

MPLS Bandwidth Protection

Overview

Cisco MPLS Bandwidth Protection enables Cisco customers to provide bandwidth safeguarding for traffic re-routed to backup tunnels in the context of MPLS Traffic Engineering (TE) Fast Reroute capability. Cisco MPLS Bandwidth Protection allows efficient sharing of the bandwidth between backup tunnels protecting against independent elements failure, resulting in a much more economical bandwidth usage than the common approach of explicit bandwidth reservations for backup tunnels.

Cisco MPLS Bandwidth Protection currently uses MPLS Tunnel Builder as the front end, and a third party technology as the back end. Tunnel Builder will be responsible for the user interface, collection of topology information and installing the backup tunnels, while Parc Technologies, Inc. develops the algorithm for the computation of the paths for the backup tunnels.

Abstract

This white paper describes a network management tool, called Tunnel Builder Pro, for computing and establishing Fast Reroute Label-Switched Paths (LSPs) that protect against node, link and Shared Risk Link Group (SRLG) failure in a manner that ensures bandwidth availability in the presence of failures. As a result, Tunnel Builder Pro provides a means for increasing the level of Quality of Service (QoS) assurances a network operator can offer.

1. Introduction

1.1. The Need for Bandwidth Assurances in the Presence of Link and Node Failure

The Service Provider community has had substantial interest in supporting QoS-sensitive traffic reliably. This interest is fueled by the desire to support delay- and jitter-sensitive applications, such as Voice over IP (VoIP) or real-time video. Typically, QoS requirements for a flow can be translated into the appropriate bandwidth requirements.¹ Therefore, as long as the end-to-end propagation delay is reasonably bounded, the problem of QoS provisioning can be effectively reduced to that of bandwidth provisioning.

A number of mechanisms such as a combination of DiffServ and MPLS Traffic Engineering (TE) can be used to achieve the desired bandwidth assurances for flows that require such assurances under the normal operating conditions. However, there are no established mechanisms for maintaining these promises when node and link failures occur. While the standard routing protocols based on Constraint-Based Shortest Path First (CSPF) algorithms are suitable for establishing new LSPs, they are too slow for dynamic rerouting of QoS-sensitive traffic in the presence of node and link failures. In

1. To ensure queuing delay and jitter guarantees, it may frequently be necessary to ensure that the bandwidth available to a flow is higher than the actual data transmission rate of this flow.



SONET and SDH-based networks, bandwidth restoration and protection occurs at the SONET and SDH level. However, where SONET and SDH -based restoration is not available, other protection mechanisms become necessary to ensure that bandwidth safeguards are still in place upon link or node failure.

To achieve this goal in the context of MPLS, three basic techniques, collectively known as Fast Reroute, have been proposed: Path Protection, Node Protection and Link Protection. The essence of Fast Reroute is in establishing pre-determined backup paths that can be immediately used to re-route traffic upon detection of a failure. The differences between the three approaches are discussed in subsequent sections. In each of the techniques, think of LSPs as belonging to one of two types:

- *Primary LSPs*—established to carry traffic under normal (non-failure) conditions
- *Backup LSPs*—carry traffic only when a link or node has failed

1.2 Difficulties with Path Protection

The essence of Path Protection is in establishing an end-to-end backup tunnel for each of the primary LSPs that require bandwidth protection. A backup LSP must typically be link- and node- disjoint with its corresponding primary LSP. It is established in conjunction with the primary LSP. When a failure occurs, the head-end of the primary LSP is notified about the failure, and the traffic of this LSP is rerouted to the backup LSP.

While this approach is quite useful, it has a number of difficulties that make it a challenge to provide hard QoS assurances in the presence of failure. In addition, it is less scaleable than the alternative approach of Node and Link Protection (described in subsequent sections).

The first difficulty with a backup LSP for path protection is that there is an unacceptably long signaling delay in notifying the head end of the end-to-end LSP of a node or link failure after it occurs. That is, by the time the head-end on the affected LSP is notified about a failure, QoS may have already degraded below acceptable levels. The first component of the delay is simply the end-to-end propagation time, which may be substantial for long-haul LSPs. The second component of the delay is related to the fact that signaling messages must be sent as fast as possible to ensure timely switchover to the backup LSP. Thus, when a link or node fails, all affected primary LSPs must signal the failure to their respective head-ends, generating a significant number of simultaneous signaling messages. The loss of any of these packets creates a much longer delay (error recovery based on timer). This creates additional instability for the network, as more LSPs will be signaled around the same time. In comparison, with Node/Link protection, the traffic is directly restored at the point of failure and hence can take more time to signal the head-ends to re-optimize their end-to-end LSPs. This enables schemes such as pacing, prioritization between signaling messages etc., to achieve a more stable/scalable signaling design.

