Technical Specification

MEF 5

Traffic Management Specification: Phase I

May 2004

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 Metro Ethernet 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:

(a) 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

(b) 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

(c) 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 Metro Ethernet 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. Abstract

This document defines the Traffic and Performance parameters that may be specified as part of an Ethernet service level specification (SLS). These parameters are defined to provide Subscribers with both a quantitative and qualitative characterization of the service they are getting from a provider, sufficient to allow configuration of CE and integration of the service into the Subscriber network. Consistent use of these parameters in the SLS will allow Subscribers to compare various service offerings. Additionally, Service Providers may use these parameters as guidelines for specifying CoS-based Ethernet services.

2. Terminology

CBS	Committed Burst Size
CF	Coupling Flag
CIR	Committed Information Rate
CM	Color Mode
Color-aware	A traffic management capability that takes into account the Service Frame color when determining whether a Service Frame is conformant or non-conformant to a Bandwidth Profile.
Color-blind	A traffic management capability that ignores the Service Frame color (if any) when determining whether a Service Frame is conformant or non-conformant to a Bandwidth Profile.
Color Mode	CM is a Bandwidth Profile parameter. The Color Mode parameter indicates whether the color-aware or color-blind mode is employed at the UNI. It takes a value of “color-blind” or “color-aware” only.
Committed Burst Size	CBS is a Bandwidth Profile parameter. It limits the maximum number of bytes available for a burst of ingress Service Frames sent at the UNI speed to remain CIR-conformant.
Committed Information Rate	CIR is a Bandwidth Profile parameter. It defines the average rate in bits/s of ingress Service Frames up to which the network delivers Service Frames and meets the performance objectives defined by the CoS Service Attribute.

Coupling Flag	CF is a Bandwidth Profile parameter. The Coupling Flag allows the choice between two modes of operations of the rate enforcement algorithm. It takes a value of 0 or 1 only.
EBS	Excess Burst Size
EIR	Excess Information Rate
Excess Burst Size	EBS is a Bandwidth Profile parameter. It limits the maximum number of bytes available for a burst of ingress Service Frames sent at the UNI speed to remain EIR-conformant.
Excess Information Rate	EIR is a Bandwidth Profile parameter. It defines the average rate in bits/s of ingress Service Frames up to which the network may deliver Service Frames without any performance objectives.
FD	Frame Delay
FDV	Frame Delay Variation
FLR	Frame Loss Ratio
Frame Delay	The time required to transmit a Service Frame from source to destination across the metro Ethernet network. Frame Delay measure is applicable to green Service Frames only.
Frame Delay Variation	A measure of the variation in the delays experienced by different Service Frames belonging to the same CoS instance. Frame Delay Variation is applicable to green Service Frames only.
Frame Loss Ratio	Frame Loss Ratio is a measure of number of lost frames inside the MEN. Frame Loss Ratio is expressed as a percentage.
SLA	Service Level Agreement
SLS	Service Level Specification
Service Level Agreement	The contract between the Subscriber and Service Provider specifying the agreed to service level commitments.
Service Level Specification	The technical specification of the service level being offered by the Service Provider to the Subscriber.

3. Scope

Ethernet technology is intended to be used for the support of different applications. The ability to support those applications is fundamentally related to the ability of the network to support the appropriate set of traffic management mechanisms and parameters.

The scope of this specification is the definition of traffic management mechanisms and parameters for Ethernet services. This helps the Service Provider and its Subscriber to mutually agree on a Service Level Specification (SLS) associated with a given service instance. In particular this document specifies a common way for specifying:

- Bandwidth Profile: The bandwidth profile consists mainly of a set of traffic parameters and rules for the disposition of each Service Frame based on its level of compliance with the Bandwidth Profile parameters.
- Performance Definitions: A set of service performance metrics in terms of Frame Delay, Frame Delay Variation, and Frame Loss Ratio that are useful in aligning a service with the application that will use it.

An SLS will typically become part of a Service Level Agreement (SLA) which includes business aspects of the agreement between the Subscriber and the Service Provider such as price and performance guarantees. To support such an SLA, it will be important for the Service Provider to have technologies and best practices to do such things as monitor and verify service performance. Although important, aspects like pricing and monitoring are beyond the scope of this Technical Specification

The parameters and service attributes defined in this document are used by the Ethernet Services Definition specification for specifying Ethernet services parameters that are defined by the Service Model [1]. The relationship between this document, the Ethernet Services Definitions document and the Service Model is illustrated in Figure 1.

