Network Working Group                          F. Le Faucheur, Editor
Request for Comments: 3270                                      L. Wu
Category: Standards Track                                   B. Davie
                                                        Cisco Systems
                                                           S. Davari
                                                      PMC-Sierra Inc.
                                                        P. Vaananen
                                                              Nokia
                                                         R. Krishnan
                                                    Axiowave Networks
                                                          P. Cheval
                                                             Alcatel
                                                        J. Heinanen
                                                       Song Networks
                                                           May 2002

## Multi-Protocol Label Switching (MPLS)
## Support of Differentiated Services

Status of this Memo

Copyright Notice

Abstract

   This document defines a flexible solution for support of
   Differentiated Services (Diff-Serv) over Multi-Protocol Label
   Switching (MPLS) networks.

   This solution allows the MPLS network administrator to select how
   Diff-Serv Behavior Aggregates (BAs) are mapped onto Label Switched
   Paths (LSPs) so that he/she can best match the Diff-Serv, Traffic
   Engineering and protection objectives within his/her particular
   network.  For instance, this solution allows the network
   administrator to decide whether different sets of BAs are to be
   mapped onto the same LSP or mapped onto separate LSPs.

Table of Contents

## 1. Introduction

   In an MPLS domain [MPLS_ARCH], when a stream of data traverses a
   common path, a Label Switched Path (LSP) can be established using
   MPLS signaling protocols.  At the ingress Label Switch Router (LSR),
   each packet is assigned a label and is transmitted downstream.  At
   each LSR along the LSP, the label is used to forward the packet to
   the next hop.

   In a Differentiated Service (Diff-Serv) domain [DIFF_ARCH] all the IP
   packets crossing a link and requiring the same Diff-Serv behavior are
   said to constitute a Behavior Aggregate (BA).  At the ingress node of
   the Diff-Serv domain, the packets are classified and marked with a
   Diff-Serv Code Point (DSCP) which corresponds to their Behavior
   Aggregate.  At each transit node, the DSCP is used to select the Per
   Hop Behavior (PHB) that determines the scheduling treatment and, in
   some cases, drop probability for each packet.

   This document specifies a solution for supporting the Diff-Serv
   Behavior Aggregates whose corresponding PHBs are currently defined
   (in [DIFF_HEADER], [DIFF_AF], [DIFF_EF]) over an MPLS network.  This
   solution also offers flexibility for easy support of PHBs that may be
   defined in the future.

   This solution relies on the combined use of two types of LSPs:

   - LSPs which can transport multiple Ordered Aggregates, so that the
     EXP field of the MPLS Shim Header conveys to the LSR the PHB to be
     applied to the packet (covering both information about the
     packet's scheduling treatment and its drop precedence).

- LSPs which only transport a single Ordered Aggregate, so that the
  packet's scheduling treatment is inferred by the LSR exclusively
  from the packet's label value while the packet's drop precedence
  is conveyed in the EXP field of the MPLS Shim Header or in the
  encapsulating link layer specific selective drop mechanism (ATM,
  Frame Relay, 802.1).

As mentioned in [DIFF_HEADER], "Service providers are not required to
use the same node mechanisms or configurations to enable service
differentiation within their networks, and are free to configure the
node parameters in whatever way that is appropriate for their service
offerings and traffic engineering objectives".  Thus, the solution
defined in this document gives Service Providers flexibility in
selecting how Diff-Serv classes of service are Routed or Traffic
Engineered within their domain (e.g., separate classes of services
supported via separate LSPs and Routed separately, all classes of
service supported on the same LSP and Routed together).

Because MPLS is path-oriented it can potentially provide faster and
more predictable protection and restoration capabilities in the face
of topology changes than conventional hop by hop routed IP systems.
In this document we refer to such capabilities as "MPLS protection".
Although such capabilities and associated mechanisms are outside the
scope of this specification, we note that they may offer different
levels of protection to different LSPs.  Since the solution presented
here allow Service Providers to choose how Diff-Serv classes of
services are mapped onto LSPs, the solution also gives Service
Providers flexibility in the level of protection provided to
different Diff-Serv classes of service (e.g., some classes of service
can be supported by LSPs which are protected while some other classes
of service are supported by LSPs which are not protected).

Furthermore, the solution specified in this document achieves label
space conservation and reduces the volume of label set-up/tear-down
signaling where possible by only resorting to multiple LSPs for a
given Forwarding Equivalent Class (FEC) [MPLS_ARCH] when useful or
required.

This specification allows support of Differentiated Services for both
IPv4 and IPv6 traffic transported over an MPLS network.  This
document only describes operations for unicast.  Multicast support is
for future study.

The solution described in this document does not preclude the
signaled or configured use of the EXP bits to support Explicit
Congestion Notification [ECN] simultaneously with Diff-Serv over
MPLS.  However, techniques for supporting ECN in an MPLS environment
are outside the scope of this document.

## 1.1  Terminology

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.

The reader is assumed to be familiar with the terminology of
[MPLS_ARCH], [MPLS_ENCAPS], [MPLS_ATM], [MPLS_FR], including the
following:

| | |
|---|---|
| FEC | Forwarding Equivalency Class |
| FTN | FEC-To-NHLFE Map |
| ILM | Incoming Label Map |
| LC-ATM | Label Switching Controlled-ATM (interface) |
| LC-FR | Label Switching Controlled-Frame Relay (interface) |
| LSP | Label Switched Path |
| LSR | Label Switch Router |
| MPLS | Multi-Protocol Label Switching |
| NHLFE | Next Hop Label Forwarding Entry |

The reader is assumed to be familiar with the terminology of
[DIFF_ARCH], [DIFF_HEADER], [DIFF_AF], [DIFF_EF], including the
following:

| | |
|---|---|
| AF | Assured Forwarding |
| BA | Behavior Aggregate |
| CS | Class Selector |
| DF | Default Forwarding |
| DSCP | Differentiated Services Code Point |
| EF | Expedited Forwarding |
| PHB | Per Hop Behavior |

The reader is assumed to be familiar with the terminology of
[DIFF_NEW], including the following:

   OA          Ordered Aggregate.  The set of Behavior Aggregates which
               share an ordering constraint.

   PSC         PHB Scheduling Class.  The set of one or more PHB(s)
               that are applied to the Behavior Aggregate(s) belonging
               to a given OA.  For example, AF1x is a PSC comprising
               the AF11, AF12 and AF13 PHBs.  EF is an example of PSC
               comprising a single PHB, the EF PHB.

The following acronyms are also used:

   CLP         Cell Loss Priority

   DE          Discard Eligibility

   SNMP        Simple Network Management Protocol

Finally, the following acronyms are defined in this specification:

   E-LSP       EXP-Inferred-PSC LSP

   L-LSP       Label-Only-Inferred-PSC LSP

## 1.2 EXP-Inferred-PSC LSPs (E-LSP)

A single LSP can be used to support one or more OAs.  Such LSPs can
support up to eight BAs of a given FEC, regardless of how many OAs
these BAs span.  With such LSPs, the EXP field of the MPLS Shim
Header is used by the LSR to determine the PHB to be applied to the
packet.  This includes both the PSC and the drop preference.

We refer to such LSPs as "EXP-inferred-PSC LSPs" (E-LSP), since the
PSC of a packet transported on this LSP depends on the EXP field
value for that packet.

The mapping from the EXP field to the PHB (i.e., to PSC and drop
precedence) for a given such LSP, is either explicitly signaled at
label set-up or relies on a pre-configured mapping.

Detailed operations of E-LSPs are specified in section 3 below.

1.3 Label-Only-Inferred-PSC LSPs (L-LSP)

   A separate LSP can be established for a single <FEC, OA> pair.  With
   such LSPs, the PSC is explicitly signaled at the time of label
   establishment, so that after label establishment, the LSR can infer
   exclusively from the label value the PSC to be applied to a labeled
   packet.  When the Shim Header is used, the Drop Precedence to be
   applied by the LSR to the labeled packet, is conveyed inside the
   labeled packet MPLS Shim Header using the EXP field.  When the Shim
   Header is not used (e.g., MPLS Over ATM), the Drop Precedence to be
   applied by the LSR to the labeled packet is conveyed inside the link
   layer header encapsulation using link layer specific drop precedence
   fields (e.g., ATM CLP).

   We refer to such LSPs as "Label-Only-Inferred-PSC LSPs" (L-LSP) since
   the PSC can be fully inferred from the label without any other
   information (e.g., regardless of the EXP field value).  Detailed
   operations of L-LSPs are specified in section 4 below.

1.4 Overall Operations

   For a given FEC, and unless media specific restrictions apply as
   identified in the sections 7, 8 and 9 below, this specification
   allows any one of the following combinations within an MPLS Diff-Serv
   domain:

      -  zero or any number of E-LSPs, and

      -  zero or any number of L-LSPs.

   The network administrator selects the actual combination of LSPs from
   the set of allowed combinations and selects how the Behavior
   Aggregates are actually transported over this combination of LSPs, in
   order to best match his/her environment and objectives in terms of
   Diff-Serv support, Traffic Engineering and MPLS Protection.  Criteria
   for selecting such a combination are outside the scope of this
   specification.

   For a given FEC, there may be more than one LSP carrying the same OA,
   for example for purposes of load balancing of the OA; However in
   order to respect ordering constraints, all packets of a given
   microflow, possibly spanning multiple BAs of a given Ordered
   Aggregate, MUST be transported over the same LSP.  Conversely, each
   LSP MUST be capable of supporting all the (active) BAs of a given OA.

   Examples of deployment scenarios are provided for information in
   APPENDIX A.

## 1.5 Relationship between Label and FEC

[MPLS_ARCH] states in section `2.1. Overview' that:  `Some routers
analyze a packet's network layer header not merely to choose the
packet's next hop, but also to determine a packet's "precedence" or
"class of service".  They may then apply different discard thresholds
or scheduling disciplines to different packets.  MPLS allows (but
does not require) the precedence or class of service to be fully or
partially inferred from the label.  In this case, one may say that
the label represents the combination of a FEC and a precedence or
class of service.'

In line with this, we observe that:

- With E-LSPs, the label represents the combination of a FEC and the
  set of BAs transported over the E-LSP.  Where all the supported
  BAs are transported over an E-LSP, the label then represents the
  complete FEC.

- With L-LSPs, the label represents the combination of a FEC and an
  OA.

## 1.6 Bandwidth Reservation for E-LSPs and L-LSPs

Regardless of which label binding protocol is used, E-LSPs and L-LSPs
may be established with or without bandwidth reservation.

Establishing an E-LSP or L-LSP with bandwidth reservation means that
bandwidth requirements for the LSP are signaled at LSP establishment
time.  Such signaled bandwidth requirements may be used by LSRs at
establishment time to perform admission control of the signaled LSP
over the Diff-Serv resources provisioned (e.g., via configuration,
SNMP or policy protocols) for the relevant PSC(s).  Such signaled
bandwidth requirements may also be used by LSRs at establishment time
to perform adjustment to the Diff-Serv resources associated with the
relevant PSC(s) (e.g., adjust PSC scheduling weight).

Note that establishing an E-LSP or L-LSP with bandwidth reservation
does not mean that per-LSP scheduling is required.  Since E-LSPs and
L-LSPs are specified in this document for support of Differentiated
Services, the required forwarding treatment (scheduling and drop
policy) is defined by the appropriate Diff-Serv PHB.  This forwarding
treatment MUST be applied by the LSR at the granularity of the BA and
MUST be compliant with the relevant PHB specification.

When bandwidth requirements are signaled at the establishment of an
L-LSP, the signaled bandwidth is obviously associated with the L-
LSP's PSC.  Thus, LSRs which use the signaled bandwidth to perform
admission control may perform admission control over Diff-Serv
resources, which are dedicated to the PSC (e.g., over the bandwidth
guaranteed to the PSC through its scheduling weight).

When bandwidth requirements are signaled at the establishment of an
E-LSP, the signaled bandwidth is associated collectively with the
whole LSP and therefore with the set of transported PSCs.  Thus, LSRs
which use the signaled bandwidth to perform admission control may
perform admission control over global resources, which are shared by
the set of PSCs (e.g., over the total bandwidth of the link).

Examples of scenarios where bandwidth reservation is not used and
scenarios where bandwidth reservation is used are provided for
information in APPENDIX B.

## 2. Label Forwarding Model for Diff-Serv LSRs and Tunneling Models

## 2.1 Label Forwarding Model for Diff-Serv LSRs

Since different Ordered Aggregates of a given FEC may be transported
over different LSPs, the label swapping decision of a Diff-Serv LSR
clearly depends on the forwarded packet's Behavior Aggregate.  Also,
since the IP DS field of a forwarded packet may not be directly
visible to an LSR, the way to determine the PHB to be applied to a
received packet and to encode the PHB into a transmitted packet, is
different than a non-MPLS Diff-Serv Router.

Thus, in order to describe Label Forwarding by Diff-Serv LSRs, we
model the LSR Diff-Serv label switching behavior, comprised of four
stages:

- Incoming PHB Determination (A)

- Outgoing PHB Determination with Optional Traffic Conditioning(B)

- Label Forwarding (C)

- Encoding of Diff-Serv information into Encapsulation Layer (EXP,
  CLP, DE, User_Priority) (D)

Each stage is described in more detail in the following sections.

Obviously, to enforce the Diff-Serv service differentiation the LSR
MUST also apply the forwarding treatment corresponding to the
Outgoing PHB.

This model is illustrated below:

```
--Inc_label(s)(*)----------------------->I===I--Outg_label(s)(&)-->
  \                                       I   I \
   \---->I===I                            I C I  \-->I===I--Encaps->
         I A I              I===I--Outg_PHB->I===I      I D I    (&)
-Encaps->I===I--Inc_PHB->I B I                 \        /->I===I
   (*)                     I===I                \-------+
                                                        \----Forwarding-->
                                                             Treatment
                                                              (PHB)
```
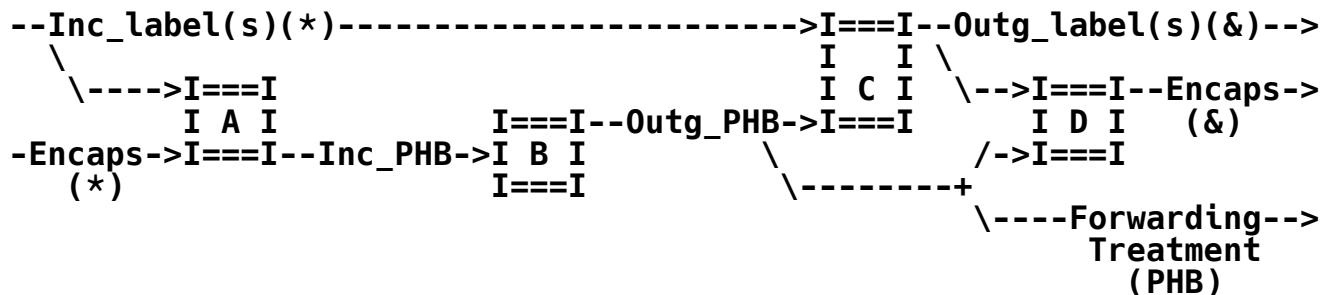
"Encaps" designates the Diff-Serv related information encoded in the
MPLS Encapsulation layer (e.g., EXP field, ATM CLP, Frame Relay DE,
802.1 User_Priority)

(*) when the LSR behaves as an MPLS ingress node, the incoming packet
may be received unlabelled.