The second difficulty with path protection is that using standard routing protocols, such as CSPF, to establish a backup LSP at the same time as a primary LSP may result in highly inefficient bandwidth usage. That is, much more bandwidth may be allocated for backup LSPs than is actually needed. The inefficiency stems from the fact that shortest-path based algorithms must allocate bandwidth for all backup LSPs sharing a link independently. If two backup LSPs share a link, CSPF routing will always allocate the sum of bandwidth requirements of these LSPs. Yet, under the assumption of a single element (node or link) failure, this may not always be necessary. If primary LSPs between the two ingress and egress nodes are link- and node-disjoint, their backup LSPs may share the bandwidth on the backup paths without compromising QoS, since only one of them can fail at any given time. Therefore, only the maximum of the bandwidth requirements of the two LSPs really need to be allocated on the shared link on the backup path instead of their sum. However, attempting to take advantage of such backup bandwidth sharing results



in a complex optimization problem, which is known to be NP-complete. A dynamic “on-the-fly” solution to such a complex problem appears to be unrealistic. In a dynamic environment when new LSP establishment requests arrive and/or old LSPs are torn down, an off-line solution is highly undesirable.

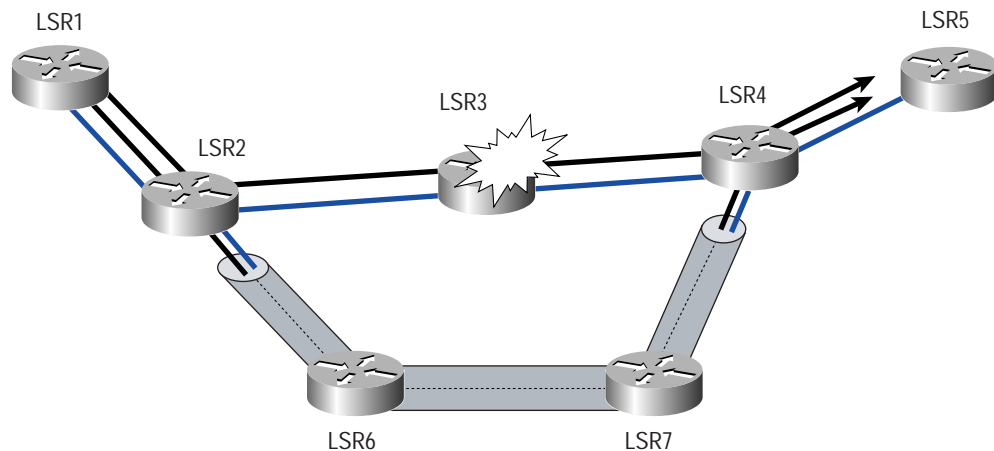
The third difficulty with Path protection is that the amount of LSP state that needs to be maintained at each node doubles in the network. This is because for each primary LSP there is also a backup LSP that needs to be maintained. In contrast, with Node and Link protection a large number of primary LSPs can share a single backup LSP.

For these reasons, the rest of this white paper will focus on the Node and Link Protection approach.

1.3 Benefits of the Node and Link Protection Approach

The essence of Node and Link Protection is in establishing local backup LSPs (also known as backup tunnels) that reroute all traffic around a failure. To protect a given element (link or node), an LSP is established between any two neighbors of this element.² When the node (or link) fails, all LSPs that traverse this node (or link) and pass through a given pair of its neighbors are rerouted through the single backup LSP that connects those two neighbors. This is illustrated in Fig.1 for the case of a node failure. Since only a single (typically short) backup LSP is used for backing up a potentially large number of primary LSPs, there is a substantial savings in the amount of LSP state relative to Path Protection. Since the reroute occurs only locally, it is possible to detect and signal the failure and reroute the affected LSPs much faster than in the case of end-to-end Path Protection. As noted above, Node and Link Protection allows for a more stable/scalable signaling of the failure to the head-ends, and thus a more stable/scalable re-optimization. These features make Node and Link Protection an attractive alternative to Path Protection.

Figure 1
Backup Tunnel Protecting LSR3



In Figure 1, a backup LSP is established between LSR2 and LSR4. If LSR3 fails, all affected LSPs are tunneled through the backup LSP via LSR6 and LSR7.

2. In principle, one can establish backup LSPs between more remote neighbors of the failed element. While it may provide more flexibility in making the backups potentially more optimal, it also substantially increases the computational difficulty of finding out where to place the backup LSPs to ensure bandwidth protection. Since for Node Protection the backup LSPs are only used for a short time, it is more important to ensure reliable QoS assurances during failure than to find the optimal placement of the backup LSP. For these reasons, this document considers backup LSPs between the immediate neighbors of a failed element only.