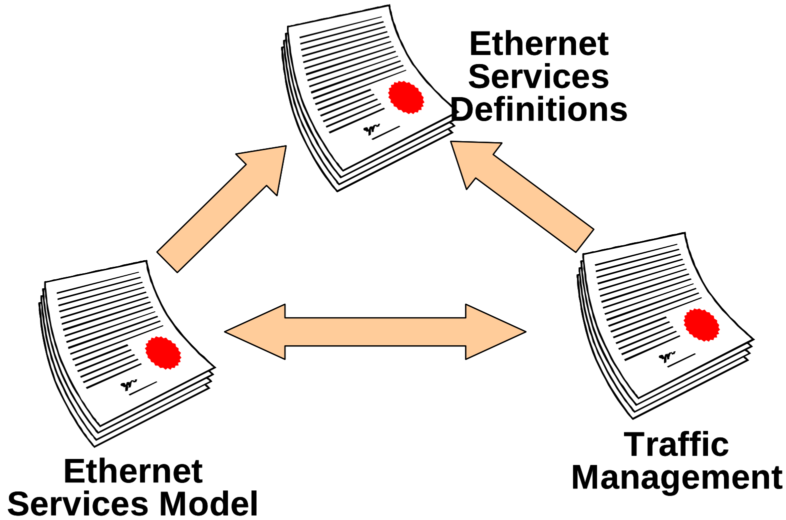

Figure 1: Relationship between Different MEF Services Group documents

4. 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 [2]. All key words must be in upper case, bold text.

5. Ethernet Traffic Management Overview

The Ethernet traffic management offers the essential components with which a Service Provider can offer a Subscriber a range of services that are differentiated based on some performance metrics, e.g. Frame Loss Ratio and Frame Delay. Subscribers request a specific performance level as stated in the SLS. Service providers achieve levels of performance assurances by employing the necessary traffic management mechanisms both at the edge and inside the metro Ethernet network (MEN).