(&) when the LSR behaves as an MPLS egress node, the outgoing packet
may be transmitted unlabelled.

This model is presented here to describe the functional operations of
Diff-Serv LSRs and does not constrain actual implementation.

## 2.2 Incoming PHB Determination

This stage determines which Behavior Aggregate the received packet
belongs to.

### 2.2.1 Incoming PHB Determination Considering a Label Stack Entry

Sections 3.3 and 4.3 provide the details on how to perform incoming
PHB Determination considering a given received label stack entry
and/or received incoming MPLS encapsulation information depending on
the incoming LSP type and depending on the incoming MPLS
encapsulation.

Section 2.6 provides the details of which label stack entry to
consider for the Incoming PHB Determination depending on the
supported Diff-Serv tunneling mode.

### 2.2.2 Incoming PHB Determination Considering IP Header

Section 2.6 provides the details of when the IP Header is to be
considered for incoming PHB determination, depending on the supported
Diff-Serv tunneling model.  In those cases where the IP header is to

be used, this stage operates exactly as with a non-MPLS IP Diff-Serv
Router and uses the DS field to determine the incoming PHB.

## 2.3 Outgoing PHB Determination With Optional Traffic Conditioning

The traffic conditioning stage is optional and may be used on an LSR
to perform traffic conditioning including Behavior Aggregate demotion
or promotion.  It is outside the scope of this specification.  For
the purpose of specifying Diff-Serv over MPLS forwarding, we simply
note that the PHB to be actually enforced and conveyed to downstream
LSRs by an LSR (referred to as "outgoing PHB"), may be different to
the PHB which had been associated with the packet by the previous LSR
(referred to as "incoming PHB").

When the traffic conditioning stage is not present, the "outgoing
PHB" is simply identical to the "incoming PHB".

## 2.4 Label Forwarding

[MPLS_ARCH] describes how label swapping is performed by LSRs on
incoming labeled packets using an Incoming Label Map (ILM), where
each incoming label is mapped to one or multiple NHLFEs.  [MPLS_ARCH]
also describes how label imposition is performed by LSRs on incoming
unlabelled packets using a FEC-to-NHLFEs Map (FTN), where each
incoming FEC is mapped to one or multiple NHLFEs.

A Diff-Serv Context for a label is comprised of:

-  `LSP type (i.e., E-LSP or L-LSP)'

-  `supported PHBs'

-  `Encaps-->PHB mapping' for an incoming label

-  `Set of PHB-->Encaps mappings' for an outgoing label

The present specification defines that a Diff-Serv Context is stored
in the ILM for each incoming label.

[MPLS_ARCH] states that the `NHLFE may also contain any other
information needed in order to properly dispose of the packet'.  In
accordance with this, the present specification defines that a Diff-
Serv Context is stored in the NHLFE for each outgoing label that is
swapped or pushed.

This Diff-Serv Context information is populated into the ILM and the
FTN at label establishment time.

   If the label corresponds to an E-LSP for which no `EXP<-->PHB
   mapping' has been explicitly signaled at LSP setup, the `supported
   PHBs' is populated with the set of PHBs of the preconfigured
   `EXP<-->PHB mapping', which is discussed below in section 3.2.1.

   If the label corresponds to an E-LSP for which an `EXP<-->PHB
   mapping' has been explicitly signaled at LSP setup, the `supported
   PHBs' is populated with the set of PHBs of the signaled `EXP<-->PHB
   mapping'.

   If the label corresponds to an L-LSP, the `supported PHBs' is
   populated with the set of PHBs forming the PSC that is signaled at
   LSP set-up.

   The details of how the `Encaps-->PHB mapping' or `Set of PHB-->Encaps
   mappings' are populated are defined below in sections 3 and 4.

   [MPLS_ARCH] also states that:

   "If the ILM [respectively, FTN] maps a particular label to a set of
   NHLFEs that contain more than one element, exactly one element of the
   set must be chosen before the packet is forwarded.  The procedures
   for choosing an element from the set are beyond the scope of this
   document.  Having the ILM [respectively, FTN] map a label
   [respectively, a FEC] to a set containing more than one NHLFE may be
   useful if, e.g., it is desired to do load balancing over multiple
   equal-cost paths."

   In accordance with this, the present specification allows that an
   incoming label [respectively FEC] may be mapped, for Diff-Serv
   purposes, to multiple NHLFEs (for instance where different NHLFEs
   correspond to egress labels supporting different sets of PHBs).  When
   a label [respectively FEC] maps to multiple NHLFEs, the Diff-Serv LSR
   MUST choose one of the NHLFEs whose Diff-Serv Context indicates that
   it supports the Outgoing PHB of the forwarded packet.

   When a label [respectively FEC] maps to multiple NHLFEs which support
   the Outgoing PHB, the procedure for choosing one among those is
   outside the scope of this document.  This situation may be
   encountered where it is desired to do load balancing of a Behavior
   Aggregate over multiple LSPs.  In such situations, in order to
   respect ordering constraints, all packets of a given microflow MUST
   be transported over the same LSP.

## 2.5 Encoding Diff-Serv Information Into Encapsulation Layer

This stage determines how to encode the fields which convey Diff-Serv information in the transmitted packet (e.g., MPLS Shim EXP, ATM CLP, Frame Relay DE, 802.1 User_Priority).

### 2.5.1 Encoding Diff-Serv Information Into Transmitted Label Entry

Sections 3.5 and 4.5 provide the details on how to perform Diff-Serv information encoding into a given transmitted label stack entry and/or transmitted MPLS encapsulation information depending on the corresponding outgoing LSP type and depending on the MPLS encapsulation.

Section 2.6 provides the details in which label stack entry to perform Diff-Serv information encoding into depending on the supported Diff-Serv tunneling mode.

### 2.5.2 Encoding Diff-Serv Information Into Transmitted IP Header

To perform Diff-Serv Information Encoding into the transmitted packet IP header, this stage operates exactly as with a non-MPLS IP Diff-Serv Router and encodes the DSCP of the Outgoing PHB into the DS field.

Section 2.6 provides the details of when Diff-Serv Information Encoding is to be performed into transmitted IP header depending on the supported Diff-Serv tunneling mode.

## 2.6 Diff-Serv Tunneling Models over MPLS

### 2.6.1 Diff-Serv Tunneling Models

[DIFF_TUNNEL] considers the interaction of Differentiated Services with IP tunnels of various forms.  MPLS LSPs are not a form of "IP tunnels" since the MPLS encapsulating header does not contain an IP header and thus MPLS LSPs are not considered in [DIFF_TUNNEL]. However, although not a form of "IP tunnel", MPLS LSPs are a form of "tunnel".

From the Diff-Serv standpoint, LSPs share a number of common characteristics with IP Tunnels:

- Intermediate nodes (i.e., Nodes somewhere along the LSP span) only see and operate on the "outer" Diff-Serv information.

- LSPs are unidirectional.

   -  The "outer" Diff-Serv information can be modified at any of the
      intermediate nodes.

   However, from the Diff-Serv standpoint, LSPs also have a distinctive
   property compared to IP Tunnels:

   -  There is generally no behavior analogous to Penultimate Hop
      Popping (PHP) used with IP Tunnels.  Furthermore, PHP results in
      the "outer" Diff-Serv information associated with the LSP not
      being visible to the LSP egress.  In situations where this
      information is not meaningful at the LSP Egress, this is obviously
      not an issue at all.  In situations where this information is
      meaningful at the LSP Egress, then it must somehow be carried in
      some other means.

   The two conceptual models for Diff-Serv tunneling over IP Tunnels
   defined in [DIFF_TUNNEL] are applicable and useful to Diff-Serv over
   MPLS but their respective detailed operations is somewhat different
   over MPLS.  These two models are the Pipe Model and the Uniform
   Model.  Their operations over MPLS are specified in the following
   sections.  Discussion and definition of alternative tunneling models
   are outside the scope of this specification.

2.6.2 Pipe Model

   With the Pipe Model, MPLS tunnels (aka LSPs) are used to hide the
   intermediate MPLS nodes between LSP Ingress and Egress from the
   Diff-Serv perspective.

   In this model, tunneled packets must convey two meaningful pieces of
   Diff-Serv information:

   -  the Diff-Serv information which is meaningful to intermediate
      nodes along the LSP span including the LSP Egress (which we refer
      to as the "LSP Diff-Serv Information").  This LSP Diff-Serv
      Information is not meaningful beyond the LSP Egress: Whether
      Traffic Conditioning at intermediate nodes on the LSP span affects
      the LSP Diff-Serv information or not, this updated Diff-Serv
      information is not considered meaningful beyond the LSP Egress and
      is ignored.

   -  the Diff-Serv information which is meaningful beyond the LSP
      Egress (which we refer to as the "Tunneled Diff-Serv
      Information").  This information is to be conveyed by the LSP
      Ingress to the LSP Egress.  This Diff-Serv information is not
      meaningful to the intermediate nodes on the LSP span.

Operation of the Pipe Model without PHP is illustrated below:
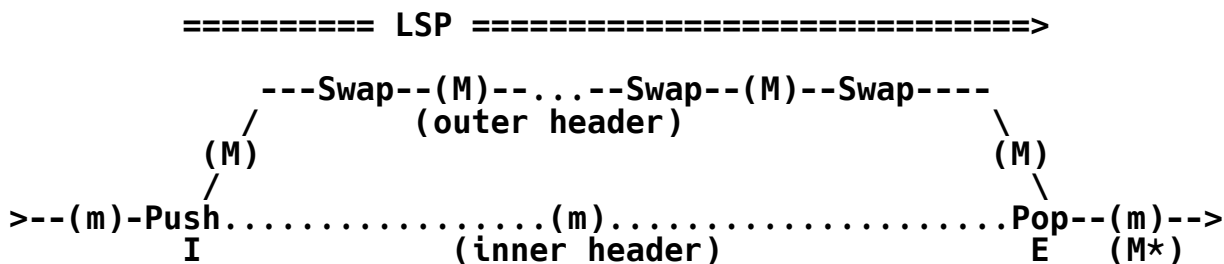
```
         ========== LSP ============================>

             ---Swap--(M)--...--Swap--(M)--Swap----
            /           (outer header)              \
         (M)                                         (M)
          /                                            \
>--(m)-Push.................(m).....................Pop--(m)-->
         I              (inner header)              E   (M*)
```

(M) represents the "LSP Diff-Serv information"
(m) represents the "Tunneled Diff-Serv information"
(*) The LSP Egress considers the LSP Diff-Serv information received
    in the outer header (i.e., before the pop) in order to apply its
    Diff-Serv forwarding treatment (i.e., actual PHB)
 I  represents the LSP ingress node
 E  represents the LSP egress node

With the Pipe Model, the "LSP Diff-Serv Information" needs to be
conveyed to the LSP Egress so that it applies its forwarding
treatment based on it.  The "Tunneled Diff-Serv information" also
needs to be conveyed to the LSP Egress so it can be conveyed further
downstream.

Since both require that Diff-Serv information be conveyed to the LSP
Egress, the Pipe Model operates only without PHP.

The Pipe Model is particularly appropriate for environments in which:

- the cloud upstream of the incoming interface of the LSP Ingress
  and the cloud downstream of the outgoing interface of the LSP
  Egress are in Diff-Serv domains which use a common set of Diff-
  Serv service provisioning policies and PHB definitions, while the
  LSP spans one (or more) Diff-Serv domain(s) which use(s) a
  different set of Diff-Serv service provisioning policies and PHB
  definitions

- the outgoing interface of the LSP Egress is in the (last) Diff-
  Serv domain spanned by the LSP.

As an example, consider the case where a service provider is offering
an MPLS VPN service (see [MPLS_VPN] for an example of MPLS VPN
architecture) including Diff-Serv differentiation.  Say that a
collection of sites is interconnected via such an MPLS VPN service.
Now say that this collection of sites is managed under a common
administration and is also supporting Diff-Serv service
differentiation.  If the VPN site administration and the Service

Provider are not sharing the exact same Diff-Serv policy (for
instance not supporting the same number of PHBs), then operation of
Diff-Serv in the Pipe Model over the MPLS VPN service would allow the
VPN Sites Diff-Serv policy to operate consistently throughout the
ingress VPN Site and Egress VPN Site and transparently over the
Service Provider Diff-Serv domain.  It may be useful to view such
LSPs as linking the Diff-Serv domains at their endpoints into a
single Diff-Serv region by making these endpoints virtually
contiguous even though they may be physically separated by
intermediate network nodes.

The Pipe Model MUST be supported.

For support of the Pipe Model over a given LSP without PHP, an LSR
performs the Incoming PHB Determination and the Diff-Serv information
Encoding in the following manner:

- when receiving an unlabelled packet, the LSR performs Incoming PHB
  Determination considering the received IP Header.

- when receiving a labeled packet, the LSR performs Incoming PHB
  Determination considering the outer label entry in the received
  label stack.  In particular, when a pop operation is to be
  performed for the considered LSP, the LSR performs Incoming PHB
  Determination BEFORE the pop.

- when performing a push operation for the considered LSP, the LSR:

  o  encodes Diff-Serv Information corresponding to the OUTGOING PHB
     in the transmitted label entry corresponding to the pushed
     label.

  o  encodes Diff-Serv Information corresponding to the INCOMING PHB
     in the encapsulated header (swapped label entry or IP header).

- when performing a swap-only operation for the considered LSP, the
  LSR encodes Diff-Serv Information in the transmitted label entry
  that contains the swapped label

- when performing a pop operation for the considered LSP, the LSR
  does not perform Encoding of Diff-Serv Information into the header
  exposed by the pop operation (i.e., the LSR leaves the exposed
  header "as is").

2.6.2.1 Short Pipe Model

The Short Pipe Model is an optional variation of the Pipe Model
described above.  The only difference is that, with the Short Pipe

Model, the Diff-Serv forwarding treatment at the LSP Egress is
applied based on the "Tunneled Diff-Serv Information" (i.e., Diff-
Serv information conveyed in the encapsulated header) rather than on
the "LSP Diff-Serv information" (i.e., Diff-Serv information conveyed
in the encapsulating header).

Operation of the Short Pipe Model without PHP is illustrated below:

```
        ========== LSP =============================>

              ---Swap--(M)--...--Swap--(M)--Swap----
             /              (outer header)           \
          (M)                                          (M)
          /                                              \
>--(m)-Push...................(m).......................Pop--(m)-->
        I              (inner header)                   E
```

(M) represents the "LSP Diff-Serv information"
(m) represents the "Tunneled Diff-Serv information"
 I  represents the LSP ingress node
 E  represents the LSP egress node

Since the LSP Egress applies its forwarding treatment based on the
"Tunneled Diff-Serv Information", the "LSP Diff-Serv information"
does not need to be conveyed by the penultimate node to the LSP
Egress.  Thus the Short Pipe Model can also operate with PHP.