1.4 Potential Issues with the Node and Link Protection Approach

One issue with Node and Link Protection is that, due to its local nature, the network may reach a sub-optimal state after failure. However, it is possible to use Node and Link Protection to reach an *acceptable* state very quickly while using mechanisms that operate over a longer time scale—the re-routing of affected LSPs using CSPF—to return to an optimal state. *The key to making the intermediate state (using a backup tunnel) acceptable is to promise that there is sufficient bandwidth available on the links traversed by the backup LSPs. If the QoS safeguards can be maintained even in the transient, possibly sub-optimal state, re-optimization can be accomplished over the longer time scale possible with CSPF, without degrading service offered to the individual LSPs.*

A more serious issue is related to the fact that ensuring protection against any node or link failure requires the path computation and establishment of a potentially large number of backup LSPs.³ It is nearly impossible to manually compute where all the backup paths should be placed without creating bandwidth or delay violations. Moreover, the mere task of setting up the appropriate backup paths after they have been computed is challenging.

A tool for both computing the backup paths and establishing the backup LSPs along those paths in advance is therefore highly desirable. Availability of such a tool will solve a significant obstacle in widespread deployment of Node and Link Protection, and will allow the operator to ensure fast restoration, while protecting QoS safeguards upon failure.

1.5 The Non-Trivial Nature of Node Protection⁴

Just as in the case of Path Protection, it is challenging to set up the backup LSPs for Node Protection in such a way that bandwidth assurance is retained during failure

The “naive” approach of using standard shortest-path routing algorithms without accounting for bandwidth requirements can easily lead to massive overload of backup links. One way to avoid this problem would be to use CSPF to ensure that a backup LSP has enough capacity to backup all the primary LSPs between the corresponding pairs of neighbors. However, CSPF is a “greedy” algorithm, which has no global optimization and therefore it is likely to fail to find a solution in many cases. Furthermore, the knowledge of bandwidth requirements of individual LSPs may not be available at the time when backup LSPs are established. The set of LSPs going through a particular node may change more quickly than is desirable to change the backup LSPs.

In principle, one can choose to use the complete information about the primary LSPs when computing the backup paths for node protection. However, this approach creates a number of difficulties:

- Necessity to maintain information about a large number of primary LSPs that may exist in the network
- Timescale over which the tool must compute the backup routes must be no larger than the timescale over which new LSPs can be established, and/or old LSPs destroyed. This is perhaps the most significant of the issues.
- A change in the set of primary LSPs may cause a cascade change in the setup of established backups, meaning that a large number of backup LSPs may need to be re-established upon every change in the set of primary LSPs.

Tunnel Builder Pro avoids these difficulties by using only aggregate information about the maximum bandwidth pool available for LSPs of a particular class type, rather than the bandwidth in use by specific LSPs. In particular, it focuses on the maximum bandwidth available for primary LSPs for which a strict bandwidth assurance is desired even in the

3. Note, however, that the number of backup LSPs is typically much smaller than the number of backup LSPs for Path Protection.

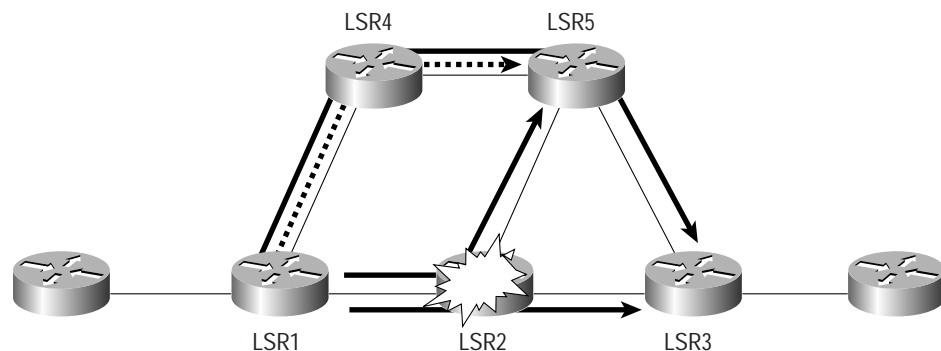
4. The SRLG Protection problems can be reduced to that of node protection, and hence the discussion in this section relates to SRLG as well. Single link protection is an easier problem to solve.



presence of node or link failure. It does not require any information about individual primary LSPs. The approach ensures that as for any set of primary LSPs whose total bandwidth does not exceed the pool allocated to them, the backup LSP connecting a given pair of neighbors of this node is sufficient to maintain the desired QoS.

It might be possible to use standard routing protocols, such as CSPF, to solve the problem of placing the backup LSPs, but the greedy nature of CSPF may prevent it from finding a solution, even if a solution does exist. Furthermore, Figure 2 illustrates that using CSPF with the maximum possible bandwidth assignment can lead to very inefficient bandwidth usage.

Figure 2
Naive CSPF routing of maximum bandwidth LSPs.



