

MPLS  
Internet-Draft  
Intended status: Standards Track  
Expires: April 10, 2013

C. Villamizar, Ed.  
Outer Cape Cod Network  
Consulting  
October 7, 2012

**Multipath Extensions for MPLS Traffic Engineering**  
**draft-villamizar-mpls-tp-multipath-te-extn-02**

Abstract

Extensions to OSPF-TE, ISIS-TE, and RSVP-TE are defined in support of carrying LSP with strict packet ordering requirements over multipath and carrying LSP with strict packet ordering requirements within LSP without violating requirements to maintain packet ordering. LSP with strict packet ordering requirements include MPLS-TP LSP.

OSPF-TE and ISIS-TE extensions defined here indicate node and link capability regarding support for ordered aggregates of traffic, multipath traffic distribution, and abilities to support multipath load distribution differently per LSP.

RSVP-TE extensions either identifies an LSP as requiring strict packet order, or identifies an LSP as carrying one or more LSP that requires strict packet order further down in the label stack, or identifies an LSP as having no restrictions on packet ordering except the restriction to avoid reordering microflows. In addition an extension indicates whether the first nibble of payload will reliably indicate whether payload is IPv4, IPv6, or other type of payload, most notably pseudowire using a pseudowire control word.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 10, 2013.

## Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.



## Table of Contents

<a href="#">1.</a>	<a href="#">Introduction . . . . .</a>	<a href="#">4</a>
<a href="#">1.1.</a>	<a href="#">Architecture Summary . . . . .</a>	<a href="#">4</a>
<a href="#">1.2.</a>	<a href="#">Requirements Language . . . . .</a>	<a href="#">5</a>
<a href="#">1.3.</a>	<a href="#">Definitions . . . . .</a>	<a href="#">5</a>
<a href="#">2.</a>	<a href="#">Protocol Extensions . . . . .</a>	<a href="#">6</a>
<a href="#">2.1.</a>	<a href="#">Multipath Node Capability sub-TLV . . . . .</a>	<a href="#">6</a>
<a href="#">2.2.</a>	<a href="#">Multipath Link Capability TLV . . . . .</a>	<a href="#">9</a>
<a href="#">2.3.</a>	<a href="#">LSP Multipath Attributes TLV . . . . .</a>	<a href="#">9</a>
<a href="#">3.</a>	<a href="#">Protocol Mechanisms . . . . .</a>	<a href="#">12</a>
<a href="#">3.1.</a>	<a href="#">OSPF-TE and ISIS-TE Advertisement . . . . .</a>	<a href="#">12</a>
<a href="#">3.1.1.</a>	<a href="#">Node Capability Advertisement . . . . .</a>	<a href="#">12</a>
<a href="#">3.1.2.</a>	<a href="#">Link Capability Advertisement . . . . .</a>	<a href="#">12</a>
<a href="#">3.1.3.</a>	<a href="#">Setting Max Depth and IP Depth . . . . .</a>	<a href="#">12</a>
<a href="#">3.1.4.</a>	<a href="#">Advertising Multipath as Link Bundling . . . . .</a>	<a href="#">13</a>
<a href="#">3.1.5.</a>	<a href="#">Hierarchical LSP Link Advertisement . . . . .</a>	<a href="#">13</a>
<a href="#">3.1.6.</a>	<a href="#">Advertisement of Legacy Multipath . . . . .</a>	<a href="#">14</a>
<a href="#">3.2.</a>	<a href="#">RSVP-TE LSP Attributes . . . . .</a>	<a href="#">15</a>
<a href="#">3.2.1.</a>	<a href="#">LSP Contained Ordered Aggregates Flags . . . . .</a>	<a href="#">15</a>
<a href="#">3.2.2.</a>	<a href="#">LSP Attributes for Ordered Aggregates . . . . .</a>	<a href="#">17</a>
<a href="#">3.2.3.</a>	<a href="#">Attributes for LSP without Packet Ordering . . . . .</a>	<a href="#">17</a>
<a href="#">3.3.</a>	<a href="#">Path Computation Constraints . . . . .</a>	<a href="#">20</a>
<a href="#">3.3.1.</a>	<a href="#">Link Multipath Capabilities and Path Computation . . . . .</a>	<a href="#">20</a>
<a href="#">3.3.1.1.</a>	<a href="#">Path Computation with Ordering Constraints . . . . .</a>	<a href="#">20</a>
<a href="#">3.3.1.2.</a>	<a href="#">Path Computation with No Ordering Constraint . . . . .</a>	<a href="#">21</a>
<a href="#">3.3.1.3.</a>	<a href="#">Path Computation for MPLS containing MPLS-TP . . . . .</a>	<a href="#">21</a>
<a href="#">3.3.2.</a>	<a href="#">Link IP Capabilities and Path Computation . . . . .</a>	<a href="#">21</a>
<a href="#">3.3.2.1.</a>	<a href="#">LSP without Packet Ordering Requirements . . . . .</a>	<a href="#">22</a>
<a href="#">3.3.2.2.</a>	<a href="#">LSP with Ordering Requirements . . . . .</a>	<a href="#">22</a>
<a href="#">3.3.3.</a>	<a href="#">Link Depth Limitations and Path Computation . . . . .</a>	<a href="#">23</a>
<a href="#">4.</a>	<a href="#">Backwards Compatibility . . . . .</a>	<a href="#">24</a>
<a href="#">4.1.</a>	<a href="#">Legacy Multipath Behavior . . . . .</a>	<a href="#">24</a>
<a href="#">4.2.</a>	<a href="#">Networks without Multipath Extensions . . . . .</a>	<a href="#">24</a>
<a href="#">4.2.1.</a>	<a href="#">Networks with MP Capability on all Multipath . . . . .</a>	<a href="#">24</a>
<a href="#">4.2.2.</a>	<a href="#">Networks with OA Capability on all Multipath . . . . .</a>	<a href="#">26</a>
<a href="#">4.2.3.</a>	<a href="#">Legacy Networks with Mixed MP and OA Links . . . . .</a>	<a href="#">26</a>
<a href="#">4.3.</a>	<a href="#">Transition to Multipath Extension Support . . . . .</a>	<a href="#">27</a>
<a href="#">4.3.1.</a>	<a href="#">Simple Transitions . . . . .</a>	<a href="#">27</a>
<a href="#">4.3.2.</a>	<a href="#">More Challenging Transitions . . . . .</a>	<a href="#">27</a>
<a href="#">5.</a>	<a href="#">IANA Considerations . . . . .</a>	<a href="#">28</a>
<a href="#">6.</a>	<a href="#">Security Considerations . . . . .</a>	<a href="#">28</a>
<a href="#">7.</a>	<a href="#">References . . . . .</a>	<a href="#">28</a>
<a href="#">7.1.</a>	<a href="#">Normative References . . . . .</a>	<a href="#">28</a>
<a href="#">7.2.</a>	<a href="#">Informative References . . . . .</a>	<a href="#">29</a>
	<a href="#">Author's Address . . . . .</a>	<a href="#">30</a>



## **1. Introduction**

Today the requirement to handle large aggregations of traffic, can be handled by a number of techniques which we will collectively call multipath. Multipath is very similar to composite link as defined in [\[ITU-T.G.800\]](#), except multipath specifically excludes inverse multiplexing. Some types of LSP, including but potentially not limited to MPLS-TP LSP, require strict packet ordering.

