Metaswitch Networks

MPLS Traffic Engineering - A Choice Of Signaling Protocols

Analysis of the similarities and differences between the two primary MPLS label distribution protocols: RSVP and CR-LDP

Paul Brittain, pjb@metaswitch.com
Adrian Farrel, afarrel@movaz.com

First issued January 2000

Table of Contents

1. Introduction 1
   1.1 Document Conventions 1
2. Background 2
   2.1 MPLS 2
   2.2 Label Distribution 4
   2.3 Explicit Routes 5
   2.4 Constrained Routes 6
   2.5 Resource Reservation 6
   2.6 Traffic Engineering 6
   2.7 Service Level Contracts 7
   2.8 Virtual Private Networks 7
   2.9 Meeting the Needs of the Modern Network 7
3. Introduction To CR-LDP 9
4. Introduction To Labels Extensions To RSVP 11
5. Comparative Analysis 13
   5.1 Availability of Transport Protocol 14
   5.2 Security 14
   5.3 Multipoint Support 15
   5.4 Scalability 15
      5.4.1 Network Flows 15
      5.4.2 Data Storage Requirements 16
      5.4.3 CPU Load 17
      5.4.4 Summary 17
   5.5 High Availability 18
   5.6 Link and Peer Failure Detection 19
   5.7 Re-routing 19
   5.8 LSP Modification 20
   5.9 LSP Protection 21
   5.10 Lambda Networking 21
   5.11 Traffic Control 22
   5.12 Policy Control 22
   5.13 Layer 3 Protocol 23
   5.14 QoS and Diff-Serv 23

5.15 Provision of VPNs 24
5.16 Voice over IP and Voice over MPLS 24
5.17 MIB Management 24
5.18 Acceptance/Availability 24
5.19 Interoperability 25
5.20 Interoperation with Other Label Distribution Methods 25
6. Summary 26
7. Glossary 27
8. References 29
9. About Metaswitch 30

1. Introduction

MPLS is a new technology that offers to open up the Internet by providing many additional services to applications using IP. MPLS forwards data using labels that are attached to each data packet. These labels must be distributed between the nodes that comprise the network.

Many of the new services that ISPs want to offer rely on Traffic Engineering functions. There are currently two label distribution protocols that provide support for Traffic Engineering: Resource ReSerVation Protocol (RSVP) and Constraint-based Routed Label Distribution Protocol (CR-LDP).

Although the two protocols provide a similar level of service, the way they operate is different, and the detailed function they offer is also not consistent. Hardware vendors and network providers need clear information to help them decide which protocol to implement in a Traffic Engineered MPLS network. Each protocol has its champions and detractors, and the specifications are still under development.

Recognizing that the choice of label distribution protocol is crucial for the success of device manufacturers and network providers, this White Paper explains the similarities and important differences between the two protocols, to help identify which protocol is the right one to use in a particular environment.

Data Connection's DC-MPLS family of portable MPLS products offers solutions for both the RSVP and CR-LDP label distribution protocols.

The structure of this white paper is shown in the table of contents. Readers who are already familiar with MPLS, CR-LDP and RSVP may prefer to skip straight to the Comparative Analysis section of this document.

1.1 Document Conventions

Throughout this document the term "RSVP" is used to indicate the Extensions to RSVP for LSP Tunnels (draft-ietf-mpls-rsvp-lsp-tunnel). Where necessary, non-labels RSVP support (RFC 2205) is explicitly referred to as "generic RSVP".

A glossary of terms and a table of references are provided at the end of the paper.

2. Background

2.1 MPLS

Multi-Protocol Label Switching (MPLS) is a new technology that will be used by many future core networks, including converged data and voice networks. MPLS does not replace IP routing, but will work alongside existing and future routing technologies to provide very high-speed data forwarding between Label-Switched Routers (LSRs) together with reservation of bandwidth for traffic flows with differing Quality of Service (QoS) requirements.

MPLS enhances the services that can be provided by IP networks, offering scope for Traffic Engineering, guaranteed QoS and Virtual Private Networks (VPNs).

The basic operation of an MPLS network is shown in the diagram below.

