress bandwidth profile policing on TCP throughput, and in particular explaining the relationship of CBS values to TCP throughput shown in Figure 13. Some details of the interaction between TCP and the bandwidth profile enforcement have been explored, and there are certainly more details to explore, but the overriding summary is that it takes large values of CBS to allow the TCP throughput to reach the configured CIR. “Large” is a relative term of course, but in this case “large” is relative to what would be expected if CBS were assumed to be approximately the same as the bandwidth delay product of the flow. Ingress bandwidth profiles with large CBS values make traffic management within the CEN more difficult as this can be directly correlated with a need for large buffers within the CEN. This provides the motivation for exploring alternatives to relying on an ingress policer alone to enforce the ingress bandwidth profile.

G.2.2 TCP Bottleneck at a Customer Edge (UNI-C) Shaper

Shaping traffic prior to the ingress bandwidth profile policer dramatically changes the TCP behavior. Figure 18 shows the TCP segments acknowledged by the receiver over time when a shaper is added prior to the policer and both are configured with CIR equals 10 Mbps and CBS equals 10 segments. With a shaper the transmit-then-timeout pattern disappears and the TCP throughput is equal to CIR. The same result occurs with any CIR and CBS values as long as the CIR of the shaper (CIR_{shaper}) is less than or equal to the CIR of the bandwidth profile policer (CIR_{BWP}), and the CBS of the shaper (CBS_{shaper}) is less than or equal to the CBS of the bandwidth profile policer (CBS_{BWP}). What is happening is that the TCP bottleneck moves from the policer to the shaper, and when TCP ramps up its transmission rate above the bandwidth profile the shaper buffers the non-conformant packets rather than discarding them. The “shape” of the traffic reaching the policer is now conformant with the ingress bandwidth profile and so the policer does not discard any packets. But there is a trade-off. Introducing a buffer at the shaper means the effect of the buffer on TCP behavior needs to be considered.