In this example, all links have the same speed C . The maximum total bandwidth of all LSPs that can share the path LSR1-LSR2-LSR3 is clearly C . Likewise, the maximum total bandwidth that can be used by all LSPs traversing the path LSR1-LSR2-LSR5 is also limited by C . If CSPF routing were used to route backup LSPs protecting against failure of LSR2, then two backup LSPs of capacity C would be required—one to protect the LSR1-LSR2-LSR3 traffic, and one to protect the LSR1-LSR2-LSR5 traffic. But clearly only a single backup LSP can be routed through the link LSR4-LSR5, because the capacity of all links in this path is only C . Therefore, LSR2 could not be fully protected in this case.

On the other hand, it is clear in this example that the total bandwidth of all LSPs traversing LSR2 is bounded by the capacity of the link LSR1-LSR2, i.e. by C . Thus the backup paths for LSR1-LSR2-LSR3 and LSR1-LSR2-LSR5 can share links, since the capacity of each link is C . In particular, both backup LSPs LSR1-LSR4-LSR5 and LSR1-LSR4-LSR5-LSR3 can share the link LSR4-LSR5. *However, such a sharing opportunity would not be detected by CSPF routing.*

A requirement for such sharing, while clearly beneficial for efficient network usage, results in a complex optimization problem that cannot be solved by standard techniques. Tunnel Builder Pro uses proprietary hybrid optimization algorithms that result in a highly efficient search for backup LSPs that efficiently share network resources.



It should be mentioned that in the worst case the problem of node protection is NP-complete. However, the algorithm used by the proposed tool is very efficient in the average case⁵ for which it is highly optimized. Moreover, the algorithm is complete, i.e. it is guaranteed to find a full set of backups (or prove that some backups do not exist). Although in the worst-case scenario, this will take exponential time to compute, in the average case the algorithm executes very efficiently. To address the especially difficult instances, Tunnel Builder Pro has a user-selectable parameter that limits the time or computing backup paths for a given element (Node, Link or Shared Risk Link Group). The tool makes an initial pass on the whole network, and finds all or most of the backup LSPs very quickly. For those elements that could not be protected within the specified timer value, the user has a choice of either letting the tool run longer, or immediately requesting a recommendation for possible changes that would result in a complete set of backups. The algorithm that produces such recommendations is very fast.

2. High-Level Tool Description

This section provides a high level description of a proprietary tool for computing and establishing backup LSPs for Node and Link Protection. Tunnel Builder Pro can also handle the simultaneous failure of multiple links resulting from a single failure at the physical layer, such as a cable cut. A group of links, all of which are at risk to such simultaneous failure, is referred to as Shared Risk Link Group (SRLG).

Tunnel Builder Pro ensures that the backup LSPs it computes satisfy both bandwidth constraints. Although the work involving the development of this tool is being scoped, statements made in this and subsequent sections are indicative of what it will support. However, the functionality available in the initial release or subsequent releases may or may not exactly match what is stated in this white paper. It is explained here for the purpose of planning and obtaining feedback from the customers.

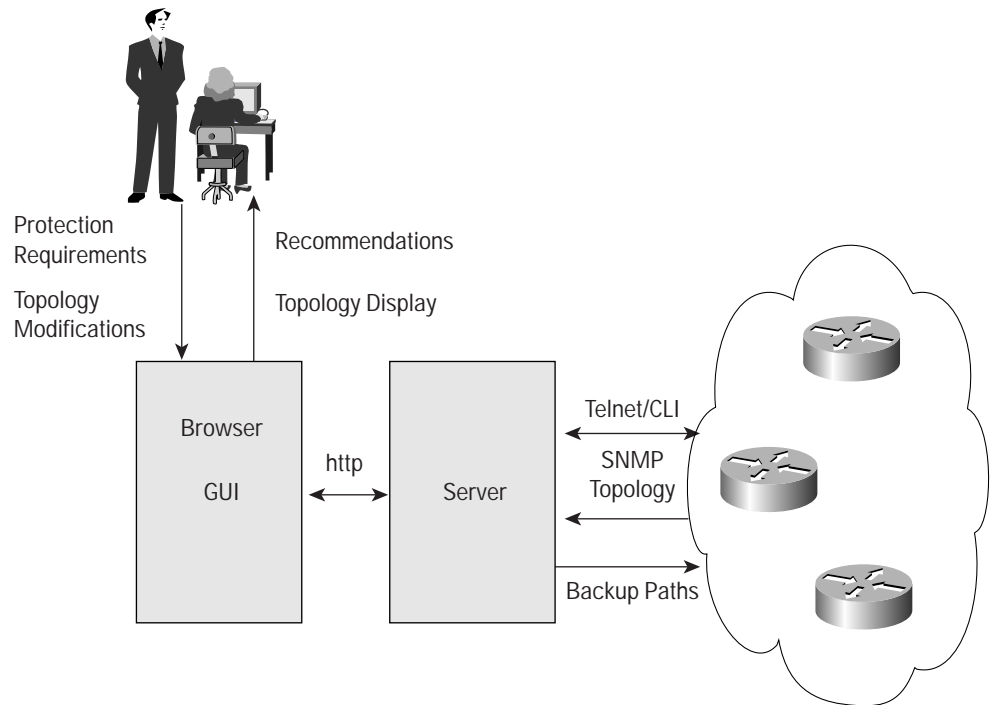
2.1 Overall System View

The overall system runs on a server that is able to communicate with all the routers in the network. The software running on the server will gather topology information from the network. The user provides other information, such as the desired amount of traffic to be backed up, and which links and nodes to protect.

