Network Working Group
Request for Comments: 4202
Category: Standards Track

K. Kompella, Ed.
Y. Rekhter, Ed.
Juniper Networks

October 2005

Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)

Status of This Memo

Copyright Notice

Abstract

This document specifies routing extensions in support of carrying link state information for Generalized Multi-Protocol Label Switching (GMPLS). This document enhances the routing extensions required to support MPLS Traffic Engineering (TE).

Table of Contents

1. Introduction

This document specifies routing extensions in support of carrying link state information for Generalized Multi-Protocol Label Switching (GMPLS). This document enhances the routing extensions [ISIS-TE], [OSPF-TE] required to support MPLS Traffic Engineering (TE).

Traditionally, a TE link is advertised as an adjunct to a "regular" link, i.e., a routing adjacency is brought up on the link, and when the link is up, both the properties of the link are used for Shortest Path First (SPF) computations (basically, the SPF metric) and the TE properties of the link are then advertised.

GMPLS challenges this notion in three ways. First, links that are not capable of sending and receiving on a packet-by-packet basis may yet have TE properties; however, a routing adjacency cannot be brought up on such links. Second, a Label Switched Path can be advertised as a point-to-point TE link (see [LSP-HIER]); thus, an advertised TE link may be between a pair of nodes that don’t have a routing adjacency with each other. Finally, a number of links may be advertised as a single TE link (perhaps for improved scalability), so again, there is no longer a one-to-one association of a regular routing adjacency and a TE link.

Thus we have a more general notion of a TE link. A TE link is a "logical" link that has TE properties. The link is logical in a sense that it represents a way to group/map the information about certain physical resources (and their properties) into the information that is used by Constrained SPF for the purpose of path computation, and by GMPLS signaling. This grouping/mapping must be done consistently at both ends of the link. LMP [LMP] could be used to check/verify this consistency.

Depending on the nature of resources that form a particular TE link, for the purpose of GMPLS signaling, in some cases the combination of <TE link identifier, label> is sufficient to unambiguously identify the appropriate resource used by an LSP. In other cases, the combination of <TE link identifier, label> is not sufficient; such cases are handled by using the link bundling construct [LINK-BUNDLE] that allows to identify the resource by <TE link identifier, Component link identifier, label>.

Some of the properties of a TE link may be configured on the advertising Label Switching Router (LSR), others which may be obtained from other LSRs by means of some protocol, and yet others which may be deduced from the component(s) of the TE link.

A TE link between a pair of LSRs doesn't imply the existence of a routing adjacency (e.g., an IGP adjacency) between these LSRs. As we mentioned above, in certain cases a TE link between a pair of LSRs could be advertised even if there is no routing adjacency at all between the LSRs (e.g., when the TE link is a Forwarding Adjacency (see [LSP-HIER])).

A TE link must have some means by which the advertising LSR can know of its liveness (this means may be routing hellos, but is not limited to routing hellos). When an LSR knows that a TE link is up, and can determine the TE link's TE properties, the LSR may then advertise that link to its (regular) neighbors.

In this document, we call the interfaces over which regular routing adjacencies are established "control channels".

[ISIS-TE] and [OSPF-TE] define the canonical TE properties, and say how to associate TE properties to regular (packet-switched) links. This document extends the set of TE properties, and also says how to associate TE properties with non-packet-switched links such as links between Optical Cross-Connects (OXCs). [LSP-HIER] says how to associate TE properties with links formed by Label Switched Paths.

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 BCP 14, RFC 2119 [RFC2119].

1.1. Requirements for Layer-Specific TE Attributes

In generalizing TE links to include traditional transport facilities, there are additional factors that influence what information is needed about the TE link. These arise from existing transport layer architecture (e.g., ITU-T Recommendations G.805 and G.806) and associated layer services. Some of these factors are:

1. The need for LSPs at a specific adaptation, not just a particular bandwidth. Clients of optical networks obtain connection services for specific adaptations, for example, a VC-3 circuit. This not only implies a particular bandwidth, but how the payload is structured. Thus the VC-3 client would not be satisfied with any LSP that offered other than 48.384 Mbit/s and with the expected

structure. The corollary is that path computation should be able to find a route that would give a connection at a specific adaptation.

2. Distinguishing variable adaptation. A resource between two OXCs (specifically a G.805 trail) can sometimes support different adaptations at the same time. An example of this is described in section 2.4.8. In this situation, the fact that two adaptations are supported on the same trail is important because the two layers are dependent, and it is important to be able to reflect this layer relationship in routing, especially in view of the relative lack of flexibility of transport layers compared to packet layers.

3. Inheritable attributes. When a whole multiplexing hierarchy is supported by a TE link, a lower layer attribute may be applicable to the upper layers. Protection attributes are a good example of this. If an OC-192 link is 1+1 protected (a duplicate OC-192 exists for protection), then an STS-3c within that OC-192 (a higher layer) would inherit the same protection property.