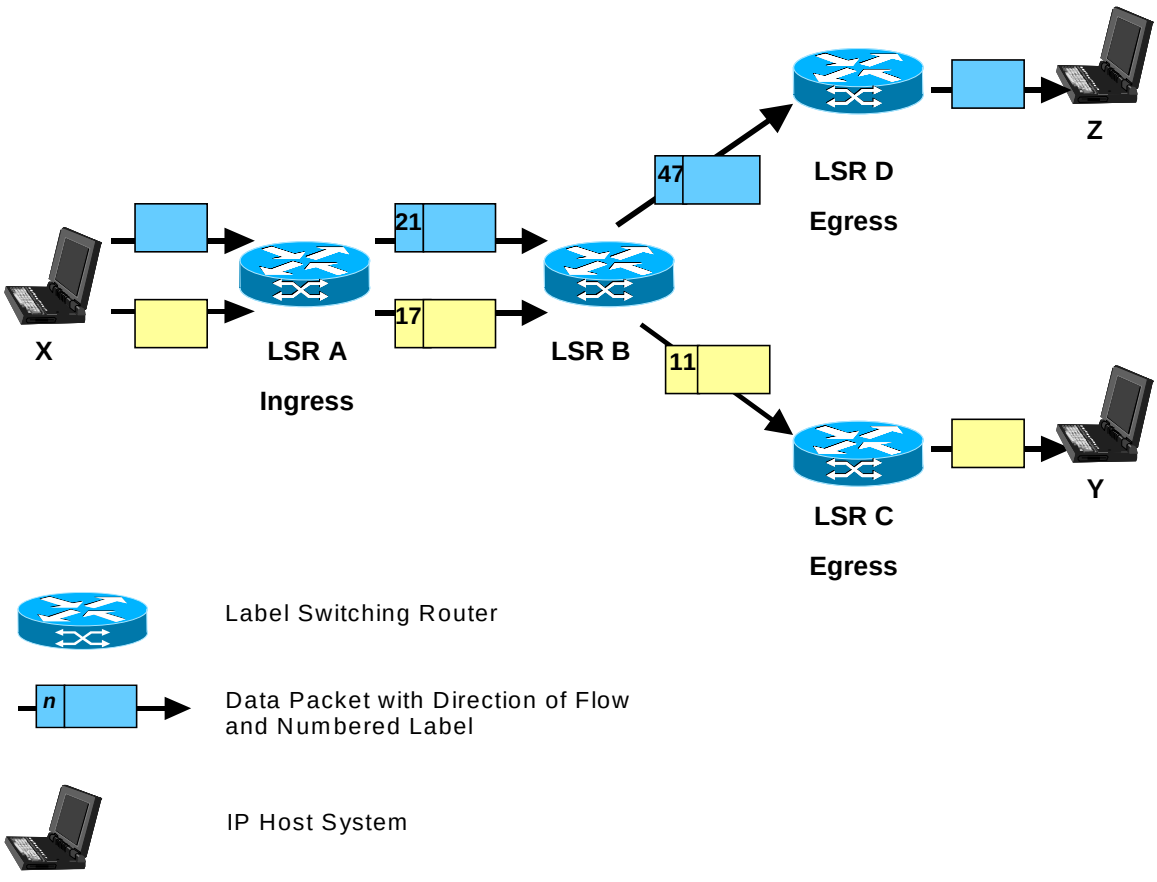

Fig.1: Two LSPs in an MPLS Network

MPLS uses a technique known as label switching to forward data through the network. A small, fixed-format label is inserted in front of each data packet on entry into the MPLS network. At each hop across the network, the packet is routed based on the value of the incoming label and dispatched to an outwards interface with a new label value.

The path that data traverses through a network is defined by the transition in label values, as the label is swapped at each LSR. Since the mapping between labels is constant at each LSR, the path is determined by the initial label value. Such a path is called a Label Switched Path (LSP).

At the ingress to an MPLS network, each packet is examined to determine which LSP it should use and hence what label to assign to it. This decision is a local matter but is likely to be based on factors including the destination address, the quality of service requirements and the current state of the network. This flexibility is one of the key elements that make MPLS so useful.

The set of all packets that are forwarded in the same way is known as a Forwarding Equivalence Class (FEC). One or more FECs may be mapped to a single LSP.

The diagram shows two data flows from host X: one to Y, and one to Z. Two LSPs are shown.

- LSR A is the ingress point into the MPLS network for data from host X. When it receives packets from X, LSR A determines the FEC for each packet, deduces the LSP to use and adds a label to the packet. LSR A then forwards the packet on the appropriate interface for the LSP.
- LSR B is an intermediate LSR in the MPLS network. It simply takes each labeled packet it receives and uses the pairing {incoming interface, label value} to decide the pairing {outgoing interface, label value} with which to forward the packet. This procedure can use a simple lookup table and, together with the swapping of label value and forwarding of the packet, can be performed in hardware. This allows MPLS networks to be built on existing label switching hardware such as ATM and Frame Relay. This way of forwarding data packets is potentially much faster than examining the full packet header to decide the next hop.

  In the example, each packet with label value 21 will be dispatched out of the interface towards LSR D, bearing label value 47. Packets with label value 17 will be re-labeled with value 11 and sent towards LSR C.
- LSR C and LSR D act as egress LSRs from the MPLS network. These LSRs perform the same lookup as the intermediate LSRs, but the {outgoing interface, label value} pair marks the packet as exiting the LSP. The egress LSRs strip the labels from the packets and forward them using layer 3 routing.

So, if LSR A identifies all packets for host Z with the upper LSP and labels them with value 21, they will be successfully forwarded through the network.

Note that the exact format of a label and how it is added to the packet depends on the layer 2 link technology used in the MPLS network. For example, a label could correspond to an ATM VPI/VCI, a Frame Relay DLCI, or a DWDM wavelength for optical networking. For other layer 2 types (such as Ethernet and PPP) the label is added to the data packet in an MPLS "shim" header, which is placed between the layer 2 and layer 3 headers.

2.2 Label Distribution

In order that LSPs can be used, the forwarding tables at each LSR must be populated with the mappings from {incoming interface, label value} to {outgoing interface, label value}. This process is called LSP setup, or Label Distribution.

The MPLS architecture document (draft-ietf-mpls-arch) does not mandate a single protocol for the distribution of labels between LSRs. In fact it specifically allows for multiple protocols for use in different scenarios.

Several different approaches to label distribution can be used depending on the requirements of the hardware that forms the MPLS network, and the administrative policies used on the network. The underlying principles are that an LSP is set up either in response to a request from the ingress LSR (downstream-on-demand), or pre-emptively by LSRs in the network, including the egress LSR (downstream unsolicited). It is possible for both to take place at once and for the LSP to meet in the middle.

In all cases, labels are allocated from the downstream direction (where downstream refers to the direction of data flow, and this means that are advertised towards the data source). Thus, in the example in Fig.1, LSR D informs LSR B that LSR B should use label 47 on all packets for host Z. LSR B allocates a new label (21), enters the mapping in its forwarding table, and informs LSR A that it should use label 21 on all packets for host Z.

