Network Working Group — Jonathan P. Lang (Calient Networks)
Internet Draft — John Drake (Calient Networks)
Category: Informational — Yakov Rekhter (Juniper Networks)
Expiration Date: January 2002 — Adrian Farrel (Movaz Networks)

July 2001

Generalized MPLS Recovery Mechanisms

draft-lang-ccamp-recovery-01.txt

Status of this Memo

Abstract

This draft discusses protection and restoration mechanisms for fault management within the GMPLS framework [GMPLS]. This draft does not propose any new additions to the GMPLS framework.

NAME OF I-D:

www.ietf.org/internet-drafts/draft-lang-ccamp-recovery-00.txt

SUMMARY

This draft discusses protection and restoration mechanisms for fault management within the GMPLS framework [GMPLS].

RELATED DOCUMENTS

www.ietf.org/internet-drafts/draft-many-gmpls-architecture-00.txt
www.ietf.org/internet-drafts/draft-ietf-mpls-generalized-signaling-04.txt
www.ietf.org/internet-drafts/draft-ietf-mpls-generalized-rsvp-te-03.txt
www.ietf.org/internet-drafts/draft-ietf-isis-gmpls-extensions-02.txt
www.ietf.org/internet-drafts/draft-kompella-ospf-gmpls-extensions-01.txt
www.ietf.org/internet-drafts/draft-ietf-mpls-lmp-02.txt
www.ietf.org/internet-drafts/draft-fredette-lmp-wdm-01.txt
www.ietf.org/internet-drafts/draft-kompella-mpls-bundle-05.txt

WHERE DOES IT FIT IN THE PICTURE OF THE SUB-IP WORK

This draft fits in the Control part of the sub-ip work.

WHY IS IT TARGETED AT THIS WG

This draft addresses the following CCAMP work items:

- Abstract link and path properties needed for link and path protection. Define signalling mechanisms for path protection, diverse routing and fast path restoration. Ensure that multi-layer path protection and restoration functions are achievable using the defined signalling and measurement protocols, either separately or in combination.

JUSTIFICATION

We believe this draft is justified for the CCAMP working group because it identifies signaling/routing mechanisms that can be used for both span and path protection.

1. Introduction

A key requirement for the development of a common control plane for both optical and electronic networks is that there must be features in the signaling, routing, and link management protocols to enable intelligent fault management. Fault management requires four steps: fault detection, fault localization, fault notification, and fault recovery. Fault detection should be handled at the layer closest to the failure; for optical networks, this is the physical (optical) layer. One measure of fault detection at the physical layer is detecting loss of light (LOL); other techniques based on, for example, OSNR, BER, dispersion, crosstalk, and attenuation are still being investigated (see, for example, [OLCP] and [LMP-DWDM]). Fault localization requires communication between nodes to determine where the failure has occurred (for example, SONET AIS is used to localize failures between SONET terminating devices). One interesting consequence of using LOL to detect failures in optical networks is that LOL propagates downstream along the connectionÆs path. The Link Management Protocol (LMP) [LMP] includes a fault localization procedure that is designed to localize failures in both transparent (all-optical) and opaque (opto-electrical) networks, and is independent of the data encoding scheme. Fault notification is the Communication of a failure between the node detecting it and a node equipped to deal with the failure. Fast fault notification is essential for rapid recovery. The Notify mechanism of [RSVP-GEN] is designed to support fast notification of non-adjacent nodes.

Once a failure has been detected and localized, and the responsible node has been notified, protection and restoration can be used to recover from the failure. We make the distinction between protection and restoration by the time scales in which they operate. Protection is designed to react to failures rapidly (say, in less than a couple hundred milliseconds) and often involves 100% resource redundancy. For example, SONET automatic protection switching (APS) is designed to switch the traffic from a primary (working) path to a secondary (protection) path in less than 50ms. This requires simultaneous transmission along both the primary and secondary paths (called 1+1 protection) with a selector at the receiving node, and uses twice as many network resources as a non-APS protected path. Restoration, on the other hand, is designed to react to failures quickly, but it typically takes an order of magnitude longer to restore the connection compared to protection switching. This is because restoration typically utilizes pools of shared resources that are more efficient in terms of the network utilization. In addition, restoration may involve rerouting connections, which can be computationally expensive if the paths are not pre-calculated or if the pre-calculated resources are no longer available.