4. Extensibility of layers. In addition to the existing defined transport layers, new layers and adaptation relationships could come into existence in the future.

5. Heterogeneous networks whose OXCs do not all support the same set of layers. In a GMPLS network, not all transport layer network elements are expected to support the same layers. For example, there may be switches capable of only VC-11, VC-12, and VC-3, and there may be others that can only support VC-3 and VC-4. Even though a network element cannot support a specific layer, it should be able to know if a network element elsewhere in the network can support an adaptation that would enable that unsupported layer to be used. For example, a VC-11 switch could use a VC-3 capable switch if it knew that a VC-11 path could be constructed over a VC-3 link connection.

From the factors presented above, development of layer specific GMPLS routing documents should use the following principles for TE-link attributes.

1. Separation of attributes. The attributes in a given layer are separated from attributes in another layer.

2. Support of inter-layer attributes (e.g., adaptation relationships). Between a client and server layer, a general mechanism for describing the layer relationship exists. For

example, "4 client links of type X can be supported by this server layer link". Another example is being able to identify when two layers share a common server layer.

3. Support for inheritable attributes. Attributes which can be inherited should be identified.

4. Layer extensibility. Attributes should be represented in routing such that future layers can be accommodated. This is much like the notion of the generalized label.

5. Explicit attribute scope. For example, it should be clear whether a given attribute applies to a set of links at the same layer.

The present document captures general attributes that apply to a single layer network, but doesn't capture inter-layer relationships of attributes. This work is left to a future document.

1.2. Excluding Data Traffic from Control Channels

The control channels between nodes in a GMPLS network, such as OXCs, SDH cross-connects and/or routers, are generally meant for control and administrative traffic. These control channels are advertised into routing as normal links as mentioned in the previous section; this allows the routing of (for example) RSVP messages and telnet sessions. However, if routers on the edge of the optical domain attempt to forward data traffic over these channels, the channel capacity will quickly be exhausted.

In order to keep these control channels from being advertised into the user data plane a variety of techniques can be used.

If one assumes that data traffic is sent to BGP destinations, and control traffic to IGP destinations, then one can exclude data traffic from the control plane by restricting BGP nexthop resolution. (It is assumed that OXCs are not BGP speakers.) Suppose that a router R is attempting to install a route to a BGP destination D. R looks up the BGP nexthop for D in its IGP's routing table. Say R finds that the path to the nexthop is over interface I. R then checks if it has an entry in its Link State database associated with the interface I. If it does, and the link is not packet-switch capable (see [LSP-HIER]), R installs a discard route for destination D. Otherwise, R installs (as usual) a route for destination D with nexthop I. Note that R need only do this check if it has packet-switch incapable links; if all of its links are packet-switch capable, then clearly this check is redundant.

In other instances it may be desirable to keep the whole address space of a GMPLS routing plane disjoint from the endpoint addresses in another portion of the GMPLS network. For example, the addresses of a carrier network where the carrier uses GMPLS but does not wish to expose the internals of the addressing or topology. In such a network the control channels are never advertised into the end data network. In this instance, independent mechanisms are used to advertise the data addresses over the carrier network.

Other techniques for excluding data traffic from control channels may also be needed.

2. GMPLS Routing Enhancements

In this section we define the enhancements to the TE properties of GMPLS TE links. Encoding of this information in IS-IS is specified in [GMPLS-ISIS]. Encoding of this information in OSPF is specified in [GMPLS-OSPF].

2.1. Support for Unnumbered Links

An unnumbered link has to be a point-to-point link. An LSR at each end of an unnumbered link assigns an identifier to that link. This identifier is a non-zero 32-bit number that is unique within the scope of the LSR that assigns it.

Consider an (unnumbered) link between LSRs A and B. LSR A chooses an idenfitier for that link. So does LSR B. From A’s perspective we refer to the identifier that A assigned to the link as the "link local identifier" (or just "local identifier"), and to the identifier that B assigned to the link as the "link remote identifier" (or just "remote identifier"). Likewise, from B’s perspective the identifier that B assigned to the link is the local identifier, and the identifier that A assigned to the link is the remote identifier.

Support for unnumbered links in routing includes carrying information about the identifiers of that link. Specifically, when an LSR advertises an unnumbered TE link, the advertisement carries both the local and the remote identifiers of the link. If the LSR doesn’t know the remote identifier of that link, the LSR should use a value of 0 as the remote identifier.

2.2. Link Protection Type

The Link Protection Type represents the protection capability that exists for a link. It is desirable to carry this information so that it may be used by the path computation algorithm to set up LSPs with appropriate protection characteristics. This information is

organized in a hierarchy where typically the minimum acceptable protection is specified at path instantiation and a path selection technique is used to find a path that satisfies at least the minimum acceptable protection. Protection schemes are presented in order from lowest to highest protection.