Some possible options for controlling how LSPs are set up, and the protocols that can be used to achieve them, are described below.

- Hop-by-hop label assignment is the process by which the LSP setup requests are routed according to the next-hop routing towards the destination of the data. LSP setup could be initiated by updates to the routing table, or in response to a new traffic flow. The IETF MPLS Working Group has specified (but not mandated) LDP as a protocol for hop-by-hop label assignment. RSVP and CR-LDP can also be used.
- In Downstream Unsolicited label distribution, the egress LSR distributes the label to be used to reach a particular host. The trigger for this will usually be new routing information received at the egress node. Additionally, if the label distribution method is Ordered Control, each upstream LSR distributes a label further upstream. This effectively builds a tree of LSPs rooted at each egress LSR. LDP is currently the only protocol suitable for this mode of label distribution.

- Once LSPs have been established across the network, they can be used to support new routes as they become available. As the routing protocols (for example BGP) distribute the new routing information upstream, they can also indicate which label (i.e. which LSP) should be used to reach the destinations to which the route refers.
- If an ingress LSR wants to set up an LSP that does not follow the next-hop routing path, it must use a label distribution protocol that allows specification of an Explicit Route. This requires downstream-on-demand label distribution. CR-LDP and RSVP are two protocols that provide this function.
- An ingress LSR may also want to set up an LSP that provides a particular level of service by, for example, reserving resources at each intermediate LSR along the path. In this case, the route of the LSP may be constrained by the availability of resources and the ability of nodes to fulfill the quality of service requirements. CR-LDP and RSVP are two protocols that allow downstream-on-demand label distribution to include requests for specific service guarantees.

2.3 Explicit Routes

An Explicit Route (ER) is most simply understood as a precise sequence of steps from ingress to egress. An LSP in MPLS can be set up to follow an explicit path, i.e. a list of IP addresses. However, it does not need to be specified this fully.

For example, the route could specify only the first few hops. After the last explicitly specified hop has been reached, routing of the LSP proceeds using hop-by-hop routing.

A component of an explicit route may also be less precisely specified. A collection of nodes, known as an Abstract Node, may be presented as a single step in the route, for example by using an IP prefix rather than a precise address. The LSP must be routed to some node within this Abstract Node as the next hop. The route may contain several hops within the Abstract Node before emerging to the next hop specified in the Explicit Route.

An Explicit Route may also contain the identifier of an Autonomous System (AS). This allows the LSP to be routed through an area of the network that is out of the administrative control of the initiator of the LSP. The route may contain several hops within the Autonomous System before emerging to the next hop specified in the Explicit Route.

An Explicit Route may be classified as “strict” or “loose”. A strict route must contain only those nodes, Abstract Nodes or Autonomous Systems specified in the Explicit Route, and must use them in the order specified. A loose route must include all of the hops specified, and must maintain the order, but it may also include additional hops as necessary to reach the hops specified.

Once a loose route has been established it can be modified (as a hop-by-hop route could be) or it can be “pinned” so that it does not change.

Explicit routing is particularly useful to force an LSP down a path that differs from the one offered by the routing protocol. It can be used to distribute traffic in a busy network, to route around network failures or hot spots, or to provide pre-allocated back-up LSPs to protect against network failures.

2.4 Constrained Routes

The route that an LSP may take can be constrained by many requirements selected at the ingress LSR. An Explicit Route is an example of a constrained route where the constraint is the order in which intermediate LSRs may be reached. Other constraints can be imposed by a description of the traffic that is to flow and may include bandwidth, delay, resource class and priority.

One approach is for the ingress LSR to calculate the entire route based on the constraints and information that it has about the current state of the network. This leads it to produce an Explicit Route that satisfies the constraints.

The other approach is a variation on hop-by-hop routing where, at each LSR, the next hop is calculated using information held at that LSR about local resource availability.

The two approaches are combined if information about part of the route is unavailable (for example, it traverses an Autonomous System). In this case the route may be loosely specified in part, and explicitly routed using the constraints where necessary.

2.5 Resource Reservation

In order to secure promised services, it is not sufficient simply to select a route that can provide the correct resources. These resources must be reserved to ensure that they are not shared or "stolen" by another LSP.

The traffic requirements can be passed during LSP setup (as with constraint-based routing). They are used at each LSR to reserve the resources required, or to fail the setup if the resources are not available.

2.6 Traffic Engineering

Traffic Engineering is the process where data is routed through the network according to a management view of the availability of resources and the current and expected traffic. The class of service and quality of service required for the data can also be factored into this process.

Traffic Engineering may be under the control of manual operators. They monitor the state of the network and route the traffic or provision additional resources to compensate for problems as they arise. Alternatively, Traffic Engineering may be driven by automated processes reacting to information fed back through routing protocols or other means.

Traffic Engineering helps the network provider make the best use of available resources, spreading the load over the layer 2 links, and allowing some links to be reserved for certain classes of traffic or for particular customers.

One of the main uses for MPLS will be to allow improved Traffic Engineering on the ISP backbone networks.

2.7 Service Level Contracts

Many uses of the Internet require particular levels of service to be supplied. For example, voice traffic requires low delay and very small delay variation. Video traffic adds the requirement for high bandwidth. Customers increasingly demand service contracts that guarantee the performance and availability of the network.

In the past, in order to meet these requirements, network providers have had to over-provision their physical networks.