A means to support a MPLS-TP client layer over a MPLS server layer using MPLS Entropy Label is defined in [\[I-D.villamizar-mpls-tp-multipath\]](#). It is noted in [\[I-D.villamizar-mpls-tp-multipath\]](#) that absent some protocol extensions, some limitations must be accepted.

This document defines protocol extensions which better supports using MPLS with multipath as a server layer for MPLS-TP, or to carry MPLS-TP directly over a network which makes use of multipath. Extensions are also applicable to MPLS when used in the presense of very large microflows or very large LSP which cannot be load split as a result of using MPLS Entropy Label [\[I-D.ietf-mpls-entropy-label\]](#).

### **1.1. Architecture Summary**

Advertisements in a link state routing protocol, such as OSPF or ISIS, support a topology map known as a link state database (LSDB). When traffic engineering information is included in the LSDB the topology map is known as a TE-LSDB or traffic engineering database (TED).

A common MPLS LSP path computation is known as a constrained shortest path first computation (CSPF) (see [\[RFC3945\]](#)). Other algorithms may be used for path computation. Constraint-based routing was first introduced in [\[RFC2702\]](#).

OSPF-TE or ISIS-TE extensions are defined in [Section 2.1](#) and [Section 2.2](#). OSPF-TE or ISIS-TE advertisements serve to populate the TE-LSDB and provide the basis for constraint-based routing path computation. [Section 3.1](#) describes the use of OSPF-TE or ISIS-TE multipath extensions in routing advertisements.

RSVP-TE extensions are defined in [Section 2.3](#). [Section 3.2](#) describes the use of RSVP-TE extensions in setting up LSP including signaling constraints on LSP which contain other LSP which specify RSVP-TE extensions.

[Section 3.3](#) describes the constraints on LSP path computation imposed by the advertised ordered aggregate and multipath capabilities of



links. [Section 3.3.2](#) describes the constraints on LSP path computation imposed by link advertisements regarding use of IP headers in multipath traffic distribution. [Section 3.3.3](#) describes the impact of label stack depth limitations.

## **[1.2.](#) Requirements Language**

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](#) [[RFC2119](#)].

## **[1.3.](#) Definitions**

Please refer to [[I-D.villamizar-mpls-tp-multipath](#)].

### **Ordered Aggregate (OA)**

An ordered aggregate (OA) requires that packets be delivered in the order in which they were received. Please refer to [[RFC3260](#)].

### **Microflow**

A microflow is a single instance of an application-to- application flow. Please refer to [[RFC2475](#)]. Reordering packets within a microflow can cause service disruption. Please refer to [[RFC2991](#)].

### **Multipath Traffic Distribution**

Multipath traffic distribution refers to the mechanism which distributes traffic among a set of component links or component lower layer paths which together comprise a multipath. No assumptions are made about the algorithms used in multipath traffic distribution. This document only discusses constraints of the type of information which can be used as the basis for multipath traffic distribution in specific circumstances.

The phrase "strict packet ordering requirements" refers to the requirement to deliver all packet in the order that they were received. The absence of strict packet ordering requirements does not imply total absence of packet ordering requirements. The requirement to avoid reordering traffic within any given microflow, as described in [[RFC2991](#)] applies to all traffic aggregates including all MPLS LSP.

The abbreviations ELI and EL are Entropy Label Indicator and Entropy Label, as defined in [[I-D.ietf-mpls-entropy-label](#)].



## 2. Protocol Extensions

This section defined protocol extensions to OSPF-TE, ISIS-TE, and RSVP-TE. Use of these extensions is described in [Section 3](#) and [Section 4](#).

Two capability sub-TLV are added to two TLV that are used in both OSPF-TE and ISIS-TE. The Multipath Node Capability sub-TLV is added to the Node Attribute TLV ( see [Section 2.1](#)). The Multipath Link Capability TLV is added to the Interface\_ID (see [Section 2.2](#)).

One TLV is added to the LSP\_ATTRIBUTES object defined in [[RFC5420](#)].

### 2.1. Multipath Node Capability sub-TLV

The Node Attribute TLV is defined in [[RFC5786](#)]. A new sub-TLV, the Multipath Node Capability sub-TLV, is defined for inclusion in the Node Attribute TLV.

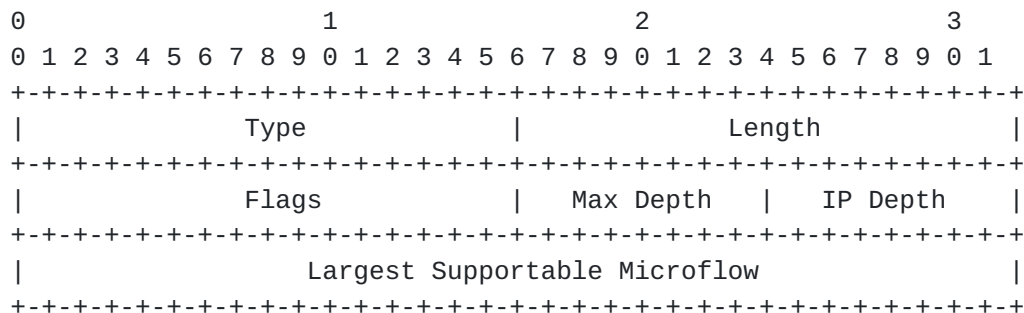


Figure 1: Multipath Capability Sub-TLV

The fields in the Multipath Capability sub-TLV are defined as follows.

#### Type

The Type field is assigned a value of IANA-TBD-1. The Type field is a two octet value.

#### Length

The Length field indicates the length of the sub-TLV in octets, excluding the Type and Length fields. The Length field is a two octet value.

#### Flags

The Flags field is a two octet (16 bit) value. The following single bit fields are assigned within this value, starting at the most significant bit, which is the bit transmitted first.



**0x8000 Ordered Aggregate Enabled**

Setting the Ordered Aggregate Enabled bit indicates that an LSP can be carried as an Ordered Aggregate Enabled on one or more links.

**0x4000 Multipath Enabled**

Setting the Multipath Enabled bit indicates that an LSP can be spread across component links at one or more multipath links.

**0x2000 IPv4 Enabled Multipath**

Setting the IPv4 Enabled Multipath bit indicates that the IPv4 header information can be used in multipath load balance. The Multipath Enabled bit must be set if the IPv4 Enabled Multipath bit is set.

**0x1000 IPv6 Enabled Multipath**

Setting the IP bit indicates that the IPv6 header information can be used in multipath load balance. The Multipath Enabled bit must be set if the IPv6 Enabled Multipath bit is set.

**0x0800 UDP/IPv4 Multipath**

Setting the UDP/IPv4 Multipath bit indicates that the UDP port numbers carried in UDP over IPv4 can be used in multipath load balance. The IPv4 Enabled Multipath bit must be set if UDP/IPv4 Multipath is set. If the IPv4 Enabled Multipath bit is set and the UDP/IPv4 Multipath bit is clear, then only source and destination IP addresses are used.