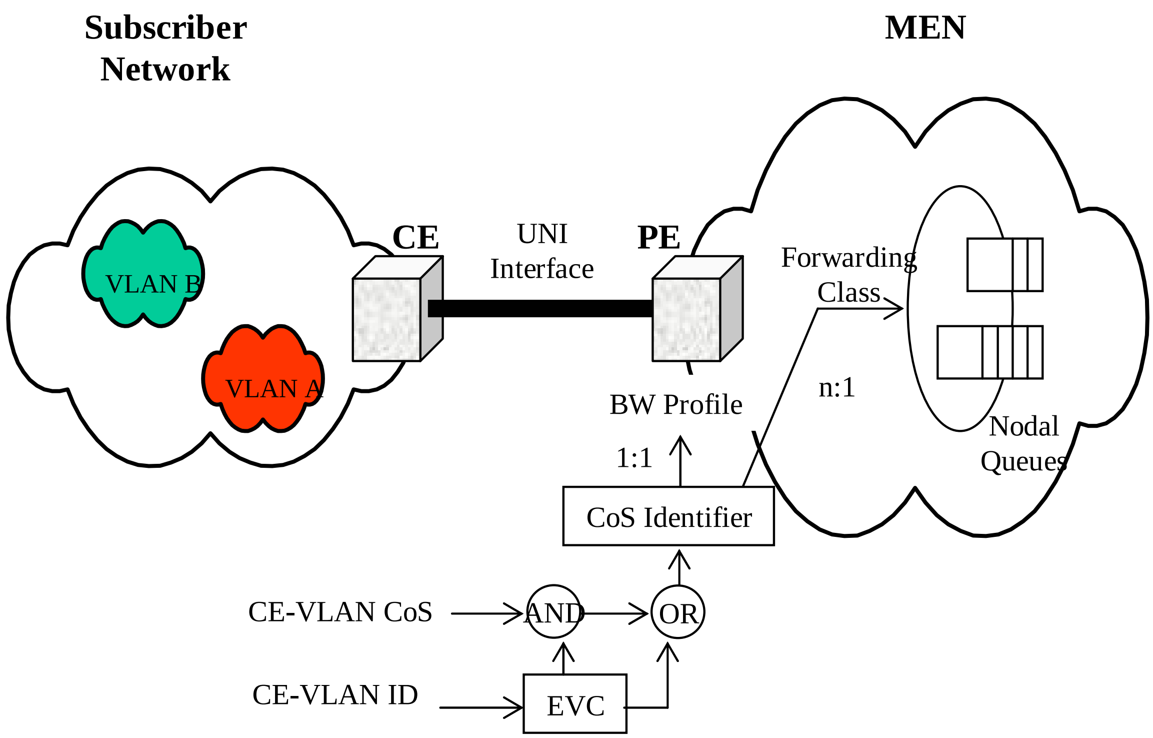

Figure 2: Ethernet Traffic Management Overview

Figure 2 shows the interaction between the Subscriber network and the provider network across the User Network Interface (UNI) between the CE and the PE. Inside the Subscriber network, a Subscriber may implement IEEE 802.1Q VLANs and traffic management mechanisms. Those mechanisms are beyond the scope of this specification. Across the UNI, CE-VLAN ID as defined in [1] is mapped to a single EVC. Mapping rules and bundling capabilities are specified in [1]. The EVC or a combination of an EVC and CE-VLAN-CoS as defined in [1] is used by the PE determine the Class of Service (CoS) instance using the CoS Idenitifier[1].

The CoS Identifier is used to identify the CoS instance applicable to a flow of Service Frames. A maximum of 32K (4095*8) CoS Identifiers **MAY** be supported at a UNI. A much smaller number of Classes of Services **MAY** be supported. CoS Identifier can be viewed as an index that points to two fundamental building blocks that together define the service level received by a Service Frame. Those building blocks are the Bandwidth Profile and the forwarding class.

The Bandwidth Profile, if present, is a set of traffic parameters that governs the expected arrival pattern of Subscriber traffic per CoS instance. It provides a deterministic upper bound to the expected volume of traffic per CoS instance. The Bandwidth Profile parameters are inputs to a metering algorithm that verifies the conformance of incoming Service Frames per CoS instance. This allows Service Providers to engineer network resources to satisfy service performance requirements. Multiple Bandwidth Profiles may exist at the same UNI.

The Bandwidth Profile, if present, **SHOULD** be enforced by the provider's network since it is part of the SLS and is agreed upon between the Subscriber and the Service Provider. QoS policy rules by way of rate enforcement algorithms are typically used to identify the specific actions taken by a Service Provider when Service Frames violate a Bandwidth Profile. The actions may be to drop or recolor the service frame to indicate high discard precedence within the MEN. The outcome of any such rate enforcement, within the Service Provider network is a set of Service Frames that are labeled as Green, Yellow, or Red based on their level of conformance to Bandwidth Profiles.

The forwarding class defines the treatment inside the provider network received by a Service Frame belonging to a particular CoS instance. For example Service Frames may be forwarded at the highest priority available at the nodal queues to assure some level of delay requirement. Forwarding class is outside the scope of this Specification.

6. Bandwidth Profile Service Attributes

The Bandwidth Profile defines the set of traffic parameters applicable to a sequence of Service Frames. Associated with the bandwidth profile is a rate enforcement algorithm to determine Service Frame compliance with the specified parameters. Rate enforcement also includes the disposition of non-compliant Service Frames by either dropping or marking.

6.1 Bandwidth Profile Parameters

The parameters comprising the Bandwidth Profile parameters are:

1. **Committed Information Rate** (CIR) expressed as bits per second. CIR **MUST** be ≥ 0.
2. **Committed Burst Size** (CBS) expressed as bytes. When CIR > 0, CBS **MUST** be greater than or equal to the maximum Service Frame size as specified in [1].

[1] CoS Identifier is normatively described in [1]

3. **Excess Information Rate** (EIR) expressed as bits per second. EIR **MUST** be ≥ 0
4. **Excess Burst Size** (EBS) expressed as bytes. When EIR > 0, EBS **MUST** be greater than or equal to the maximum Service Frame size as specified in [1].
5. **Coupling Flag** (CF) **MUST** have only one of two possible values, 0 or 1.
6. **Color Mode** (CM) **MUST** have only one of two possible values, "color-blind" and "color-aware"

Since the coupling Flag has negligible effect in color blind mode, a service definition that uses color blind operation **MAY** be defined without specifying the value of the coupling flag.

