Network Working Group
Internet Draft
Expiration Date: August 2001

Kireeti Kompella
Juniper Networks
Yakov Rekhter
Juniper Networks
Lou Berger
Movaz Networks

Link Bundling in MPLS Traffic Engineering

draft-kompella-mpls-bundle-05.txt

1. Status of this Memo

2. Abstract

In some cases a pair of Label Switching Routers (LSRs) may be connected by several (parallel) links. From the MPLS Traffic Engineering point of view for reasons of scalability it may be desirable to advertise all these links as a single link into OSPF and/or IS-IS. This document describes a mechanism to accomplish this. This document also defines corresponding signaling (RSVP-TE and CR-LDP) support.

3. Link Bundling

When a pair of LSRs are connected by multiple links, then for the purpose of MPLS Traffic Engineering it is possible to advertise several (or all) of these links as a single link into OSPF and/or IS-IS. We refer to this process as "link bundling", or just "bundling". We refer to the link that is advertised into OSPF/IS-IS as a "bundled link". We refer to the links associated with that bundled link as "component links".

The purpose of link bundling is to improve routing scalability by reducing the amount of information that has to be handled by OSPF and/or IS-IS. This reduction is accomplished by performing information aggregation/abstraction. As with any other information aggregation/abstraction, this results in losing some of the information. To limit the amount of losses one need to restrict the type of the information that can be aggregated/abstracted.

3.1. Restrictions on Bundling

All component links in a bundle must begin and end on the same pair of LSRs, have the same Link Type (i.e., point-to-point or multi-access), the same Traffic Engineering metric, the same set of resource classes, and the same Link Multiplex Capability (see [LSP-HIER]). A Forwarding Adjacency may be a component link; in fact, a bundle can consist of a mix of point-to-point links and FAs.

If the component links are all multi-access links, the set of IS-IS or OSPF routers connected to each component link must be the same, and the Designated Router for each component link must be the same. If these conditions cannot be enforced, multi-access links must not be bundled.

3.2. Routing Considerations

A bundled link is just another kind of Traffic Engineering (TE) link (see [GMPLS-ISIS] and [GMPLS-OSPF]). The "liveness" of the bundled link is determined by the liveness of each of the component links within the bundled link. The liveness of a component link can be determined by any of several means: IS-IS or OSPF hellos over the component link, or RSVP Hello, or LMP hellos (see [LMP]), or from layer 1 or layer 2 indications.

Once a bundled link is determined to be alive, it can be advertised as a TE link and the TE information can be flooded. If IS-IS/OSPF hellos are run over the component links, IS-IS/OSPF flooding can be

restricted to just one of the component links [ZININ] [MOY].

Note that advertising a (bundled) TE link between a pair of LSRs doesn't imply that there is an IGP adjacency between these LSRs that is associated with just that link. In fact, in certain cases a TE link between a pair of LSRs could be advertised even if there is no IGP adjacency at all between the LSR (e.g., when the TE link is an FA).

In the future, as new Traffic Engineering parameters are added to IS-IS and OSPF, they should be accompanied by descriptions as to how they can be bundled, and possible restrictions on bundling.

3.3. Signaling Considerations

Typically, an LSP's ERO will choose the bundled link to be used for the LSP, but not the component link(s), since information about the bundled link is flooded, but information about the component links is kept local to the LSR. If the ERO chooses the component links by means outside the scope of this document, neither this section nor section 5.2 applies. Otherwise, the choice of the component link(s) for the LSP is a local matter between the two LSRs at each end of the bundled link.

The choice of the component link to use is always made by the sender of the Path/REQUEST message. Three mechanisms for indicating this choice to the receiver of the Path/REQUEST message are discussed below; which of these mechanisms is used SHOULD be configurable by the user, preferably on a per-bundle basis.

3.3.1. Mechanism 1: Implicit Indication

This mechanism requires that each component link has a dedicated signaling channel (for example, the link is packet-switch capable; or the link is a SONET link with an in-band channel for signaling). The sender of the Path/REQUEST message tells the receiver which component link to use by sending the message over the chosen component link's dedicated signaling channel.

3.3.2. Mechanism 2: Explicit Indication by IP Address

This mechanism requires that each component link has a unique remote IP address. The sender can either send the Path/REQUEST message addressed to the remote IP address for the component link or encapsulate the message in an IP header whose destination address is

the remote IP address. This mechanism does not require each component link to have its own control channel. In fact, it doesn't even require the whole (bundled) link to have its own control channel.