**0x0400 UDP/IPv6 Multipath**

Setting the UDP/IPv6 Multipath bit indicates that the UDP port numbers carried in UDP over IPv6 can be used in multipath load balance. The IPv6 Enabled Multipath bit must be set if UDP/IPv6 Multipath is set. The IPv6 Enabled Multipath bit must be set if UDP/IPv6 Multipath is set. If the IPv6 Enabled Multipath bit is set and the UDP/IPv6 Multipath bit is clear, then only source and destination IP addresses are used.

**0x0200 TCP/IPv4 Multipath**

Setting the TCP/IPv4 Multipath bit indicates that the TCP port numbers carried in TCP over IPv4 can be used in multipath load balance. The IPv4 Enabled Multipath bit must be set if TCP/IPv4 Multipath is set. If the IPv4 Enabled Multipath bit is set and the TCP/IPv4 Multipath bit is clear, then only source and destination IP addresses are used.



**0x0100 TCP/IPv6 Multipath**

Setting the TCP/IPv6 Multipath bit indicates that the TCP port numbers carried in TCP over IPv6 can be used in multipath load balance. The IPv6 Enabled Multipath bit must be set if TCP/IPv6 Multipath is set. The IPv6 Enabled Multipath bit must be set if TCP/IPv6 Multipath is set. If the IPv6 Enabled Multipath bit is set and the TCP/IPv6 Multipath bit is clear, then only source and destination IP addresses are used.

**0x0080 Default to Multipath**

Setting the Default to Multipath bit indicates that for an LSP which does not signal a desired behavior the traffic for that LSP will be spread across component links at one or more multipath links. If the Default to Multipath bit is not set, then an LSP which does not signal otherwise will be treated as an ordered aggregate.

**0x0040 Default to IP/MPLS Multipath**

Setting the Default to IP/MPLS Multipath indicates that for an LSP which does not signal a desired behavior, the IP header information will be used in the multipath load distribution. If the Default to IP/MPLS Multipath is clear it indicates that the the IP header information will not be used by default.

**0x0020 Entropy Label Multipath**

Setting the Entropy Label Multipath bit indicates that when multipath is enabled for a given LSP, if an Entropy Label Indicator (ELI) is found in the label stack, information below the Entropy Label (EL) will not be used in multipath load distribution.

**0x0010 IP Optional Multipath**

Setting the IP Optional Multipath bit indicates that when multipath is enabled for a given LSP, whether the IP header information is used in the multipath load distribution can be set on a per LSP basis.

The remaining bits in the Flags field are reserved.

**Max Depth**

The Max Depth field is a one octet field indicating the maximum label stack depth beyond which the multipath load distribution cannot make use of further label stack entries.

**IP Depth**

The IP Depth field is a one octet field indicating the maximum label stack depth beyond which the multipath load distribution cannot make use of IP information.



#### Largest Supportable Microflow

The Largest Supportable Microflow field is a IEEE 32 bit floating point value expressing in bytes/second. Any microflow which exceeds this capacity may experience either packet loss, or out-of-order delivery, or both.

The reserved bits in the Flags field MUST be set to zero and MUST be ignored unless implementing an extension which redefines one or more of the reserved bits. Any further extension which redefines one or more reserved Flags bit should maintain backwards compatibility with prior implementations.

### 2.2. Multipath Link Capability TLV

The Interface\_ID is defined in [RFC3471]. The Interface\_ID is updated in [RFC4201] to support a form of multipath known as Link Bundling.

A new TLV, the Multipath Link Capability TLV, is defined here. The Multipath Link Capability TLV is optionally included in the Interface\_ID. The format of the Multipath Link Capability TLV is identical to the Multipath Node Capability sub-TLV defined in [Section 2.1](#), with one exception. In the Multipath Link Capability TLV the Type field value is IANA-TBD-2.

If a Multipath Link Capability TLV is advertised for any link, then a Multipath Node Capability sub-TLV MUST be advertised for the node.

### 2.3. LSP Multipath Attributes TLV

The LSP\_ATTRIBUTES object is defined in [RFC5420]. A new LSP Multipath Attributes TLV is defined for the LSP\_ATTRIBUTES object. The TLV Type is IANA\_TBD\_3. The format is described below.

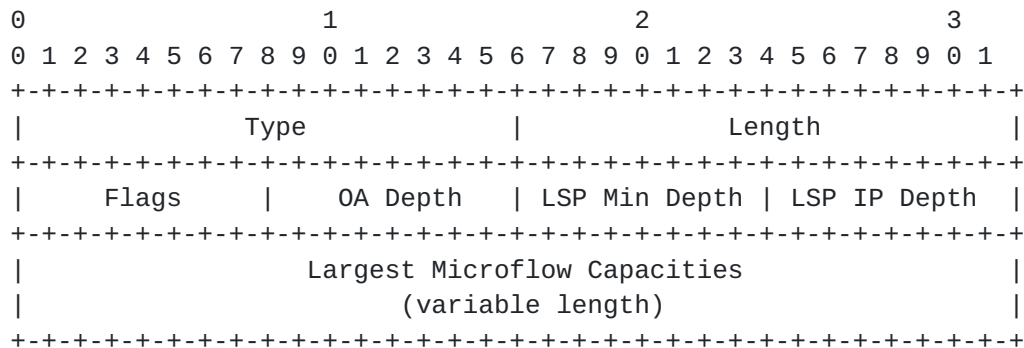


Figure 2: LSP Multipath Attributes TLV

The fields in the LSP Multipath Attributes TLV are defined as



follows.

#### Type

The Type field is assigned a value of IANA-TBD-3. The Type field is a two octet value.

#### Length

The Length field indicates the length of the sub-TLV in octets, excluding the Type and Length fields. The Length field is a two octet value.

#### Flags

The Flags field is a one octet (8 bit) value. The following single bit fields are assigned within this value, starting at the most significant bit, which is the bit transmitted first.

##### 0x80 IP Multipath Allowed

Setting the IP Multipath Allowed bit indicates that it is safe to enable the use of a potential IP payload in the multipath traffic distribution.

##### 0x40 May Contain IPv4

Setting the May Contain IPv4 bit indicates that IPv4 traffic may be contained within this LSP.

##### 0x20 May Contain IPv6

Setting the May Contain IPv6 bit indicates that IPv6 traffic may be contained within this LSP.

##### 0x02 Entropy Label Required

Setting the Entropy Label Used bit indicates that midpoint LSR MUST support ELI and EL in order to not violate packet ordering constraints of the LSP or of contained LSP.

##### 0x01 Entropy Label Used

Setting the Entropy Label Used bit indicates that an ELI and EL is present in some or all label stack entries within this LSP.

The remaining bits in the Flags field are reserved.

#### OA Depth

The OA Depth field is set as follows

- 0 An OA Depth value of zero indicates that no ordered aggregates are carried within the LSP, except those which are protected from out of order delivery using Entropy Label.



- 1 An OA Depth value of one indicates that the LSP is an ordered aggregate of traffic (the LSP requires strict ordering of packets) and has protected packet ordering using ELI and EL.
- >1 An OA Depth value greater than one indicates that the LSP does not have strict packet ordering requirements but contains ordered aggregates at the label stack depth indicated or deeper and that the ordered aggregates are not protected using ELI and EL.