Operation of the Short Pipe Model with PHP is illustrated below:
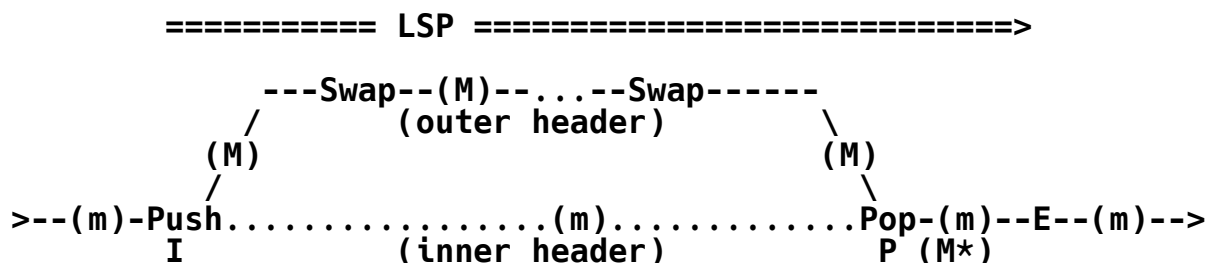
```
        ========== LSP =============================>

              ---Swap--(M)--...--Swap------
             /              (outer header)    \
          (M)                                  (M)
          /                                      \
>--(m)-Push...................(m).............Pop-(m)--E--(m)-->
        I              (inner header)         P (M*)
```

(M) represents the "LSP Diff-Serv information"
(m) represents the "Tunneled Diff-Serv information"
(*) The Penultimate LSR considers the LSP Diff-Serv information
    received in the outer header (i.e., before the pop) in order to
    apply its Diff-Serv forwarding treatment (i.e., actual PHB)
 I  represents the LSP ingress node
 P  represents the LSP penultimate node
 E  represents the LSP egress node

The Short Pipe Model is particularly appropriate for environments in which:

- the cloud upstream of the incoming interface of the LSP Ingress and the cloud downstream of the outgoing interface of the LSP Egress are in Diff-Serv domains which use a common set of Diff-Serv service provisioning policies and PHB definitions, while the LSP spans one (or more) Diff-Serv domain(s) which use(s) a different set of Diff-Serv service provisioning policies and PHB definitions

- the outgoing interface of the LSP Egress is in the same Diff-Serv domain as the cloud downstream of it.

Since each outgoing interface of the LSP Egress is in the same Diff-Serv domain as the cloud downstream of it, each outgoing interface may potentially be in a different Diff-Serv domain, and the LSP Egress needs to be configured with awareness of every corresponding Diff-Serv policy.  This operational overhead is justified in some situations where the respective downstream Diff-Serv policies are better suited to offering service differentiation over each egress interface than the common Diff-Serv policy used on the LSP span.  An example of such a situation is where a Service Provider offers an MPLS VPN service and where some VPN users request that their own VPN Diff-Serv policy be applied to control service differentiation on the dedicated link from the LSP Egress to the destination VPN site, rather than the Service Provider's Diff-Serv policy.

The Short Pipe Model MAY be supported.

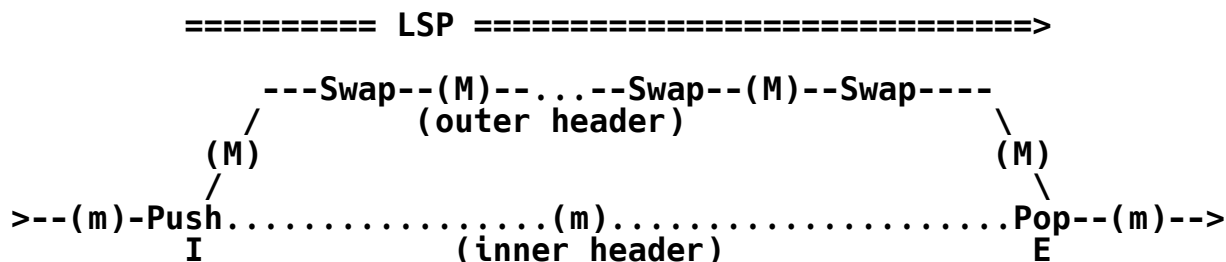For support of the Short Pipe Model over a given LSP without PHP, an LSR performs the Incoming PHB Determination and the Diff-Serv information Encoding in the same manner as with the Pipe Model with the following exception:

- when receiving a labeled packet, the LSR performs Incoming PHB Determination considering the header (label entry or IP header) which is used to do the actual forwarding.  In particular, when a pop operation is to be performed for the considered LSP, the LSR performs Incoming PHB Determination AFTER the pop.

For support of the Short Pipe Model over a given LSP with PHP, an LSR performs Incoming PHB Determination and Diff-Serv information Encoding in the same manner as without PHP with the following exceptions:

        - the Penultimate LSR performs Incoming PHB Determination
          considering the outer label entry in the received label stack.  In
          other words, when a pop operation is to be performed for the
          considered LSP, the Penultimate LSR performs Incoming PHB
          Determination BEFORE the pop.

     Note that the behavior of the Penultimate LSR in the Short Pipe Mode
     with PHP, is identical to the behavior of the LSP Egress in the Pipe
     Mode (necessarily without PHP).

## 2.6.3 Uniform Model

     With the Uniform Model, MPLS tunnels (aka LSPs) are viewed as
     artifacts of the end-to-end path from the Diff-Serv standpoint.  MPLS
     Tunnels may be used for forwarding purposes but have no significant
     impact on Diff-Serv.  In this model, any packet contains exactly one
     piece of Diff-Serv information which is meaningful and is always
     encoded in the outer most label entry (or in the IP DSCP where the IP
     packet is transmitted unlabelled for instance at the egress of the
     LSP).  Any Diff-Serv information encoded somewhere else (e.g., in
     deeper label entries) is of no significance to intermediate nodes or
     to the tunnel egress and is ignored.  If Traffic Conditioning at
     intermediate nodes on the LSP span affects the "outer" Diff-Serv
     information, the updated Diff-Serv information is the one considered
     meaningful at the egress of the LSP.

     Operation of the Uniform Model without PHP is illustrated below:

```
              ========== LSP =============================>

                  ---Swap--(M)--...-Swap--(M)--Swap----
                 /          (outer header)              \
              (M)                                        (M)
              /                                            \
   >--(M)--Push...............(x)........................Pop--(M)->
           I              (inner header)                   E
```

     (M) represents the Meaningful Diff-Serv information encoded in the
         corresponding header.
     (x) represents non-meaningful Diff-Serv information.
      I  represents the LSP ingress node
      E  represents the LSP egress node

Operation of the Uniform Model with PHP is illustrated below:
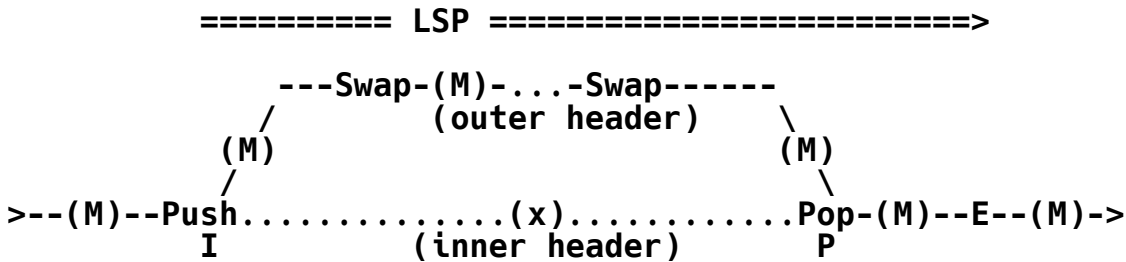
```
        ========== LSP =========================>

          ---Swap-(M)-...-Swap------
         /          (outer header)      \
       (M)                              (M)
       /                                  \
>--(M)--Push..............(x)............Pop-(M)--E--(M)->
        I              (inner header)      P
```

(M) represents the Meaningful Diff-Serv information encoded in the
    corresponding header.
(x) represents non-meaningful Diff-Serv information.
 I  represents the LSP ingress node
 P  represents the LSP penultimate node
 E  represents the LSP egress node

The Uniform Model for Diff-Serv over MPLS is such that, from the
Diff-Serv perspective, operations are exactly identical to the
operations if MPLS was not used.  In other words, MPLS is entirely
transparent to the Diff-Serv operations.

Use of the Uniform Model allows LSPs to span Diff-Serv domain
boundaries without any other measure in place than an inter-domain
Traffic Conditioning Agreement at the physical boundary between the
Diff-Serv domains and operating exclusively on the "outer" header,
since the meaningful Diff-Serv information is always visible and
modifiable in the outmost label entry.

The Uniform Model MAY be supported.

For support of the Uniform Model over a given LSP, an LSR performs
Incoming PHB Determination and Diff-Serv information Encoding in the
following manner:

- when receiving an unlabelled packet, the LSR performs Incoming PHB
  Determination considering the received IP Header.

- when receiving a labeled packet, the LSR performs Incoming PHB
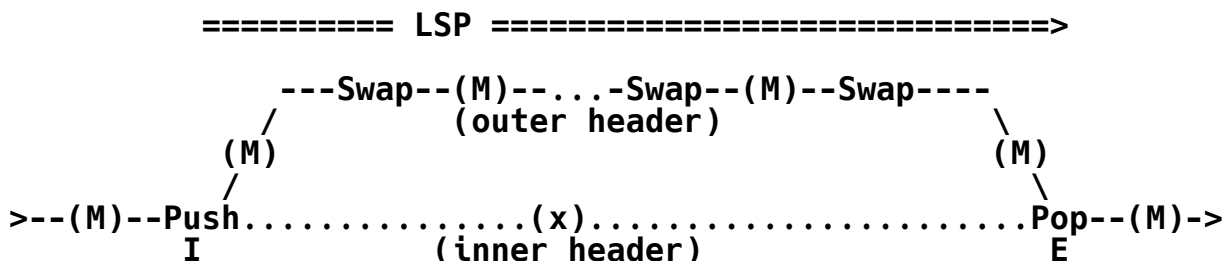  Determination considering the outer label entry in the received
  label stack.  In particular, when a pop operation is to be
  performed for the considered LSP, the LSR performs Incoming PHB
  Determination BEFORE the pop.

- when performing a push operation for the considered LSP, the LSR encodes Diff-Serv Information in the transmitted label entry corresponding to the pushed label.  The Diff-Serv Information encoded in the encapsulated header (swapped label entry or IP Header) is of no importance.

- when performing a swap-only operation for the considered LSP, the LSR encodes Diff-Serv Information in the transmitted label entry that contains the swapped label.

- when PHP is used, the Penultimate LSR needs to be aware of the "Set of PHB-->Encaps mappings" for the label corresponding to the exposed header (or the `PHB-->DSCP mapping') in order to perform Diff-Serv Information Encoding.  Methods for providing this mapping awareness are outside the scope of this specification.  As an example, the "PHB-->DSCP mapping" may be locally configured. As another example, in some environments, it may be appropriate for the Penultimate LSR to assume that the "Set of PHB-->Encaps mappings" to be used for the outgoing label in the exposed header is the "Set of PHB-->Encaps mappings" that would be used by the LSR if the LSR was not doing PHP.  Note also that this specification assumes that the Penultimate LSR does not perform label swapping over the label entry exposed by the pop operation (and in fact that it does not even look at the exposed label). Consequently, restrictions may apply to the Diff-Serv Information Encoding that can be performed by the Penultimate LSR.  For example, this specification does not allow situations where the Penultimate LSR pops a label corresponding to an E-LSP supporting two PSCs, while the header exposed by the pop contains label values for two L-LSPs each supporting one PSC, since the Diff-Serv Information Encoding would require selecting one label or the other.

   Note that LSR behaviors for the Pipe, the Short Pipe and the Uniform Model only differ when doing a push or a pop.  Thus, Intermediate LSRs which perform swap only operations for an LSP, behave in exactly the same way, regardless of whether they are behaving in the Pipe, Short Pipe or the Uniform model.  With a Diff-Serv implementation supporting multiple Tunneling Models, only LSRs behaving as LSP Ingress, Penultimate LSR or LSP Egress need to be configured to operate in a particular Model.  Signaling to associate a Diff-Serv tunneling model on a per-LSP basis is not within the scope of this specification.

## 2.6.4 Hierarchy

Through the label stack mechanism, MPLS allows LSP tunneling to nest
to any depth.  We observe that with such nesting, the push of level
N+1 takes place on a subsequent (or the same) LSR to the LSR doing
the push for level N, while the pop of level N+1 takes place on a
previous (or the same) LSR to the LSR doing the pop of level N.  For
a given level N LSP, the Ingress LSR doing the push and the LSR doing
the pop (Penultimate LSR or LSP Egress) must operate in the same
Tunneling Model (i.e., Pipe, Short Pipe or Uniform).  However, there
is no requirement for consistent tunneling models across levels so
that LSPs at different levels may be operating in different Tunneling
Models.

Hierarchical operations are illustrated below in the case of two
levels of tunnels:

```
        +--------Swap--...---+
       /     (outmost header) \
      /                        \
    Push(2)..................(2)Pop
     / (outer header)          \
    /                           \
>>---Push(1)........................(1)Pop-->>
        (inner header)
```
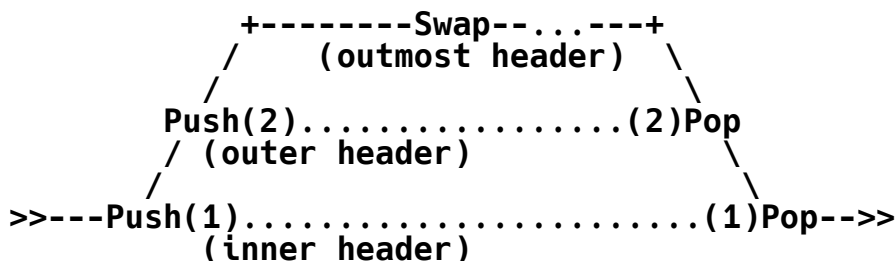
(1) Tunneling Model 1
(2) Tunneling Model 2

Tunneling Model 2 may be the same as or may be different from
Tunneling Model 1.

For a given LSP of level N, the LSR must perform the Incoming PHB
Determination and the Diff-Serv information Encoding as specified in
section 2.6.2, 2.6.2.1 and 2.6.3 according to the Tunneling Model of
this level N LSP and independently of the Tunneling Model of other
level LSPs.

## 3. Detailed Operations of E-LSPs

## 3.1 E-LSP Definition

E-LSPs are defined in section 1.2.

Within a given MPLS Diff-Serv domain, all the E-LSPs relying on the
pre-configured mapping are capable of transporting the same common
set of 8, or fewer, BAs.  Each of those E-LSPs may actually transport
this full set of BAs or any arbitrary subset of it.

For a given FEC, two given E-LSPs using a signaled `EXP<-->PHB
mapping' can support the same or different sets of Ordered
Aggregates.

## 3.2 Populating the `Encaps-->PHB mapping' for an incoming E-LSP

This section defines how the `Encaps-->PHB mapping' of the Diff-Serv
Context is populated for an incoming E-LSP in order to allow Incoming
PHB determination.

The `Encaps-->PHB mapping' for an E-LSP is always of the form
`EXP-->PHB mapping'.

If the label corresponds to an E-LSP for which no `EXP<-->PHB
mapping' has been explicitly signaled at LSP setup, the `EXP-->PHB
mapping' is populated based on the Preconfigured `EXP<-->PHB mapping'
which is discussed below in section 3.2.1.

If the label corresponds to an E-LSP for which an `EXP<-->PHB
mapping' has been explicitly signaled at LSP setup, the `EXP-->PHB
mapping' is populated as per the signaled `EXP<-->PHB mapping'.