This document defines the following protection capabilities:

Extra Traffic
If the link is of type Extra Traffic, it means that the link is protecting another link or links. The LSPs on a link of this type will be lost if any of the links it is protecting fail.

Unprotected
If the link is of type Unprotected, it means that there is no other link protecting this link. The LSPs on a link of this type will be lost if the link fails.

Shared
If the link is of type Shared, it means that there are one or more disjoint links of type Extra Traffic that are protecting this link. These Extra Traffic links are shared between one or more links of type Shared.

Dedicated 1:1
If the link is of type Dedicated 1:1, it means that there is one dedicated disjoint link of type Extra Traffic that is protecting this link.

Dedicated 1+1
If the link is of type Dedicated 1+1, it means that a dedicated disjoint link is protecting this link. However, the protecting link is not advertised in the link state database and is therefore not available for the routing of LSPs.

Enhanced
If the link is of type Enhanced, it means that a protection scheme that is more reliable than Dedicated 1+1, e.g., 4 fiber BLSR/MS-SPRING, is being used to protect this link.

The Link Protection Type is optional, and if a Link State Advertisement doesn't carry this information, then the Link Protection Type is unknown.

2.3. Shared Risk Link Group Information

A set of links may constitute a 'shared risk link group' (SRLG) if they share a resource whose failure may affect all links in the set. For example, two fibers in the same conduit would be in the same SRLG. A link may belong to multiple SRLGs. Thus the SRLG Information describes a list of SRLGs that the link belongs to. An SRLG is identified by a 32 bit number that is unique within an IGP domain. The SRLG Information is an unordered list of SRLGs that the link belongs to.

The SRLG of a LSP is the union of the SRLGs of the links in the LSP. The SRLG of a bundled link is the union of the SRLGs of all the component links.

If an LSR is required to have multiple diversely routed LSPs to another LSR, the path computation should attempt to route the paths so that they do not have any links in common, and such that the path SRLGs are disjoint.

The SRLG Information may start with a configured value, in which case it does not change over time, unless reconfigured.

The SRLG Information is optional and if a Link State Advertisement doesn't carry the SRLG Information, then it means that SRLG of that link is unknown.

2.4. Interface Switching Capability Descriptor

In the context of this document we say that a link is connected to a node by an interface. In the context of GMPLS interfaces may have different switching capabilities. For example an interface that connects a given link to a node may not be able to switch individual packets, but it may be able to switch channels within an SDH payload. Interfaces at each end of a link need not have the same switching capabilities. Interfaces on the same node need not have the same switching capabilities.

The Interface Switching Capability Descriptor describes switching capability of an interface. For bi-directional links, the switching capabilities of an interface are defined to be the same in either direction. I.e., for data entering the node through that interface and for data leaving the node through that interface.

A Link State Advertisement of a link carries the Interface Switching Capability Descriptor(s) only of the near end (the end incumbent on the LSR originating the advertisement).

An LSR performing path computation uses the Link State Database to determine whether a link is unidirectional or bidirectional.

For a bidirectional link the LSR uses its Link State Database to determine the Interface Switching Capability Descriptor(s) of the far-end of the link, as bidirectional links with different Interface Switching Capabilities at its two ends are allowed.

For a unidirectional link it is assumed that the Interface Switching Capability Descriptor at the far-end of the link is the same as at the near-end. Thus, an unidirectional link is required to have the same interface switching capabilities at both ends. This seems a reasonable assumption given that unidirectional links arise only with packet forwarding adjacencies and for these both ends belong to the same level of the PSC hierarchy.

This document defines the following Interface Switching Capabilities:

Packet-Switch Capable-1	(PSC-1)
Packet-Switch Capable-2	(PSC-2)
Packet-Switch Capable-3	(PSC-3)
Packet-Switch Capable-4	(PSC-4)
Layer-2 Switch Capable	(L2SC)
Time-Division-Multiplex Capable	(TDM)
Lambda-Switch Capable	(LSC)
Fiber-Switch Capable	(FSC)

If there is no Interface Switching Capability Descriptor for an interface, the interface is assumed to be packet-switch capable (PSC-1).

Interface Switching Capability Descriptors present a new constraint for LSP path computation.

Irrespective of a particular Interface Switching Capability, the Interface Switching Capability Descriptor always includes information about the encoding supported by an interface. The defined encodings are the same as LSP Encoding as defined in [GMPLS-SIG].

An interface may have more than one Interface Switching Capability Descriptor. This is used to handle interfaces that support multiple switching capabilities, for interfaces that have Max LSP Bandwidth values that differ by priority level, and for interfaces that support discrete bandwidths.