#### LSP Min Depth

The LSP Min Depth field indicates a minimal acceptable number of label used in multipath traffic distribution for the stated Largest Microflow Capacities field to be valid. If the LSP Min Depth field is set to zero this value is unknown. See [Section 3.3.3](#).

#### LSP IP Depth

The LSP IP Depth field indicates a minimal label stack depth where using an IP header is necessary for the stated Largest Microflow Capacities field to be valid. If the LSP IP Depth field is set to zero this value is unknown. See [Section 3.3.3](#).

#### Largest Microflow Capacities

The Largest Microflow Capacities field contains zero, one, two, or three IEEE 32 bit floating point values. Each value is a capacity expressed in bytes per second.

#### Largest LSE Microflow

The first value, the Largest LSE Microflow, is the capacity of the largest microflow if only the label stack entries are used in multipath traffic distribution. If a Largest LSE Microflow is not included, the LSP bandwidth request MUST be used.

#### Largest IP Microflow

The second value, the Largest IP Microflow, if present, is the capacity of the largest microflow if the label stack entries and any potential IP source and destination address are used in multipath traffic distribution. If the Largest IP Microflow is not included, the value of the Largest LSE Microflow MUST be used.

#### Largest L4 Microflow

The third, the Largest L4 Microflow, if present, is the capacity of the largest microflow if the label stack entries and any potential IP addresses and TCP or UDP port numbers are used in multipath traffic distribution. If a Largest L4 Microflow is not included, the value of the Largest IP



Microflow MUST be used.

### **3. Protocol Mechanisms**

#### **3.1. OSPF-TE and ISIS-TE Advertisement**

Every compliant node MUST advertise exactly one Multipath Node Capability sub-TLV and MAY advertise zero or more Multipath Link Capability sub-TLV as needed.

Procedures for accommodating legacy forwarding capabilities and non-compliant nodes are discussed in [Section 4](#).

##### **3.1.1. Node Capability Advertisement**

Every LSR which is adjacent to one or more multipath link MUST advertise a Multipath Node Capability sub-TLV (see [Section 2.1](#)). The capabilities advertised for the node SHOULD reflect the capabilities of the majority of multipath links adjacent to the node.

Every LSR which is not adjacent to any multipath links MUST advertise a Multipath Node Capability sub-TLV with both the Ordered Aggregate Enabled bit in Flags set and all other Flags bits clear. Both Max Depth and IP Depth can be set to zero. This advertisement identifies the LSR as one which is conformant but has no multipath links, allowing it to be distinguished from a non-conformant LSR with LAG or other multipath which may have to be avoided when determining a path for some LSP.

##### **3.1.2. Link Capability Advertisement**

For all of the links whose capability does not exactly match the Multipath Node Capability sub-TLV advertised by that same LSR, the LSR MUST advertise a Multipath Link Capability sub-TLV (see [Section 2.2](#)).

For all of the links whose capability does exactly match the Multipath Node Capability sub-TLV advertised by that same LSR, the LSR SHOULD NOT advertise a Multipath Link Capability sub-TLV (see [Section 2.2](#)). In this case the Multipath Link Capability TLV is redundant, but harmless.

##### **3.1.3. Setting Max Depth and IP Depth**

The Max Depth and IP Depth field are intended to capture architectural limits. Most forwarding hardware will only use a limited number of label entries in the multipath traffic



distribution. This limit is reflected in the Max Depth field. Most forwarding hardware will limit the number of label entries that it will look past before looking for an IP header to be used in the multipath traffic distribution. This limit is reflected in the IP Depth field.

#### **3.1.4. Advertising Multipath as Link Bundling**

All multipath links and FA for PSC LSP which traverse multipath links MAY be advertised as Link Bundles as defined in [RFC4201]. The settings of the Ordered Aggregate Enabled and Multipath Enabled fields indicate key capabilities of the multipath. Advertising the multipath as a link bundle can provide additional information, such as the capability to place LSP on individual components.

If the Multipath Enabled bit is set in the Multipath Link Capability TLV Flags, then the Maximum LSP Bandwidth in the Interface Identification TLV can be set in one of two ways.

1. If the desired behavior for legacy LSR acting as ingress is to limit LSP to the capacity of a single component link, then Maximum LSP Bandwidth SHOULD be set as specified in [RFC4201].
2. If the desired behavior for legacy LSR acting as ingress is to allow LSP to exceed the capacity of a single component link, then Maximum LSP Bandwidth MAY be set to a higher value, up to the value of Maximum Reservable Bandwidth. This would normally be done if the legacy LSR were known to be carrying traffic which is easily load split, such as typical Internet traffic.

LSR acting as ingress SHOULD ignore the Maximum LSP Bandwidth and MAY set up LSP with capacity up to the Maximum Reservable Bandwidth as long as microflows within the LSP will not exceed the Largest Supportable Microflow capacity.

#### **3.1.5. Hierarchical LSP Link Advertisement**

A PSC LSP, as defined in [RFC4206] and updated in [RFC6107], may carry other LSP. When signaling a PSC LSP expected characteristics of the contained traffic must be estimated. The FA advertised for the PSC LSP MUST reflect the broadest set of requirements the PSC LSP can carry. If the setup of an additional reservation would exceed current capacity, a PSC LSP may be resigaled using make-before-break semantics ([RFC3209]).

For example, if it is expected that a PSC LSP will carry MPLS-TP LSP or other LSP with strict packet reordering requirements at some label depth, the minimum label stack depth at which an LSP with strict



packet reordering requirements can be carried must be signaled in the OA Depth field of the LSP Multipath Attributes TLV (see [Section 2.3](#)).

When the Forwarding Adjacency (FA) is advertised, the advertised Max Depth and IP Depth MUST be one less than the minimum of the Max Depth and IP Depth of any link that the PSC LSP traverses. The Max Depth and IP Depth are considered independently of each other. The reduction by one takes into account the PSC label. The Flags advertised for the FA MUST reflect the capabilities of the links along the path.

#### **3.1.6. Advertisement of Legacy Multipath**

An Ethernet LAG with no support for Entropy Label MUST have the Ordered Aggregate Enabled bit cleared and the Multipath Enabled bit set in the Multipath Link Capability TLV Flags. This indicates that a MPLS-TP compliant server layer cannot be supported and that LSP larger than the component links (LAG members) can be supported.

A Link Bundle that does not support the all-ones component defined in [\[RFC4201\]](#) MUST have the Ordered Aggregate Enabled bit set and the Multipath Enabled bit cleared in the Multipath Link Capability TLV Flags. This indicates that a MPLS-TP compliant server layer can be supported and that LSP larger than the component links cannot be supported.

A link bundle that can support either the pinning of a LSP to a single component link or the spreading of traffic across multiple component links MUST have the Ordered Aggregate Enabled bit set and the Multipath Enabled bit set in the Multipath Link Capability TLV Flags. This indicates that a MPLS-TP compliant server layer can be supported and that LSP larger than the component links can also be supported.

Any form of multipath that supports Entropy Label MUST have the Ordered Aggregate Enabled bit set and the Multipath Enabled bit set and the Entropy Label Multipath bit set in the Multipath Link Capability TLV Flags. Any form of multipath that does not support Entropy Label MUST have the Entropy Label Multipath bit cleared in the Multipath Link Capability TLV Flags.