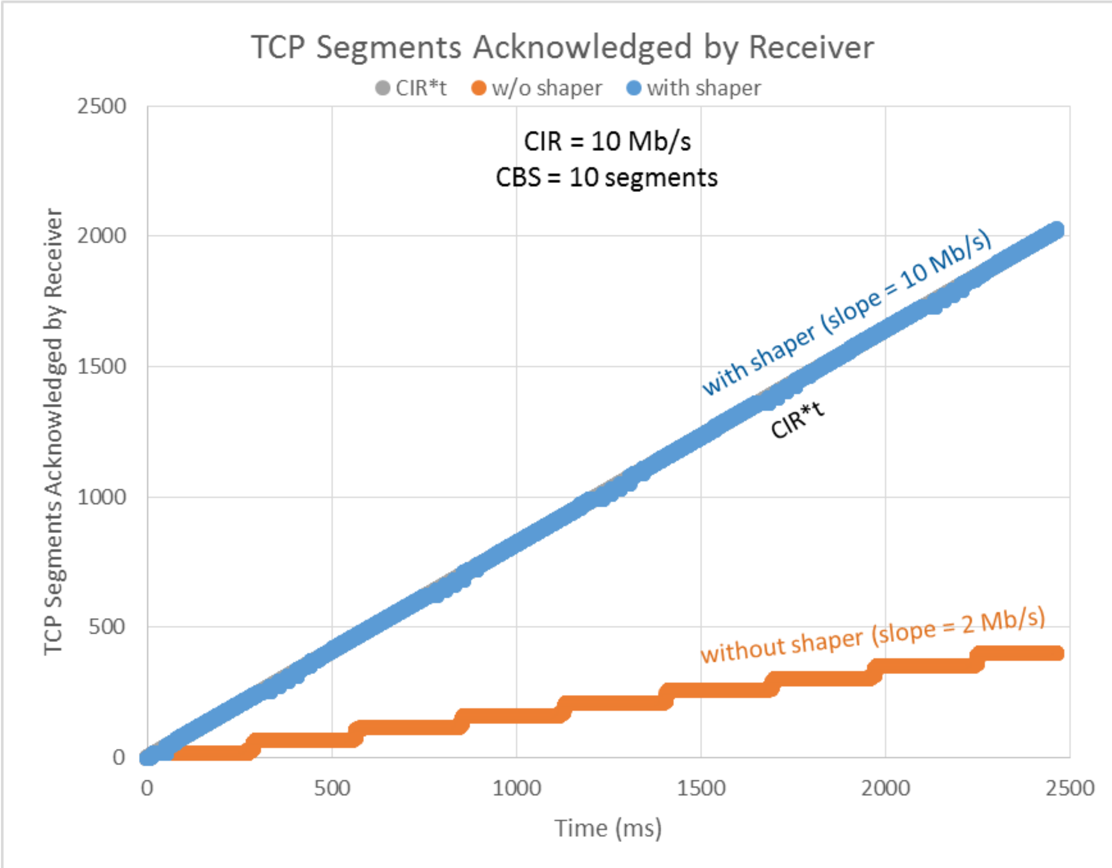


Figure 18: TCP behavior with and without a shaper

When the TCP bottleneck is at a shaper, TCP behavior is dominated by the capacity of the shaper buffer. In most cases the effect of the buffer is seen in the delay of the packets, not throughput, although for small buffers the TCP throughput may be affected. (Conceptually, in the limit, as the size of the buffer approaches zero the behavior of a shaper becomes indistinguishable from a policer.) The detailed behavior depends upon whether the shaper buffer capacity is greater than or less than the sum of the TCP window size of all the TCP flows using the EVC. Both cases will be discussed.

When the bottleneck bandwidth is at a shaper and the capacity of the shaper buffer is greater than sum of the receiver window size for all TCP flows, each flow will ramp up the amount of data in flight to the limit imposed by the receiver window size. The majority of the data in flight will be stored in the shaper buffer. (The actual amount of data stored is the sum of the window sizes of all flows minus the capacity of the TCP flows themselves.) In steady state the shaper buffer depth remains constant, which imposes a delay of the buffer depth times CIR_{shaper} on each packet. No packets will be lost, however, and the aggregate TCP throughput will be equal to CIR_{shaper}. An example is shown in Figure 19. The upper right plot shows the TCP segments acknowledged by the receivers of four TCP flows multiplexed into a single shaper and then an ingress bandwidth profile policer. The start times of the TCP flows are staggered. In this case the CIR_{shaper} is 95 Mb/s and the CBS_{shaper} is 8 segments. The green line is the sum of the segments received on each flow, and its slope matches CIR_{shaper}. The lower right plot shows the shaper buffer depth, which increases by the receiver window size (in this case equal to 25 segments on each flow) as each flow begins transmitting. The lower left plot shows the actual RTT of each packet, and it increases as the buffer depth increases. The advantage to sizing the buffer to be greater than the sum of the window sizes is that no packets are lost. The disadvantage is that the buffers can get to be very large introducing a large fixed delay, and knowing how large requires knowing the number of flows and the window size of each flow.