Protection and restoration methods have traditionally been addressed

using two techniques: path-level recovery, where the failure is addressed at the end nodes (i.e., the initiating and terminating nodes of the path); and span-level recovery, where the failure is

addressed at an intermediate or transit node. Path-level recovery can be further subdivided into path protection, where secondary (or protection) paths are pre-allocated, and path restoration, where connections are rerouted, either dynamically or using pre-calculated (but not pre-allocated) paths. Span-level recovery can be subdivided into span protection, where traffic is switched to an alternate channel or link connecting the same two nodes, and span restoration, where traffic is switched to an alternate route between the two nodes (this involves passing through additional intermediate nodes).

To effectively use protection, there must be mechanisms to configure protected links on a span between nodes, advertise the protection bandwidth of a link so that it may be used by a class of traffic that has different availability requirements, establish secondary (protection) LSPs to protect primary LSPs, allow the resources of secondary LSPs to be used by lower priority traffic until a switchover occurs, and signal protection switchover when necessary. In this draft, we discuss protection and restoration in the context of GMPLS signaling. Specifically, we address these issues in the context of RSVP signaling and OSPF and IS-IS routing.

2. Protection Mechanisms

Protection is designed to react to failures in the fastest timescale and typically involves pre-provisioning protection resources. In this section we discuss both span and path protection and present mechanisms within GMPLS to implement both protection schemes.

2.1 Protection Levels

The level of protection available is a function of the protection resources available for protecting a failed resource.

- 1+1 Protection
  Two pre-provisioned resources are used in parallel. For example, data is transmitted simultaneously on two parallel links and a selector is used at the receiving node to choose the best signal.

- 1:1 Protection
  Two resources (1 primary, 1 backup) are pre-provisioned. If the primary resource fails, then the data is switched to the backup resource.

- 1:n Protection
  n+1 resources (n primary, 1 backup) are pre-provisioned. If there is a failure on any one of the primary resources, then the data is switched to the backup resource. At this point, the remaining n-1 primaries are no longer protected.

- m:n Protection

n+m resources (n primary, m backup) are pre-provisioned. If there is a failure on any one of the primary resources, then data is switched to the backup resource.

Note that 1:n and 1:1 are special cases of m:n protection.

2.1 Span Protection

A span consists of a number of channels between two adjacent nodes that are grouped together into a single link, often called a traffic engineered (TE) link (see [LMP]). Span protection involves switching to a protection channel when a failure occurs on a working channel. At the span level, both dedicated (1+1, 1:1) and shared (M:N) protection may be implemented. The protection type supported by a TE link (LPT) is advertised throughout the network using an IGP so that intelligent routing decisions can be made (see Section 4). The desired protection for a path is signaled as part of the Generalized Label Request in GMPLS signaling. This is needed in signaling if a link supports multiple protection types or if loose routing is used.

For dedicated 1+1 span protection, each node must replicate the data onto two separate channels (possibly using separate component links of a bundled link or separate ports of a TE link) and the adjacent node must select the data from only one channel based on the signal integrity. This is the fastest protection mechanism, however, it requires using twice the LSP bandwidth between each pair of nodes and the ability to replicate the data on two separate channels.

For shared M:N protection, M protection links are shared between N primary links. Since data is not replicated on both the primary and secondary links, failures must first be localized before the switchover can occur. LMP can be used for fault localization, and the upstream node (upstream in terms of the direction an RSVP Path message traverses) will initiate the local span protection. To initiate span protection, the upstream node SHOULD send an RSVP Path message with a Label Set object including the labels for the available secondary links. If more than one label is included in the Label Set object, the Suggested Label object should be used to indicate the preferred secondary label.

If the failure affected a bi-directional LSP, a new Upstream Label may also need to be transmitted. If the reverse direction of the bi-directional LSP uses a distinct component link from the failed forwards direction there is no need to re-signal the reverse path label unless there is a close correspondence between the label values chosen for the two directions. If the failed component link is bi-directional the failure might affect only one direction, but could affect both directions. If both directions fail then both labels must be re-signaled for use on new links. If the component

link carrying the reverse path fails, but the forward path is unaffected, the reverse path label must be Resignaled.