MPLS offers a good way to avoid this issue by allocating the network resources to particular flows using constraint-based routing of LSPs.

2.8 Virtual Private Networks

A Virtual Private Network (VPN) allows a customer to extend their private network across a wider public network in a secure way.

ISPs offer this service by ensuring that entry points to their network can exchange data only if they are configured as belonging to the same VPN.

MPLS LSPs provide an excellent way to offer this service over an IP network.

2.9 Meeting the Needs of the Modern Network

VPNs have been addressed with additions to the BGP routing protocol, but IP has not provided good solutions to the requirements set out in the previous three sections. There has been no way of providing a guarantee of service, because the network is connectionless. Destination-based routing along shortest path routes tends to overload some links and leave others unused.

A popular solution is to use an overlay network, for example running IP over ATM PVCs. This is notoriously hard to manage, because many resources must be configured at each router in the network, and because there are two distinct protocols to be configured. It also leads to scaling issues, with an order of n^2 connections needed in a network with n nodes.

MPLS allows the use of just one set of protocols in the network. Using MPLS to meet the aims described in the previous three sections while avoiding the problems described above requires a label distribution protocol that supports Explicit Routes and constraint-based routing.

There are currently two label distribution protocols that meet this definition: CR-LDP and RSVP. There is a debate about which of these protocols is preferable, which is most suitable for particular scenarios, and whether it is necessary to implement both of the protocols in an MPLS network.

Since the LSPs set up to support Traffic Engineering, Service Contracts and VPNs are all configured in the same way for RSVP and CR-LDP (through the Traffic Engineering MIB), they are referred to as Traffic Engineered LSPs.

3. Introduction To CR-LDP

CR-LDP is a set of extensions to LDP specifically designed to facilitate constraint-based routing of LSPs. Like LDP, it uses TCP sessions between LSR peers and sends label distribution messages along the sessions. This allows it to assume reliable distribution of control messages. The basic flow for LSP setup using CR-LDP is as shown below.

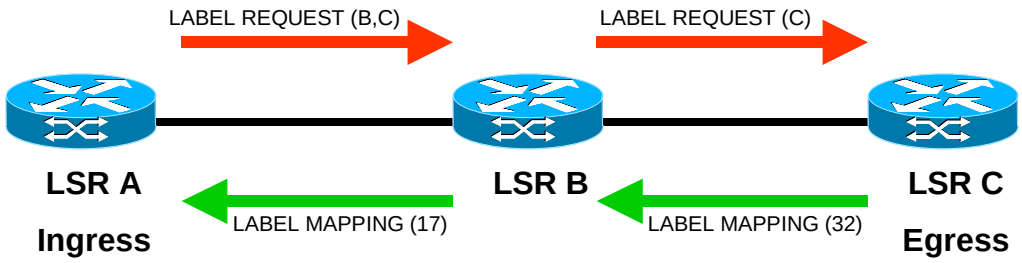

Figure.2 CR-LDP LSP Setup Flow

- The Ingress LSR, LSR A, determines that it needs to set up a new LSP to LSR C. The traffic parameters required for the session or administrative policies for the network enable LSR A to determine that the route for the new LSP should go through LSR B, which might not be the same as the hop-by-hop route to LSR C. LSR A builds a LABEL_REQUEST message with an explicit route of (B,C) and details of the traffic parameters requested for the new route. LSR A reserves the resources it needs for the new LSP, and then forwards the LABEL_REQUEST to LSR B on the TCP session.
- LSR B receives the LABEL_REQUEST message, determines that it is not the egress for this LSP, and forwards the request along the route specified in the message. It reserves the resources requested for the new LSP, modifies the explicit route in the LABEL_REQUEST message, and passes the message to LSR C. If necessary, LSR B may reduce the reservation it makes for the new LSP if the appropriate parameters were marked as negotiable in the LABEL_REQUEST.
- LSR C determines that it is the egress for this new LSP. It performs any final negotiation on the resources, and makes the reservation for the LSP. It allocates a label to the new LSP and distributes the label to LSR B in a LABEL_MAPPING message, which contains details of the final traffic parameters reserved for the LSP.
- LSR B receives the LABEL_MAPPING and matches it to the original request using the LSP ID contained in both the LABEL_REQUEST and LABEL_MAPPING messages. It finalizes the

- The processing at LSR A is similar, but it does not have to allocate a label and forward it to an upstream LSR because it is the ingress LSR for the new LSP.

4. Introduction To Labels Extensions To RSVP

Generic RSVP uses a message exchange to reserve resources across a network for IP flows. The Extensions to RSVP for LSP Tunnels enhances generic RSVP so that it can be used to distribute MPLS labels.

RSVP is a separate protocol at the IP level. It uses IP datagrams (or UDP at the margins of the network) to communicate between LSR peers. It does not require the maintenance of TCP sessions, but as a consequence of this it must handle the loss of control messages.

The basic flow for setting up an LSP using RSVP for LSP Tunnels is shown in Fig.3 below.