Depending on a particular Interface Switching Capability, the Interface Switching Capability Descriptor may include additional information, as specified below.

2.4.1. Layer-2 Switch Capable

If an interface is of type L2SC, it means that the node receiving data over this interface can switch the received frames based on the layer 2 address. For example, an interface associated with a link terminating on an ATM switch would be considered L2SC.

2.4.2. Packet-Switch Capable

If an interface is of type PSC-1 through PSC-4, it means that the node receiving data over this interface can switch the received data on a packet-by-packet basis, based on the label carried in the "shim" header [RFC3032]. The various levels of PSC establish a hierarchy of LSPs tunneled within LSPs.

For Packet-Switch Capable interfaces the additional information includes Maximum LSP Bandwidth, Minimum LSP Bandwidth, and interface MTU.

For a simple (unbundled) link, the Maximum LSP Bandwidth at priority p is defined to be the smaller of the unreserved bandwidth at priority p and a "Maximum LSP Size" parameter which is locally configured on the link, and whose default value is equal to the Max Link Bandwidth. Maximum LSP Bandwidth for a bundled link is defined in [LINK-BUNDLE].

The Maximum LSP Bandwidth takes the place of the Maximum Link Bandwidth ([ISIS-TE], [OSPF-TE]). However, while Maximum Link Bandwidth is a single fixed value (usually simply the link capacity), Maximum LSP Bandwidth is carried per priority, and may vary as LSPs are set up and torn down.

Although Maximum Link Bandwidth is to be deprecated, for backward compatibility, one MAY set the Maximum Link Bandwidth to the Maximum LSP Bandwidth at priority 7.

The Minimum LSP Bandwidth specifies the minimum bandwidth an LSP could reserve.

Typical values for the Minimum LSP Bandwidth and for the Maximum LSP Bandwidth are enumerated in [GMPLS-SIG].

On a PSC interface that supports Standard SDH encoding, an LSP at priority p could reserve any bandwidth allowed by the branch of the SDH hierarchy, with the leaf and the root of the branch being defined by the Minimum LSP Bandwidth and the Maximum LSP Bandwidth at priority p.

On a PSC interface that supports Arbitrary SDH encoding, an LSP at priority p could reserve any bandwidth between the Minimum LSP Bandwidth and the Maximum LSP Bandwidth at priority p, provided that the bandwidth reserved by the LSP is a multiple of the Minimum LSP Bandwidth.

The Interface MTU is the maximum size of a packet that can be transmitted on this interface without being fragmented.

2.4.3. Time-Division Multiplex Capable

If an interface is of type TDM, it means that the node receiving data over this interface can multiplex or demultiplex channels within an SDH payload.

For Time-Division Multiplex Capable interfaces the additional information includes Maximum LSP Bandwidth, the information on whether the interface supports Standard or Arbitrary SDH, and Minimum LSP Bandwidth.

For a simple (unbundled) link the Maximum LSP Bandwidth at priority p is defined as the maximum bandwidth an LSP at priority p could reserve. Maximum LSP Bandwidth for a bundled link is defined in [LINK-BUNDLE].

The Minimum LSP Bandwidth specifies the minimum bandwidth an LSP could reserve.

Typical values for the Minimum LSP Bandwidth and for the Maximum LSP Bandwidth are enumerated in [GMPLS-SIG].

On an interface having Standard SDH multiplexing, an LSP at priority p could reserve any bandwidth allowed by the branch of the SDH hierarchy, with the leaf and the root of the branch being defined by the Minimum LSP Bandwidth and the Maximum LSP Bandwidth at priority p.

On an interface having Arbitrary SDH multiplexing, an LSP at priority p could reserve any bandwidth between the Minimum LSP Bandwidth and the Maximum LSP Bandwidth at priority p, provided that the bandwidth reserved by the LSP is a multiple of the Minimum LSP Bandwidth.

Interface Switching Capability Descriptor for the interfaces that support sub VC-3 may include additional information. The nature and the encoding of such information is outside the scope of this document.

A way to handle the case where an interface supports multiple branches of the SDH multiplexing hierarchy, multiple Interface Switching Capability Descriptors would be advertised, one per branch. For example, if an interface supports VC-11 and VC-12 (which are not part of same branch of SDH multiplexing tree), then it could advertise two descriptors, one for each one.

2.4.4. Lambda-Switch Capable

If an interface is of type LSC, it means that the node receiving data over this interface can recognize and switch individual lambdas within the interface. An interface that allows only one lambda per interface, and switches just that lambda is of type LSC.

The additional information includes Reservable Bandwidth per priority, which specifies the bandwidth of an LSP that could be supported by the interface at a given priority number.