6.2 Enforcement of Bandwidth Profile Parameters

Each incoming Service Frame is classified to determine which, if any, Bandwidth Profile is applicable to the Service Frame[2]. Operation of the Bandwidth Profile algorithm is governed by the six parameters, <CIR, CBS, EIR, EBS, CF, CM>. The algorithm declares Service Frames as compliant or non-compliant relative to the Bandwidth Profile parameters. The level of conformance is expressed as one of three colors, Green, Yellow, or Red[3].

The Bandwidth Profile algorithm is said to be in color aware mode when each incoming Service Frame already has a level of conformance color associated with it and that color is taken into account in determining the level of conformance to the bandwidth profile parameters. The Bandwidth Profile algorithm is said to be in color blind mode when level of conformance color (if any) already associated with each incoming Service Frame is ignored in determining the level of conformance. Color blind mode support is **REQUIRED** at the UNI. Color aware mode is **OPTIONAL** at the UNI. The color mode of operation **MUST** be determined using the parameter CM.

Metered Service Frames that are not declared Green are marked, remarked, or dropped according to QoS and SLS policies.

6.2.1 Bandwidth Profile Algorithm

The Bandwidth Profile algorithm is shown in Figure 3. For a sequence of ingress Service Frames, $\{t_j,\lambda_j\}_{j\geq 0}$, with arrival times t_j and lengths λ_j, the level of conformance color assigned to each Service Frame **MUST** be defined according to the algorithm in Figure 3. For this algorithm, $B_c(t_0)$ = CBS and $B_e(t_0)$ = EBS. $B_c(t)$ and $B_e(t)$ are the number of bytes in the Committed and Excess token buckets respectively at a given time t.

[2] Recall that in [1], a Service Frame is defined as any Ethernet Frame transmitted across the UNI and thus a Layer 2 Control Protocol Ethernet frame is a Service Frame.

[3] The categorization of a Service Frame does not imply any change to the content of the frame. Certain approaches to network implementation may "mark" frames internal to the MEN but such procedures are beyond the scope of this Technical Specification.