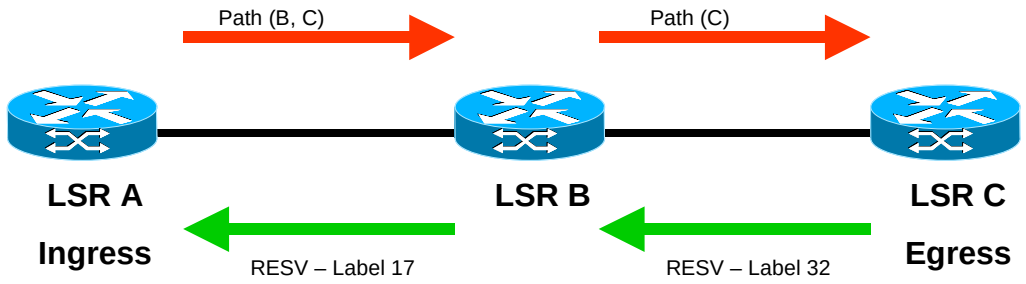

Figure.3 RSVP LSP Setup Flow

- The Ingress LSR, LSR A, determines that it needs to set up a new LSP to LSR C. The traffic parameters required for the session or administrative policies for the network enable LSR A to determine that the route for the new LSP should go through LSR B, which might not be the same as the hop-by-hop route to LSR C. LSR A builds a Path message with an explicit route of (B,C) and details of the traffic parameters requested for the new route. LSR A then forwards the Path to LSR B as an IP datagram.
- LSR B receives the Path request, determines that it is not the egress for this LSP, and forwards the request along the route specified in the request. It modifies the explicit route in the Path message and passes the message to LSR C.
- LSR C determines that it is the egress for this new LSP, determines from the requested traffic parameters what bandwidth it needs to reserve and allocates the resources required. It selects a label for the new LSP and distributes the label to LSR B in a Resv message, which also contains actual details of the reservation required for the LSP.

- LSR B receives the Resv message and matches it to the original request using the LSP ID contained in both the Path and Resv messages. It determines what resources to reserve from the details in the Resv message, allocates a label for the LSP, sets up the forwarding table, and passes the new label to LSR A in a Resv message.
- The processing at LSR A is similar, but it does not have to allocate a new label and forward this to an upstream LSR because it is the ingress LSR for the new LSP.

5. Comparative Analysis

The key differences between CR-LDP and RSVP are the reliability of the underlying transport protocol and whether the resource reservations are done in the forward or reverse direction. From these points come many of the other functional differences.

The table below summarizes the main technical similarities and differences between CR-LDP and RSVP for LSP Tunnels. The sections that follow explain in greater detail the implications of these technical differences between the protocols.

	CR-LDP support	RSVP Support
Transport	TCP	Raw IP
Security	Yes[1]	Yes[1]
Multipoint-to-Point	Yes	Yes
Multicast Support	No[2]	No[2]
LSP Merging	Yes[3]	Yes[3]
LSP State	Hard	Soft
LSP Refresh	Not needed	Periodic, hop-by-hop
High Availability	No	Yes
Re-routing	Yes	Yes
Explicit Routing	Strict and loose	Strict and loose
Route Pinning	Yes	Yes, by recording path
LSP Pre-emption	Yes, priority based	Yes, priority based
LSP Protection	Yes	Yes
Shared Reservations	No[4]	Yes
Traffic Parm Exchange	Yes	Yes
Traffic Control	Forward Path	Reverse Path
Policy Control	Implicit	Explicit
Layer 3 Protocol Indicated	No	Yes
Resource Class Constraint	Yes	No

Notes:

1. CR-LDP inherits any security applied to TCP. RSVP cannot use IPSEC but has its own authentication. See "Security" below.
2. Multicast support is currently not defined for any of the existing label distribution protocols.
3. See "Multipoint support" below for more details.
4. CR-LDP does not allow explicit sharing, but see "LSP Modification" below for details of changing the allocated resources.

5.1 Availability of Transport Protocol

The most obvious difference between CR-LDP and RSVP is the choice of transport protocol used to distribute the label requests. RSVP uses connectionless raw IP (or UDP packets at the margins of the network). CR-LDP uses UDP to discover MPLS peers, and uses connection-oriented TCP sessions to distribute label requests.

- Many operating systems are packaged with full IP stacks including UDP and TCP, but sometimes TCP is not available. On some platforms access to raw IP is restricted.

  Some existing ATM switches might not already incorporate an IP stack at all and one must be added to support either CR-LDP or RSVP.

  The availability and accessibility of the transport protocols may dictate which label distribution protocol is used, but is unlikely to be a major factor in the choice made by most MPLS equipment suppliers.

- RSVP requires that all received IP packets carrying RSVP messages are delivered to the RSVP protocol code without reference to the actual destination IP address in the packet. This feature may require a minor modification to the IP implementation.

See the "Security", "Scalability", "High Availability" and "Failure detection" sections below for details of how the choice of transport protocol affects other function provided in an MPLS system.

5.2 Security

TCP is vulnerable to denial of service attacks, where the performance of the TCP session can be seriously impacted by unauthorized access to the network. This could impact CR-LDP.

Authentication and policy control are specified for RSVP. This allows the originator of the messages to be verified (for example using MD5) and makes it possible to police unauthorized or malicious

reservation of resources. Similar features could be defined for CR-LDP but the connection-oriented nature of the TCP session makes this less of a requirement. TCP itself could make use of MD5.

IPSEC is a series of drafts from the IETF to provide authentication and encryption security for packets transported over IP. If IPSEC support is available in the IP stack it can be used by CR-LDP simply as part of the normal TCP/IP processing.

RSVP targets its Path messages at the egress LSR, not at the intermediate LSRs. This means that IPSEC cannot be used because the intermediate LSRs would find themselves unable to access the information in the Path messages.

5.3 Multipoint Support

Multipoint-to-point LSPs allow label switched paths to merge at intermediate LSRs, reducing the number of labels required in the system and sharing downstream resources. This approach works particularly well in packet-switched networks, but requires non-standard hardware in cell-switched networks such as ATM to prevent interleaving of cells. CR-LDP and RSVP support multipoint-to-point LSPs.