A way to handle the case of multiple data rates or multiple encodings within a single TE Link, multiple Interface Switching Capability Descriptors would be advertised, one per supported data rate and encoding combination. For example, an LSC interface could support the establishment of LSC LSPs at both STM-16 and STM-64 data rates.

2.4.5. Fiber-Switch Capable

If an interface is of type FSC, it means that the node receiving data over this interface can switch the entire contents to another interface (without distinguishing lambdas, channels or packets). I.e., an interface of type FSC switches at the granularity of an entire interface, and can not extract individual lambdas within the interface. An interface of type FSC can not restrict itself to just one lambda.

2.4.6. Multiple Switching Capabilities per Interface

An interface that connects a link to an LSR may support not one, but several Interface Switching Capabilities. For example, consider a fiber link carrying a set of lambdas that terminates on an LSR interface that could either cross-connect one of these lambdas to some other outgoing optical channel, or could terminate the lambda, and extract (demultiplex) data from that lambda using TDM, and then cross-connect these TDM channels to some outgoing TDM channels. To support this a Link State Advertisement may carry a list of Interface Switching Capabilities Descriptors.

2.4.7. Interface Switching Capabilities and Labels

Depicting a TE link as a tuple that contains Interface Switching Capabilities at both ends of the link, some examples links may be:

[PSC, PSC] - a link between two packet LSRs
[TDM, TDM] - a link between two Digital Cross Connects
[LSC, LSC] - a link between two OXCs
[PSC, TDM] - a link between a packet LSR and Digital Cross Connect
[PSC, LSC] - a link between a packet LSR and an OXC
[TDM, LSC] - a link between a Digital Cross Connect and an OXC

Both ends of a given TE link has to use the same way of carrying label information over that link. Carrying label information on a given TE link depends on the Interface Switching Capability at both ends of the link, and is determined as follows:

[PSC, PSC] - label is carried in the "shim" header [RFC3032]
[TDM, TDM] - label represents a TDM time slot [GMPLS-SONET-SDH]
[LSC, LSC] - label represents a lambda
[FSC, FSC] - label represents a port on an OXC
[PSC, TDM] - label represents a TDM time slot [GMPLS-SONET-SDH]
[PSC, LSC] - label represents a lambda
[PSC, FSC] - label represents a port
[TDM, LSC] - label represents a lambda
[TDM, FSC] - label represents a port
[LSC, FSC] - label represents a port

2.4.8. Other Issues

It is possible that Interface Switching Capability Descriptor will change over time, reflecting the allocation/deallocation of LSPs. For example, assume that VC-3, VC-4, VC-4-4c, VC-4-16c and VC-4-64c LSPs can be established on a STM-64 interface whose Encoding Type is SDH. Thus, initially in the Interface Switching Capability Descriptor the Minimum LSP Bandwidth is set to VC-3, and Maximum LSP Bandwidth is set to STM-64 for all priorities. As soon as an LSP of VC-3 size at priority 1 is established on the interface, it is no longer capable of VC-4-64c for all but LSPs at priority 0. Therefore, the node advertises a modified Interface Switching Capability Descriptor indicating that the Maximum LSP Bandwidth is no longer STM-64, but STM-16 for all but priority 0 (at priority 0 the Maximum LSP Bandwidth is still STM-64). If subsequently there is another VC-3 LSP, there is no change in the Interface Switching Capability Descriptor. The Descriptor remains the same until the node can no longer establish a VC-4-16c LSP over the interface (which

means that at this point more than 144 time slots are taken by LSPs on the interface). Once this happened, the Descriptor is modified again, and the modified Descriptor is advertised to other nodes.

2.5. Bandwidth Encoding

Encoding in IEEE floating point format [IEEE] of the discrete values that could be used to identify Unreserved bandwidth, Maximum LSP bandwidth and Minimum LSP bandwidth is described in Section 3.1.2 of [GMPLS-SIG].

3. Examples of Interface Switching Capability Descriptor

3.1. STM-16 POS Interface on a LSR

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = PSC-1
   Encoding = SDH
   Max LSP Bandwidth[p] = 2.5 Gbps, for all p
```

If multiple links with such interfaces at both ends were to be advertised as one TE link, link bundling techniques should be used.

3.2. GigE Packet Interface on a LSR

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = PSC-1
   Encoding = Ethernet 802.3
   Max LSP Bandwidth[p] = 1.0 Gbps, for all p
```

If multiple links with such interfaces at both ends were to be advertised as one TE link, link bundling techniques should be used.