5. Initial testing of the algorithm on a variety of networks derived from real ISP maps completed the computation in a matter of a few minutes.



Figure 3
Overall system view



The tool's output is either a complete set of backup LSPs if they can be found, or a set of recommendations if they cannot. Recommendations may include suggested places to add link capacity, or suggested lower level of backed up traffic, and/or increased bandwidth pool available for backups of a subset of links.

When the set of backup LSPs is generated, the user may examine the LSPs and then request that the tool configure them. Tunnel Builder Pro will then communicate with the routers to establish the appropriate backup LSPs.

2.2 Applicability to Bandwidth Protection for DS-TE Tunnels

The primary application of Tunnel Builder Pro is in establishing QoS assurances for traffic that requires strict bandwidth safeguards, such as real-time traffic that MPLS DiffServ aware Traffic Engineering (DS-TE) tunnels are expected to carry.

In order to provide stringent delay assurances to delay-sensitive traffic such as VoIP in routers implementing aggregate class-based scheduling, the total amount of high priority traffic on any link should be limited to a pre-determined fraction of the link bandwidth. If the same delay assurances are desired even in the presence of failures, the total amount of primary and backup high priority traffic should not exceed the chosen fraction of link bandwidth when failure occurs.

An alternative approach might be to relax delay requirements in the case of node or link failure, while still demanding complete bandwidth protection. In that case, instead of limiting the sum of primary and backup high priority traffic to a fixed fraction of the link bandwidth, a user may decide to allow backup DS-TE tunnels to use all or part of the bandwidth of any link not already used by primary DS-TE tunnels.



Tunnel Builder Pro can support both of these approaches. Conceptually, in either case, it needs to know two link parameters:

1. Amount of bandwidth available for primary DS-TE tunnels
2. Amount of bandwidth available for backing up DS-TE tunnels

However, there is a subtlety related to the existing implementation of DS-TE tunnels. In the current implementation, a single value only is available to specify the bandwidth pool available to DS-TE tunnels. This bandwidth pool value is what signaling uses to decide how many DS-TE tunnels can be admitted on a given link. Furthermore, in order to enable Best Effort (BE) traffic to use bandwidth allocated for DS-TE backup tunnels under normal operating conditions, the backup LSPs are established with zero bandwidth, and hence are effectively not consuming bandwidth of the specified DS-TE pool. Therefore, signaling would allow the entire DS-TE pool to be allocated to primary DS-TE tunnels, and hence the value of DS-TE pool has the meaning of primary *DS-TE bandwidth pool only*. The amount of bandwidth desired for backup DS-TE tunnels is not currently available through the CLI, and therefore must either be manually provided to the tool, or be inferred from the other known parameters.

To circumvent this difficulty, the tool supports two modes of operation.

- In the normal (default) mode, the tool assumes that each link has the amount of bandwidth available for backup traffic equal to DS-TE pool. The implication of this is that in the presence of failure, the total amount of primary and backup DS-TE traffic on any link may be as high as twice the specified DS-TE pool. That means that if a network operator desires to limit the total amount of bandwidth of primary and backup traffic to, say, 40% of the link bandwidth, the value of (primary) DS-TE pool should be set (via CLI) to 20%.
- If more flexibility is required, Tunnel Builder Pro can operate in the second mode. In this mode it is explicitly provided with the two values: the amount of bandwidth available for primary traffic, and the amount of bandwidth available for backup traffic. Since the latter is not currently available through the CLI, it has to be provided by the operator.⁶ This mode is more general, but requires user input. The normal, default mode is more constrained, but can be used if automatic collection of relevant link information is desirable.

Once Tunnel Builder Pro is notified about the primary bandwidth pool and the amount of bandwidth available for backups, it either returns the placement of backup DS-TE tunnels satisfying the specified bandwidth constraints, or provides recommendations to change the bandwidth pools and/or upgrade bandwidth on some links.

2.3 Applicability to Best-Effort Traffic Engineering

In the context of networks where all traffic belongs to the same class, the tool can ensure that the degree of maximum overbooking of any link upon any node or link failure is bounded by a desired factor.