Point-to-multipoint (multicast) IP traffic is not addressed by the current version of the MPLS Architecture, so it is not supported by CR-LDP or Labels RSVP. Generic RSVP was originally designed to include resource reservation for IP multicast trees, so it may be easier to extend to support multicast traffic in the future. However, this is an area for further study in both protocols.

5.4 Scalability

The scalability of a protocol should be considered in terms of the network flows it uses, the resources needed to maintain the protocol state at each node, and the CPU load on each node.

All of this must be considered in the context of the way in which MPLS is to be used in the network. If trunk LSPs are to be used across the network to connect key edge points, there will be less demand on scalability than using one LSP per flow, or setting up LSPs based on the routing topology. The ability to merge LSPs also has a clear impact on scalability requirements, because data flows may be able to share resource allocations, and the number of labels needed in the network is reduced.

5.4.1 Network Flows

Both protocols have similar flows for label setup, sending an end-to-end request and replying with an end-to-end response.

RSVP is a *soft state* protocol. This means that it must periodically refresh the state of each LSP between adjacent nodes. This allows RSVP to pick up changes to the routing tree automatically. RSVP uses IP datagrams as its transport, meaning that control messages may be lost and that an

adjacent node may fail without notification. State refreshes help to make sure that LSP state is properly synchronized between adjacent nodes.

The network hit from this periodic refresh depends on the sensitivity to failure that is chosen by configuring the refresh timer. An RSVP Path message will be of the order of 128 bytes, increasing by 16 bytes per hop if an explicit route is used. A Resv message will be of the order of 100 bytes. With 10,000 LSPs on a link (a reasonably high number) and a refresh period of 30 seconds, this consumes over 600 kbits per second of link bandwidth. Whether this is significant depends on the link and what this is as a fraction of the traffic carried.

CR-LDP, however, does not require the LSRs to refresh each LSP after setup. This is achieved by using TCP as the transport for control messages. CR-LDP can assume reliable delivery of LABEL_REQUEST and LABEL_MAPPING messages. The use of TCP on the control path adds no overhead to the data path (where it is not used) and only 20 bytes to each control message.

In order to maintain connectivity with adjacent nodes, CR-LDP uses HELLO messages to check that the adjacent nodes are still active, and KEEPALIVE messages to monitor the TCP connections. These relatively small messages are exchanged periodically on a per link basis rather than per LSP, and so have virtually no impact on the throughput of the link.

Thus, CR-LDP should present a lower load to the network than RSVP in its present form.

At the time of writing, an Internet draft (draft-ietf-rsvp-refresh-reduct) is being prepared to document the latest ideas for reducing the refresh messages required by RSVP. The process described in the draft relies on refreshing many LSP states in a single RSVP BUNDLE message. This, together with the ability to indicate that nothing has changed on a given Path or Resv rather than having to send the entire normal payload, reduces the network refresh flows for RSVP so that they are closer to per LSR than per LSP.

The RSVP BUNDLE messages will still typically be larger than a single CR-LDP HELLO or KEEPALIVE, because they have to list the message IDs for each Path or Resv refreshed by the bundle. This is not a significant difference compared with the number of messages that would flow without this extension. This draft is likely to progress to RFC status quite quickly.

CR-LDP therefore currently presents a lower signaling load on the network itself than RSVP, but once refresh reduction is implemented in RSVP this will not be significant.

5.4.2 Data Storage Requirements

All connection-oriented protocols require that a certain amount of data is stored to maintain the connection state, both at the end points and to some extent at the intermediate nodes.

For RSVP the requirements are much the same across the network because the state information must be kept at each LSR to be periodically refreshed. This data must include the traffic parameters, resource reservations and explicit routes. It amounts to something of the order of 500 bytes per LSP.

CR-LDP requires the Ingress and Egress LSRs to maintain a similar amount of state information, including the traffic parameters and explicit routes. The total size of the state information required for CR-LDP is also around 500 bytes at the end points. At intermediate LSRs it is possible to reduce the storage requirements to around 200 bytes by not offering support for LSP modification (re-routing or changing resource requirements).

Note that the data-forwarding buffers required to guarantee QoS for LSPs are likely to be much larger than the storage needed to maintain state. Thus, the difference between RSVP and CR-LDP in an MPLS network where LSP modification is not required is made less significant.

5.4.3 CPU Load

The CPU load on the LSRs is determined by the number of messages they must parse and act upon, and by the complexity of the processing required for each message. The initial LSP setup flows are similar for both protocols, so the CPU load for this phase of an LSP's life will not differ greatly. However, RSVP's need to refresh state presents an additional load per LSP.

Even with refresh reduction, RSVP requires the exchange of complex aggregated refresh messages, each of which requires processing through the stored state information for a number of LSPs.

For example, an RSVP LSR that handles 10,000 LSPs concurrently needs to be able to parse and process the aggregated LSP refreshes at the rate of around 300 refreshes per second, if they are to be issued every 30 seconds. If each refresh message ID in the aggregated refresh messages requires several hundred source code instructions, this might represent well under 1% CPU load on a modern processor to maintain the existing LSPs. This is not a significant CPU load.

Note that the CPU load to re-route LSPs is likely to be even higher than the requirements for LSP setup, and would be incurred equally by CR-LDP and RSVP. In a network that is designed to handle failures without disrupting new connections, re-routing LSPs after network failure may become the limiting factor on network size well before the steady-state RSVP refreshes become an issue.