## 3.2.1 Preconfigured `EXP<-->PHB mapping'

LSRs supporting E-LSPs which use the preconfigured `EXP<-->PHB
mapping' must allow local configuration of this `EXP<-->PHB mapping'.
This mapping applies to all the E-LSPs established on this LSR
without a mapping explicitly signaled at set-up time.

The preconfigured `EXP<-->PHB mapping' must either be consistent at
every E-LSP hop throughout the MPLS Diff-Serv domain spanned by the
LSP or appropriate remarking of the EXP field must be performed by
the LSR whenever a different preconfigured mapping is used on the
ingress and egress interfaces.

In case, the preconfigured `EXP<-->PHB mapping' has not actually been
configured by the Network Administrator, the LSR should use a default
preconfigured `EXP<-->PHB mapping' which maps all EXP values to the
Default PHB.

## 3.3 Incoming PHB Determination On Incoming E-LSP

This section defines how Incoming PHB Determination is carried out
when the considered label entry in the received label stack
corresponds to an E-LSP.  This requires that the `Encaps-->PHB
mapping' is populated as defined in section 3.2.

When considering a label entry corresponding to an incoming E-LSP for
Incoming PHB Determination, the LSR:

- determines the `EXP-->PHB mapping' by looking up the `Encaps-->PHB
  mapping' of the Diff-Serv Context associated in the ILM with the
  considered incoming E-LSP label.

- determines the incoming PHB by looking up the EXP field of the
  considered label entry in the `EXP-->PHB mapping' table.

## 3.4 Populating the `Set of PHB-->Encaps mappings' for an outgoing E-LSP

This section defines how the `Set of PHB-->Encaps mappings' of the
Diff-Serv Context is populated at label setup for an outgoing E-LSP
in order to allow Encoding of Diff-Serv information in the
Encapsulation Layer.

### 3.4.1 `PHB-->EXP mapping'

An outgoing E-LSP must always have a `PHB-->EXP mapping' as part of
the `Set of PHB-->Encaps mappings' of its Diff-Serv Context.

If the label corresponds to an E-LSP for which no `EXP<-->PHB
mapping' has been explicitly signaled at LSP setup, this `PHB-->EXP
mapping' is populated based on the Preconfigured `EXP<-->PHB mapping'
which is discussed above in section 3.2.1.

If the label corresponds to an E-LSP for which an `EXP<-->PHB
mapping' has been explicitly signaled at LSP setup, the `PHB-->EXP
mapping' is populated as per the signaled `EXP<-->PHB mapping'.

### 3.4.2 `PHB-->CLP mapping'

If the LSP is egressing over an ATM interface which is not label
switching controlled, then one `PHB-->CLP mapping' is added to the
`Set of PHB-->Encaps mappings' for this outgoing LSP.  This
`PHB-->CLP mapping' is populated in the following way:

- it is a function of the PHBs supported on this LSP, and may use
  the relevant mapping entries for these PHBs from the Default
  `PHB-->CLP mapping' defined in section 3.4.2.1.  Mappings other
  than the one defined in section 3.4.2.1 may be used.  In
  particular, if a mapping from PHBs to CLP is standardized in the
  future for operations of Diff-Serv over ATM, such a standardized
  mapping may then be used.

For example if the outgoing label corresponds to an LSP supporting
the AF1 PSC, then the `PHB-->CLP mapping' may be populated with:

| PHB | | CLP Field |
|-----|---|-----------|
| AF11 | ----> | 0 |
| AF12 | ----> | 1 |
| AF13 | ----> | 1 |
| EF | ----> | 0 |

Notice that in this case the `Set of PHB-->Encaps mappings' contains
both a `PHB-->EXP mapping' and a `PHB-->CLP mapping'.

## 3.4.2.1 Default `PHB-->CLP mapping'

| PHB | | CLP Bit |
|-----|---|---------|
| DF | ----> | 0 |
| CSn | ----> | 0 |
| AFn1 | ----> | 0 |
| AFn2 | ----> | 1 |
| AFn3 | ----> | 1 |
| EF | ----> | 0 |

## 3.4.3 `PHB-->DE mapping'

If the LSP is egressing over a Frame Relay interface which is not
label switching controlled, one `PHB-->DE mapping' is added to the
`Set of PHB-->Encaps mappings' for this outgoing LSP and is populated
in the following way:

- it is a function of the PHBs supported on this LSP, and may use
  the relevant mapping entries for these PHBs from the Default
  `PHB-->DE mapping' defined in section 3.4.3.1.  Mappings other
  than the one defined in section 3.4.3.1 may be used.  In
  particular, if a mapping from PHBs to DE is standardized in the
  future for operations of Diff-Serv over Frame Relay, such a
  standardized mapping may then be used.

Notice that in this case the `Set of PHB-->Encaps mappings' contains
both a `PHB-->EXP mapping' and a `PHB-->DE mapping'.

3.4.3.1 `Default PHB-->DE mapping'

         PHB                    DE Bit

          DF       ---->          0
          CSn      ---->          0
          AFn1     ---->          0
          AFn2     ---->          1
          AFn3     ---->          1
          EF       ---->          0

3.4.4 `PHB-->802.1 mapping'

   If the LSP is egressing over a LAN interface on which multiple 802.1
   Traffic Classes are supported as per [IEEE_802.1], then one
   `PHB-->802.1 mapping' is added to the `Set of PHB-->Encaps mappings'
   for this outgoing LSP.  This `PHB-->802.1 mapping' is populated in
   the following way:

   -  it is a function of the PHBs supported on this LSP, and uses the
      relevant mapping entries for these PHBs from the Preconfigured
      `PHB-->802.1 mapping' defined in section 3.4.4.1.

   Notice that the `Set of PHB-->Encaps mappings' then contains both a
   `PHB-->EXP mapping' and a `PHB-->802.1 mapping'.

3.4.4.1 Preconfigured `PHB-->802.1 Mapping'

   At the time of producing this specification, there are no
   standardized mapping from PHBs to 802.1 Traffic Classes.
   Consequently, an LSR supporting multiple 802.1 Traffic Classes over
   LAN interfaces must allow local configuration of a `PHB-->802.1
   mapping'.  This mapping applies to all the outgoing LSPs established
   by the LSR on such LAN interfaces.

3.5 Encoding Diff-Serv information into Encapsulation Layer On Outgoing
    E-LSP

   This section defines how to encode Diff-Serv information into the
   MPLS encapsulation Layer for a given transmitted label entry
   corresponding to an outgoing E-LSP.  This requires that the `Set of
   PHB-->Encaps mappings' be populated as defined in section 3.4.

   The LSR first determines the `Set of PHB-->Encaps mappings' of the
   Diff-Serv Context associated with the corresponding label in the
   NHLFE.

3.5.1 `PHB-->EXP mapping'

   If the `Set of PHB-->Encaps mappings' contains a mapping of the form
   `PHB-->EXP mapping', then the LSR:

   -  determines the value to be written in the EXP field of the
      corresponding level label entry by looking up the "outgoing PHB"
      in this `PHB-->EXP mapping' table.

3.5.2 `PHB-->CLP mapping'

   If the `Set of PHB-->Encaps mappings' contains a mapping of the form
   `PHB-->CLP mapping', then the LSR:

   -  determines the value to be written in the CLP field of the ATM
      encapsulation header, by looking up the "outgoing PHB" in this
      `PHB-->CLP mapping' table.

3.5.3 `PHB-->DE mapping'

   If the `Set of PHB-->Encaps mappings' contains a mapping of the form
   `PHB-->DE mapping', then the LSR:

   -  determines the value to be written in the DE field of the Frame
      Relay encapsulation header, by looking up the "outgoing PHB" in
      this `PHB-->DE mapping' table.

3.5.4 `PHB-->802.1 mapping'

   If the `Set of PHB-->Encaps mappings' contains a mapping of the form
   `PHB-->802.1 mapping', then the LSR:

   -  determines the value to be written in the User_Priority field of
      the Tag Control Information of the 802.1 encapsulation header
      [IEEE_802.1], by looking up the "outgoing PHB" in this 'PHB--
      >802.1 mapping' table.

3.6 E-LSP Merging

   In an MPLS domain, two or more LSPs can be merged into one LSP at one
   LSR.  E-LSPs are compatible with LSP Merging under the following
   condition:

      E-LSPs can only be merged into one LSP if they support the exact
      same set of BAs.

For E-LSPs using a signaled `EXP<-->PHB mapping', the above merge
condition MUST be enforced by LSRs through explicit checking at label
setup that the exact same set of PHBs is supported on the merged
LSPs.

For E-LSPs using the preconfigured `EXP<-->PHB mapping', since the
PHBs supported over an E-LSP is not signaled at establishment time,
an LSR can not rely on signaling information to enforce the above
merge.  However all E-LSPs using the preconfigured `EXP<-->PHB
mapping' are required to support the same set of Behavior Aggregates
within a given MPLS Diff-Serv domain.  Thus, merging of E-LSPs using
the preconfigured `EXP<-->PHB mapping' is allowed within a given MPLS
Diff-Serv domain.

## 4.  Detailed Operation of L-LSPs

## 4.1 L-LSP Definition

L-LSPs are defined in section 1.3.

## 4.2 Populating the `Encaps-->PHB mapping' for an incoming L-LSP

This section defines how the `Encaps-->PHB mapping' of the Diff-Serv
Context is populated at label setup for an incoming L-LSP in order to
allow Incoming PHB determination.

## 4.2.1 `EXP-->PHB mapping'

If the LSR terminates the MPLS Shim Layer over this incoming L-LSP
and the L-LSP ingresses on an interface which is not ATM nor Frame
Relay, then the `Encaps-->PHB mapping' is populated in the following
way:

- it is actually a `EXP-->PHB mapping'

- this mapping is a function of the PSC which is carried on this
  LSP, and must use the relevant mapping entries for this PSC from
  the Mandatory `EXP/PSC-->PHB mapping' defined in Section 4.2.1.1.

For example if the incoming label corresponds to an L-LSP supporting
the AF1 PSC, then the `Encaps-->PHB mapping' will be populated with:

      EXP Field              PHB

        001        ---->     AF11
        010        ---->     AF12
        011        ---->     AF13

An LSR, supporting L-LSPs over PPP interfaces and LAN interfaces, is
an example of an LSR terminating the Shim layer over ingress
interfaces which are not ATM nor Frame Relay.

If the LSR terminates the MPLS Shim Layer over this incoming L-LSP
and the L-LSP ingresses on an ATM or Frame Relay interface, then the
`Encaps-->PHB mapping' is populated in the following way:

- it should actually be a `EXP-->PHB mapping'.  Alternative optional
  ways of populating the `Encaps-->PHB mapping' might be defined in
  the future (e.g., using a 'CLP/EXP--> PHB mapping' or a
  'DE/EXP-->PHB mapping') but are outside the scope of this
  document.

- when the `Encaps-->PHB mapping' is an `EXP-->PHB mapping', this
  `EXP-->PHB mapping' mapping is a function of the PSC which is
  carried on the L-LSP, and must use the relevant mapping entries
  for this PSC from the Mandatory `EXP/PSC-->PHB mapping' defined in
  Section 4.2.1.1.

An Edge-LSR of an ATM-MPLS domain or of a FR-MPLS domain is an
example of an LSR terminating the shim layer over an ingress ATM/FR
interface.

4.2.1.1 Mandatory `EXP/PSC --> PHB mapping'

| EXP Field | PSC | | PHB |
|-----------|-----|------|-----|
| 000 | DF | ----> | DF |
| 000 | CSn | ----> | CSn |
| 001 | AFn | ----> | AFn1 |
| 010 | AFn | ----> | AFn2 |
| 011 | AFn | ----> | AFn3 |
| 000 | EF | ----> | EF |

4.2.2 `CLP-->PHB mapping'

If the LSR does not terminate an MPLS Shim Layer over this incoming
label and uses ATM encapsulation (i.e., it is an ATM-LSR), then the
`Encaps-->PHB mapping' for this incoming L-LSP is populated in the
following way:

- it is actually a `CLP-->PHB mapping'

- the mapping is a function of the PSC, which is carried on this
  LSP, and should use the relevant mapping entries for this PSC from
  the Default `CLP/PSC-->PHB mapping' defined in Section 4.2.2.1.

For example if the incoming label corresponds to an L-LSP supporting
the AF1 PSC, then the `Encaps-->PHB mapping' should be populated
with:

```
    CLP Field              PHB

     0          ---->    AF11
     1          ---->    AF12
```

## 4.2.2.1 Default `CLP/PSC --> PHB mapping'

```
    CLP Bit      PSC            PHB

     0          DF    ---->    DF
     0          CSn   ---->    CSn
     0          AFn   ---->    AFn1
     1          AFn   ---->    AFn2
     0          EF    ---->    EF
```

## 4.2.3 `DE-->PHB mapping'

If the LSR does not terminate an MPLS Shim Layer over this incoming
label and uses Frame Relay encapsulation (i.e., it is a FR-LSR), then
the `Encaps-->PHB mapping' for this incoming L-LSP is populated in
the following way:

- it is actually a `DE-->PHB mapping'

- the mapping is a function of the PSC which is carried on this LSP,
  and should use the relevant mapping entries for this PSC from the
  Default `DE/PSC-->PHB mapping' defined in Section 4.2.3.1.

## 4.2.3.1 Default `DE/PSC --> PHB mapping'

```
    DE Bit       PSC            PHB

     0          DF    ---->    DF
     0          CSn   ---->    CSn
     0          AFn   ---->    AFn1
     1          AFn   ---->    AFn2
     0          EF    ---->    EF
```

## 4.3 Incoming PHB Determination On Incoming L-LSP

This section defines how Incoming PHB determination is carried out
when the considered label entry in the received label stack
corresponds to an L-LSP.  This requires that the `Encaps-->PHB
mapping' is populated as defined in section 4.2.

When considering a label entry corresponding to an incoming L-LSP
for Incoming PHB Determination, the LSR first determines the
`Encaps-->PHB mapping' associated with the corresponding label.

## 4.3.1 `EXP-->PHB mapping'

If the `Encaps-->PHB mapping' is of the form `EXP-->PHB mapping',
then the LSR:

- determines the incoming PHB by looking at the EXP field of the
  considered label entry and using the `EXP-->PHB mapping'.

## 4.3.2 `CLP-->PHB mapping'

If the `Encaps-->PHB mapping' is of the form `CLP-->PHB mapping',
then the LSR:

- determines the incoming PHB by looking at the CLP field of the
  ATM Layer encapsulation and using the `CLP-->PHB mapping'.

## 4.3.3 `DE-->PHB mapping'

If the `Encaps-->PHB mapping' is of the form `DE-->PHB mapping',
then the LSR:

- determines the incoming PHB by looking at the DE field of the
  Frame Relay encapsulation and by using the `DE-->PHB mapping'.

## 4.4 Populating the `Set of PHB-->Encaps mappings' for an outgoing L-LSP

This section defines how the `Set of PHB-->Encaps mappings' of the
Diff-Serv Context is populated at label setup for an outgoing L-LSP
in order to allow Encoding of Diff-Serv Information.