In addition, new LinkId, PHOP, and modified ERO may also need to be included based on the shared protection configuration. Note that the benefit of exchanging the shared protection configuration in advance using LMP is that it minimizes the potential label conflict when protection switching. When the downstream node receives the Path message with the new objects, it MUST verify the parameters, update the RSVP Path state, and respond with either an RSVP Resv message with a new label or it should generate a PathError message if the resources are not available.

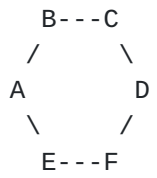
2.2 Path Protection

Path protection is addressed at the end nodes of an LSP (i.e., LSP initiator and terminator) and requires switching to an alternate path when a failure occurs. For 1+1 path protection, a signal is transmitted simultaneously over two disjoint paths and a selector is used at the receiving node to choose the better signal. For M:N path protection, N primary signals are transmitted along disjoint paths, and M secondary paths are pre-established for shared protection switching among the N primary paths.

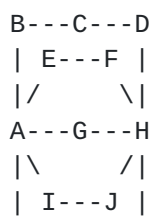
2.2.1 Simple Path Protection

There are a number of path protection variations that may be implemented that provide different levels of protection. The most common notion of path protection is to select two disjoint paths, one primary and one secondary, where each link along both paths is unprotected. This protects against a single link or node failure, depending on how the two paths are disjoint.

For example, in the network below it is possible to have a primary path A, B, C, D and a backup path A, E, F, D. These paths are entirely disjoint and are suitable for 1+1 or 1:1 protection.



m:n path protection is also possible in simple topologies. Consider, for example:




K---L---M

In this network there are five disjoint paths: {A, G, H}, {A, E, F, H}, {A, B, C, D, H}, {A, I, J, H} and {A, K, L , M, H}. These can be assigned as primary and backup resources to provide anything from 1:4 through 4:1 path protection.

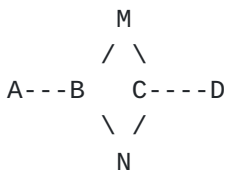
2.2.2 1+1 Path Protection Using Span Protection

One variation of 1+1 path protection is to select a single path where each link individually supports 1+1 span protection as discussed in Subsection 2.1. This protects against a single link failure, but not a node failure. One may also combine the two approaches by ensuring that for every contiguous segment of the path that includes only the links that don't support 1+1 span protection, the head-end LSR has to compute a link-disjoint segment, with the constraint that none of the links in the newly computed segments have link protection.

After the two paths are computed, the head-end LSR will originate two LSPs with dedicated 1+1 and unprotected bits set in the LPT. The setup will indicate that these two paths request Shared-Explicit reservations (see [TUNNEL]). At each node where the two paths branch out, the node must replicate the data into both branches. At each node where the two paths merge, the node must select the data from only one path based on the integrity of the signal.

For bi-directional LSPs, each branching point is also a merging point and vice versa.

As an example consider the following:



Only links A-B and C-D support 1+1 span protection. Node A wants to establish a 1+1 protected path to D. In this case, A computes a primary path, A, B, M, C, D where the segment B, M, C has links that do not support 1+1 protection. Therefore, A computes a link-disjoint segment, B, N, C, and uses it to construct a secondary path, A, B, N, C, D. A initiates a setup of two LSPs indicating the desire for Shared Explicit (SE) reservations - the first path is routed along A, B, M, C, D, and the second path is routed along A, B, N, C, D.

Since the two LSPs branch out at node B, B sends the data it receives from A to both M and N. At node C, the two LSPs merge and C selects the data received over one of these LSPs (based on the integrity of the signal), and forwards this data to D.

When the LSP from A to D is bi-directional, then C must also send the data it receives from D to both M and N, and B must select the data received from either M or N, and forward it the to A.

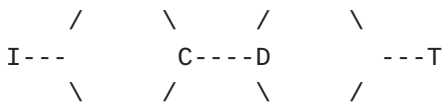
2.2.3 M:N Path Protection

There are a number of M:N path protection variations that may be implemented to provide different levels of protection and to address different network configurations. The most common notion of M:N path protection is to route N node-disjoint primary paths and pre-establish M backup paths that are node disjoint from the primary paths. This protects against M path failures. Another variation of M:N path protection is to select a single path where each link individually supports M:N span protection. This protects against M link failures over each span, but is not robust to node failures. One may also combine the two approaches by ensuring that for every contiguous segment of the path that includes only the links that donÆt support M:N span protection, the head-end node has to compute a node- or link-disjoint segment, with the constraint that none of the links in the newly computed segments need to be protected.