5.4.4 Summary

When the proposed RSVP refresh reduction extensions are adopted by the IETF, and implemented by the MPLS equipment vendors, the scalability of both RSVP and CR-LDP is largely determined by the number of LSR peers in a network. Without these extensions, RSVP scalability is determined by the number of LSPs that transit a node.

It is therefore important to determine the likely number of LSPs transiting the LSRs in a network. Where LSPs are set up per data flow or per IGP route this is much more likely to be an issue than in

networks which use Traffic Engineering to set up a smaller number of large LSP tunnels. The amount of LSP merging in the network also makes a considerable difference.

Both protocols provide solutions that should scale to accommodate the largest of networks in use today. Ultimately, RSVP scalability is the more suspect, even with refresh reduction, if the number of LSPs transiting a single node is very large. It is too early to tell whether this problem will actually be encountered in practice.

5.5 High Availability

Availability is a measure of the percentage of time that a node is in service. Equipment vendors typically claim high availability for their boxes when they attain availability levels in the region of 99.999% ("5-nines").

High Availability is a matter of detecting failures and handling them in a timely manner without any – or with only minimal – disruption to service. Detection and survival of link failures is covered in the following sections. This section is concerned with detection of and recovery from local failures, specifically hardware and software fault-tolerance and the use of online software upgrades to minimize system downtime.

Survival of LSPs across software failure, and provision of online software upgrades in an MPLS system, are software implementation issues and should be addressed by any vendor serious about the provision of networking solutions. Tolerance of hardware faults relies on hardware detection and reporting of failures, on the availability of backup hardware, and on a suitably designed software implementation.

Because RSVP is designed to run over a connectionless transport, it lends itself well to a system that must survive hardware failures or online software upgrades. Any control steps that are lost during the failover to the replacement backup system can be recovered by the state refresh processing that is built into RSVP.

CR-LDP, on the other hand, assumes reliable delivery of control messages and so is not well placed to survive failover. Additionally, it is particularly hard to make TCP fault tolerant (a problem familiar to BGP implementers), with the result that a failover to a backup TCP stack results in the loss of the TCP connections. This is interpreted by CR-LDP as a failure in all of the associated LSPs, which must subsequently be re-established from the ingress LSR.

Metaswitchis researching ways to extend CR-LDP to allow it to survive online software upgrades and hardware faults.

Until such extensions are added to CR-LDP, RSVP implementations will be able to provide better solutions for highly available MPLS networks.

5.6 Link and Peer Failure Detection

Where two LSRs are directly connected using a point-to-point link technology, such as ATM, the failure of LSPs can usually be detected by monitoring the state of the interfaces for the LSP. If, for example, an ATM link suffers loss of signal, both CR-LDP and RSVP can use the interface failure notification to detect the failure of the LSP.

If two LSRs are connected over a shared medium, such as Ethernet, or are indirectly connected over a WAN cloud, for example using an ATM PVC, they may not necessarily receive a link failure notification from the link hardware. LSP failure detection then relies on techniques inherent in the signaling protocols. So long as normal signaling traffic is flowing nothing else is necessary, but in stable state, additional processing is required to detect a failure.

- CR-LDP uses an exchange of LDP HELLO and KEEPALIVE messages to validate that the LSR peer and link are still active. Although TCP has a built-in keepalive system, this is typically too slow to respond to link and peer failures for the demands of MPLS LSPs.
- In RSVP, Path and Resv refresh messages serve to provide background traffic that indicates that the link is still active. However, to keep the per-LSP refresh traffic in a relatively stable network to a minimum, the refresh timer would be set quite high. To address this problem, an extension has been added to Labels RSVP so that RSVP HELLO messages can be exchanged to prove that the link and peer LSR are still active.

The failure detection techniques and speed are therefore similar for both CR-LDP and RSVP, provided that RSVP uses the HELLO extensions. MPLS failure detection is much faster for directly attached LSRs.

5.7 Re-routing

This section discusses the provision of a new route for an LSP after notification of a failure or a topology change. Pre-programming of alternate paths for an LSP is known as LSP protection and is discussed in the next section.

A strictly specified explicit route cannot be re-routed except by the ingress LSR (initiator). Consequently, failure at some point of an LSP must be reported to the ingress, effectively bringing down the whole LSP. However, a loosely specified portion of an explicit routed LSP, and any part of a hop-by-hop routed LSP, may be re-routed if

- a failure of a link or neighbor is detected (this is called Local Recovery)
- a better route becomes available
- the resources for the LSP are required for a new, higher priority LSP (this is called Pre-emption).

Re-routing is most easily managed from the ingress (including re-routing of strictly specified LSPs) and is supported by both CR-LDP and RSVP, though with slightly different characteristics.

- An LSR using RSVP can install a new route by simply refreshing the Path for an LSP to a different next-hop as soon as the alternate route is available/required. The old path can be left to time out because refreshes will no longer be sent. However, this wastes resources on the old path.
- "Make-before-break" is a mechanism whereby the old path is used (and refreshed) while the new path is set up, and then the LSR performing the re-routing swaps to using the new path and tears down the old path. This basic technique can be used to avoid double reservation of resources in both CR-LDP (using the **modify** value for the action flag on the LABEL_REQUEST) and RSVP (using shared explicit filters).