3.3.3. Mechanism 3: Explicit Indication by Component Interface ID

This mechanism requires that each component link is assigned a unique Interface Identifier per [UNNUM-RSVP] or [UNNUM-CRLDP] and that the assigned identifiers be exchanged by the two LSRs at each end of the bundled link. This identifier is referred to as "component interface identifier". The choice of the component link is indicated by the sender of the Path/REQUEST message by including the component link's interface identifier in the message, as described in section 5.2.

3.4. Unnumbered Bundled Links

Note that a bundled link may itself be numbered or unnumbered independent of whether the component links are numbered or not. This affects how the bundled link is advertised in IS-IS/OSPF, and the format of LSP EROs that traverse the bundled link. Furthermore, unnumbered Interface Identifiers for all unnumbered outgoing links of a given LSR (whether component links, Forwarding Adjacencies or bundled links) MUST be unique in the context of that LSR.

4. Traffic Engineering Parameters for Bundled Links

In this section, we define the Traffic Engineering parameters to be advertised for a bundled link, based on the configuration of the component links and of the bundled link. The definition of these parameters for component links was undertaken in [ISIS] and [OSPF]; we use the terminology from [OSPF].

4.1. OSPF Link Type

The Link Type of a bundled link is the (unique) Link Type of the component links. (Note: this parameter is not present in IS-IS.)

4.2. OSPF Link ID

For point-to-point links, the Link ID of a bundled link is the (unique) Router ID of the neighbor. For multi-access links, this is the interface address of the (unique) Designated Router. (Note: this parameter is not present in IS-IS.)

4.3. Local and Remote Interface IP Address

(Note: in IS-IS, these are known as IPv4 Interface Address and IPv4 Neighbor Address, respectively.)

If the bundled link is numbered, the Local Interface IP Address is the local address of the bundled link; similarly, the Remote Interface IP Address is the remote address of the bundled link.

4.4. Outgoing and Incoming Interface Identifiers

If the bundled link is unnumbered, the Outgoing Interface Identifier is set to the outgoing interface identifier chosen for the bundle by the advertising LSR. The Incoming Interface Identifier is set to the outgoing interface identifier chosen by the neighboring LSR for the reverse link corresponding to this bundle, if known; otherwise, this is set to 0.

4.5. Traffic Engineering Metric

The Traffic Engineering Metric for a bundled link is that of the component links.

4.6. Maximum Link Bandwidth

This TLV is not used. The maximum LSP Bandwidth (as described below) replaces the maximum link bandwidth for bundled links. For backward compatibility, one MAY advertise the Maximum LSP Bandwidth at priority 7 of the bundle as the Maximum Link Bandwidth.

4.7. Maximum Reservable Bandwidth

We assume that for a given bundled link either each of its component links is configured with the maximum reservable bandwidth, or the bundled link is configured with the maximum reservable bandwidth. In the former case, the Maximum Reservable Bandwidth of the bundled link

is set to the sum of the maximum reservable bandwidths of all component links associated with the bundled link.

4.8. Unreserved Bandwidth

The unreserved bandwidth of a bundled link at priority p is the sum of the unreserved bandwidths at priority p of all the component links associated with the bundled link.

4.9. Resource Classes (Administrative Groups)

The Resource Classes for a bundled link are the same as those of the component links.

4.10. Maximum LSP Bandwidth

The Maximum LSP Bandwidth takes the place of the Maximum Link Bandwidth. 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.

The Maximum LSP Bandwidth of a bundled link at priority p is defined to be the maximum of the Maximum LSP Bandwidth at priority p of each component link.

If a component link is a simple (unbundled) link, define its Maximum LSP Bandwidth at priority p to be the smaller of its unreserved bandwidth at priority p and its maximum link bandwidth.

Since bundling may be applied recursively, a component link may itself be a bundled link. In this case, its Maximum LSP Bandwidth as a component link is the same as its Maximum LSP Bandwidth as a bundled link.

The details of how Maximum LSP Bandwidth is carried in IS-IS is given in [GMPLS-ISIS]. The details of how Maximum LSP Bandwidth is carried in OSPF is given in [GMPLS-OSPF].

5. Procedures

5.1. Bandwidth Accounting

The RSVP (or CR-LDP) Traffic Control module, or its equivalent, on an LSR with bundled links must apply admission control on a per-component link basis. An LSP with a bandwidth requirement b and setup priority p fits in a bundled link if at least one component link has maximum LSP bandwidth >= b at priority p. If there are several such links, the choice of which link is used for the LSP is up to the implementation.