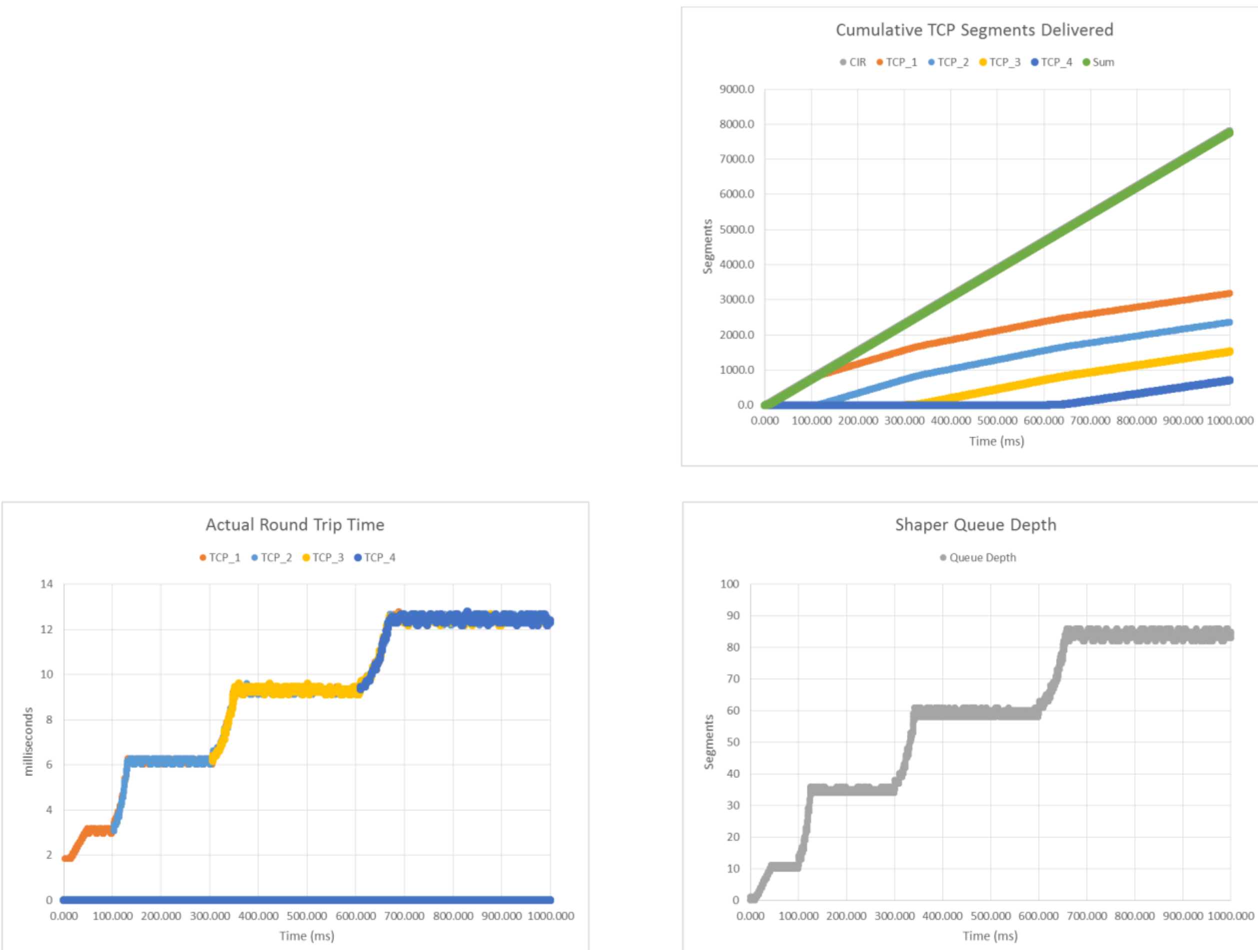


Figure 19: Multiple TCP flows with a shaper

When the bottleneck bandwidth is at a shaper and the capacity of the shaper buffer is less than the sum of the receiver window size for all TCP flows, each flow will ramp up the amount of data in flight until the shaper buffer fills and packets are dropped. TCP responds to the lost packets by reducing the amount of data allowed to be in flight by a factor of two, which forces the transmitter to wait until the number of packets currently in flight is reduced by a factor of two before transmitting new data. This gives the shaper buffer an opportunity to drain. When TCP resumes transmission it will ramp up the amount of data in flight again until the shaper buffer fills again and packets are dropped again. The cycle repeats, creating the "sawtooth" pattern in the buffer depth over time that is characteristic of TCP.

For a single TCP flow, throughput is maximized when the shaper buffer depth equals the capacity of the TCP flow. The capacity is the bandwidth delay product given by the bottleneck bandwidth times the round trip time of the TCP flow. With this buffer depth the TCP transmitter will resume transmitting new data just as the shaper buffer empties. A larger shaper buffer depth results in there being residual data in the shaper buffer when TCP transmissions resume. This establishes a minimum buffer depth that adds delay to each packet but does not increase throughput. A smaller buffer depth results in the shaper buffer emptying before the TCP transmitter resumes. This means the TCP flow is not kept filled to capacity, which reduces throughput. Therefore, given an estimate of the flow round trip time (RTT), the ideal shaper buffer depth can be calculated as