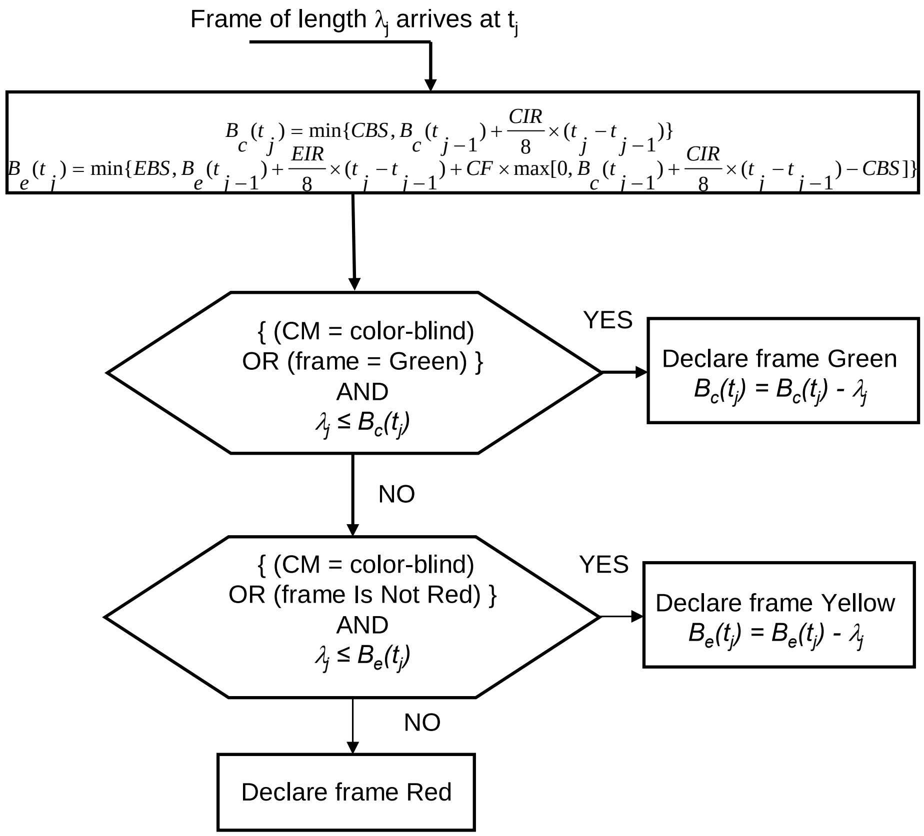

Figure 3: The Bandwidth Profile Algorithm

Note that the algorithm in Figure 3 does not define an implementation of any network equipment. In fact, since the behavior is described with real numbers for representing time, exactly implementing the behavior is theoretically impossible. However, an implementation should be such that over any time interval $[t_j, t_k]$ the amount of traffic accepted as green, W_G and the amount of traffic accepted as yellow, W_Y are lower bounded by the values:

$$W_G \geq B_c(t_j) + CIR \times (t_k - t_j)$$

$$W_Y \geq B_e(t_j) + EIR \times (t_k - t_j)$$

provided that the ingress traffic is greater than these values.

The Coupling Flag CF is set to either 0 or 1. The choice of the value for CF has the effect of controlling the volume of the yellow Service Frames admitted to the network. When CF is set to 0, the long term average bit rate of bytes in yellow service frames admitted to the network is bounded by EIR. When CF is set to 1, the long term average bit rate of bytes in yellow Service Frames admitted to the network is bounded by CIR + EIR depending on volume of the offered green Service Frames. In both cases the burst size of the yellow Service Frames admitted to the network is bounded by EBS.

6.2.2 Service Frame Disposition

The second step of Bandwidth Profile rate algorithm is the disposition of the ingress Service Frame based on its level of conformance and **MUST** be as described in Table 1.

Level of Conformance	Service Frame Disposition
Red	Discard
Yellow	Deliver according to the Service Attributes of the service instance but SLS performance objectives do not apply.
Green	Deliver according to the Service Attributes of the service instance and SLS performance objectives apply.

Table 1: Service Frame Disposition

6.3 Bandwidth Profile Applicability

There are three ways in which a Bandwidth Profile can be applied:

- Per Ingress UNI,
- Per EVC, and
- Per EVC and CE-VLAN CoS .

Normative descriptions are included in [1]. Each is described below for convenience.

The Ingress Bandwidth Profile per Ingress UNI provides a bandwidth profile that applies to all ingress Service Frames at the UNI. In the Example of Figure 4, ingress Service Frames for the three EVCs would all be subject to a single Bandwidth Profile. The Bandwidth Profile per Ingress UNI manages bandwidth non-discriminately for all EVCs at the UNI, i.e. some EVCs may get more bandwidth while others may get less.

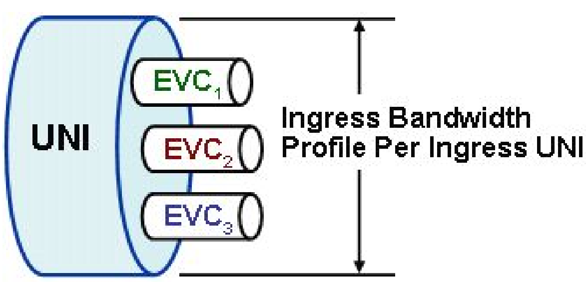

Figure 4: Ingress Bandwidth Profile per Ingress UNI

The Ingress Bandwidth Profile Per EVC provides a bandwidth profile that applies to all Service Frames for a given EVC as is illustrated in the example in Figure 5. In this example, EVC_1 could have CIR=15 Mbps, EVC_2 could have CIR = 10 Mbps, and EVC_3 could have CIR = 20 Mbps.

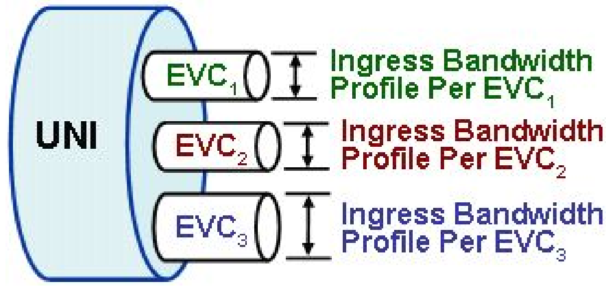

Figure 5: Ingress Bandwidth Profile per EVC

The Ingress bandwidth profile per EVC and CE-VLAN CoS provides a bandwidth profile that applies to all Service Frames for a given EVC that are identified via the CoS Identifier. Refer to the Example in Figure 6. In this example, there are three Class of Service Identifiers, each with a separate Bandwidth Profile