## 4.4.1 `PHB-->EXP mapping'

If the LSR uses an MPLS Shim Layer over this outgoing L-LSP, then
one `PHB-->EXP mapping' is added to the `Set of
PHB-->Encaps mappings' for this outgoing
L-LSP.  This `PHB-->EXP mapping' is populated in the following way:

- it is a function of the PSC supported on this LSP, and must use
  the mapping entries relevant for this PSC from the Mandatory
  `PHB-->EXP mapping' defined in section 4.4.1.1.

For example, if the outgoing label corresponds to an L-LSP supporting
the AF1 PSC, then the following `PHB-->EXP mapping' is added into
the `Set of PHB-->Encaps mappings':

```
           PHB                 EXP Field

           AF11     ---->       001
           AF12     ---->       010
           AF13     ---->       011
```

## 4.4.1.1 Mandatory `PHB-->EXP mapping'

```
           PHB                 EXP Field

           DF       ---->       000
           CSn      ---->       000
           AFn1     ---->       001
           AFn2     ---->       010
           AFn3     ---->       011
           EF       ---->       000
```

## 4.4.2 `PHB-->CLP mapping'

If the L-LSP is egressing on an ATM interface (i.e., it is an ATM-LSR or it is a frame-based LSR sending packets on an LC-ATM interface or on an ATM interface which is not label switching controlled), then one `PHB-->CLP mapping' is added to the `Set of PHB-->Encaps mappings' for this outgoing L-LSP.

If the L-LSP is egressing over an ATM interface which is not label-controlled, the `PHB-->CLP mapping' is populated as per section 3.4.2.

If the L-LSP is egressing over an LC-ATM interface, the `PHB-->CLP mapping' is populated in the following way:

-   it is a function of the PSC supported on this LSP, and should use the relevant mapping entries for this PSC from the Default `PHB-->CLP mapping' defined in section 3.4.2.1.

Notice that if the LSR is a frame-based LSR supporting an L-LSP egressing over an ATM interface, then the `Set of PHB-->Encaps mappings' contains both a `PHB-->EXP mapping' and a `PHB-->CLP mapping'.  If the LSR is an ATM-LSR supporting an L-LSP, then the `Set of PHB-->Encaps mappings' only contains a `PHB-->CLP mapping'.

4.4.3 `PHB-->DE mapping'

   If the L-LSP is egressing over a Frame Relay interface (i.e., it is
   an LSR sending packets on an LC-FR interface or on a Frame Relay
   interface which is not label switching controlled), one `PHB-->DE
   mapping' is added to the `Set of PHB-->Encaps mappings' for this
   outgoing L-LSP.

   If the L-LSP is egressing over a FR interface which is not label
   switching controlled, the `PHB-->DE mapping' is populated as per
   section 3.4.3.

   If the L-LSP is egressing over an LC-FR interface, the `PHB-->DE
   mapping' is populated in the following way:

   -  it is a function of the PSC supported on this LSP, and should use
      the relevant mapping entries for this PSC from the Default
      `PHB-->DE mapping' defined in section 3.4.3.1.

   Notice that if the LSR is an Edge-LSR supporting an L-LSP egressing
   over a LC-FR interface, then the `Set of PHB-->Encaps mappings'
   contains both a `PHB-->EXP mapping' and a `PHB-->DE mapping'.  If the
   LSR is a FR-LSR supporting an L-LSP, then the `Set of PHB-->Encaps
   mappings' only contains a `PHB-->DE mapping'.

4.4.4 `PHB-->802.1 mapping'

   If the LSP is egressing over a LAN interface on which multiple 802.1
   Traffic Classes are supported, as defined in [IEEE_802.1], then one
   `PHB-->802.1 mapping' is added as per section 3.4.4.

4.5 Encoding Diff-Serv Information into Encapsulation Layer on Outgoing
    L-LSP

   This section defines how to encode Diff-Serv information into the
   MPLS encapsulation Layer for a transmitted label entry corresponding
   to an outgoing L-LSP.  This requires that the `Set of PHB-->Encaps
   mappings' is populated as defined in section 4.4.

   The LSR first determines the `Set of PHB-->Encaps mappings' of the
   Diff-Serv Context associated with the corresponding label in the
   NHLFE and then performs corresponding encoding as specified in
   sections 3.5.1, 3.5.2, 3.5.3 and 3.5.4.

## 4.6 L-LSP Merging

In an MPLS domain, two or more LSPs can be merged into one LSP at one LSR.  L-LSPs are compatible with LSP Merging under the following condition:

L-LSPs can only be merged into one L-LSP if they support the same PSC.

The above merge condition MUST be enforced by LSRs, through explicit checking at label setup, that the same PSC is supported on the merged LSPs.

Note that when L-LSPs merge, the bandwidth that is available for the PSC downstream of the merge point must be sufficient to carry the sum of the merged traffic.  This is particularly important in the case of EF traffic.  This can be ensured in multiple ways (for instance via provisioning, or via bandwidth signaling and explicit admission control).

## 5. RSVP Extension for Diff-Serv Support

The MPLS architecture does not assume a single label distribution protocol.  [RSVP_MPLS_TE] defines the extension to RSVP for establishing LSPs in MPLS networks.  This section specifies the extensions to RSVP, beyond those defined in [RSVP_MPLS_TE], to establish LSPs supporting Differentiated Services in MPLS networks.

## 5.1 Diff-Serv related RSVP Messages Format

One new RSVP Object is defined in this document: the DIFFSERV Object.  Detailed description of this Object is provided below.  This new Object is applicable to Path messages.  This specification only defines the use of the DIFFSERV Object in Path messages used to establish LSP Tunnels in accordance with [RSVP_MPLS_TE] and thus containing a Session Object with a C-Type equal to LSP_TUNNEL_IPv4 and containing a LABEL_REQUEST object.

Restrictions defined in [RSVP_MPLS_TE] for support of the establishment of LSP Tunnels via RSVP are also applicable to the establishment of LSP Tunnels supporting Diff-Serv: for instance, only unicast LSPs are supported and Multicast LSPs are for further study.

This new DIFFSERV object is optional with respect to RSVP so that general RSVP implementations not concerned with MPLS LSP set up do not have to support this object.

The DIFFSERV Object is optional for support of LSP Tunnels as defined
in [RSVP_MPLS_TE].  A Diff-Serv capable LSR supporting E-LSPs using
the preconfigured `EXP<-->PHB mapping' in compliance with this
specification MAY support the DIFFSERV Object.  A Diff-Serv capable
LSR supporting E-LSPs using a signaled `EXP<-->PHB mapping' in
compliance with this specification MUST support the DIFFSERV Object.
A Diff-Serv capable LSR supporting L-LSPs in compliance with this
specification MUST support the DIFFSERV Object.

## 5.1.1 Path Message Format

The format of the Path message is as follows:

```
        <Path Message> ::=          <Common Header> [ <INTEGRITY> ]
                                    <SESSION> <RSVP_HOP>
                                    <TIME_VALUES>
                                    [ <EXPLICIT_ROUTE> ]
                                    <LABEL_REQUEST>
                                    [ <SESSION_ATTRIBUTE> ]
                                    [ <DIFFSERV> ]
                                    [ <POLICY_DATA> ... ]
                                    [ <sender descriptor> ]

        <sender descriptor> ::=  <SENDER_TEMPLATE> <SENDER_TSPEC>
                                    [ <ADSPEC> ]
                                    [ <RECORD_ROUTE> ]
```

## 5.2 DIFFSERV Object

The DIFFSERV object formats are shown below.  Currently there are two
possible C_Types.  Type 1 is a DIFFSERV object for an E-LSP.  Type 2
is a DIFFSERV object for an L-LSP.

5.2.1. DIFFSERV object for an E-LSP:

   class = 65, C_Type = 1

```
        0                   1                   2                   3
        0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                     Reserved                      | MAPnb |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                          MAP (1)                             |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                                                              |
      //                           ...                              //
       |                                                              |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                        MAP (MAPnb)                           |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
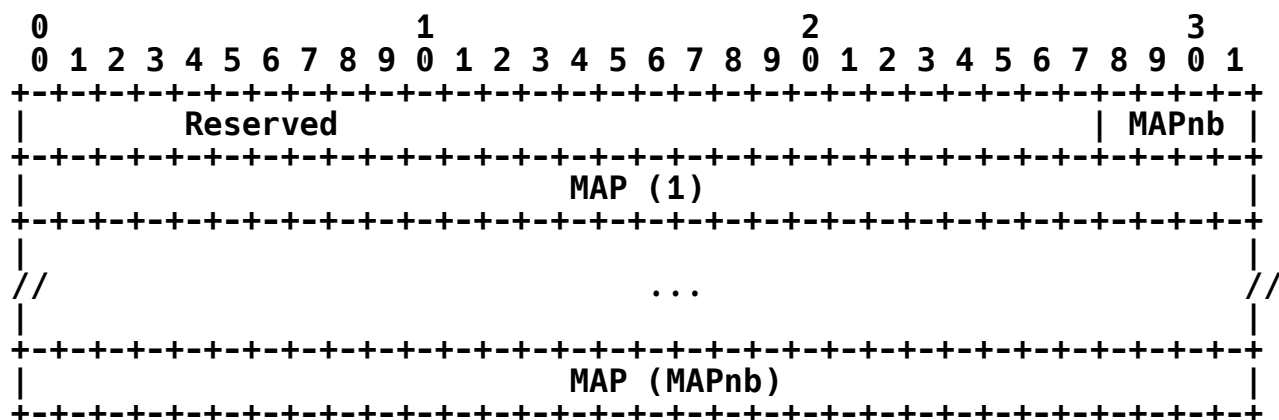
      Reserved : 28 bits
         This field is reserved.  It must be set to zero on transmission
         and must be ignored on receipt.

      MAPnb : 4 bits
         Indicates the number of MAP entries included in the DIFFSERV
         Object.  This can be set to any value from 0 to 8.

      MAP : 32 bits
         Each MAP entry defines the mapping between one EXP field value
         and one PHB.  The MAP entry has the following format:

```
        0                   1                   2                   3
        0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |              Reserved         | EXP |          PHBID          |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

      Reserved : 13 bits
         This field is reserved.  It must be set to zero on transmission
         and must be ignored on receipt.

      EXP : 3 bits
         This field contains the value of the EXP field for the
         `EXP<-->PHB mapping' defined in this MAP entry.

      PHBID : 16 bits
         This field contains the PHBID of the PHB for the `EXP<-->PHB
         mapping' defined in this MAP entry.  The PHBID is encoded as
         specified in [PHBID].

5.2.2 DIFFSERV object for an L-LSP:

   class = 65, C_Type = 2

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |            Reserved            |               PSC             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   Reserved : 16 bits
      This field is reserved.  It must be set to zero on transmission
      and must be ignored on receipt.

   PSC : 16 bits
      The PSC indicates a PHB Scheduling Class to be supported by the
      LSP.  The PSC is encoded as specified in [PHBID].

5.3 Handling DIFFSERV Object

   To establish an LSP tunnel with RSVP, the sender creates a Path
   message with a session type of LSP_Tunnel_IPv4 and with a
   LABEL_REQUEST object as per [RSVP_MPLS_TE].

   To establish an E-LSP tunnel with RSVP, which uses the Preconfigured
   `EXP<-->PHB mapping', the sender creates a Path message:

   -  with a session type of LSP_Tunnel_IPv4,

   -  with the LABEL_REQUEST object, and

   -  without the DIFFSERV object.

   To establish an E-LSP tunnel with RSVP, which uses the Preconfigured
   `EXP<-->PHB mapping', the sender MAY alternatively create a Path
   message:

   -  with a session type of LSP_Tunnel_IPv4,

   -  with the LABEL_REQUEST object, and

   -  with the DIFFSERV object for an E-LSP containing no MAP entries.

   To establish an E-LSP tunnel with RSVP, which uses a signaled
   `EXP<-->PHB mapping', the sender creates a Path message:

   -  with a session type of LSP_Tunnel_IPv4,

- with the LABEL_REQUEST object,

- with the DIFFSERV object for an E-LSP containing one MAP entry for
  each EXP value to be supported on this E-LSP.

To establish with RSVP an L-LSP tunnel, the sender creates a Path
message:

- with a session type of LSP_Tunnel_IPv4,

- with the LABEL_REQUEST object,

- with the DIFFSERV object for an L-LSP containing the PHB
  Scheduling Class (PSC) supported on this L-LSP.

If a path message contains multiple DIFFSERV objects, only the first
one is meaningful; subsequent DIFFSERV object(s) must be ignored and
not forwarded.

Each LSR along the path records the DIFFSERV object, when present, in
its path state block.

If a DIFFSERV object is not present in the Path message, the LSR
SHOULD interpret this as a request for an E-LSP using the
Preconfigured `EXP<-->PHB mapping'.  However, for backward
compatibility purposes, with other non-Diff-Serv Quality of Service
options allowed by [RSVP_MPLS_TE] such as Integrated Services
Controlled Load or Guaranteed Services, the LSR MAY support a
configurable "override option".  When this "override option" is
configured, the LSR interprets a path message without a Diff-Serv
object as a request for an LSP with such non-Diff-Serv Quality of
Service.

If a DIFFSERV object for an E-LSP containing no MAP entry is present
in the Path message, the LSR MUST interpret this as a request for an
E-LSP using the Preconfigured `EXP<-->PHB mapping'.  In particular,
this allows an LSR with the "override option" configured to support
E-LSPs with Preconfigured `EXP<-->PHB mapping', simultaneously with
LSPs with non-Diff-Serv Quality of Service.

If a DIFFSERV object for an E-LSP containing at least one MAP entry
is present in the Path message, the LSR MUST interpret this as a
request for an E-LSP with signaled `EXP<-->PHB mapping'.

If a DIFFSERV object for an L-LSP is present in the Path message, the
LSR MUST interpret this as a request for an L-LSP.

The destination LSR of an E-LSP or L-LSP responds to the Path message containing the LABEL_REQUEST object by sending a Resv message:

-   with the LABEL object

-   without a DIFFSERV object.

Assuming the label request is accepted and a label is allocated, the Diff-Serv LSRs (sender, destination, intermediate nodes) must:

-   update the Diff-Serv Context associated with the established LSPs in their ILM/FTN as specified in previous sections (incoming and outgoing label),

-   install the required Diff-Serv forwarding treatment (scheduling and dropping behavior) for this NHLFE (outgoing label).

An LSR that recognizes the DIFFSERV object and that receives a path message which contains the DIFFSERV object but which does not contain a LABEL_REQUEST object or which does not have a session type of LSP_Tunnel_IPv4, sends a PathErr towards the sender with the error code `Diff-Serv Error' and an error value of `Unexpected DIFFSERV object'.  Those are defined below in section 5.5.

An LSR receiving a Path message with the DIFFSERV object for E-LSP, which recognizes the DIFFSERV object but does not support the particular PHB encoded in one, or more, of the MAP entries, sends a PathErr towards the sender with the error code `Diff-Serv Error' and an error value of `Unsupported PHB'.  Those are defined below in section 5.5.

An LSR receiving a Path message with the DIFFSERV object for E-LSP, which recognizes the DIFFSERV object but determines that the signaled `EXP<-->PHB mapping' is invalid, sends a PathErr towards the sender with the error code `Diff-Serv Error' and an error value of Invalid `EXP<-->PHB mapping'.  Those are defined below in section 5.5.  `The EXP<-->PHB mapping' signaled in the DIFFSERV Object for an E-LSP is invalid when:

-   the MAPnb field is not within the range 0 to 8 or

-   a given EXP value appears in more than one MAP entry, or

-   the PHBID encoding is invalid.