$$(4)\ \mathrm{CBS}^{*} = \mathrm{CIR} \times \mathrm{RTT}$$

For multiple TCP flows, all flows will experience packet loss when the shaper buffer is full. This tends to synchronize the TCP behavior resulting in a sawtooth pattern in the buffer depth over time with the same amplitude but higher frequency as that for a single TCP flow. This means that equation (4) can be applied to situations with multiple TCP flows, where the value of RTT is the average round trip time of all the flows. An example of this is shown in Figure 20. This shows a case that begins with a single TCP flow, and three more are added one second later. The three new flows immediately synchronize with the first. The shaper settings in this case are CIR_{shaper} equal 50 Mb/s, CBS_{shaper} equal 1 segment, and a maximum depth of 32 segments. The TCP flows have an RTT of 7.87 ms corresponding to a flow capacity (bandwidth delay product) of 32 segments.

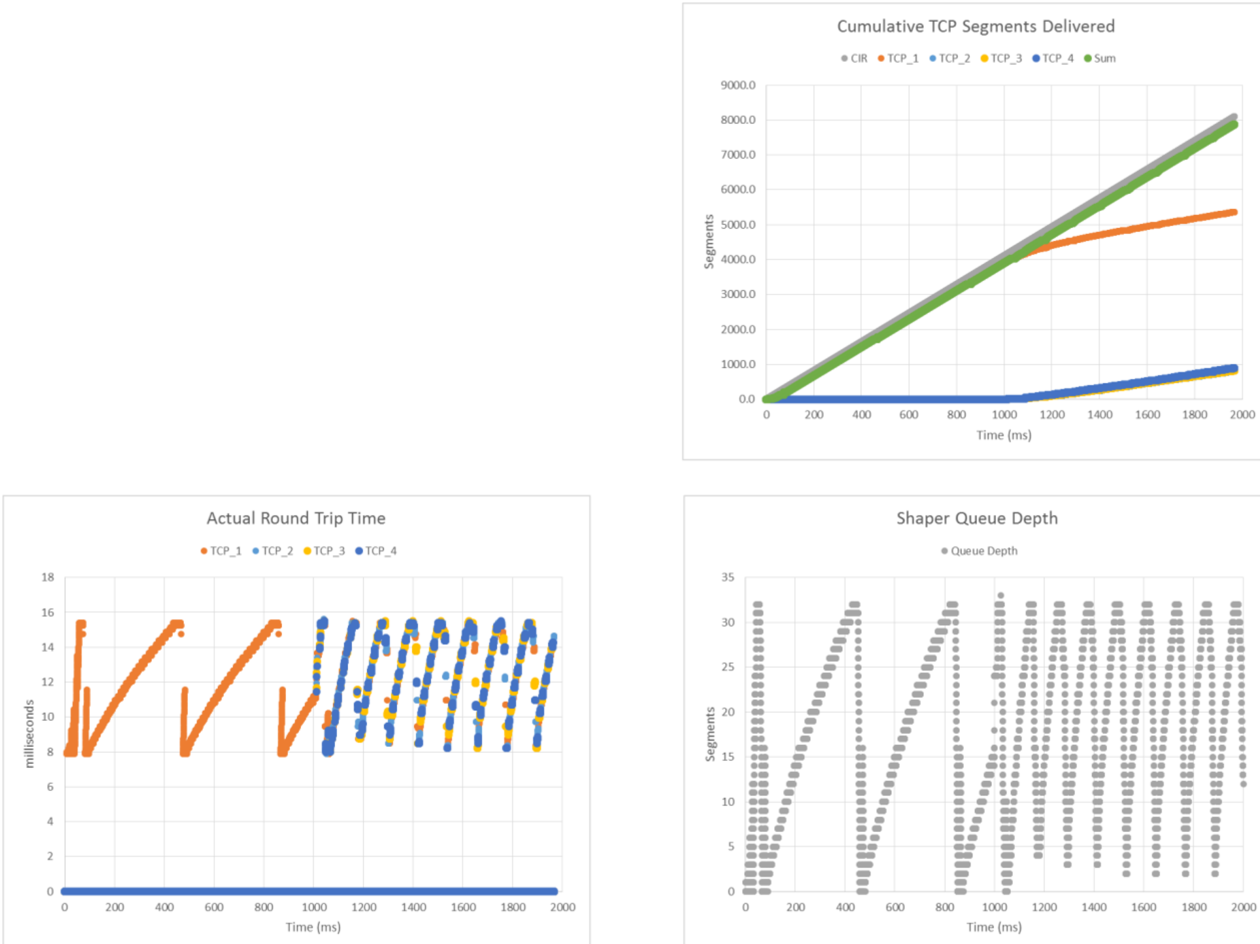


Figure 20: Multiple TCP flows with a shaper

The difficulty in applying equation (4) is estimating the RTT value. This needs to represent the average RTT of all the flows sharing the EVC, but the RTT of each flow includes the total delay between the transmitter and receiver, not just the delay contribution of the EVC. If the maximum shaper buffer depth turns out to be lower than the average RTT times CIR, then the aggregate throughput of the EVC will be lower than CIR. If the maximum shaper depth turns out to be higher than the average RTT times CIR, then the aggregate throughput of the EVC will equal CIR but there will be an incremental fixed delay on all packets (a phenomenon known as “buffer bloat”). The conclusion in equation (4) that the buffer size should be determined by the bandwidth delay product of the flow is familiar from studies of router buffer sizes. This is a well-researched area, and the research results for router buffers can be applied to the shaper buffer.

One example is that research suggests equation (4) holds for “small” numbers of flows (N < 100), but tends to overestimate the ideal shaper buffer depth for “large” numbers of flows (N > 500) [19]. The rationale is that as the number of flows gets large it is not possible for all flows to have a packet arriving at the shaper at the point when the sawtooth pattern reaches the maximum shaper depth. The result is that flows cannot synchronize so the sawtooth pattern breaks down and the difference between the maximum and minimum shaper buffer depth decreases. The conclusion is that ideal shaper depth can be reduced by the square root of the number of TCP flows when there is reason to assume a large number of long-lived flows:

$$(5)\ CBS^{*} = \frac{CIR \times RTT_{avg}}{\sqrt{N}} \mathrm{for large N}$$