The remaining bits in the Multipath Link Capability TLV Flags MUST be set according to the capability and configuration of the multipath or LSP.



### **3.2. RSVP-TE LSP Attributes**

All LSR SHOULD advertise a LSP Multipath Attributes TLV with the RSVP-TE signaling for each LSP for which it is serving as ingress. If any legacy LSR remain in the network, it is easier to assign an acceptable default treatment for LSP signaled by those legacy LSR if the conforming LSR always send a LSP Multipath Attributes TLV.

There are two general cases, an LSP requires strict ordering of packets, or it doesn't. In the latter case the LSP may contain other LSP that require strict ordering and those must be protected from reordering using an Entropy Label as described in [[I-D.villamizar-mpls-tp-multipath](#)]. These two cases are briefly described below.

#### **Ordered Aggregates**

LSP with strict packet order requirements MUST set the OA Depth field to one to indicate that the LSP MUST be treated as ordered aggregate. See [Section 3.2.2](#).

#### **LSP without Packet Ordering**

LSP which do not have strict packet order requirements MUST only carry LSP whose requirements are reflected in the containing LSP Multipath Attributes TLV. See [Section 3.2.3](#).

If an attempt is made to signal a contained LSP whose Ordered Aggregate Attributes TLV and LSP Multipath Attributes TLV cannot be supported by the containing (PSC) LSP, one of the two following actions MUST be taken.

1. The containing (PSC) LSP MAY be resigaled with appropriate TLVs to carry the new contained LSP using make-before-break semantics, then continue to forward the contained LSP PATH message if the containing (PSC) LSP is successfully updated.
2. The LSR MAY reject the contained LSP signaling by sending a PathErr message.

#### **3.2.1. LSP Contained Ordered Aggregates Flags**

The Flags field in the LSP Multipath Attributes TLV MUST be set as follows.

1. If the LSP may directly contain IPv4 traffic, then the May Contain IPv4 bit in the Flags field MUST be set.
2. If the LSP may directly contain IPv6 traffic, then the May Contain IPv6 bit in the Flags field MUST be set.



3. If the LSP contains an LSP which has the May Contain IPv4 bit in the Flags field then the May Contain IPv4 bit in the Flags field MUST be set in the containing LSP.
4. If the LSP contains an LSP which has the May Contain IPv6 bit in the Flags field then the May Contain IPv6 bit in the Flags field MUST be set in the containing LSP.
5. If the LSP may contain pseudowires that do not use a pseudowire control word [[RFC4385](#)], and may contain IPv4 or IPv6 traffic, then the IP Multipath Allowed bit in the Flags field MUST be cleared.
6. If the LSP is known to contain no pseudowires that do not use a pseudowire control word, then the IP Multipath Allowed bit in the Flags field SHOULD be set unless disallowed due to a contained LSP.
7. If an Entropy Label is added below the LSP label, then the Entropy Label Used bit MUST be set.
8. If the LSP contains any LSP with the IP Multipath Allowed bit in the Flags field clear, then the IP Multipath Allowed bit in the Flags field MUST be clear.

If the LSP does not contain other LSP, it may contain IP traffic and/or pseudowire that terminate on that LSR. If the LSP does not contain other LSP. The LER will know whether the LSP is used in an IP LER capacity. The LER will also know whether it terminates any pseudowire for a given LSP. The LER will also know if it is using Entropy Label for a given LSP and if it requires strict packet ordering for a given LSP. Therefore, when a LSP does not contain other LSP, then it is possible to accurately set the Flags field in the LSP Multipath Attributes TLV, as well the OA Depth, and LSP IP Depth fields.

If an LSP contains other LSP, and all of the contained include a LSP Multipath Attributes TLV, then it is still possible to accurately set the Flags field in the LSP Multipath Attributes TLV, as well the OA Depth, and LSP IP Depth fields. It is only when LSP contains other LSP that do not have a LSP Multipath Attributes TLV where default behavior has to be configured based on assumptions about LSP originated by the legacy LSR where there is a potential for those configured assumptions to be inaccurate.

See [Section 4](#) for guidelines for handling LSP which contain LSP that do not have a LSP Multipath Attributes TLV. The most conservative approach in this case is to clear the IP Multipath Allowed bit and



set the May Contain IPv4 bit and the May Contain IPv6 bit, however this may not always be necessary.

### **3.2.2. LSP Attributes for Ordered Aggregates**

An LSP with strict packet order requirements MUST always include a LSP Multipath Attributes TLV.

Most of the Flags in the LSP Multipath Attributes TLV can be set as described in [Section 3.2.1](#). There are two cases which affect the setting of the remaining Flags bits.

Entropy Label not used

If an Entropy Label is not used, then the Entropy Label Used bit, the Entropy Label Required bit, and the IP Multipath Allowed bit MUST be cleared.

Entropy Label is used If an Entropy Label is used, then the Entropy Label Used bit, and the Entropy Label Required bit, and the IP Multipath Allowed bit MUST be set.

The OA Depth field MUST be set to one. The Min Depth MUST be set to one. The LSP IP Depth SHOULD be set to zero. The Largest Microflow Capacities field SHOULD be empty. The entire LSP is one microflow. The Largest Microflow Capacities field MAY contain one entry if there is some reason to do so, such as an LSP which may peak at capacity higher than the bandwidth reserved for the LSP. The Largest Microflow Capacities for an LSP SHOULD be configurable independently of the LSP bandwidth reservation.

### **3.2.3. Attributes for LSP without Packet Ordering**

If an LSP does not have strict packet order constraints, then the LSR\_ATTRIBUTE object SHOULD always include a LSP Multipath Attributes TLV.

Most of the Flags in the LSP Multipath Attributes TLV can be set as described in [Section 3.2.1](#). There are two cases which affect the setting of the remaining Flags bits, the OA Depth field, the LSP Min Depth, and the LSP IP Depth field.

Entropy Label not used

- \* The OA Depth MUST be either set to zero or set to a configured value that is greater than one, or set based on the contained LSP.



- \* If the OA Depth is set to a configured value, then any setup attempt for a contained LSP with a depth greater than or equal to that value SHOULD be rejected and a PathErr message sent. Otherwise, if a setup attempt for a contained LSP with a depth greater than the current value included in the containing LSP OA Depth field, then the containing LSP MUST be rerouted with a OA Depth field value greater than any of the contained OA Depth field values.
- \* The Entropy Label Used bit MUST be set if any contained LSP has the Entropy Label Used bit set.
- \* The Entropy Label Required bit MUST be set if any contained LSP has the Entropy Label Required bit set.

#### Entropy Label is used

- \* The OA Depth MUST be set to zero.
- \* The Entropy Label Used bit MUST be set.
- \* The Entropy Label Required bit MUST be set if any contained LSP has the Entropy Label Required bit set.
- \* The Entropy Label Required bit MUST be set if any contained LSP has the OA Depth field set to a non-zero value.
- \* The Entropy Label Required bit SHOULD be clear if there are no contained LSP has the OA Depth field set to a non-zero value and no contained LSP with the Entropy Label Required bit set. In this case the Entropy Label Required bit MAY be set by configuration, generally in anticipation of needing it in the future to carry other LSP.
- \* LSP Min Depth field MUST be set to three if the Entropy Label Required bit is set.
- \* If the Entropy Label Required bit is not set, then the LSP Min Depth field and LSP IP Depth field SHOULD be set to three if there are no contained LSP. The LSP Min Depth field and LSP IP Depth MAY be configured to a minimum value greater than three, generally in anticipation of needing it in the future to carry other LSP.
- \* If the Entropy Label Required bit is not set, and there are contained LSP, then the LSP Min Depth field MUST be set to a value greater than three.