An LSR receiving a Path message with the DIFFSERV object for L-LSP,
which recognizes the DIFFSERV object but does not support the
particular PSC encoded in the PSC field, sends a PathErr towards the
sender with the error code `Diff-Serv Error' and an error value of
`Unsupported PSC'.  Those are defined below in section 5.5.

An LSR receiving a Path message with the DIFFSERV object, which
recognizes the DIFFSERV object but that is unable to allocate the
required per-LSP Diff-Serv context sends a PathErr with the error
code "Diff-Serv Error" and the error value "Per-LSP context
allocation failure".  Those are defined below in section 5.5.

A Diff-Serv LSR MUST handle the situations where the label request
can not be accepted for reasons other than those already discussed in
this section, in accordance with [RSVP_MPLS_TE] (e.g., reservation
rejected by admission control, a label can not be associated).

## 5.4 Non-support of the DIFFSERV Object

An LSR that does not recognize the DIFFSERV object Class-Num MUST
behave in accordance with the procedures specified in [RSVP] for an
unknown Class-Num whose format is 0bbbbbbb i.e., it must send a
PathErr with the error code `Unknown object class' toward the sender.

An LSR that recognize the DIFFSERV object Class-Num but does not
recognize the DIFFSERV object C-Type, must behave in accordance with
the procedures specified in [RSVP] for an unknown C-type i.e., it
must send a PathErr with the error code `Unknown object C-Type'
toward the sender.

In both situations, this causes the path set-up to fail.  The sender
should notify management that a L-LSP cannot be established and
should possibly take action to retry LSP establishment without the
DIFFSERV object (e.g., attempt to use E-LSPs with Preconfigured
`EXP<-->PHB mapping' as a fall-back strategy).

## 5.5 Error Codes For Diff-Serv

In the procedures described above, certain errors must be reported as
a `Diff-Serv Error'.  The value of the `Diff-Serv Error' error code
is 27.

The following defines error values for the Diff-Serv Error:

    Value      Error

      1        Unexpected DIFFSERV object
      2        Unsupported PHB
      3        Invalid `EXP<-->PHB mapping'
      4        Unsupported PSC
      5        Per-LSP context allocation failure

## 5.6 Intserv Service Type

Both E-LSPs and L-LSPs can be established with or without bandwidth
reservation.

As specified in [RSVP_MPLS_TE], to establish an E-LSP or an L-LSP
with bandwidth reservation, Int-Serv's Controlled Load service (or
possibly Guaranteed Service) is used and the bandwidth is signaled in
the SENDER_TSPEC (respectively FLOWSPEC) of the path (respectively
Resv) message.

As specified in [RSVP_MPLS_TE],to establish an E-LSP or an L-LSP
without bandwidth reservation, the Null Service specified in [NULL]
is used.

Note that this specification defines usage of E-LSPs and L-LSPs for
support of the Diff-Serv service only.  Regardless of the Intserv
service (Controlled Load, Null Service, Guaranteed Service,...) and
regardless of whether the reservation is with or without bandwidth
reservation, E-LSPs and L-LSPs are defined here for support of Diff-
Serv services.  Support of Int-Serv services over an MPLS Diff-Serv
backbone is outside the scope of this specification.

Note also that this specification does not concern itself with the
DCLASS object defined in [DCLASS], since this object conveys
information on DSCP values, which are not relevant inside the MPLS
network.

## 6. LDP Extensions for Diff-Serv Support

The MPLS architecture does not assume a single label distribution
protocol.  [LDP] defines the Label Distribution Protocol and its
usage for establishment of label switched paths (LSPs) in MPLS
networks.  This section specifies the extensions to LDP to establish
LSPs supporting Differentiated Services in MPLS networks.

One new LDP TLV is defined in this document:

- the Diff-Serv TLV

Detailed description of this TLV is provided below.

The new Diff-Serv TLV is optional with respect to LDP.  A Diff-Serv
capable LSR supporting E-LSPs which uses the Preconfigured `EXP<--
>PHB mapping' in compliance with this specification MAY support the
Diff-Serv TLV.  A Diff-Serv capable LSR supporting E-LSPs which uses
the signaled `EXP<-->PHB mapping' in compliance with this
specification MUST support the Diff-Serv TLV.  A Diff-Serv capable
LSR supporting L-LSPs in compliance with this specification MUST
support the Diff-Serv TLV.

## 6.1 Diff-Serv TLV

The Diff-Serv TLV has the following formats:

Diff-Serv TLV for an E-LSP:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |U|F|   Diff-Serv (0x0901)        |            Length            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |T|           Reserved                                | MAPnb |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                            MAP (1)                            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                 ...

   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                          MAP (MAPnb)                          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
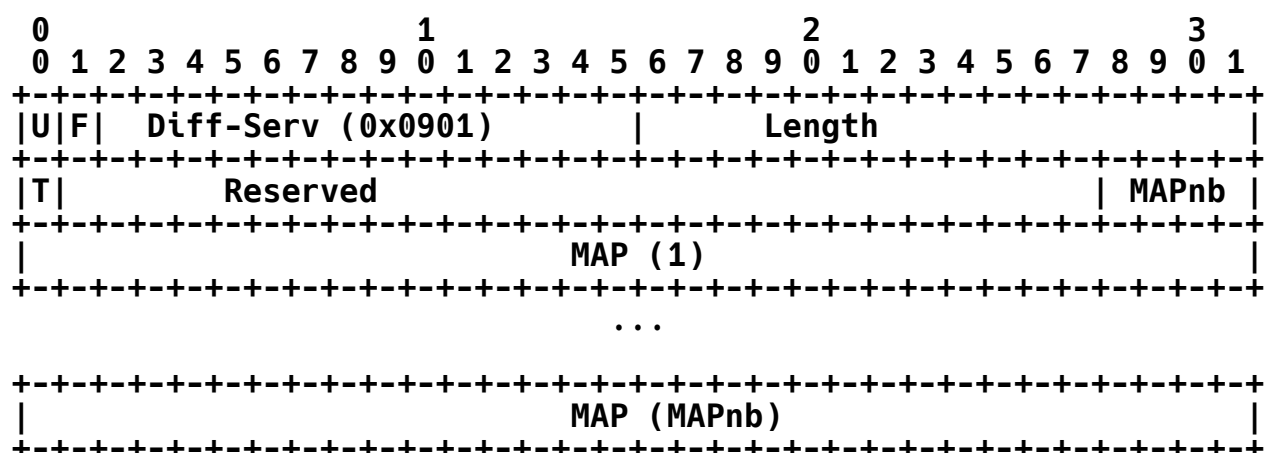
T:1 bit
   LSP Type.  This is set to 0 for an E-LSP

Reserved : 27 bits
   This field is reserved.  It must be set to zero on transmission
   and must be ignored on receipt.

MAPnb : 4 bits
   Indicates the number of MAP entries included in the DIFFSERV
   Object.  This can be set to any value from 1 to 8.

MAP : 32 bits
    Each MAP entry defines the mapping between one EXP field value
    and one PHB.  The MAP entry has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Reserved           | EXP |            PHBID        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
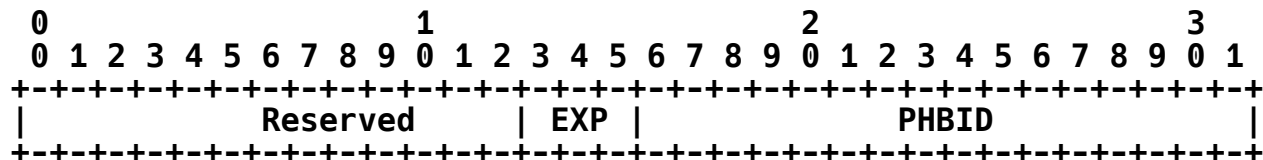
Reserved : 13 bits
    This field is reserved.  It must be set to zero on transmission
    and must be ignored on receipt.

EXP : 3 bits
    This field contains the value of the EXP field for the
    `EXP<-->PHB mapping' defined in this MAP entry.

PHBID : 16 bits
    This field contains the PHBID of the PHB for the `EXP<-->PHB
    mapping' defined in this MAP entry.  The PHBID is encoded as
    specified in [PHBID].

Diff-Serv TLV for an L-LSP:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|U|F|  Type = PSC (0x0901)         |            Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|        Reserved                |              PSC            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

T:1 bit
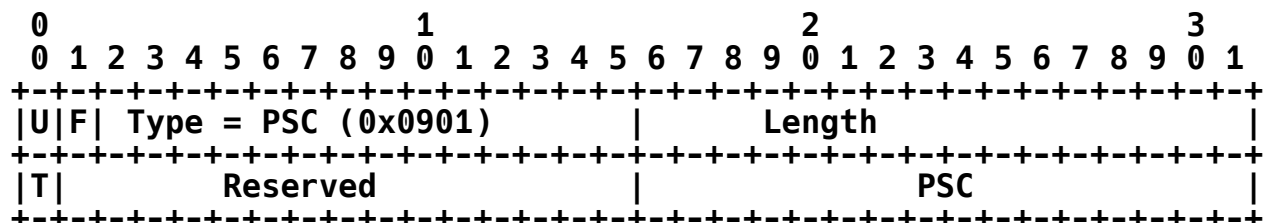    LSP Type.  This is set to 1 for an L-LSP

Reserved : 15 bits
    This field is reserved.  It must be set to zero on transmission
    and must be ignored on receipt.

PSC : 16 bits
    The PSC indicates a PHB Scheduling Class to be supported by the
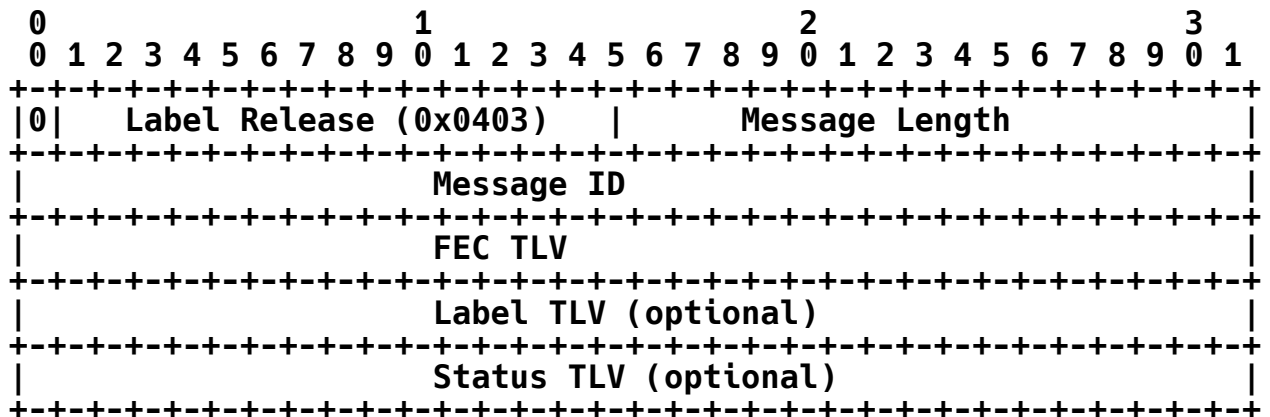    LSP.  The PSC is encoded as specified in [PHBID].

## 6.2 Diff-Serv Status Code Values

The following values are defined for the Status Code field of the
Status TLV:

| Status Code | E | Status Data |
|---|---|---|
| Unexpected Diff-Serv TLV | 0 | 0x01000001 |
| Unsupported PHB | 0 | 0x01000002 |
| Invalid `EXP<-->PHB mapping' | 0 | 0x01000003 |
| Unsupported PSC | 0 | 0x01000004 |
| Per-LSP context allocation failure | 0 | 0x01000005 |

## 6.3 Diff-Serv Related LDP Messages

## 6.3.1 Label Request Message

The format of the Label Request message is extended as follows, to
optionally include the Diff-Serv TLV:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|    Label Request (0x0401)    |          Message Length       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Message ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           FEC TLV                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Diff-Serv TLV (optional)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

## 6.3.2 Label Mapping Message

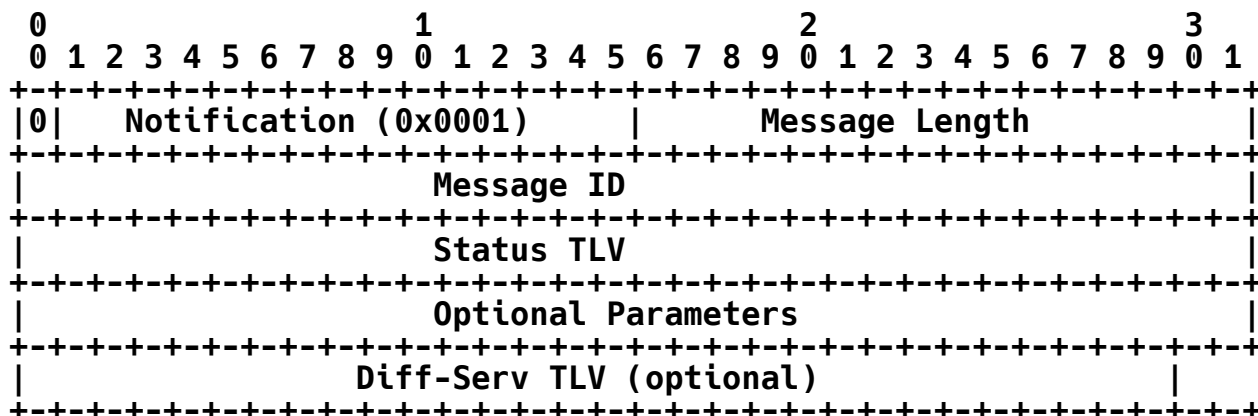The format of the Label Mapping message is extended as follows, to optionally include the Diff-Serv TLV:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|   Label Mapping (0x0400)    |        Message Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Message ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            FEC TLV                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Label TLV                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Diff-Serv TLV (optional)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

## 6.3.3 Label Release Message

The format of the Label Release message is extended as follows, to optionally include the Status TLV:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|    Label Release (0x0403)   |        Message Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Message ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            FEC TLV                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Label TLV (optional)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Status TLV (optional)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

6.3.4 Notification Message

   The format of the Notification message is extended as follows, to
   optionally include the Diff-Serv TLV:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|   Notification (0x0001)     |        Message Length          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Message ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Status TLV                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Optional Parameters                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Diff-Serv TLV (optional)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

6.4 Handling of the Diff-Serv TLV

6.4.1 Handling of the Diff-Serv TLV in Downstream Unsolicited Mode

   This section describes operations when the Downstream Unsolicited
   Mode is used.

   When allocating a label for an E-LSP which is to use the
   preconfigured `EXP<-->PHB mapping', a downstream Diff-Serv LSR issues
   a Label Mapping message without the Diff-Serv TLV.

   When allocating a label for an E-LSP which is to use a signaled
   `EXP<-->PHB mapping', a downstream Diff-Serv LSR issues a Label
   Mapping message with the Diff-Serv TLV for an E-LSP which contains
   one MAP entry for each EXP value to be supported on this E-LSP.

   When allocating a label for an L-LSP, a downstream Diff-Serv LSR
   issues a Label Mapping message with the Diff-Serv TLV for an L-LSP
   which contains the PHB Scheduling Class (PSC) to be supported on this
   L-LSP.