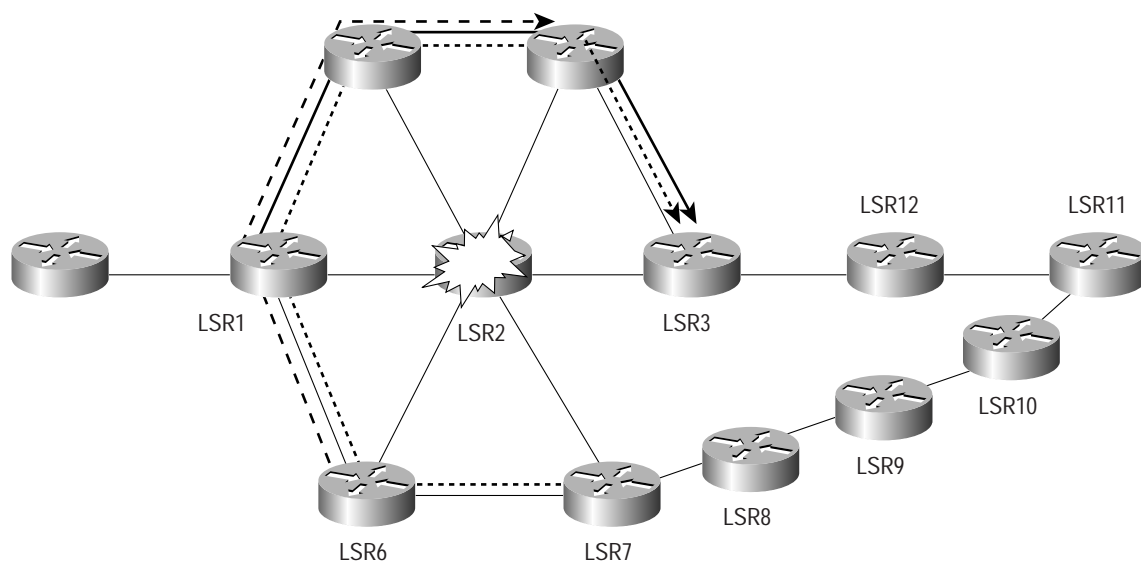
If the operator limits utilization of any link to some fixed factor (such as 80% during normal operation), and the tool that establishes the backup LSPs is provided with 80% bandwidth of any link for establishing the backup LSPs, then a link or node failure will result in the overload not exceeding 1.6. If no overload is desired, the operator can limit maximum bandwidth utilization of primary traffic to 50% of the link bandwidth.

6. An option to automatically set the amount of bandwidth available to backup LSPs to the link bandwidth minus the primary bandwidth pool can also be provided.



Note that limiting the utilization of primary traffic to 50% does not guarantee complete elimination of overload if CSPF is used to determine the backup paths. If no careful bandwidth accounting is done upon establishing backup paths, even a 50% loaded network may become overloaded if a node fails. This can be seen in Figure 4. Suppose that the utilization of all links in this network is restricted to 50% (all links are of the same capacity C) in the absence of failure. If CSPF routing is used to place backup LSPs protecting LSR2, then LSPs protecting paths LSR1-LSR2-LSR3, LSR6-LSR2-LSR4 and LSR7-LSR2-LSR5 will all be rerouted through the link LSR1-LSR4. If all link capacities are the same, then the combined requirement for backup traffic rerouted through link LSR1-LSR4 will be $3/2C$ when LSR2 fails. In addition, this link can also carry $C/2$ of its own primary traffic. Hence, the link will be overloaded by the factor of 2, even though the amount of primary traffic on any link is limited to 50%. In more complicated configurations the overload may be quite large.

Figure 4
Overload of backup paths with 50% utilization of primary traffic under SPF



2.4 Simultaneous Support for Multiple bandwidth Pools for Different Class Types

The tool will initially support bandwidth protection for a single class type. This means that the tool will be capable of one of the following two modes of operation:

1. TE, along with conventional TE, Tunnel Builder Pro will establish backup LSPs that will assure bandwidth for all DS-TE tunnels during failures. The same backup LSP will be used for backing up conventional TE, but those will not be provided any bandwidth assurances.
2. In a network that supports only conventional TE, Tunnel Builder Pro will operate as described in section 2.3.

Future releases of Tunnel Builder Pro may be targeted to support DiffServ-aware TE tunnels of multiple class types simultaneously. This can be achieved by applying it recursively to all class types, starting with traffic requiring the greatest safeguards.



2.5 Multiple Backups for a Single Link or Node failure

As described in the previous section, the initial release of Tunnel Builder Pro will support bandwidth protection for tunnels of a single class type. When the network supports LSPs of two class types (DS-TE tunnels and conventional TE LSPs), all traffic (DS-TE and regular) between the two routers adjacent to a failure will be tunneled through the same backup tunnel. In a later release of the product, it may be possible to setup two different backup LSPs connecting the same pair of nodes adjacent to the failure - one LSP per class type supported by the element protected by these LSPs.

The initial release will support multiple backup LSPs between the same pair of nodes for a given class type. This capability will enable load balancing between multiple backup LSPs, and will potentially lead to more efficient bandwidth usage.