Other research has targeted methods of reducing the average buffer occupancy, and therefore reduce the average delay, but not necessarily reducing the maximum queue depth necessary to maintain TCP throughput at CIR. This includes Random Early Detection (RED, or variations such as Weighted Random Early Detection (WRED)), which also attempt to break the synchronization of multiple TCP flows and therefore disrupt the “sawtooth” pattern. This also includes current research into Active Queue Management (AQM) based on measuring the time a packet spends in the buffer rather than the buffer depth.

The conclusion is that the use of shaping prior to an ingress bandwidth profile is highly recommended as a means of allowing TCP to get a high utilization of the bandwidth profile CIR while limiting the size of the CBS value.

G.2.3 TCP Bottleneck at an Egress Shaper

TCP behavior is dominated by the characteristics of the device at the bottleneck bandwidth regardless of where that device is located in the path relative to other devices that present less constraint on the bandwidth. A somewhat counter-intuitive consequence of this is that any buffering in the TCP path, whether before or after an ingress bandwidth profile policer, will have the same effect as a Customer Edge (UNI-C) shaper if the rate at which the buffer is serviced is more constraining than the CIR of the policer. Specifically this means that a shaper at the egress UNI has basically the same effect on TCP behavior as a Customer Edge (UNI-C) shaper at the ingress UNI if the CIR of the egress bandwidth profile is less than the CIR of the ingress bandwidth profile, and the CBS of the egress bandwidth profile is less than the CBS of the ingress bandwidth profile. The potential benefit of this to a CEN Operator or Service Provider is that egress shaping can be used to avoid the pathological interaction between TCP and an ingress bandwidth profile policer in situations where a Customer Edge (UNI-C) shaper is not available or is not within the Operator or Service Provider’s control.

It should be noted that even though the TCP behavior is the same whether shaping occurs at an ingress or egress UNI, there will be a significant difference in the performance metrics of the EVC. An ingress shaper or ingress bandwidth profile policer will delay or discard frames before they are admitted to the EVC and before they are declared to be qualified frames for performance monitoring. Therefore the performance metrics of the EVC will not reflect this delay or frame loss. With the bottleneck bandwidth at an egress shaper, however, it will be qualified frames that are delayed or discarded as TCP adapts its transmission rate, and this will affect the EVC performance metrics.