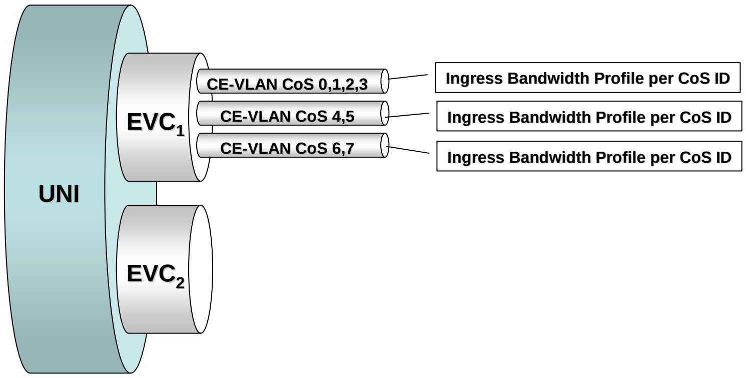

Figure 6: Ingress Bandwidth Profile per EVC and CE-VLAN CoS

As specified in [1], Bandwidth Profile can be applied in any fashion so long as any given ingress Service Frame is subject to at most one Bandwidth Profile. For example, if there is a Bandwidth Profile for the ingress UNI, then there cannot be any Bandwidth Profile per EVC or per EVC and CE-VLAN CoS Identifier. As another example, in the configuration of Figure 6, there cannot be a Bandwidth Profile for EVC_1. Note also for the configuration in Figure 6, that it is possible to configure a per EVC Bandwidth Profile for EVC_2 but there happens to not be a Bandwidth Profile for EVC_2 in the Figure.

7. Class of Service Performance Attributes (Point-to-Point EVC)

The CoS Performance Attributes specify the Service Frame delivery performance. Three performance attributes are considered in this specification. Those are Frame Delay Performance, Frame Delay Variation Performance, and Frame Loss Ratio. If defined, the Performance Attributes **MUST** apply to all Service Frames with their level of Bandwidth Profile conformance determined as Green, and with a particular Class of Service Identifier on a Point-to-Point EVC. Performance Attributes **MUST NOT** apply to Service Frames with level of conformance determined as Yellow or Red. Performance Attributes for a Multipoint-to-Multipoint EVC are beyond the scope of Phase 1 of this Technical Specification. Service Performance Attributes are applicable to Green Service Frames that arrive at the ingress UNI during a time interval **T**, and are successfully delivered to the egress UNI. A successfully delivered frame is a frame that is received not in duplication and its FCS is valid.

7.1 Frame Delay Performance

The Frame Delay for a Service Frame is defined as the time elapsed from reception at the ingress UNI of the first bit of the ingress Service Frame until the transmission of the last bit of the Service Frame at the egress UNI. This delay is illustrated in Figure 7. Note that this definition of Frame Delay for a Service Frame is the one-way[4] delay that includes the delays encountered at the ingress and egress UNI as well as that introduced by the MEN.

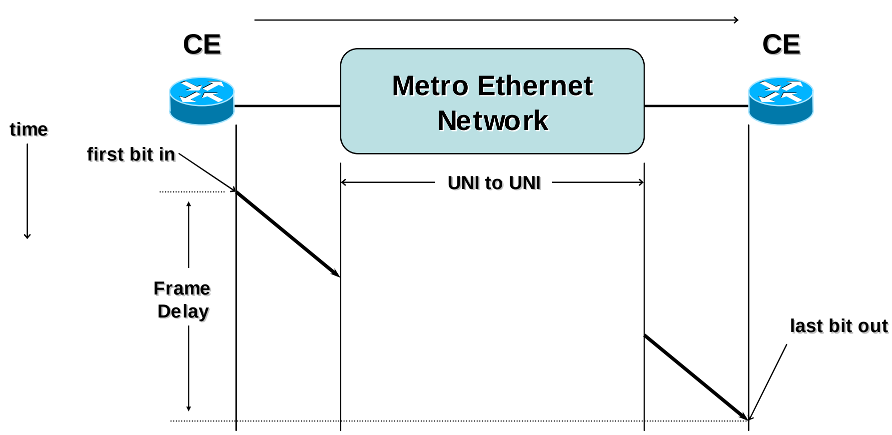

Figure 7: Frame Delay for Service Frame

Frame Delay Performance for a particular Class of Service instance on a Point-to-Point EVC for a time interval **T SHALL** be defined as the *P*-Percentile of the delay for all Green Service Frames successfully delivered between the UNI pairs during the interval **T**. The Unicast, Broadcast, Multicast and Layer 2 Control Protocol Service Frame Delivery service attributes define which Service Frames should be successfully delivered. Refer to [1] for the definitions of these service attributes.