An important feature of the GMPLS work is that it allows pre-configuring secondary (backup) LSPs to protect primary LSPs. This is done by indicating the LSP is of type Secondary in the protection field of the Generalized Label Request. Secondary LSPs are used for fast switchover when primary LSPs fail. Although the resources for the secondary LSPs are pre-allocated, lower priority traffic may use the resources with the caveat that the lower priority traffic will be preempted if the primary LSP fails. If lower priority traffic is using resources along the secondary LSPs, the end nodes may need to be notified of the failure in order to complete the switchover.

The setup of the primary LSP SHOULD indicate that the LSP initiator and terminator wish to receive Notify messages using the Notify Request object. If a failure occurs, LMP can be used to isolate the failure. Once the failure is isolated, the upstream node (upstream in terms of the direction an RSVP Path message traverses) SHOULD send an RSVP Notify message to the LSP initiator, and the downstream node SHOULD send an RSVP Notify message to the LSP terminator. Upon receipt of the Notify messages, the source and destination nodes MUST switch the traffic from the primary LSP to the pre-configured secondary LSP. Note that if a common initiator-terminator is used for all N primary paths sharing the secondary path (assuming 1:N protection), no further notification is required to indicate that the N primary LSPs are no longer protected.

As an example consider the following:



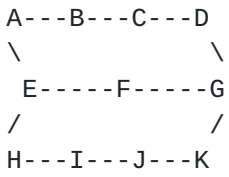
Two node-disjoint routes from initiator I to terminator T cannot be found; however, two node-disjoint routes can be found from node I to node C and from node D to node T. Furthermore, the link from node C to node D is protected using dedicated 1:1 protection. In this case, I computes the primary route R1={I,A,B,C,D,E,F,T} and secondary route R2={I,J,K,C,D,L,M,T} where the segment {C,D} supports 1:1 span protection. A initiates a setup of two LSPs indicating the desire for Shared Explicit reservations; the primary LSP is routed along R1 and the secondary LSP is routed along R2.

3. Shared Resources

The protection mechanisms described above are expensive in the resources that need to be dedicated. For example, if all of the LSPs in a network were afforded 1+1 or 1:1 protection, only half of the available network resources (bandwidth) would available for actual data traffic. Since the events that necessitate switching from primary span/path to backup span/path are supposedly rare or at least infrequent, this is a high price to pay for protection.

3.1 Merged Backups

A popular way to reduce the pre-provisioned resource requirements is to have backup paths share network resources when the paths that they protect have different ingress points but share an egress. Consider the following topology:



The path A,B,C,D,G can be protected by the path A,E,F,G. Similarly, the path H,I,J,K,G can be protected by the path H,E,F,G. However, to achieve this level of protection the links EF and FG need to have available and provisioned the sum of the resources used on the paths A,B,C,D,G and H,I,J,K,G (that is the sum of the bandwidth). This may be impractical if the resources are unavailable, and is undesirable since it ties up excessive resources given that it is unlikely that both of the entirely distinct paths A to G and H to G will fail at the same time.

In order to allow the backup paths to share resources using the standard features of GMPLS signaling, they must be signaled requesting Shared-Explicit reservations. Additionally, the LSPs must be identically identified so that the paths can be merged at

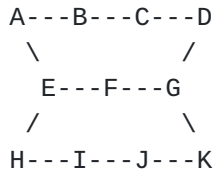
node E. To achieve this, the Extended Tunnel Id must be set to the same value on both paths (usually zero) and the Tunnel Id must be

set to the same value on the two paths. This requires external co-ordination between the ingress points of the two paths.

When a failure is detected on one primary path (say at B), the error is propagated to the ingress (A) which re-routes the data down the backup path. At this point, it is important that a failure on the other path (say at J) does not cause the other ingress (H) to send the data down the backup path since the labels and resources are already in use. This can be achieved by having a Notify message sent to H when B reports the failure. The Notify could be sent direct from B (by specifying H as the Notify recipient in the Notify Request object), could be sent from A after the error has been reported to A, or from E when the backup path starts to be used.

3.2 Sharing Resources without LSP Merging

A further variant (shown below) occurs when the two paths to be protected have different ingress points and different egress points.