G.2.4 Summary of Key Findings

This section has presented the results of simulations of the interaction between TCP and a Carrier Ethernet Service Bandwidth Profile, both when the Bandwidth Profile is enforced only by an ingress Bandwidth Profile policer and when that policer is complemented with a shaping function in the CE or the CEN. The key findings are as follows:

1. When the TCP transmission rate is constrained only by an ingress Bandwidth Profile policer, the TCP transmitter will send as fast as possible until the maximum burst tolerance of the Bandwidth Profile (CBS) is exceeded. At this point the sudden loss of packets causes the TCP transmitter to cease transmissions for a timeout interval before rapidly increasing the transmission rate until the maximum burst tolerance is exceeded again.
    a. This limits the TCP throughput to approximately CBS worth of data transmitted every timeout interval. With CBS values that are reasonable (from the viewpoint of buffer management and delay management within the CEN) this typically results in TCP throughput well below the Service rate (the Bandwidth Profile CIR).
    b. To get TCP throughput close to the Service rate, CBS must be increased to be greater than or equal to CIR multiplied by the timeout interval. With a minimum timeout value typically at 250ms, this results in exceedingly large CBS values.
2. When the TCP transmission rate is contrained by a shaper, which can enforce a Bandwidth Profile by delaying frames until they are conformant rather than by immediately discarding them, TCP can adapt to the Service rate without incurring timeouts. The result is that when the ingress Bandwidth Profile policer is complemented by a shaper, either in the CE or the CEN, TCP throughput close to the Service rate is achievable with significantly reduced values of CBS.
    a. A perhaps surprising finding is that as long as TCP is most constrained by the shaper, it doesn't matter whether the shaper is before the policer or after the policer. The egress burst size configuration of the shaper is not very significant to the TCP behavior, and can be set as low as the shaper implementation allows. What is significant is that the sum of the egress burst size of the shaper and maximum queue depth of the shaper is greater than or equal to the bandwidth delay product of the TCP flow(s). In this case the "bandwidth" is the Service rate (CIR) and the "delay" is the best approximation of the average end-to-end round trip time of the TCP connection.
    b. When the shaping is done before the ingress Bandwidth Profile policer (i.e. in the CE), the CBS value can be reduced to match the egress burst size of the shaper. In theory this can be just one to three times the maximum frame size. In practice the minimum CBS value is determined by the minimum configurable egress burst size of the shaper.

When the shaping is done after (or in conjunction with) the ingress Bandwidth Profile policer (i.e. in the CEN), the CBS value of the ingress Bandwidth Profile policer needs to be greater than or equal to the TCP bandwidth delay product (otherwise the policer becomes the primary constraint on TCP rather than the shaper).

Appendix H Guidelines for choosing value for CBS (Informative)

H.1 General Guidelines

It is difficult to provide specific values for setting CBS since the need for CBS depends on several factors such as the Subscriber's application and traffic characteristic, whether traffic shaping is being done either by the Customer Edge equipment or the CEN[11], the Class of Service and Performance Objectives, etc. Each of these has a different impact on the CBS.

It is commonly assumed that increasing the value of CBS is a simple solution to Subscriber performance issues and that providing large values for CBS has no network or cost impact. This is not true. CBS is, in essence, the right to transmit at line speed for a period of time. If multiple Subscribers have the ability to present a large bursts of traffic to the CEN at rates substantially higher than their CIRs, and the network traffic engineering did not take this into account, internal network links can be overwhelmed with large queues, causing increased delay, delay variation, and frame loss. Therefore, care should be taken when configuring CBS for a Bandwidth Profile Flow that has tight objectives for frame delay (FD or MFD) and/or frame delay variation (FDR or IFDV).

This appendix provides some general guidelines and direction for the selection of an appropriate value for CBS.