- \* If the Entropy Label Required bit is not set, the LSP Min Depth field MUST be set to a value that is at least the sum of three plus the LSP Min Depth field in any contained LSP.
- \* If the Entropy Label Required bit is not set, and either the May Contain IPv4 bit or the May Contain IPv6 bit is set, then the LSP IP Depth field MUST be set to a value of one or greater.
- \* If the Entropy Label Required bit is not set, and any contained LSP has the May Contain IPv4 bit or the May Contain IPv6 bit is set, then the LSP IP Depth field MUST be set to a value of two or greater.
- \* If the Entropy Label Required bit is not set, and any contained LSP has the LSP IP Depth field set to a value greater than one, then the LSP IP Depth field MUST be set to a value greater than the highest LSP IP Depth value set in any contained LSP.

The values of the LSP Min Depth field and the LSP IP Depth field MAY be constrained to upper limits by configuration. If an attempt to setup a contained LSP would result in exceeding one of these limits, then the LSR SHOULD reject the signaling attempt and send a PathErr message.

If Entropy Label is not used on the signaled LSP and there are no contained LSP, then the Largest LSE Microflow in the Largest Microflow Capacities field MUST be set to the requested bandwidth of the LSP. The optional Largest IP Microflow and Largest L4 Microflow SHOULD be included and MAY be set to configured minimum values.

If Entropy Label is not used on the signaled LSP an LSP that does not have strict packet order constraints contains other LSP, then the LSP Multipath Attributes TLV advertised by the set of contained LSP MUST be used to set the LSP Multipath Attributes TLV Largest Microflow Capacities values for LSP Multipath Attributes TLV. The value of Largest LSE Microflow, Largest IP Microflow, and Largest L4 Microflow in the LSP Multipath Attributes TLV of the containing LSP cannot be less than the maximum effective value of the same parameter for any contained LSP Multipath Attributes TLV.

If Entropy Label is used on the signaled LSP then the Largest LSE Microflow field is set according to the largest microflow that can result from computing the Entropy Label. This value is the greatest of a set of sources of entropy. A configured value MAY be used for IP, or it MAY be assumed that the largest interface carrying IP could carry a single microflow. For contained LSP which have the Entropy Label Used bit clear, the value in the Largest IP Microflow can be



used if the IP Multipath Allowed bit is set for that LSP and the LSR can make use of the IP information in the label stack. For contained LSP which have the Entropy Label Used bit set, the Largest LSE Microflow value already reflects any prior hashing of IP fields.

If the Entropy Label Required bit is set and the LSP being set up is using Entropy Label, then the Largest IP Microflow and Largest L4 Microflow SHOULD be omitted. All midpoint LSR SHOULD not look for entropy beyond the Entropy Label.

If the LSP being set up is not using Entropy Label, then the Largest IP Microflow and Largest L4 Microflow SHOULD be included unless the Entropy Label Used bit is set for every contained LSP. The Largest IP Microflow and Largest L4 Microflow SHOULD be omitted if hashing on the IP addresses or IP addresses and ports would yield no greater entropy than hashing on the label stack alone.

### **3.3. Path Computation Constraints**

The RSVP-TE extensions provides a set of requirements to be met by the links which the LSP is to traverse. This set of requirements also serves as the basis for path computation constraints and for admission control constraints.

#### **3.3.1. Link Multipath Capabilities and Path Computation**

Three cases are considered. An LSP may have strict ordering constraints. An MPLS-TP LSP is an example of an LSP with strict ordering constraints. This first type of LSP is covered in [Section 3.3.1.1](#). An LSP may have no ordering constraints at all other than the constraint that microflows cannot be reordered. This second case is covered in [Section 3.3.1.2](#). The remaining case is where an LSP has no ordering constraints but carries traffic for other LSP which do have ordering constraints. This third case is covered in [Section 3.3.1.3](#).

##### **3.3.1.1. Path Computation with Ordering Constraints**

For an MPLS-TP LSP or other LSP with a strict packet ordering constraint, any link or FA for which the Ordered Aggregate Enabled bit and Entropy Label Multipath are both clear MUST be excluded from the path computation. If the Default to Multipath bit is set on a link, then setting the OA Depth field to one will override that default.

Link with the Ordered Aggregate Enabled bit clear and the Entropy Label Multipath bit set MAY be included in the path computation if the LSR is capable of encapsulating an LSP with a strict packet



ordering constraint with a fixed Entropy Label. If the LSR is not capable of adding an ELI and EL, then these links MUST be excluded from the path computation.

#### **3.3.1.2. Path Computation with No Ordering Constraint**

For an MPLS LSP which has no constraint on packet ordering except that microflows must remain in order and does not contain other LSP with ordering constraints, any link for which the Multipath Enabled bit is set can be used. If a link is advertised as a Link Bundle and the Multipath Enabled bit is set for the link, the available bandwidth SHOULD be taken from the "Unreserved Bandwidth" rather than the "Maximum LSP Bandwidth" (see [[RFC4201](#)]).

For most LSP, the bandwidth requirement of the largest microflow is not known but an upper bound is known. For example if the LSP aggregates pseudowires or other LSP of no more than some maximum capacity or LSP which have signaled a microflow upper bound, then an upper bound on the largest microflow is known. If this upper bound exceeds the "Maximum LSP Bandwidth" of a given link, then that link MUST be excluded from the path computation.

#### **3.3.1.3. Path Computation for MPLS containing MPLS-TP**

To carry LSP which have strict packet ordering requirements and do not have the Entropy Label Used flag set as a client within a server LSP that do not have strict packet ordering requirements, Entropy Label must be added at the server layer LSP to traverse any link or FA that has the Multipath Enabled bit set. For these LSP links which have the Multipath Enabled bit set and the Entropy Label Multipath bit clear MUST be excluded from the path computation.

If the LSR is not capable of adding an Entropy Label, then to carry any client LSP with the Entropy Label Used clear and the OA Depth set to a non-zero value, the server LSP SHOULD exclude any link or FA that has the Multipath Enabled bit set. For these LSP, any link or FA that has the Multipath Enabled bit set and cannot carry a microflow as large as the entire LSP MUST be excluded from the path computation. These LSP MAY be signaled as having strict packet ordering requirements if they can be carried as a single microflow, but this practice is NOT RECOMMENDED.

#### **3.3.2. Link IP Capabilities and Path Computation**

An MPLS-TP LSP cannot be reordered. There may be other types of LSP with strict packet ordering requirements. If LSP with strict packet ordering requirements carry IP, using IP headers in the multipath load distribution would violate the packet ordering requirements.



Some LSP cannot be reordered but do not carry IP, and do not carry payloads which could be mistaken as IP. For example, any LSP carrying only pseudowire traffic, where all pseudowires are using a control word carries no payloads which could be mistaken as IP. These type of LSP can be carried within MPLS LSP that allow use of IP header information in multipath load distribution.