3.3. STM-64 SDH Interface on a Digital Cross Connect with Standard SDH

Consider a branch of SDH multiplexing tree : VC-3, VC-4, VC-4-4c, VC-4-16c, VC-4-64c. If it is possible to establish all these connections on a STM-64 interface, the Interface Switching Capability Descriptor of that interface can be advertised as follows:

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = TDM [Standard SDH]
   Encoding = SDH
   Min LSP Bandwidth = VC-3
   Max LSP Bandwidth[p] = STM-64, for all p
```

If multiple links with such interfaces at both ends were to be advertised as one TE link, link bundling techniques should be used.

3.4. STM-64 SDH Interface on a Digital Cross Connect with Two Types of SDH Multiplexing Hierarchy Supported

```
Interface Switching Capability Descriptor 1:
   Interface Switching Capability = TDM [Standard SDH]
   Encoding = SDH
   Min LSP Bandwidth = VC-3
   Max LSP Bandwidth[p] = STM-64, for all p

Interface Switching Capability Descriptor 2:
   Interface Switching Capability = TDM [Arbitrary SDH]
   Encoding = SDH
   Min LSP Bandwidth = VC-4
   Max LSP Bandwidth[p] = STM-64, for all p
```

If multiple links with such interfaces at both ends were to be advertised as one TE link, link bundling techniques should be used.

3.5. Interface on an Opaque OXC (SDH Framed) with Support for One Lambda per Port/Interface

An "opaque OXC" is considered operationally an OXC, as the whole lambda (carrying the SDH line) is switched transparently without further multiplexing/demultiplexing, and either none of the SDH overhead bytes, or at least the important ones are not changed.

An interface on an opaque OXC handles a single wavelength, and cannot switch multiple wavelengths as a whole. Thus, an interface on an opaque OXC is always LSC, and not FSC, irrespective of whether there is DWDM external to it.

Note that if there is external DWDM, then the framing understood by the DWDM must be same as that understood by the OXC.

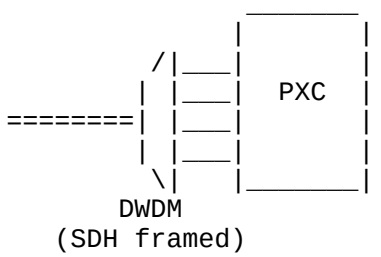
A TE link is a group of one or more interfaces on an OXC. All interfaces on a given OXC are required to have identifiers unique to that OXC, and these identifiers are used as labels (see 3.2.1.1 of [GMPLS-SIG]).

The following is an example of an interface switching capability descriptor on an SDH framed opaque OXC:

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = LSC
   Encoding = SDH
   Reservable Bandwidth = Determined by SDH Framer (say STM-64)
```

3.6. Interface on a Transparent OXC (PXC) with External DWDM That Understands SDH Framing

This example assumes that DWDM and PXC are connected in such a way that each interface (port) on the PXC handles just a single wavelength. Thus, even if in principle an interface on the PXC could switch multiple wavelengths as a whole, in this particular case an interface on the PXC is considered LSC, and not FSC.



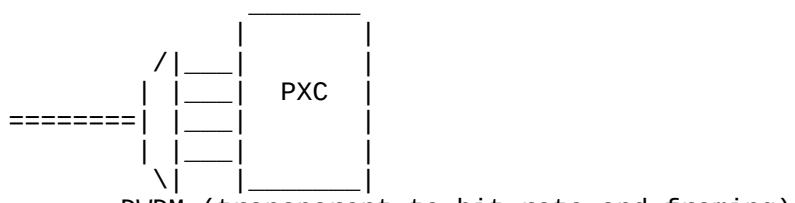
A TE link is a group of one or more interfaces on the PXC. All interfaces on a given PXC are required to have identifiers unique to that PXC, and these identifiers are used as labels (see 3.2.1.1 of [GMPLS-SIG]).

The following is an example of an interface switching capability descriptor on a transparent OXC (PXC) with external DWDM that understands SDH framing:

Interface Switching Capability Descriptor:
Interface Switching Capability = LSC
Encoding = SDH (comes from DWDM)
Reservable Bandwidth = Determined by DWDM (say STM-64)

3.7. Interface on a Transparent OXC (PXC) with External DWDM That Is Transparent to Bit-Rate and Framing

This example assumes that DWDM and PXC are connected in such a way that each interface (port) on the PXC handles just a single wavelength. Thus, even if in principle an interface on the PXC could switch multiple wavelengths as a whole, in this particular case an interface on the PXC is considered LSC, and not FSC.



DWDM (transparent to bit-rate and framing)

A TE link is a group of one or more interfaces on the PXC. All interfaces on a given PXC are required to have identifiers unique to that PXC, and these identifiers are used as labels (see 3.2.1.1 of [GMPLS-SIG]).

The following is an example of an interface switching capability descriptor on a transparent OXC (PXC) with external DWDM that is transparent to bit-rate and framing:

Interface Switching Capability Descriptor:
Interface Switching Capability = LSC
Encoding = Lambda (photonic)
Reservable Bandwidth = Determined by optical technology limits

3.8. Interface on a PXC with No External DWDM

The absence of DWDM in between two PXCs, implies that an interface is not limited to one wavelength. Thus, the interface is advertised as FSC.