The paths A,B,C,D and H,I,J,K could be protected by A,E,F,G,D and H,E,F,G,K, respectively. The signaling to allow these backups to share resources cannot be done as described above since in order to achieve resource merging, the LSPs must have the same Session Ids, but the Session Id includes the target (egress) IP address. These addresses are not the same in this example.

Resource sharing along E,F,G can only be achieved if the nodes E, F and G recognize the LSP type setting of Secondary in the protection field of the Generalized Label Request and act accordingly. In this case the backup LSPs are not merged (which is useful since the paths diverge at G), but the resources can be shared.

When a primary path fails the other primary path ceases to be protected and must be sent a notification as described above.

4. Restoration Mechanisms

Restoration is designed to react to failures quickly and use bandwidth efficiently, but typically involves dynamic resource establishment and may also require route calculation, and therefore, takes more time to switch to an alternate path than protection techniques. Restoration can be implemented at the intiator node or at an intermediate node once the responsible node has been notified. Failure notification can be done using the Notify procedures of

[GMPLS] or using the standard RSVP PathError messages. In this

section, we briefly discuss span and path restoration and highlight the RSVP mechanisms that can be used to implement them.

4.1 Span Restoration

To support span restoration, where traffic is switched to an alternate route around a failure, a new LSP is established at an intermediate node that involves passing through additional intermediate nodes. Span restoration may be beneficial for LSPs that span multiple hops and/or large distances because the latency incurred for failure notification may be significantly reduced and only segments of the LSP are rerouted instead of the entire path.

If the protected part of the LSP is a single span, then error detection is sufficient to trigger restoration. If, however, protection is required over a series of more than one span a mechanism is required to notify to the point of repair that an error has occurred and that restoration is required.

The RSVP Notify Request object can be used by an intermediate node to request that it be the target of an RSVP Notify message. Span restoration may break traffic-engineering (TE) requirements if a strict-hop route is defined for the connection. Furthermore, the constraints used for routing the connection must be forwarded so that an intermediate node doing span restoration is able to calculate an appropriate alternate route; this is similar to the problems when establishing/maintaining TE requirements that span mult-areas (see [MULTI] for a proposed mechanism).

4.1.1 Local Repair

Local repair is a special case of span protection supported by the base RSVP-TE draft [TUNNEL]. The node that detects the failure may, make an alternate routing decision and attempt to re-signal the LSP. This approach may be considered too slow since it could rely on convergence of the routing table at the repair node. However, if there is a close link between routing and path computation components, Local Repair may be equivalent to span protection.

4.2 Path Restoration

Path restoration, on the other hand, switches traffic to an alternate route around a failure, where the new route is selected at the LSP initiator and may reuse intermediate nodes used by the original LSP and it may include additional intermediate nodes. For strict-hop routing, TE requirements can be directly applied to the route calculation, and the filed node or link can be avoided. However, if the failure occurred within a loose-routed hop, the source node may not have enough information to reroute the

connection around the failure.

The backup route may be calculated on demand (that is, when the failure occurs) or may be pre-calculated and stored for use when the failure is reported. This offers faster restoration time. There is, however, a risk that the backup route will become out of date through other changes in the network - this can be mitigated to some extent by periodic recalculation of idle backup routes.

Restoration (span or path) will be initiated by the node that has isolated the failure or by the node that has received either an RSVP Notify message or an RSVP Path Error message indicating that a failure has occurred. The new resources can be established in a make-before-break fashion, where the new LSP is setup before the old LSP is torn down, using the mechanisms of the LSP_Tunnel Session object (see [TUNNEL]) and the Shared-Explicit reservation style. Both the new and old LSPs share resources at nodes common to both LSPs. The Tunnel end point addresses, Tunnel Id, Extended Tunnel Id, Tunnel sender address, and LSP Id are all used to uniquely identify both the old and new LSPs; this ensures new resources are established without double counting resource requirements along common segments. Note that make-before-break is not used to avoid disruption to the data flow (this has already been broken by the failure that is being repaired), but is valuable to retain the resources allocated on the original primary path that will be re-used by the new primary path.