This section focuses on Cases in which links or FA must be excluded from path computation based on the settings of the IP related Flags bits in the Multipath Link Capability TLV.

#### **3.3.2.1. LSP without Packet Ordering Requirements**

Many LSP carry only IP or predominantly IP, use no hierarchy or have little diversity in the MPLS label stack, and carry far more traffic than can be carried over a single component link in a multipath. Many LSP due to their high capacity, must traverse only multipath which will use IP header information in the multipath traffic distribution.

For these LSP, links must be excluded from the path computation which do not have the IPv4 Enabled Multipath and IPv6 Enabled Multipath bit set (if carrying both IPv4 and IPv6) and do not have either the Default to IP/MPLS Multipath bit set or the IP Optional Multipath bit set.

Hierarchical PSC LSP which require the use IP header information in the multipath traffic distribution MUST NOT set the Ordered Aggregate Enabled bit, MUST set the Default to IP/MPLS Multipath bit, and MUST NOT set the IP Optional Multipath bit in the FA advertisement. The IP Optional Multipath bit MUST be clear because the LSP cannot change the behavior of midpoint LSR, except perhaps in the case of single hop LSP.

#### **3.3.2.2. LSP with Ordering Requirements**

The only time that links or FA with both the Ordered Aggregate Enabled bit and the Entropy Label Multipath bit clear can be used is a special case for MPLS-TP LSP that carry only IP traffic. This case does not apply if the MPLS\_TP LSP is carrying other LSP or if it is carrying pseudowires.

Where MPLS-TP LSP are carrying only IP, any link or FA with both the Ordered Aggregate Enabled bit and the Entropy Label Multipath bit clear for which the use of IP header information is not disabled or cannot be disabled on a per LSP basis, that link MUST be excluded from the path computation.



Where MPLS-TP LSP are carrying only IP, links MAY be included in the path computation have the IPv4 Enabled Multipath and IPv6 Enabled Multipath bits clear, or have the Default to IP/MPLS Multipath bit clear, or have the IP Optional Multipath bit set. Links with the IP Optional Multipath set, MUST disable use of IP in the load balance for LSP with the IP Multipath Allowed bit clear.

An MPLS-TP LSP are carrying only IP MUST have OA Depth set to one and LSP Min Depth set to one and the IP Multipath Allowed bit cleared. Call admission SHOULD NOT reject an LSP on the basis of OA Depth being set to one if use of IP headers is always disabled, or can be disabled for the specific LSP. If an MPLS-TP is set up this way end then does start to carry other LSP or carry pseudowires, then reordering within the MPLS-TP LSP will occur.

### **3.3.3. Link Depth Limitations and Path Computation**

For any LSP which does not have strict packet ordering constraints, LSP configuration SHOULD include the following parameters.

#### **LSP Min Depth**

a minimal acceptable number of label used in multipath traffic distribution,

#### **LSP IP Depth**

a minimal label stack depth where the IP header can be used in multipath traffic distribution

For example, if a PSC LSP will carry LSP which in turn carry very high capacity pseudowires using the pseudowire flow label (see [[RFC6391](#)]), the flow label is four labels deep. In this case, LSP Min Depth should be four or higher.

For example, if the same PSC LSP will carry LSP which carry IP traffic with no additional labels, then the IP header is two labels deep. In this case, LSP IP Depth should be two or higher.

For an LSP with non-zero LSP Min Depth, all links with Max Depth set to a value below LSP Min Depth MUST be excluded from the LSP Path Computation.

For an LSP with non-zero LSP IP Depth, all links with IP Depth set to a value below LSP IP Depth MUST be excluded from the LSP Path Computation.

If all LSP carried an accurate LSP Min Depth and LSP IP Depth then neither of these parameters would ever have to be configured. In a network with legacy LSR, it may be necessary to configure these



parameters for some LSP. A per-LSP minimum value of each parameter SHOULD be configurable.

#### **4. Backwards Compatibility**

Networks today use three forms of multipath.

1. IP ECMP, including IP ECMP at LER using more than one LSP.
2. Ethernet Link Aggregation [[IEEE-802.1AX](#)].
3. MPLS Link Bundling [[RFC4201](#)].

##### **4.1. Legacy Multipath Behavior**

IP ECMP and Ethernet Link Aggregation always distribute traffic over the entire multipath either using information in the MPLS label stack, or using information in a potential IP header, or using both types of information.

One of two behaviors is assumed for link bundles. Either the link bundles place each LSP in its entirety on a single link bundle component link for all LSP, or link bundles distribute traffic over the entire link bundle using the same techniques used for ECMP and Ethernet Link Aggregation. This second behavior is known as the "all ones" component link (see [[RFC4201](#)]).

##### **4.2. Networks without Multipath Extensions**

Networks exist that are comprised entirely of LSR which do not support these multipath extensions. In these networks there is no way of telling how multipath links will behave. Since an Ethernet Link Aggregation Group (LAG) is advertised as an ordinary link, there is no way to tell that it is a LAG and not an ordinary link.

###### **4.2.1. Networks with MP Capability on all Multipath**

Most large core network today rely heavily on the use of multipath. Ethernet Link Aggregation is heavily used and LSR are configured to use the "all ones" component link for all LSP. The "all ones" component link is the default for many Link Bundling implementations used in core networks.

This is equivalent to the following setting in the Multipath Node Capabilities sub-TLV or Multipath Link Capabilities sub-TLV.



1. Clear the Ordered Aggregate Enabled bit and the IP Optional Multipath bit.
2. Set the Multipath Enabled bit, set the Default to Multipath bit, and clear the Entropy Label Multipath bit.
3. If the label stack is used in the multipath traffic distribution, set Max Depth to the number of label stack entries supported, otherwise set it to zero.
4. Since Entropy Label support is not yet widespread, most LSR would behave as if Entropy Label Multipath were clear.
5. If an IP packet under the label stack can be used in the multipath traffic distribution (very common, almost universal in core LSR), set IPv4 Enabled Multipath, set IPv6 Enabled Multipath, set Default to IP/MPLS Multipath, and set IP Depth to the maximum number of label stack entries which can be skipped over before finding the IP stack. Otherwise clear IPv4 Enabled Multipath, clear IPv6 Enabled Multipath and clear Default to IP/MPLS Multipath.
6. On core networks where UDP and TCP ports are rarely used in multipath, clear all UDP and TCP related bits. On networks where multipath is configured to use TCP and UDP port numbers, these bits would be set.

These networks can support very large LSP but cannot support LSP which require strict packet ordering with other labels below such an LSP, such as pseudowire labels. They may also misroute OAM packet which use GAL (see [[RFC5586](#)]) if they use the GAL label in determining the load balance. Generally the Link Bundle advertisements indicate a "Maximum LSP Bandwidth" that is equal to the "Unreserved Bandwidth" if Link Bundling is used with the all-ones component link.

Good or bad, if the behavior is consistent, defaults can be configured in other LSR, such that an accurate guess can be made when no Multipath Link Capability TLV is available for a given link.

For example, in many networks, any link of 10 Gb/s or less can be assumed to be a plain link (no multipath) while any link with a capacity greater than 10 Gb/s can be assumed to be a multipath. These assumptions would hold if no 40 Gb/s or 100 Gb/s links are used.