In order to know the maximum LSP bandwidth (per priority) of each component link, the Traffic Control module must track the unreserved bandwidth (per priority) for each component link. This is done as follows. If an LSP with bandwidth b and holding priority p is set up through a component link, that component link's unreserved bandwidth at priority p and lower is reduced by b. If an LSP with bandwidth b and holding priority p that is currently set up through a component link is torn down, the unreserved bandwidth at priority p and lower for that component link is increased by b.

A change in the unreserved bandwidth of a component link results in a change in the unreserved bandwidth of the bundled link. It also potentially results in a change in the maximum LSP bandwidth of the bundle; thus, the maximum LSP bandwidth should be recomputed.

If one of the component links goes down, the associated bundled link remains up and continues to be advertised, provided that at least one component link associated with the bundled link is up. The unreserved bandwidth of the component link that is down is set to zero, and the unreserved bandwidth and maximum LSP bandwidth of the bundle must be recomputed. If all the component links associated with a given bundled link are down, the bundled link MUST not be advertised into OSPF/IS-IS.

5.2. Signaling

Signaling must identify both the component link to use and the label to use. The sender of the Path/REQUEST message always chooses the component link(s) to be used for the LSP (if the LSP is bidirectional [GMPLS-SIG], the sender chooses a component link in each direction).

For unidirectional LSPs and the forward direction of a bidirectional LSP, the sender of a MAPPING/Resv message chooses the label (if needed). For the reverse direction of a bidirectional LSP, the

sender of the Path/REQUEST message selects the upstream label (if needed).

As mentioned above, there are three methods for communicating the selected component link, implicit indication and explicit indication by IP address and by component interface identifier. The first two are described in sections 3.3.1 and 3.3.2. In this section, we define the objects needed to indicate the component link by component interface identifier.

In explicit indication by component interface identifier, the sender of the Path (REQUEST) message communicates the selected component link via the COMPONENT_INTERFACE_ID object class (Component Interface ID TLV) defined below. Doing this assumes that an LSR connected to a component link knows the component interface identifier assigned to that link by the LSR at the other end of the link. Exchanging the identity of a component link between the LSRs connected by that link may be accomplished by configuration, by means of a protocol such as [LMP], by means of RSVP/CR-LDP (especially in the case where a component link is a Forwarding Adjacency), or by means of IS-IS or OSPF extensions.

In both RSVP and CR-LDP, if a Component Interface Identifier has the special value of 0xFFFFFFFF, this means that the same label is to be valid across all component links.

5.2.1. RSVP-TE COMPONENT_INTERFACE_ID Object Class

A new object class, the COMPONENT_INTERFACE_ID object class, is defined. The Length field is set to 8. The Class Num (TBD) is of the form 0bbbbbbb. The DOWNSTREAM_COMPONENT_INTERFACE_ID object, which has a C_Type of 1, is used to indicate the component interface to be used for traffic flowing in the downstream direction. The UPSTREAM_COMPONENT_INTERFACE_ID object, which has a C_Type of 2, is used to indicate the component interface to be used for traffic flowing in the upstream direction. Both objects have the same format and carry a 32-bit Component Interface Identifier. The format of the objects are:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Length             |Class Num (TBD)|  C_Type (1|2) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Component Interface Identifier               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

5.2.2. COMPONENT_INTERFACE_ID Object Class Usage

The COMPONENT_INTERFACE_ID objects are carried in RSVP messages as part of the sender descriptor. They are optional with respect to the protocol, and are only used when component links are being identified using the COMPONENT_INTERFACE_ID objects. There are two formats for the sender descriptor, one for traditional LSPs and one for bidirectional LSPs.

The format of the sender descriptor for unidirectional LSPs is:

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

The format of the sender descriptor for bidirectional LSPs is:

```
       <sender descriptor> ::=  <SENDER_TEMPLATE> <SENDER_TSPEC>
                                [ <ADSPEC> ]
                                [ <RECORD_ROUTE> ]
                                <DOWNSTREAM_COMPONENT_INTERFACE_ID>
                                [ <SUGGESTED_LABEL> ]
                                <UPSTREAM_COMPONENT_INTERFACE_ID>
                                <UPSTREAM_LABEL>
```

We introduce a new error value for the error code "Routing problem", namely "Unknown Component Interface ID" with error value 11.