To restate the definition mathematically, let S_T be the set of Frame Delay values for all successfully delivered Green Service Frames. S_T can be expressed as, $S_T = \{ d_1, d_2, ..., d_N)$ where d_i is the delay of the i^{th} Service Frame. Then the Frame Delay Performance, $\bar{d}_T$ can be expressed as

[4] One-way delay is difficult to estimate and therefore one way delay estimates may be approximated from two way estimates, however these techniques are beyond the scope of this document.

$$\bar{d}_T = \min\left\{ d_i \mid P \le \frac{100}{N} \sum_{j=1}^{N} I\left(d_i, d_j\right) \right\}$$

Where,

$$I(d_i, d_j\} = \begin{cases} 1 \text{ if } d_i > d_j \\ 0 \text{ otherwise} \end{cases}$$

The parameters of the Frame Delay Performance and are given in Table 2.

Parameter	Description
T	The interval on time over which the population of interest of green frames arrives at the ingress UNI
P	The percentile of the Frame Delay performance
$\hat{d}$	Frame Delay performance objective

Table 2: Frame Delay Performance Parameters

Given **T**, *P*, and a Frame Delay performance objective $\hat{d}$, expressed in time units, the Frame Delay performance **SHALL** be defined as met over the time interval **T** if and only if $\bar{d}_T \le \hat{d}$.

7.2 Frame Delay Variation Performance

Frame Delay Variation (FDV) is a measure of the variations in the Frame Delay performance among a set of Service Frames. FDV is applicable to all successfully delivered Green Service Frames for a particular Class of Service Identifier on a Point-to-Point EVC for a time interval **T**. The Unicast, Broadcast, Multicast, and Layer 2 Control Protocol Service Frame Delivery service attributes define which Service Frames should be successfully delivered. Refer to [1] for the definitions of these service attributes.

The Frame Delay Variation performance **SHALL** be defined as the *P*-percentile of the difference of the one way delay of Green frame pairs that arrive at the ingress UNI within a time interval **T** and that have an inter-frame interval time of Δt seconds. This definition is in agreement with the IP packet delay variation definition given in [3] where the delay variation is defined as the difference between the one-way delay of two packets selected according to some selection function and are within a given interval [I_1, I_2].

The choice of the value for Δt is related to the application timing information. As an example for voice applications where voice frames are generated at regular intervals, Δt may be chosen to be few multiples of the inter-frame time.

Let a_i be the time of the arrival of the first bit of the i^{th} Green frame at the UNI, then the two frames i and j are selected according to the selection criterion:

$$\{a_j - a_i = \Delta t \quad and \quad j > i\}$$

Let r_i be the time frame i is successfully received (last bit of the frame) at the destination, then the difference in the delays encountered by frame i and frame j is given by:

$$\Delta T_{ij} = (a_j - a_i) - (r_j - r_i) = (r_i - a_i) - (r_j - a_j) = d_i - d_j$$

With d_j being the delay of the j^th frame, a positive value for ΔT_{ij} implies that the two frames are lumped together at the destination while a negative value for ΔT_{ij} implies that the two frames are further apart at the destination. If either or both frames are lost or not delivered due to, for example, FCS violation, then the value ΔT_{ij} is not computed and does not contribute to the evaluation of the Frame Delay Variation.

Figure 8 shows a depiction of the different times that are related to Frame Delay Variation performance.