1. The absolute minimum value for CBS is the Maximum Frame Size (MFS) for the EVC or OVC. Since the Bandwidth Profile algorithm attempts to remove L bytes from the token bucket (where L is the frame size), it is not possible to admit a maximum-sized frame if CBS is less than MFS. That being said, CBS should never be equal to MFS since this does not allow for the common case of back-to-back MFS frames. It is suggested that CBS be ≥ 3*MFS (this would be about 5KB for 1522 byte MFS and 6KB for 2000 byte MFS).
2. Providing guidance for an upper bound for CBS is more difficult since it is application dependent.

- For constant bit rate (or near CBR) applications (e.g., Circuit Emulation), a maximum CBS value of 3*MFS or 4*MFS, is suggested. These applications do not burst so a small value of CBS is sufficient.
- For applications requiring tight control of delay and delay variation, a small CBS is suggested in order to avoid long queues at egress interfaces so a maximum value of 8*MFS is suggested.
- The same guidance (i.e. a maximum of 8*MFS) would be acceptable for heavily interactive applications which alternately send a few packets and then wait for a response (e.g., telnet, ssh, database transactions, etc.).

[11] Although it is common to assume that traffic shaping is done at the Customer Edge (CE), the results in Appendix G indicate that the benefits of shaping are seen if it is done anywhere in the path of the flow, i.e., at the CE or somewhere in the CEN.

- For TCP/IP sessions with heavy data transfer in either or both directions (e.g., FTP file transfer) see the guidance in bullets 4 and 5 below.

3. CBS should, in most cases, be inversely proportional to CIR / <line rate>. When the ratio is small (low CIR and/or high line rate) the token bucket can be emptied quickly but is filled slowly. As the ratio increases the bucket is filled more quickly. In the limit, when the ratio =1, the bucket can't be emptied faster than it is being filled and a minimal CBS is sufficient.

4. File-transfer types of applications in which one side is attempting to achieve full CIR throughput for a long transaction require more CBS in order to allow TCP to open its transmit window sufficiently. The behavior described here is discussed at great length in Appendix G, but there are two cases depending on whether shaping is being done on the TCP flow:

    a. Shaping is not being done. In this case TCP performance for file transfer-type applications will be poor (sometimes as low as 10% of CIR) unless CBS is very large (on the order of ¼ second of traffic at CIR). This problem is the cause of many Subscribers requesting large values for CBS.

    b. Shaping is being done. In this case TCP can achieve very good performance (>90% of CIR) even with modest values of CBS (e.g., 3 – 8*MFS) as long as the shaper parameters are matched to the ingress Bandwidth Profile Flow parameter values (see section H.2).

    Note that MEF 13 [18], the UNI Type 1 IA, recommends that the UNI-C "shape its traffic to the contracted BWP in order to receive the contracted QoS commitments" (section 6.1.2 for UNI type 1.1 and section 6.2.3 for UNI type 1.2).

5. If the Bandwidth Profile Flow is supporting multiple simultaneous (but statistically multiplexed) flows at a UNI or ENNI (a common situation), it is appropriate to multiply these recommendations by a factor such as 4 to 8 to account for this. In general, except in the TCP/non-shaping case discussed in point #4a above, it is suggested that CBS not exceed about 50*MFS (this is about 80KB at a UNI with MFS=1522).

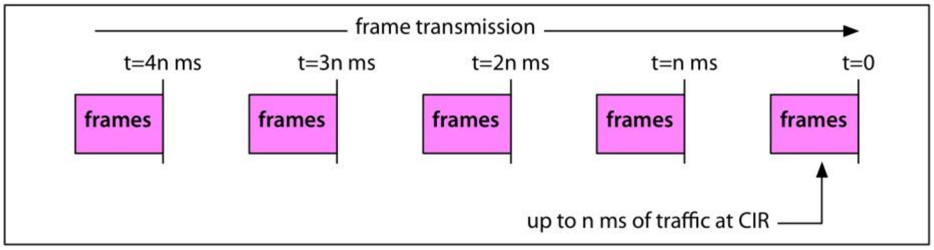
H.2 Practical Considerations for Burst Size and CE Shapers

As noted in Appendix G, due to the interaction between the policer and TCP's congestion control algorithm, reasonable performance (for both the Service Provider and the Subscriber) can only be achieved with the use of a Shaper in the path of the session, either at the Customer Edge (CE) or within the CEN.