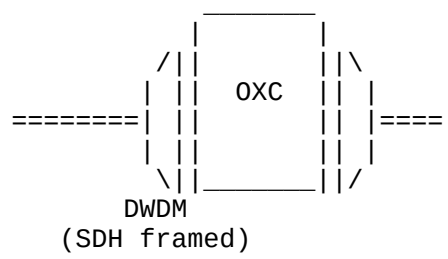
A TE link is a group of one or more interfaces on the PXC. All interfaces on a given PXC are required to have identifiers unique to that PXC, and these identifiers are used as port labels (see 3.2.1.1 of [GMPLS-SIG]).

Interface Switching Capability Descriptor:
Interface Switching Capability = FSC
Encoding = Lambda (photonic)
Reservable Bandwidth = Determined by optical technology limits

Note that this example assumes that the PXC does not restrict each port to carry only one wavelength.

3.9. Interface on a OXC with Internal DWDM That Understands SDH Framing

This example assumes that DWDM and OXC are connected in such a way that each interface on the OXC handles multiple wavelengths individually. In this case an interface on the OXC is considered LSC, and not FSC.



A TE link is a group of one or more of the interfaces on the OXC. All lambdas associated with a particular interface are required to have identifiers unique to that interface, and these identifiers are used as labels (see 3.2.1.1 of [GMPLS-SIG]).

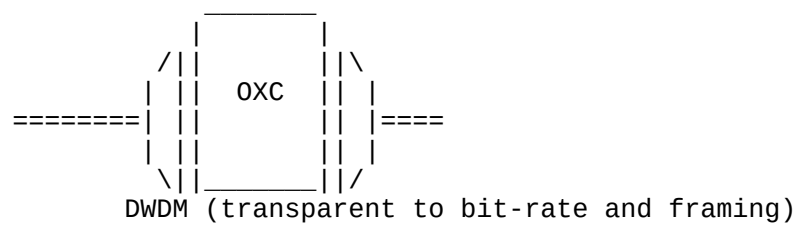
The following is an example of an interface switching capability descriptor on an OXC with internal DWDM that understands SDH framing and supports discrete bandwidths:

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = LSC
   Encoding = SDH (comes from DWDM)
   Max LSP Bandwidth = Determined by DWDM (say STM-16)

   Interface Switching Capability = LSC
   Encoding = SDH (comes from DWDM)
   Max LSP Bandwidth = Determined by DWDM (say STM-64)
```

3.10. Interface on a OXC with Internal DWDM That Is Transparent to Bit-Rate and Framing

This example assumes that DWDM and OXC are connected in such a way that each interface on the OXC handles multiple wavelengths individually. In this case an interface on the OXC is considered LSC, and not FSC.



A TE link is a group of one or more of the interfaces on the OXC. All lambdas associated with a particular interface are required to have identifiers unique to that interface, and these identifiers are used as labels (see 3.2.1.1 of [GMPLS-SIG]).

The following is an example of an interface switching capability descriptor on an OXC with internal DWDM that is transparent to bit-rate and framing:

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = LSC
   Encoding = Lambda (photonic)
   Max LSP Bandwidth = Determined by optical technology limits
```

4. Example of Interfaces That Support Multiple Switching Capabilities

There can be many combinations possible, some are described below.

4.1. Interface on a PXC+TDM Device with External DWDM

As discussed earlier, the presence of the external DWDM limits that only one wavelength be on a port of the PXC. On such a port, the attached PXC+TDM device can do one of the following. The wavelength may be cross-connected by the PXC element to other out-bound optical channel, or the wavelength may be terminated as an SDH interface and SDH channels switched.

From a GMPLS perspective the PXC+TDM functionality is treated as a single interface. The interface is described using two Interface descriptors, one for the LSC and another for the TDM, with appropriate parameters. For example,

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = LSC
   Encoding = SDH (comes from WDM)
   Reservable Bandwidth = STM-64

and

Interface Switching Capability Descriptor:
   Interface Switching Capability = TDM [Standard SDH]
   Encoding = SDH
   Min LSP Bandwidth = VC-3
   Max LSP Bandwidth[p] = STM-64, for all p
```

4.2. Interface on an Opaque OXC+TDM Device with External DWDM

An interface on an "opaque OXC+TDM" device would also be advertised as LSC+TDM much the same way as the previous case.

4.3. Interface on a PXC+LSR Device with External DWDM

As discussed earlier, the presence of the external DWDM limits that only one wavelength be on a port of the PXC. On such a port, the attached PXC+LSR device can do one of the following. The wavelength may be cross-connected by the PXC element to other out-bound optical channel, or the wavelength may be terminated as a Packet interface and packets switched.

From a GMPLS perspective the PXC+LSR functionality is treated as a single interface. The interface is described using two Interface descriptors, one for the LSC and another for the PSC, with appropriate parameters. For example,

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = LSC
   Encoding = SDH (comes from WDM)
   Reservable Bandwidth = STM-64

and

Interface Switching Capability Descriptor:
   Interface Switching Capability = PSC-1
   Encoding = SDH
   Max LSP Bandwidth[p] = 10 Gbps, for all p
```

4.4. Interface on a TDM+LSR Device

On a TDM+LSR device that offers a channelized SDH interface the following may be possible:

- A subset of the SDH channels may be uncommitted. That is, they are not currently in use and hence are available for allocation.