#### **4.2.2. Netowrks with OA Capability on all Multipath**

Some networks, particularly edge networks which tend to be lower capacity, do not use Link Aggregation, and if they use Link Bundling at all, configure each LSR to place each LSP in its entirety on a single link bundle component link for all LSP. Some edge equipment only supports this link bundle behavior.

This is equivalent to the following setting in the Multipath Node Capabilities sub-TLV or Multipath Link Capabilities sub-TLV.

Set the Ordered Aggregate Enabled bit,

Clear the Multipath Enabled bit.

All remaining bits in the Flags field should be clear.

The Max Depth and IP Depth should be set to zero.

These networks can support LSP which require strict packet ordering, but cannot support very large LSP.

Like the behavior described in [Section 4.2.1](#), whether this behavior is good or bad, defaults can be configured which accurately guess the capabilities of links for which no Multipath Link Capability TLV is available.

#### **4.2.3. Legacy Netowrks with Mixed MP and OA Links**

Some network may support Ethernet Link Aggregation and all or a subset of LSR which place each LSP in its entirety on a single link bundle component link for all LSP.

If the "Maximum LSP Bandwidth" is set as described in [Section 4.2.1](#), then very large LSP can be supported over link bundles. Very large LSP cannot be supported on LSR which place each LSP in its entirety on a single link bundle component link for all LSP, but these are clearly indicated in signaling,

In these mixed networks it may not be possible to reliably support LSP which require strict packet ordering. It is not possible to know where Ethernet Link Aggregation is used and it is not possible to accurately determine Link Bundling behavior on link bundles where "Maximum LSP Bandwidth" is equal to "Unreserved Bandwidth".

Upgrading LSR to support Entropy Label on both LAG and Link Bundles would improve the ability to carry LSP which require strict packet ordering. To gain any benefit the LSP ingress would have to add ELI



and EL.

If not all LSR are upgraded, then the MPLS TE multipath extensions identify those LSR and multipath that have been upgraded.

#### **4.3. Transition to Multipath Extension Support**

If a Multipath Node Capability sub-TLV is not advertised (see [Section 2.1](#)), then the LSR does not support these multipath extensions. This allows detection of such nodes and if necessary application of defaults to cover legacy multipath such as typical Ethernet Link Aggregation Behavior.

##### **4.3.1. Simple Transitions**

For networks with LSR that do not support multipath extensions, transition is easiest if all legacy LSR support and are configured with a common link bundling behavior. If Ethernet Link Aggregation is not used, a single configured default is needed to cover LSR that do not advertise a Multipath Node Capability sub-TLV.

If Ethernet Link Aggregation had been previously used on Legacy LSR, if possible LAG should be disabled and the members of the former LAG configured and advertised as a link bundle which uses the equivalent "all ones" behavior.

If Ethernet Link Aggregation remains but can be identified in some way, such as links with capacity in excess of some value (for example: greater than 10 Gb/s), then defaults can be set up for these LAG. Alternately administrative attributes could be used [[RFC3209](#)].

The transition network in this case lacks the ability to determine the largest microflow that can pass through legacy nodes, but this was the case prior to transition for the entire network prior to transition.

##### **4.3.2. More Challenging Transitions**

Transition is made more difficult if legacy LSR in a network support Ethernet Link Aggregation but do not support Link Bundle and cannot be identified by simple means, or the newer LSR lack sufficient configuration capability to support conditional defaults.

This situation is most easily handled if a small upgrade to such an LSR can advertise a fixed Multipath Node Capability sub-TLV giving the characteristics of the Ethernet Link Aggregation implementation on that node. Absent of such cooperation, the problem can be solved by configuration on newer LSR which allows association of a Multipath



Node Capability sub-TLV with a specific legacy router ID and possibly a legacy router ID and link.

LSR supporting Multipath Extensions will need to assign default values through configuration to these legacy LSR running Ethernet Link Aggregation. These default values serve to allow LSP which require strict packet ordering to avoid these legacy LSR.

LSR which do not support [\[RFC4201\]](#) may be sufficiently rare that the ability to assign default values per legacy LSR or per [\[RFC3209\]](#) administrative attribute may not be needed in practice.

## **5. IANA Considerations**

[ ... to be completed ... ]

The symbolic constants IANA-TBD-1 through IANA-TBD-3 need to be replaced. Complete instructions, including identification of the number space for each of these will be added to a later version of this internet-draft.

## **6. Security Considerations**

The combination of MPLS, MPLS-TP, and multipath does not introduce any new security threats. The security considerations for MPLS/GMPLS and for MPLS-TP are documented in [\[RFC5920\]](#) and [\[I-D.ietf-mpls-tp-security-framework\]](#).

## **7. References**

### **[7.1.](#) Normative References**

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", [RFC 3471](#), January 2003.
- [RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", [RFC 4201](#), October 2005.
- [RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic



Engineering (RSVP-TE)", [RFC 5420](#), February 2009.

- [RFC5786] Aggarwal, R. and K. Kompella, "Advertising a Router's Local Addresses in OSPF Traffic Engineering (TE) Extensions", [RFC 5786](#), March 2010.

## 7.2. Informative References

- [I-D.ietf-mpls-entropy-label]  
Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", [draft-ietf-mpls-entropy-label-06](#) (work in progress), September 2012.
- [I-D.ietf-mpls-tp-security-framework]  
Fang, L., Niven-Jenkins, B., Mansfield, S., and R. Graveman, "MPLS-TP Security Framework", [draft-ietf-mpls-tp-security-framework-04](#) (work in progress), July 2012.
- [I-D.villamizar-mpls-tp-multipath]  
Villamizar, C., "Use of Multipath with MPLS-TP and MPLS", [draft-villamizar-mpls-tp-multipath-03](#) (work in progress), October 2012.
- [IEEE-802.1AX]  
IEEE Standards Association, "IEEE Std 802.1AX-2008 IEEE Standard for Local and Metropolitan Area Networks - Link Aggregation", 2006, <<http://standards.ieee.org/getieee802/download/802.1AX-2008.pdf>>.
- [ITU-T.G.800]  
ITU-T, "Unified functional architecture of transport networks", 2007, <<http://www.itu.int/rec/T-REC-G/recommendation.asp?parent=T-REC-G.800>>.
- [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", [RFC 2475](#), December 1998.
- [RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", [RFC 2702](#), September 1999.
- [RFC2991] Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", [RFC 2991](#), November 2000.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,



and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.

- [RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", [RFC 3260](#), April 2002.
- [RFC3945] Mannie, E., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", [RFC 3945](#), October 2004.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", [RFC 4206](#), October 2005.
- [RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", [RFC 4385](#), February 2006.
- [RFC5586] Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic Associated Channel", [RFC 5586](#), June 2009.
- [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", [RFC 5920](#), July 2010.
- [RFC6107] Shiimoto, K. and A. Farrel, "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", [RFC 6107](#), February 2011.
- [RFC6391] Bryant, S., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network", [RFC 6391](#), November 2011.

#### Author's Address

Curtis Villamizar (editor)  
Outer Cape Cod Network Consulting

Email: [curtis@ocnc.com](mailto:curtis@ocnc.com)