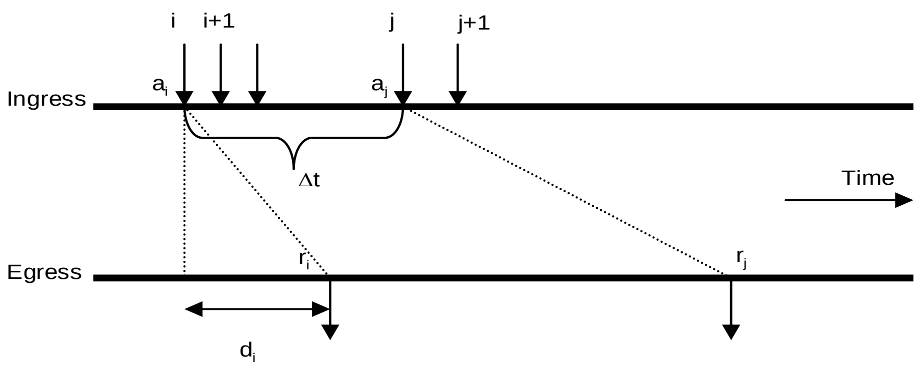

Figure 8: Frame Delay Variation Parameters

Let $S_{\Delta T} = \{\Delta T_{ij} : \forall i, j \text{ such that } a_j - a_i = \Delta t, a_i \in T, and \ a_j \in T\}$ be the set of all delay variations for all eligible pairs of Service Frames. Define $\tilde{d}_{\Delta T}$ to be the P-percentile of the set such that

$$\tilde{d}_{\Delta T} = \min\{d : P \le \frac{100}{K} \sum I(d, \Delta T_{ij})$$

where;

$$I(d, \Delta T_{ij}\} = \begin{cases} 1 \text{ if } d > \Delta T_{ij} \\ 0 \text{ otherwise} \end{cases}$$, and the sum is carried out over all the values in the set $S_{\Delta T}$.

Frame Delay Variation performance depends on the choice of the value for Δt. Values for both Δt and T typically should be chosen to achieve a reasonable level of statistical accuracy.

For the SLS, the Frame Delay Variation entry **MUST** specify a set of parameters and an objective. The parameters of the Frame Delay Variations performance are given in Table 3.

Parameter	Description
T	The interval on time over which the population of interest of green frames arrives at the ingress UNI
P	Frame Delay Variation percentile
Δt	the maximum separation between frame pairs for which Frame Delay Variation is computed
$\hat{d}$	Frame Delay Variation performance Objective

Table 3: Frame Delay Variation Parameters

Given **T**, P, Δt, and $\hat{d}$, the Frame Delay Variation performance **SHALL** be defined as met over the time interval T if and only if $\tilde{d}_{\Delta T} \le \hat{d}$.

7.3 Frame Loss Ratio

The Frame Loss Ratio for a particular Class of Service instance on a Point-to-Point EVC for a time interval **T SHALL** be defined as the ratio, expressed as percentage, of the number of lost Green Service Frames during the time interval **T** to the number of ingress Green Service Frames that should be successfully delivered during the time interval **T**. L2CP frames that are peered at the ingress UNI are not counted as lost frames. The Unicast, Broadcast, Multicast, and Layer 2 Control Protocol Service Frame Delivery service attributes define which Service Frames should be successfully delivered. Refer to [1] for the definitions of these service attributes.

Frame Loss Ratio can be expressed mathematically as follows. Let I_T be the number of Green ingress Service Frames during the interval T that should be delivered. Let E_T be the number of green egress Service Frames that are delivered during the interval T. Then

$$FLR_T = \left(\frac{I_T - E_T}{I_T} \right) \times 100\% .$$

For the SLS, the Frame Loss Ratio entry **MUST** specify a set of parameters and an objective. The parameters of the Frame Delay Variations performance are given in Table 4.

Parameter	Description
T	The interval on time over which the population of interest of green frames arrives at the ingress UNI
L	Frame Loss performance objective

Table 4: Frame Loss Ratio Performance Parameters

Given **T**, the Frame Loss Ratio performance **SHALL** be defined as met over the time interval **T** if and only if FLR_T is less or equal to L.

8. References

[1] **MEF 1**, *MEF Service Model, Phase I*, Metro Ethernet Forum, November 2003.

[2] **Bradner, S.**, *Key words for use in RFCs to Indicate Requirement Levels*, Internet Engineering Task Force, RFC 2119, March 1997

[3] **C. Demichelis and P. Chimento,** *IP Packet Delay Variation Metric for IP Performance Metric (IPPM)*, Internet Engineering Task Force (IETF), RFC 3393, November 2002.