- A second subset of channels may already be committed for transit purposes. That is, they are already cross-connected by the SDH cross connect function to other out-bound channels and thus are not immediately available for allocation.

- Another subset of channels could be in use as terminal channels. That is, they are already allocated by terminate on a packet interface and packets switched.

From a GMPLS perspective the TDM+PSC functionality is treated as a single interface. The interface is described using two Interface descriptors, one for the TDM and another for the PSC, with appropriate parameters. For example,

```
Interface Switching Capability Descriptor:
   Interface Switching Capability = TDM [Standard SDH]
   Encoding = SDH
   Min LSP Bandwidth = VC-3
   Max LSP Bandwidth[p] = STM-64, for all p

and

Interface Switching Capability Descriptor:
   Interface Switching Capability = PSC-1
   Encoding = SDH
   Max LSP Bandwidth[p] = 10 Gbps, for all p
```

5. Acknowledgements

The authors would like to thank Suresh Katukam, Jonathan Lang, Zhi-Wei Lin, and Quaizar Vohra for their comments and contributions to the document. Thanks too to Stephen Shew for the text regarding "Representing TE Link Capabilities".

6. Security Considerations

There are a number of security concerns in implementing the extensions proposed here, particularly since these extensions will potentially be used to control the underlying transport infrastructure. It is vital that there be secure and/or authenticated means of transferring this information among the entities that require its use.

While this document proposes extensions, it does not state how these extensions are implemented in routing protocols such as OSPF or IS-IS. The documents that do state how routing protocols implement these extensions [GMPLS-OSPF, GMPLS-ISIS] must also state how the information is to be secured.

7. References

7.1. Normative References

[GMPLS-OSPF] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[GMPLS-SIG] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[GMPLS-SONET-SDH] Mannie, E. and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 3946, October 2004.

[IEEE] IEEE, "IEEE Standard for Binary Floating-Point Arithmetic", Standard 754-1985, 1985 (ISBN 1-5593-7653-8).

[LINK-BUNDLE] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.

[LMP] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204, October 2005.

[LSP-HIER] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE))", RFC 4206, October 2005.

[OSPF-TE] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

7.2. Informative References

[GMPLS-ISIS] Kompella, K., Ed. and Y. Rekhter, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4205, October 2005.

[ISIS-TE] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

8. Contributors

Ayan Banerjee
Calient Networks
5853 Rue Ferrari
San Jose, CA 95138

Phone: +1.408.972.3645
EMail: abanerjee@calient.net

John Drake
Calient Networks
5853 Rue Ferrari
San Jose, CA 95138

Phone: (408) 972-3720
EMail: jdrake@calient.net

Greg Bernstein
Ciena Corporation
10480 Ridgeview Court
Cupertino, CA 94014

Phone: (408) 366-4713
EMail: greg@ciena.com

Don Fedyk
Nortel Networks Corp.
600 Technology Park Drive
Billerica, MA 01821

Phone: +1-978-288-4506
EMail: dwfedyk@nortelnetworks.com

Eric Mannie
Libre Exaministe

EMail: eric_mannie@hotmail.com

Debanjan Saha
Tellium Optical Systems
2 Crescent Place
P.O. Box 901
Ocean Port, NJ 07757

Phone: (732) 923-4264
EMail: dsaha@tellium.com

Vishal Sharma
Metanoia, Inc.
335 Elan Village Lane, Unit 203
San Jose, CA 95134-2539

Phone: +1 408-943-1794
EMail: v.sharma@ieee.org

Debashis Basak
AcceLight Networks,
70 Abele Rd, Bldg 1200
Bridgeville PA 15017

EMail: dbasak@accelight.com

Lou Berger
Movaz Networks, Inc.
7926 Jones Branch Drive
Suite 615
McLean VA, 22102

EMail: lberger@movaz.com

Authors' Addresses

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave
Sunnyvale, CA 94089

EMail: kireeti@juniper.net

Yakov Rekhter
Juniper Networks, Inc.
1194 N. Mathilda Ave
Sunnyvale, CA 94089

EMail: yakov@juniper.net

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.