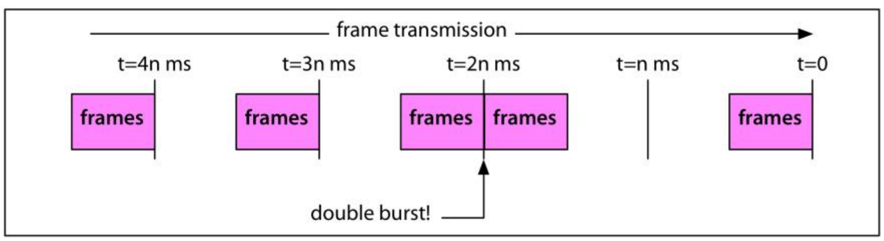
When a shaper is used at the CE edge it must be configured to schedule traffic in bursts that don't deplete the policer's token bucket since that will result in dropped frames[12]. Point number 5 in the previous section indicates that a policer CBS value in the vicinity of 80KB (50*MFS) is a reasonable upper bound for most situations.

[12] In reality the burst can be larger than CBS, especially if CIR is high compared to the line rate, since the token bucket is being replenished at CIR during the burst.

Many popular bridges and routers do not allow the shaper burst to be configured that low. A common approach is to specify configuration of the shaper burst in the time domain based on CIR (e.g., 1ms at CIR). As noted in bullet 3 in the previous section, this is counter-intuitive (since a time-based burst increases with CIR). In some devices the burst can be set to as low as 1ms, but in other devices the minimum burst is 4ms. This could result in a pattern that looks like this:



Depending on relationship of CIR to the line rate, this can create large bursts of traffic (the shaded boxes in the diagram). The problem is compounded by the fact that many shapers will accept traffic for transmission any time during the n ms window. So if there are no frames to transmit at the beginning of the window and a burst shows up near the end of the window, the result would be a pattern that looks like this:



This results in a burst that is twice as large. In both cases, appropriate TCP performance will be achieved only if the policer's CBS is configured to be consistent with the scheduling approach of the shaper. In the second diagram, the policer must be prepared to absorb a burst that is twice as large as expected if only the burst interval is considered.

The following tables demonstrate this with different values of CIR on a 1G physical interface with an MFS of 1522 bytes. The first table shows a configuration with a 1ms shaper burst and the second one with a 4ms shaper burst.

CIR in Mbps	MFS frames in 1ms	Replenished 1ms	CBS Needed (#MFS)
100	8.1	0.8	15

200	16.2	3.2	26
300	24.3	7.3	35
400	32.4	13.0	39
500	40.5	20.3	41
600	48.6	29.2	39
700	56.7	39.7	35
800	64.8	51.8	26
900	72.9	65.6	15

CIR in Mbps	MFS frames in 4ms	Replenished 4ms	CBS Needed (#MFS)
100	32.4	3.2	59
200	64.8	13.0	104
300	97.2	29.2	137
400	129.6	51.8	156
500	162	81.0	163
600	194.4	116.6	156
700	226.8	158.8	137
800	259.2	207.4	104
900	291.6	262.4	59

The first column is the CIR. The second column is the number of Maximum Frame Size (MFS) frames that the shaper can burst at 1ms (top chart) and 4ms (bottom chart). The third column is the number of MFS frames that are replenished in the token bucket at the CIR rate in the time it took to transfer those frames at line rate (1G). The fourth column is the number of MFS frames needed in the policer's CBS in order to handle the burst assuming the possible of a double burst.

From this table, it is clear that in most cases with a 1 ms shaping interval all of the rates up to 1G can be achieved within the guidance provided in H.1 (i.e., ≤ 50*MFS). However with a 4 ms shaping interval substantially larger values of CBS are required at the policer in order to minimize frame loss.

In all cases the Subscriber and the Service Provider should work together to determine the appropriate configuration of the shaper and the policer in order to achieve acceptable results. Network Operators should be aware of the traffic engineering implications of having Subscribers with large CBS allowance and design accordingly.