2.6 Physical Layer Dependencies

Tunnel Builder Pro will allow protection against simultaneous link failure caused by the dependency of these links at the physical layer. The tool will allow several layer 3 links to be grouped into a Shared Risk Link Group that may fail simultaneously, for example due to a fiber cut. Initially, the operator will manually provide information to Tunnel Builder Pro about which links should be grouped together. In future releases the information about the underlying physical topology will be collected automatically.

For any SRLG, the tool will find a set of backup LSPs connecting the endpoints of each link in the group. This can be used if all the links in the group fail simultaneously.

2.7 Topology Changes

Whenever an element (node or link) comes up or goes down, the failure will be signaled to the tool automatically, triggering a new computation of the backup paths for the affected nodes and/or links. After the new set of backup LSPs is computed, Tunnel Builder Pro will communicate with the routers to install a new backup LSP state, and if necessary, tear down some backup LSPs that have become obsolete as a result of failure.

The following two paragraphs describe what happens when a new element comes up or an element goes down in the network.

2.7.1 New Element Comes Up

When a link or a node comes up, none of the already established backup LSPs is affected, so Tunnel Builder Pro simply needs to compute new tunnels protecting the new element.⁷

2.7.2 Element Goes Down

More work may be required when an element X goes down, because a number of backup LSPs protecting some other links or nodes may have traversed X, and now need to be re-established. Note also that re-routing a backup LSP that used to go through the failed node may trigger a re-computation of other backup LSPs that did not traverse X at all, since it may no longer be possible to satisfy aggregate bandwidth constraints on all links.

The tool will compute the new set of backup LSPs and communicate with the routers to install them and potentially tear down some of the old backup LSPs rendered obsolete by the failure.

7. The situation is slightly more complicated when a new link belongs to an already existing link group, in which case the set of backup LSPs protecting other links in the group may need to be changed.



2.9. Coexistence with other vendor equipment

To get the full benefit of the tool, all nodes in the network should implement the Node and Link Protection schemes described in this document and as implemented by Cisco.⁸ When some equipment does not provide such support, certain failures of the equipment or adjacent links may still be protected, but complete protection may not be possible.

3. Detailed Tool Description

The tool is a Java based client/server application with a GUI⁹ that can run either as standalone application or as an applet within a web browser. Only the Java server will interact with the routers. The server will receive commands from the GUI expressing the user's intent, and translate these into specific CLI commands that are then sent to the routers. Topologies (MPLS and CDP) and other information are gathered from the routers via CLI commands in Telnet sessions. The GUI allows the user the choice of just generating config files as output or having the server issue CLI config commands over a telnet session to configure the routers.

3.1 User Interface

The GUI includes a topology display and forms-based tabs for textual input and output. These operate during each of five basic stages of operation:

1. Topology collection and specification
2. Topology analysis (e.g. detecting changes from the previous topology)
3. Protection requirement specification (e.g. amount of bandwidth to be protected)
4. Computation and analysis of backup paths
5. Configuration of backup paths

The GUI forms-based input screens allow the user to request that the tool collect topology information from a specified router, modify the topology that the tool uses (i.e.: add bandwidth capacity), specify protection requirements, and calculate recommended backup paths. The GUI allows the user the choice of generating config files as output or having the server issue CLI config commands over a telnet session to configure the routers.

In all stages of operation, MPLS Bandwidth Protection allows the user to visualize the layout of the network in different topology views, showing nodes and links, whether MPLS TE is enabled, and where tunnels are configured. The user can request detailed information about each node, link, or tunnel by clicking or specifying the element name. The user can, for example, request the following types of configuration information:

- Link attributes—reserved and available bandwidth, affinity, interface type
- Primary tunnel attributes—bandwidth, affinity, tunnel type
- Backup tunnel attributes—bandwidth, class, protected element
- SRLG names for a network, and the membership of each group
- Primary and backup tunnels going through a particular link or node
- Backup LSPs configured to protect against the failure of a particular router

8. These mechanisms have been specified in Internet drafts.

9. Full GUI support is not planned for the initial release.



Thus the tool provides a graphical visualization of bandwidth assurances on the network, and the backup tunnels in place to protect that bandwidth.

3.2 Input

The Java server will gather the following information from the network and/or from the user input:

1. Network topology (set of nodes and links)
2. Link bandwidth
3. Set of TE class types supported in the network¹⁰
4. For each class type and each link
5. Maximum bandwidth available to primary traffic of this class type
6. Maximum bandwidth available for backup traffic of this class type
7. A set of link groups sharing the risk of a single failure at the physical layer (Shared Risk Link Groups)
8. A set of already established backup LSPs¹¹
9. Various control parameters

The input data may be collected automatically, or can be entered or modified manually.