5. Routing Enhancements

The GMPLS extensions to OSPF [OSPF-GE] and IS-IS [ISIS-GE] include the advertisement of the LPT. The LPT field is a bit vector that indicates the protection capabilities that are supported for the link. The LPT field may be configured with Dedicated 1+1, Dedicated 1:1, Shared M:N, and Enhanced protection, as well as Unprotected. For a link that has dedicated 1+1 protection or is unprotected, this advertisement provides a complete description of the link capabilities and the usable bandwidth. However, a key argument for using dedicated 1:1 or shared M:N is the efficiency gained by reusing the protection bandwidth for lower priority traffic when the bandwidth would otherwise be idle.

To advertise the protection bandwidth for a link that has dedicated 1:1 or shared M:N protection, a link with LPT field Extra Traffic should be advertised. This indicates that bandwidth can be used by LSPs, with the caveat that any LSPs routed over this link will be preempted if the resources are needed as a result of a failure over the primary link.

When a failure occurs on a dedicated 1:1 or shared M:N link, the LSPs routed over the link will automatically be switched to the Extra Traffic link that is protecting it.

To support the routing of Secondary LSPs for M:N path protection (as described in Section 2.2.2), new extensions must be added to the

current GMPLS routing extensions. In particular, there must be a mechanism to advertise secondary bandwidth and processing rules must be defined for bandwidth accounting when LSP requests arrive at a node. See [BWAcct] for a proposal addressing these issues.

6. Acknowledgments

We would like to thank Kireeti Kompella and Ayan Banerjee for their comments and fruitful discussions.

7. Author's Addresses

Jonathan P. Lang
Calient Networks
25 Castilian Drive
Goleta, CA 93117
email: jplang@calient.net

John Drake
Calient Networks
5853 Rue Ferrari
San Jose, CA 95138
email: jdrake@calient.net

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
email: yakov@juniper.net

Adrian Farrel
Movaz Networks Inc.
7926 Jones Branch Drive
Suite 615
McLean, VA 22102
email: afarrel@movaz.net

7. References

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision 3," BCP 9, RFC 2026, October 1996.

[GMPLS] Ashwood-Smith, P., Banerjee, A., Berger, L., et al, "Generalized MPLS - signaling functional description," Internet Draft, draft-ietf-mpls-generalized-mpls-signaling-04.txt, (work in progress).

[OLCP] Chiu, A., Strand, J., Tkach, R., ôUnique Features and Requirements for The Optical Layer Control Plane, Internet Draft, draft-chiu--strand-unique-OLCP-01.txt, (work in progress).

[LMP-DWDM] Fredette, A., Snyder, E., Shantigram, J., et al, ôLink Management Protocol (LMP) for WDM Transmission Systems,ö Internet Draft, draft-fredette-lmp-wdm-00.txt, (work in progress).

[LMP] Lang, J. P., Mitra, K., Drake, J., Kompella, K., et al, ôLink Management Protocol (LMP),ö Internet Draft, draft-ietf-mpls-lmp-02.txt, (work in progress).

[RSVP-GEN] Ashwood-Smith, P., Banerjee, A., Berger, L., et al, " Generalized MPLS Signaling - RSVP-TE Extensions," Internet Draft, draft-ietf-mpls-generalized-rsvp-te-03.txt, (work

in progress).

[TUNNEL] Awduche, D., Berger, L., Gan, D-H., Li. T., Srinivasan, V., Swallow, G., ôRSVP-TE: Extensions to RSVP for LSP Tunnels,ö Internet Draft, draft-ietf-mpls-rsvp-lsp-tunnel-08.txt, (work in progress).

[MULTI] Kompella, K., Rekhter, Y., ôMulti-area MPLS Traffic Engineering,ö Internet Draft, draft-kompella-mpls-multiarea-te-00.txt, (work in progress).

[OSPF-GE] Kompella, K., Rekhter, Y., Banerjee, A., Drake, J., et al, ôOSPF Extensions in Support of MPL(ambda)S," Internet Draft, draft-kompella-ospf-gmpls-extensions-01.txt, (work in progress).

[ISIS-GE] Kompella, K., Rekhter, Y., Banerjee, A., Drake, J., et al, ôIS-IS Extensions in Support of Generalized MPLS,ö Internet Draft, draft-ietf-isis-gmpls-extensions-0.2txt, (work in progress).

[BWAcct] Kompella, K., Lang, J.P., Drake, J., ôBandwidth Accouting in Support of Secondary LSPs,ö Internet Draft, (work in progress).