If the receiver doesn't recognize the COMPONENT_INTERFACE_ID object class, per [RSVP], it SHOULD send an error message with an "Unknown Object Class". If the class is recognized but the C-Type is not, per [RSVP], the receiver SHOULD send an "Unknown Object C-Type" error. A node that recognizes either COMPONENT_INTERFACE_ID objects, but that is unable to support it (possibly because of a failure to allocate labels) SHOULD send an error message with the error code "Routing problem" and the error value "MPLS label allocation failure." If LMP or some other link identification protocol is not running, or there is no component link with the Component Interface Identifier in either object, the receiver SHOULD send an error message with the error code "Routing problem" and the error value "Unknown Component Interface ID".

5.2.3. CR-LDP Component Interface ID TLVs

Two new TLVs are introduced to support bundling in CR-LDP. Both TLVs are carried in LDP REQUEST messages. The TLVs share a common format and differ in the direction of the component link being identified. The Downstream Component Interface ID TLV, which has a Type value to be determined by IETF consensus, is used to indicate the component interface to be used for traffic flowing in the downstream direction. The Upstream Component Interface ID TLV, which has a Type value to be determined by IETF consensus, is used to indicate the component interface to be used for traffic flowing in the upstream direction. Both TLVs have the same format and carry a 32-bit Component Interface Identifier. The format of the TLVs are:

```
 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 = (TBD | TBD)      |             Length (8)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Component Interface Identifier               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

We introduce a new status code "Unknown Component Interface ID" with value 0x1A.

If the receiver doesn't recognize either Component Interface ID TLV class, per [LDP], it SHOULD send a Notification message with an "Unknown TLV" Status Code. A node that recognizes either Component Interface ID TLV, but that is unable to support it (possibly because of a failure to allocate labels) SHOULD send a Notification message with a "No Label Resources" Status Code. If LMP or some other link identification protocol is not running, or there is no component link with the Component Interface Identifier in either TLV, the receiver SHOULD send a Notification message with an "Unknown Component Interface ID" Status Code.

6. Security Considerations

This document raises no new security issues for RSVP or CR-LDP.

7. References

[GMPLS-ISIS] Kompella, K., Rekhter, Y., Banerjee, A. et al, "IS-IS Extensions in Support of Generalized MPLS", draft-ietf-isis-gmpls-extensions-02.txt (work in progress)

[GMPLS-OSPF] Kompella, K., Rekhter, Y., Banerjee, A. et al, "OSPF Extensions in Support of Generalized MPLS", draft-kompella-ospf-gmpls-extensions-01.txt (work in progress)

[GMPLS-SIG] Ashwood, P., et al., "Generalized MPLS - Signalling Functional Description", draft-ietf-generalized-mpls-signalling-02.txt

[ISIS] Smit, H., Li, T., "IS-IS extensions for Traffic Engineering", draft-ietf-isis-traffic-02.txt (work in progress)

[LDP] Andersson, L. et al, "LDP Specification", RFC 3036, January 2001.

[LMP] Lang, J., Mitra, K., et al., "Link Management Protocol (LMP)", draft-ietf-mpls-lmp-02.txt (work in progress)

[LSP-HIER] Kompella, K., Rekhter, Y., "LSP Hierarchy with MPLS TE", draft-ietf-mpls-lsp-hierarchy-01.txt (work in progress)

[MOY] Moy, J., draft-ietf-ospf-ppp-flood-00.txt (work in progress)

[OSPF] Katz, D., Yeung, D., "Traffic Engineering Extensions to OSPF", draft-katz-yeung-ospf-traffic-04.txt (work in progress)

[RSVP] Braden, Ed., et. al., "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC2205, September 1997.

[RSVP-TE] Awduche, D., Berger, L., Gan, D., et al, "Extensions to RSVP for LSP Tunnels", draft-ietf-mpls-rsvp-lsp-tunnel-08.txt (work in progress)

[UNNUM-CRLDP] Kompella, K., Rekhter, Y., Kullberg, A., "Signalling Unnumbered Links in CR-LDP", draft-ietf-mpls-crldp-unnum-01.txt (work in progress)

[UNNUM-RSVP] Kompella, K., Rekhter, Y., "Signalling Unnumbered Links

in RSVP-TE", draft-ietf-mpls-rsvp-unnum-01.txt (work in progress)

[ZININ] Zinin, A., Shand, M., "Flooding optimizations in link-state routing protocols", draft-ietf-ospf-isis-flood-opt-00.txt (work in progress)

8. Author Information

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

Lou Berger
Movaz Networks, Inc.
Voice: +1 301 468 9228
Email: lberger@movaz.com