3.2.1 Link parameters

Each link in the network is bi-directional, and is associated with the two routers it connects (RouterA and RouterB). Several links may connect the same pair of routers. Each link has a capacity (Capacity), and a unique name (LinkName).¹²

3.2.2 Class Type Parameters

Each class type has a name (ClassTypeName). For each link in the network two parameters must be specified for each ClassTypeName enabled on that link: the total bandwidth pool available for its primary traffic (ClassPrimaryBandwidth), and the total pool allowed for backup traffic (ClassBackupBandwidth).

ClassPrimaryBandwidth is either collected from the network, or is explicitly specified by user.

ClassBackupBandwidth is by default set equal to ClassPrimaryBandwidth, or is explicitly specified by the user.

In addition, for each class type optional global parameters GlobalPrimary% and GlobalBackup% may be specified. If these parameters are specified, then the values ClassPrimaryBandwidth (or ClassBackupBandwidth) on each link in the network are computed as Capacity times GlobalPrimary% (or GlobalBackup%).

3.2.3. Shared Risk Link Groups

Links can be grouped together by the user into a single link group with a unique name (SharedLinkGroupName). In the initial release of Tunnel Builder Pro, a given link can belong to a single SRLG. This limitation will be removed in subsequent releases.

Note that different links in the same SRLG do not have to support the same set of class types.

10. Only a single class type may be supported in the initial release, two or more class types may be supported in the subsequent releases

11. This is required to enable incremental updates

12. If Delay is not known, it may be assigned the value zero.



3.3 Output

For each class-type enabled on the element and for each element (node, link or shared risk link group), Tunnel Builder Pro attempts to compute the complete set of backup paths protecting this element. There are two cases to consider:

1. The tool *succeeds* in obtaining valid backup paths for all the reserved bandwidth that would be affected by the failure of the given element.
2. The tool fails to find valid backup paths satisfying bandwidth constraints for the given element.

For each element on which Tunnel Builder Pro successfully computes all the backup tunnels that satisfy bandwidth constraints, the tool outputs the set of backup tunnels for each ClassType enabled on the element.

If some elements could not be protected, Tunnel Builder Pro outputs the set of such unprotected elements. For each of them, it indicates whether the problem was proven to be insolvable, (i.e. all the reserved bandwidth cannot be backed up), or whether the tool timed-out before it showed whether the problem is solvable or not.

If the problem is insolvable, Tunnel Builder Pro indicates which of the following conditions caused that condition:

1. One or more pairs of nodes directly adjacent to the failure become disconnected due to the failure.
2. For some pair of nodes directly adjacent to the failure there does not exist a backup path of sufficient capacity to carry all the traffic between the neighbors that previously went through the failed node.
3. Sufficient capacity does not exist to reroute all traffic between all pairs of neighbors of the failed element simultaneously.

The tool may also be requested to provide network change suggestions that would enable backups that satisfy all the constraints to be found. The suggestions can be:

1. Suggested decrease in ClassPrimaryBandwidth or increase in ClassBackupBandwidth on certain links that would result in successful computation of backups;
2. Suggested increase in bandwidth of some links, and the corresponding increase of ClassBackupBandwidth on those links;

The user will then be allowed to make changes to the network, and Tunnel Builder Pro will re-compute the new backup paths.

In the event that backup paths could not be computed because the tool timed out, the user will be given an option to request recommendations for changes that will yield a solution, as described above, or to increase the value of timeout and retry the computation of the backup paths.

The initial release of the tool may output the set of router configuration commands required to set up the LSPs that have been computed, but may not have the capability to set up all backup LSPs automatically. Subsequent releases may enable Tunnel Builder Pro to communicate with the routers and set up all backup LSPs automatically.

4 Problem Scale

The tool is initially targeted to handle networks within the following limits:

- Up to 1,000 routers in network core
- Degree of connectivity up to 16
- Up to 32 links in any SRLG
- Up to 8 parallel links connecting the same router pair

These limits will be relaxed in subsequent releases. It is important to note that violation of these limitations does not cause the tool to fail, but may result in longer computation time.

5 Summary

Tunnel Builder Pro enables Service Provider to provide strict QoS assurances in the presence of a single link, node or Shared Risk Link Group failure. It is intended for the use in conjunction with the Cisco IOS Node and Link Protection feature of Fast Reroute for MPLS TE.

Tunnel Builder Pro, which runs on a standalone server, is capable of collecting network topology information. It computes backup paths that bypass any single failed element, satisfying bandwidth constraints necessary to maintain QoS assurances of traffic re-routed around the failure, and communicating with the routers to establish backup LSPs along these paths. It can also monitor topology changes and update the set of backup LSPs in response to these changes.

The primary benefit of Tunnel Builder Pro is that it provides a means for increasing the level of QoS assurances a network operator can provide.



Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 317 7777
Fax: +65 317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland
Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland
Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

All contents are Copyright © 1992–2002, Cisco, Cisco IOS, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company.
(0208R) xxxxxx/ETMG 9/02