   Assuming the label set-up is successful, the downstream and upstream
   LSRs must:

   -  update the Diff-Serv Context associated with the established LSPs
      in their ILM/FTN as specified in previous sections (incoming and
      outgoing label),

- install the required Diff-Serv forwarding treatment (scheduling
  and dropping behavior) for this NHLFE (outgoing label).

An upstream Diff-Serv LSR receiving a Label Mapping message with
multiple Diff-Serv TLVs only considers the first one as meaningful.
The LSR must ignore and not forward the subsequent Diff-Serv TLV(s).

An upstream Diff-Serv LSR which receives a Label Mapping message,
with the Diff-Serv TLV for an E-LSP and does not support the
particular PHB encoded in one or more of the MAP entries, must reject
the mapping by sending a Label Release message which includes the
Label TLV and the Status TLV with a Status Code of `Unsupported PHB'.

An upstream Diff-Serv LSR receiving a Label Mapping message with the
Diff-Serv TLV for an E-LSP and determining that the signaled
`EXP<-->PHB mapping' is invalid, must reject the mapping by sending a
Label Release message which includes the Label TLV and the Status TLV
with a Status Code of Invalid `EXP<-->PHB mapping'.  The
`EXP<-->PHB mapping' signaled in the DIFFSERV Object for an E-LSP is
invalid when:

- the MAPnb field is not within the range 1 to 8, or

- a given EXP value appears in more than one MAP entry, or

- the PHBID encoding is invalid

An upstream Diff-Serv LSR receiving a Label Mapping message with the
Diff-Serv TLV for an L-LSP containing a PSC value which is not
supported, must reject the mapping by sending a Label Release message
which includes the Label TLV and the Status TLV with a Status Code of
`Unsupported PSC'.

## 6.4.2 Handling of the Diff-Serv TLV in Downstream on Demand Mode

This section describes operations when the Downstream on Demand Mode
is used.

When requesting a label for an E-LSP which is to use the
preconfigured `EXP<-->PHB mapping', an upstream Diff-Serv LSR sends a
Label Request message without the Diff-Serv TLV.

When requesting a label for an E-LSP which is to use a signaled
`EXP<-->PHB mapping', an upstream Diff-Serv LSR sends a Label Request
message with the Diff-Serv TLV for an E-LSP which contains one MAP
entry for each EXP value to be supported on this E-LSP.

When requesting a label for an L-LSP, an upstream Diff-Serv LSR sends
a Label Request message with the Diff-Serv TLV for an L-LSP which
contains the PSC to be supported on this L-LSP.

A downstream Diff-Serv LSR sending a Label Mapping message in
response to a Label Request message for an E-LSP or an L-LSP must not
include a Diff-Serv TLV in this Label Mapping message.  Assuming the
label set-up is successful, the downstream and upstream LSRs must:

-  update the Diff-Serv Context associated with the established LSPs
   in their ILM/FTN as specified in previous sections (incoming and
   outgoing label),

-  install the required Diff-Serv forwarding treatment (scheduling
   and dropping behavior) for this NHLFE (outgoing label).

An upstream Diff-Serv LSR receiving a Label Mapping message
containing a Diff-Serv TLV in response to its Label Request message,
must reject the label mapping by sending a Label Release message
which includes the Label TLV and the Status TLV with a Status Code of
`Unexpected Diff-Serv TLV'.

A downstream Diff-Serv LSR receiving a Label Request message with
multiple Diff-Serv TLVs only considers the first one as meaningful.
The LSR must ignore and not forward the subsequent Diff-Serv TLV(s).

A downstream Diff-Serv LSR which receives a Label Request message
with the Diff-Serv TLV for an E-LSP and does not support the
particular PHB encoded in one (or more) of the MAP entries, must
reject the request by sending a Notification message which includes
the Status TLV with a Status Code of `Unsupported PHB'.

A downstream Diff-Serv LSR receiving a Label Request message with the
Diff-Serv TLV for an E-LSP and determining that the signaled
`EXP<-->PHB mapping' is invalid, must reject the request by sending a
Notification message which includes the Status TLV with a Status Code
of Invalid `EXP<-->PHB mapping'.  The `EXP<-->PHB mapping' signaled
in the DIFFSERV TLV for an E-LSP is invalid when:

-  the MAPnb field is not within the range 1 to 8, or

-  a given EXP value appears in more than one MAP entry, or

-  the PHBID encoding is invalid

A downstream Diff-Serv LSR receiving a Label Request message with the
Diff-Serv TLV for an L-LSP containing a PSC value which is not
supported, must reject the request by sending a Notification message
which includes the Status TLV with a Status Code of `Unsupported
PSC'.

A downstream Diff-Serv LSR that recognizes the Diff-Serv TLV Type in
a Label Request message but is unable to allocate the required per-
LSP context information, must reject the request sending a
Notification message which includes the Status TLV with a Status Code
of `Per-LSP context allocation failure'.

A downstream Diff-Serv LSR that recognizes the Diff-Serv TLV Type in
a Label Request message and supports the requested PSC but is not
able to satisfy the label request for other reasons (e.g., no label
available), must send a Notification message in accordance with
existing LDP procedures [LDP] (e.g., with a `No Label Resource'
Status Code).  This Notification message must include the requested
Diff-Serv TLV.

## 6.5 Non-Handling of the Diff-Serv TLV

An LSR that does not recognize the Diff-Serv TLV Type, on receipt of
a Label Request message or a Label Mapping message containing the
Diff-Serv TLV, must behave in accordance with the procedures
specified in [LDP] for an unknown TLV whose U Bit and F Bit are set
to 0 i.e., it must ignore the message, return a Notification message
with `Unknown TLV' Status.

## 6.6 Bandwidth Information

Bandwidth information may also be signaled at the establishment time
of E-LSP and L-LSP, for instance for the purpose of Traffic
Engineering, using the Traffic Parameters TLV as described in [MPLS
CR LDP].

## 7. MPLS Support of Diff-Serv over PPP, LAN, Non-LC-ATM and Non-LC-FR
   Interfaces

The general operations for MPLS support of Diff-Serv, including label
forwarding and LSP setup operations are specified in the previous
sections.  This section describes the specific operations required
for MPLS support of Diff-Serv over PPP interfaces, LAN interfaces,
ATM Interfaces which are not label controlled and Frame Relay
interfaces which are not label controlled.

On these interfaces, this specification allows any of the following
LSP combinations per FEC:

- Zero or any number of E-LSP, and

- Zero or any number of L-LSPs.

A Diff-Serv capable LSR MUST support E-LSPs which use preconfigured `EXP<-->PHB mapping' over these interfaces.

A Diff-Serv capable LSR MAY support E-LSPs which use signaled `EXP<-->PHB mapping' and L-LSPs over these interfaces.

## 8. MPLS Support of Diff-Serv over LC-ATM Interfaces

This section describes the specific operations required for MPLS support of Diff-Serv over label switching controlled ATM (LC-ATM) interfaces.

This document allows any number of L-LSPs per FEC within an MPLS ATM Diff-Serv domain.  E-LSPs are not supported over LC-ATM interfaces.

## 8.1 Use of ATM Traffic Classes and Traffic Management mechanisms

The use of the "ATM service categories" specified by the ATM Forum, of the "ATM Transfer Capabilities" specified by the ITU-T or of vendor specific ATM traffic classes is outside of the scope of this specification.  The only requirement for compliant implementation is that the forwarding behavior experienced by a Behavior Aggregate forwarded over an L-LSP by the ATM LSR MUST be compliant with the corresponding Diff-Serv PHB specifications.

Since there is only one bit (CLP) for encoding the PHB drop precedence value over ATM links, only two different drop precedence levels are supported in ATM LSRs.  Sections 4.2.2 and 4.4.2 define how the three drop precedence levels of the AFn Ordered Aggregates are mapped to these two ATM drop precedence levels.  This mapping is in accordance with the requirements specified in [DIFF_AF] for the case when only two drop precedence levels are supported.

To avoid discarding parts of the packets, frame discard mechanisms, such as Early Packet Discard (EPD) (see [ATMF_TM]) SHOULD be enabled in the ATM-LSRs for all PHBs described in this document.

## 8.2 LSR Implementation With LC-ATM Interfaces

A Diff-Serv capable LSR MUST support L-LSPs over LC-ATM interfaces. This specification assumes that Edge-LSRs of the ATM-LSR domain use the "shim header" encapsulation method defined in [MPLS_ATM]. Operations without the "shim header" encapsulation are outside the scope of this specification.

## 9. MPLS Support of Diff-Serv over LC-FR Interfaces

This section describes the specific operations required for MPLS
support of Diff-Serv over label switching controlled Frame Relay
(LC-FR) interfaces.

This document allows any number of L-LSPs per FEC within an MPLS
Frame Relay Diff-Serv domain.  E-LSPs are not supported over LC-FR
interfaces.

## 9.1 Use of Frame Relay Traffic parameters and Traffic Management mechanisms

The use of the Frame Relay traffic parameters as specified by ITU-T
and Frame Relay-Forum or of vendor specific Frame Relay traffic
management mechanisms is outside of the scope of this specification.
The only requirement for compliant implementation is that the
forwarding behavior experienced by a Behavior Aggregate forwarded
over an L-LSP by the Frame Relay LSR MUST be compliant with the
corresponding Diff-Serv PHB specifications.

Since there is only one bit (DE) for encoding the PHB drop precedence
value over Frame Relay links, only two different drop precedence
levels are supported in Frame Relay LSRs.  Sections 4.2.3 and 4.4.3
define how the three drop precedence levels of the AFn Ordered
Aggregates are mapped to these two Frame Relay drop precedence
levels.  This mapping is in accordance with the requirements
specified in [DIFF_AF] for the case when only two drop precedence
levels are supported.

## 9.2 LSR Implementation With LC-FR Interfaces

A Diff-Serv capable LSR MUST support L-LSPs over LC-Frame Relay
interfaces.

This specification assumes that Edge-LSRs of the FR-LSR domain use
the "generic encapsulation" method as recommended in [MPLS_FR].
Operations without the "generic encapsulation" are outside the scope
of this specification.

## 10. IANA Considerations

This document defines a number of objects with implications for IANA.

This document defines in section 5.2 a new RSVP object, the DIFFSERV object.  This object required a number from the space defined in [RSVP] for those objects which, if not understood, cause the entire RSVP message to be rejected with an error code of "Unknown Object Class".  Such objects are identified by a zero in the most significant bit of the class number.  Within that space, this object required a number from the "IETF Consensus" space. "65" has been allocated by IANA for the DIFFSERV object.

This document defines in section 5.5 a new RSVP error code, "Diffserv Error".  Error code "27" has been assigned by IANA to the "Diffserv Error".  This document defines values 1 through 5 of the value field to be used within the ERROR_SPEC object for this error code.  Future allocations of values in this space should be handled by IANA using the First Come First Served policy defined in [IANA].

This document defines in section 6.1 a new LDP TLV, the Diffserv TLV. The number for this TLV has been assigned by working group consensus according to the policies defined in [LDP].

This document defines in section 6.2 five new LDP Status Code values for Diffserv-related error conditions.  The values for the Status Code have been assigned by working group consensus according to the policies defined in [LDP].

## 11. Security Considerations

This document does not introduce any new security issues beyond those inherent in Diff-Serv, MPLS and RSVP, and may use the same mechanisms proposed for those technologies.

## 12. Acknowledgments

This document has benefited from discussions with Eric Rosen, Angela Chiu and Carol Iturralde.  It has also borrowed from the work done by D. Black regarding Diff-Serv and IP Tunnels interaction.

APPENDIX A. Example Deployment Scenarios

   This section does not provide additional specification and is only
   here to provide examples of how this flexible approach for Diff-Serv
   support over MPLS may be deployed.  Pros and cons of various
   deployment options for particular environments are beyond the scope
   of this document.

A.1 Scenario 1: 8 (or fewer) BAs, no Traffic Engineering, no MPLS
     Protection

   A Service Provider running 8 (or fewer) BAs over MPLS, not performing
   Traffic engineering, not using MPLS protection and using MPLS Shim
   Header encapsulation in his/her network, may elect to run Diff-Serv
   over MPLS using a single E-LSP per FEC established via LDP.
   Furthermore the Service Provider may elect to use the preconfigured
   `EXP<-->PHB mapping'.

   Operations can be summarized as follows:

   -  the Service Provider configures at every LSR, the bi-directional
      mapping between each PHB and a value of the EXP field
      (e.g., 000<-->AF11, 001<-->AF12, 010<-->AF13)

   -  the Service Provider configures at every LSR, and for every
      interface, the scheduling behavior for each PSC (e.g., bandwidth
      allocated to AF1) and the dropping behavior for each PHB (e.g.,
      drop profile for AF11, AF12, AF13)

   -  LSRs signal establishment of a single E-LSP per FEC using LDP in
      accordance with the specification above (i.e., no Diff-Serv TLV in
      LDP Label Request/Label Mapping messages to implicitly indicate
      that the LSP is an E-LSP and that it uses the preconfigured
      mapping)

A.2 Scenario 2: More than 8 BAs, no Traffic Engineering, no MPLS
     Protection

   A Service Provider running more than 8 BAs over MPLS, not performing
   Traffic Engineering, not using MPLS protection and using MPLS Shim
   encapsulation in his/her network may elect to run Diff-Serv over MPLS
   using for each FEC:

   -  one E-LSP established via LDP and using the preconfigured mapping
      to support a set of 8 (or less) BAs, AND

   -  one L-LSP per <FEC,OA> established via LDP for support of the
      other BAs.

Operations can be summarized as follows:

- the Service Provider configures at every LSR the bi-directional
  mapping between each PHB and a value of the EXP field for the BAs
  transported over the E-LSP

- the Service Provider configures at every LSR, and for every
  interface, the scheduling behavior for each PSC supported over the
  E-LSP and the dropping behavior for each corresponding PHB

- the Service Provider configures at every LSR, and for every
  interface, the scheduling behavior for each PSC supported over the
  L-LSPs and the dropping behavior for each corresponding PHB

- LSRs signal establishment of a single E-LSP per FEC for the set of
  E-LSP transported BAs using LDP as specified above (i.e., no
  Diff-Serv TLV in LDP Label Request/Label Mapping messages to
  implicitly indicate that the LSP is an E-LSP and that it uses the
  preconfigured mapping)

- LSRs signal establishment of one L-LSP per <FEC,OA> for the other
  BAs using LDP as specified above (i.e., Diff-Serv TLV in LDP Label
  Request/Label Mapping messages to indicate the L-LSP's PSC).

A.3 Scenario 3: 8 (or fewer) BAs, Aggregate Traffic Engineering,
     Aggregate MPLS Protection

   A Service Provider running 8 (or fewer) BAs over MPLS, performing
   aggregate Traffic Engineering (i.e., performing a single common path
   selection for all BAs), using aggregate MPLS protection (i.e.,
   restoring service to all PSCs jointly) and using MPLS Shim Header
   encapsulation in his/her network, may elect to run Diff-Serv over
   MPLS using a single E-LSP per FEC established via RSVP [RSVP_MPLS_TE]
   or CR-LDP [CR-LDP_MPLS_TE] and using the preconfigured mapping.

   Operations can be summarized as follows:

   - the Service Provider configures at every LSR the bi-directional
     mapping between each PHB and a value of the EXP field
     (e.g., 000<-->AF11, 001<-->AF12, 010<-->AF13)

   - the Service Provider configures at every LSR, and for every
     interface, the scheduling behavior for each PSC (e.g., bandwidth
     allocated to AF1) and the dropping behavior for each PHB (eg drop
     profile for AF11, AF12, AF13)

   - LSRs signal establishment of a single E-LSP per FEC which will use
     the preconfigured mapping:

    *   using the RSVP protocol as specified above (i.e., no DIFFSERV
        RSVP Object in the PATH message containing the LABEL_REQUEST
        Object), OR

    *   using the CR-LDP protocol as specified above (i.e., no Diff-
        Serv TLV in LDP Label Request/Label Mapping messages).

-   protection is activated on all the E-LSPs in order to achieve MPLS
    protection via mechanisms outside the scope of this document.

A.4 Scenario 4: per-OA Traffic Engineering/MPLS Protection

    A Service Provider running any number of BAs over MPLS, performing
    per-OA Traffic Engineering (i.e., performing a separate path
    selection for each OA) and performing per-OA MPLS protection (i.e.,
    performing protection with potentially different levels of protection
    for the different OAs) in his/her network, may elect to run Diff-Serv
    over MPLS using one L-LSP per <FEC,OA> pair established via RSVP or
    CR-LDP.

    Operations can be summarized as follows:

    -   the Service Provider configures at every LSR, and for every
        interface, the scheduling behavior for each PSC (e.g., bandwidth
        allocated to AF1) and the dropping behavior for each PHB (e.g.,
        drop profile for AF11, AF12, AF13)

    -   LSRs signal establishment of one L-LSP per <FEC,OA>:

        *   using the RSVP as specified above to signal the L-LSP's PSC
            (i.e., DIFFSERV RSVP Object in the PATH message containing the
            LABEL_REQUEST), OR

        *   using the CR-LDP protocol as specified above to signal the L-
            LSP PSC (i.e., Diff-Serv TLV in LDP Label Request/Label Mapping
            messages).

    -   the appropriate level of protection is activated on the different
        L-LSPs (potentially with a different level of protection for each
        PSC) via mechanisms outside the scope of this document.

A.5 Scenario 5: 8 (or fewer) BAs, per-OA Traffic Engineering/MPLS
    Protection

   A Service Provider running 8 (or fewer) BAs over MPLS, performing
   per-OA Traffic Engineering (i.e., performing a separate path
   selection for each OA) and performing per-OA MPLS protection (i.e.,
   performing protection with potentially different levels of protection
   for the different OAs) in his/her network, may elect to run Diff-Serv
   over MPLS using one E-LSP per <FEC,OA> pair established via RSVP or
   CR-LDP.  Furthermore, the Service Provider may elect to use the
   preconfigured mapping on all the E-LSPs.

   Operations can be summarized as follows:

   -  the Service Provider configures at every LSR the bi-directional
      mapping between each PHB and a value of the EXP field
      (e.g., 000<-->AF11, 001<-->AF12, 010<-->AF13)

   -  the Service Provider configures at every LSR, and for every
      interface, the scheduling behavior for each PSC (e.g., bandwidth
      allocated to AF1) and the dropping behavior for each PHB (eg drop
      profile for AF11, AF12, AF13)

   -  LSRs signal establishment of one E-LSP per <FEC,OA>:

      *  using the RSVP protocol as specified above to signal that the
         LSP is an E-LSP which uses the preconfigured mapping (i.e., no
         DIFFSERV RSVP Object in the PATH message containing the
         LABEL_REQUEST), OR

      *  using the CR-LDP protocol as specified above to signal that the
         LSP is an E-LSP which uses the preconfigured mapping (i.e., no
         Diff-Serv TLV in LDP Label Request/Label Mapping messages)

   -  the Service Provider configures, for each E-LSP, at the head-end
      of that E-LSP, a filtering/forwarding criteria so that only the
      packets belonging to a given OA are forwarded on the E-LSP
      established for the corresponding FEC and corresponding OA.

   -  the appropriate level of protection is activated on the different
      E-LSPs (potentially with a different level of protection depending
      on the PSC actually transported over each E-LSP) via mechanisms
      outside the scope of this document.

A.6 Scenario 6: no Traffic Engineering/MPLS Protection on 8 BAs, per-OA
    Traffic Engineering/MPLS Protection on other BAs.

   A Service Provider not performing Traffic Engineering/MPLS Protection
   on 8 (or fewer) BAs, performing per-OA Traffic Engineering/MPLS
   Protection on the other BAs (i.e., performing a separate path
   selection for each OA corresponding to the other BAs and performing
   MPLS Protection with a potentially different policy for each of these
   OA) and using the MPLS Shim encapsulation in his/her network may
   elect to run Diff-Serv over MPLS, using for each FEC:

   -  one E-LSP using the preconfigured mapping established via LDP to
      support the set of 8 (or fewer) non-traffic-engineered/non-
      protected BAs, AND

   -  one L-LSP per <FEC,OA> pair established via RSVP or CR-LDP for
      support of the other BAs.

   Operations can be summarized as follows:

   -  the Service Provider configures at every LSR the bi-directional
      mapping between each PHB and a value of the EXP field for the BAs
      supported over the E-LSP

   -  the Service Provider configures at every LSR, and for every
      interface, the scheduling behavior for each PSC supported over the
      E-LSP and the dropping behavior for each corresponding PHB

   -  the Service Provider configures at every LSR, and for every
      interface, the scheduling behavior for each PSC supported over the
      L-LSPs and the dropping behavior for each corresponding PHB

   -  LSRs signal establishment of a single E-LSP per FEC for the non-
      traffic engineered BAs using LDP as specified above (i.e., no
      Diff-Serv TLV in LDP Label Request/Label Mapping messages)

   -  LSRs signal establishment of one L-LSP per <FEC,OA> for the other
      BAs:

      *  using the RSVP protocol as specified above to signal the L-LSP
         PSC (i.e., DIFFSERV RSVP Object in the PATH message containing
         the LABEL_REQUEST Object), OR

      *  using the CR-LDP protocol as specified above to signal the L-
         LSP PSC (i.e., Diff-Serv TLV in LDP Label Request/Label Mapping
         messages).

- protection is not activated on the E-LSPs.

- the appropriate level of protection is activated on the different
  L-LSPs (potentially with a different level of protection depending
  on the L-LSP's PSC) via mechanisms outside the scope of this
  document.

A.7 Scenario 7: More than 8 BAs, no Traffic Engineering, no MPLS
    Protection

   A Service Provider running more than 8 BAs over MPLS, not performing
   Traffic engineering, not performing MPLS protection and using MPLS
   Shim Header encapsulation in his/her network, may elect to run Diff-
   Serv over MPLS using two E-LSPs per FEC established via LDP and using
   signaled `EXP<-->PHB mapping'.

   Operations can be summarized as follows:

   - the Service Provider configures at every LSR, and for every
     interface, the scheduling behavior for each PSC (e.g., bandwidth
     allocated to AF1) and the dropping behavior for each PHB (e.g.,
     drop profile for AF11, AF12, AF13)

   - LSRs signal establishment of two E-LSPs per FEC using LDP in
     accordance with the specification above (i.e., Diff-Serv TLV in
     LDP Label Request/Label Mapping messages to explicitly indicate
     that the LSP is an E-LSP and its `EXP<-->PHB mapping').  The
     signaled mapping will indicate the subset of 8 (or less) BAs to be
     transported on each E-LSP and what EXP values are mapped to each
     BA on each E-LSP.

APPENDIX B. Example Bandwidth Reservation Scenarios

B.1 Scenario 1: No Bandwidth Reservation

   Consider the case where a network administrator elects to:

   - have Diff-Serv resources entirely provisioned off-line (e.g., via
     Command Line Interface, via SNMP, via COPS,...)

   - have Shortest Path Routing used for all the Diff-Serv traffic.

   This is the closest model to provisioned Diff-Serv over non-MPLS IP.
   In that case, E-LSPs and/or L-LSPs would be established without
   signaled bandwidth.

B.2 Scenario 2: Bandwidth Reservation for per-PSC Admission Control

   Consider the case where a network administrator elects to:

   -  have Diff-Serv resources entirely provisioned off-line (e.g., via
      Command Line Interface, via SNMP, via COPS,...)

   -  use L-LSPs

   -  have Constraint Based Routing performed separately for each PSC,
      where one of the constraints is availability of bandwidth from the
      bandwidth allocated to the relevant PSC.

   In that case, L-LSPs would be established with signaled bandwidth.
   The bandwidth signaled at L-LSP establishment would be used by LSRs
   to perform admission control at every hop to ensure that the
   constraint on availability of bandwidth for the relevant PSC is met.

B.3 Scenario 3: Bandwidth Reservation for per-PSC Admission Control and
     per-PSC Resource Adjustment

   Consider the case where a network administrator elects to:

   -  use L-LSPs

   -  have Constraint Based Routing performed separately for each PSC,
      where one of the constraints is availability of bandwidth from the
      bandwidth allocated to the relevant PSC.

   -  have Diff-Serv resources dynamically adjusted

   In that case, L-LSPs would be established with signaled bandwidth.
   The bandwidth signaled at L-LSP establishment would be used by LSRs
   to attempt to adjust the resources allocated to the relevant PSC
   (e.g., scheduling weight) and then perform admission control to
   ensure that the constraint on availability of bandwidth for the
   relevant PSC is met after the adjustment.

References

   [ANSI/IEEE]        ANSI/IEEE Std 802.1D, 1993 Edition, incorporating
                      IEEE supplements P802.1p, 802.1j-1996, 802.6k-1992,
                      802.11c-1998, and P802.12e).

   [ATMF_TM]          ATM Forum, "Traffic Management Specification Version
                      4.1", March 1999.

   [CR-LDP_MPLS_TE]   Jamoussi, B., Editor, Andersson, L., Callon, R. and
                      R. Dantu, "Constraint-Based LSP Setup using LDP",
                      RFC 3212, January 2002.

   [DCLASS]           Bernet, Y., "Format of the RSVP DCLASS Object", RFC
                      2996, November 2000.

   [DIFF_AF]          Heinanen, J., Baker, F., Weiss, W. and J.
                      Wroclawski, "Assured Forwarding PHB Group", RFC
                      2597, June 1999.

   [DIFF_ARCH]        Blake, S., Black, D., Carlson, M., Davies, E., Wang,
                      Z. and W. Weiss, "An Architecture for Differentiated
                      Services", RFC 2475, December 1998.

   [DIFF_EF]          Davie, B., Charny, A., Baker, F., Bennet, J.,
                      Benson, K., Boudec, J., Chiu, A., Courtney, W.,
                      Davari, S., Firoiu, V., Kalmanek, C., Ramakrishnam,
                      K. and D. Stiliadis, "An Expedited Forwarding PHB
                      (Per-Hop Behavior)", RFC 3246, March 2002.

   [DIFF_HEADER]      Nichols, K., Blake, S., Baker, F. and D. Black,
                      "Definition of the Differentiated Services Field (DS
                      Field) in the IPv4 and IPv6 Headers", RFC 2474,
                      December 1998.

   [DIFF_NEW]         Grossman, D., "New Terminology and Clarifications
                      for Diffserv", RFC 3260, April 2002.

   [DIFF_TUNNEL]      Black, D., "Differentiated Services and Tunnels",
                      RFC 2983, October 2000.

   [ECN]              Ramakrishnan, K., Floyd, S. and D. Black, "The
                      Addition of Explicit Congestion Notification (ECN)
                      to IP", RFC 3168, September 2001.

   [IANA]             Narten, T. and H. Alvestrand, "Guidelines for
                      Writing an IANA Considerations Section in RFCs", BCP
                      26, RFC 2434, October 1998.

   [IEEE_802.1]        ISO/IEC 15802-3: 1998 ANSI/IEEE Std 802.1D, 1998
                       Edition (Revision and redesignation of ISO/IEC
                       10038:98.

   [LDP]               Andersson, L., Doolan, D., Feldman, N., Fredette, A.
                       and B. Thomas, "LDP Specification", RFC 3036,
                       January 2001.

   [MPLS_ARCH]         Rosen, E., Viswanathan, A. and R. Callon,
                       "Multiprotocol Label Switching Architecture", RFC
                       3031, January 2001.

   [MPLS_ATM]          Davie, B., Lawrence, J., McCloghrie, K., Rosen, E.,
                       Swallow, G., Rekhter, Y. and P. Doolan, "MPLS using
                       LDP and ATM VC Switching", RFC 3035, January 2001.

   [MPLS_ENCAPS]       Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y.,
                       Farinacci, D., Li, T. and A. Conta, "MPLS Label
                       Stack Encoding", RFC 3032, January 2001.

   [MPLS_FR]           Conta, A., Doolan, P. and A. Malis, "Use of Label
                       Switching on Frame Relay Networks Specification",
                       RFC 3034, January 2001.

   [MPLS_VPN]          Rosen, E., "BGP/MPLS VPNs", Work in Progress.

   [NULL]              Bernet, Y., Smith, A. and B. Davie, "Specification
                       of the Null Service Type", RFC 2997, November 2000.

   [PHBID]             Black, D., Brim, S., Carpenter, B. and F. Le
                       Faucheur, "Per Hop Behavior Identification Codes"
                       RFC 3140, June 2001.

   [RSVP]              Braden, R., Zhang, L., Berson, S., Herzog, S. and S.
                       Jamin, "Resource ReSerVation Protocol (RSVP) -
                       Version 1 Functional Specification", RFC 2205,
                       September 1997.

   [RSVP_MPLS_TE]      Awduche, D., Berger, L., Gan, D., Li, T.,
                       Srinivasan, V. and G. Swallow, "Extensions to RSVP
                       for LSP Tunnels", RFC 3209, December 2001.

Authors' Addresses

    Francois Le Faucheur
    Cisco Systems
    Village d'Entreprise Green Side - Batiment T3
    400, Avenue de Roumanille
    06410 Biot-Sophia Antipolis
    France

    Phone: +33 4 97 23 26 19
    EMail: flefauch@cisco.com


    Liwen Wu
    Cisco Systems
    3550 Cisco Way
    San Jose, CA 95134
    USA

    Phone: +1 (408) 853-4065
    EMail: liwwu@cisco.com


    Bruce Davie
    Cisco Systems
    250 Apollo Drive, Chelmsford, MA 01824
    USA

    Phone: +1 (978) 244-8000
    EMail: bsd@cisco.com


    Shahram Davari
    PMC-Sierra Inc.
    411 Legget Drive
    Kanata, Ontario K2K 3C9
    Canada

    Phone: +1 (613) 271-4018
    EMail: davari@ieee.org

Pasi Vaananen
Nokia
3 Burlington Woods Drive, Suit 250
Burlington, MA 01803
USA

Phone +1 (781) 993-4900
EMail: pasi.vaananen@nokia.com


Ram Krishnan
Axiowave Networks
200 Nickerson Road
Marlboro, MA 01752

EMail: ram@axiowave.com


Pierrick Cheval
Alcatel
5 rue Noel-Pons
92737 Nanterre Cedex
France
EMail: pierrick.cheval@space.alcatel.fr


Juha Heinanen
Song Networks, Inc.
Hallituskatu 16
33200 Tampere, Finland

EMail: jh@song.fi

Full Copyright Statement

Acknowledgement