Re-routing of loosely specified parts of LSPs at intermediate LSRs when a "better" route becomes available can lead to thrashing in unstable networks. To prevent this, a loosely specified part of a route may be "pinned":

- In CR-LDP this is simply a matter of flagging the loose part of the explicit route as pinned. This means that once the route has been set up, it is treated as though it had been strictly specified and cannot be changed.
- In RSVP, pinning requires some additional processing. The initial route is specified with a loose hop. The Record Route object is used on the Path and Resv messages to feed back the selected route to the ingress. The ingress can use this information to re-issue the Path message with a strictly specified explicit route.

Both RSVP and CR-LDP offer flexible approaches to re-routing and make-before-break provisioning of LSPs. CR-LDP relies on a recent addition to the specification for make-before-break processing, while RSVP requires additional message exchanges to pin a route.

5.8 LSP Modification

LSP modification, for example, to change the traffic parameters for an LSP, is an equivalent operation to re-routing, though the change of route is optional for LSP modification. This means that the function is always present in RSVP and will be present in CR-LDP provided that the **modify** value of the action flag on a LABEL_REQUEST is supported by the implementations (this is a relatively recent addition to the CR-LDP drafts and so early implementations might not support modification of LSPs).

Note that support for the **modify** value of the action flag in CR-LDP leads to increased data occupancy, bringing intermediate LSR occupancy up to a figure similar to that required at RSVP intermediate LSRs.

See the re-routing section for details of how RSVP and CR-LDP handle this function.

5.9 LSP Protection

LSP protection is the pre-programming of alternate paths for an LSP with automatic switching to an alternate path if the primary path fails. Though conceptually similar to re-routing, LSP protection is normally assumed to be a much more time-critical operation, in which the aim is to switch over to the new path with absolute minimal disruption (less than 50 ms is a common target) to the data traffic on the LSP.

Several levels of LSP protection are possible in both protocols.

- The simplest form of LSP protection is for the Ingress or an intermediate LSR to immediately attempt to re-route the LSP when it is notified of a failure. This is possible in both protocols, but may result in a relatively slow failover (typically at least several seconds) as the failure must be propagated and the new LSP signaled. This will not be fast enough for applications such as voice.
- Much faster LSP protection can be achieved if the link between two LSRs is protected by a layer 2 protection scheme such as SONET, or APS built on top of ATM. Such protection is transparent to the LSP and could be deployed with either protocol.
- Layer 2 protection can be expensive to implement and is localized to a single hop in the LSP. Link protection of this sort is, in any case, no protection against the failure of an individual LSR. At the time of writing this white paper the MPLS working group is considering schemes to provide pre-programmed alternate routes for an LSP across a wider portion of the LSP path, and automatic switching of traffic to one of the alternate routes after a failure. Protection switching may be performed by intermediate routers in the LSP path, not just the Ingress LSR. These extensions to MPLS for protection switching are not yet fully specified but should be available for both CR-LDP and RSVP.

RSVP and CR-LDP will probably both be good protocols for providing LSP protection.

5.10 Lambda Networking

Lambda networking presents an interesting set of problems for an MPLS implementation. The full advantages of wavelength switching can only be encompassed if LSPs are switched in hardware without recourse to software. In this respect the lambda network is similar to an ATM network; the MPLS labels are identified with individual wavelengths.

The number of wavelengths is, however, very small – too small for the likely number of LSPs transiting any one link. Additionally, the capabilities of an individual wavelength are far in excess of the normal requirements of an LSP, so that such a one-to-one mapping would be highly wasteful of network resources.

The MPLS working group is currently considering some early drafts that address these issues. The approach being looked at involves sharing a wavelength between multiple LSPs, and is equally applicable to RSVP and CR-LDP.

5.11 Traffic Control

Significantly, CR-LDP and RSVP perform resource reservation at different times in the process of LSP setup.

CR-LDP carries the full traffic parameters on the LABEL_REQUEST. This allows each hop to perform traffic control on the forward portion of LSP setup. The traffic parameters can be negotiated as the setup progresses, and the final values are passed back on the LABEL_MAPPING allowing the admission control and resource reservation to be updated at each LSR. This approach means that an LSP will not be set up on a route where there are insufficient resources.

RSVP carries a set of traffic parameters, the Tspec on the Path message. This describes the data that is likely to use the LSP. Intermediate LSPs can examine this information and could make routing decisions based on it. However, it is not until the egress LSR is reached that the Tspec is converted to a Flowspec returned on the Resv message, which gives details of the resource reservation required for the LSP. This means that the reservation does not take place until the Resv passes through the network, with the result that LSP set up may fail on the selected route because of resource shortage.

RSVP includes an optional function (adspec) whereby the available resources on a link can be reported on the Path message. This allows the egress LSR to know what resources are available, and modify the Flowspec on the Resv accordingly. Unfortunately, not only does this function require that all the LSRs on the path support the option, but it has an obvious window where resources reported on a Path message may already have been used by another LSP by the time the Resv is received.

A partial solution for RSVP LSRs lies within the implementation, which could make a provisional reservation of resources as it processes the Path message. This reservation can only be approximate since it is based on the Tspec not the Flowspec, but it can considerably ease the problem.

CR-LDP offers a slightly tighter approach to traffic control especially in heavily used networks, but individual RSVP implementations can provide a solution that is almost as good.

5.12 Policy Control

RSVP is specified to allow the Path and Resv messages to carry a policy object with opaque content. This data is used when processing messages to perform policy-based admission control. This allows Labels RSVP to be tied closely to policy policing protocols such as COPS (Common Open Policy Service) using the Internet draft “COPS Usage for RSVP”.

By contrast, CR-LDP currently only carries implicit policy data in the form of the destination addresses, and the administrative resource class in the traffic parameters.

5.13 Layer 3 Protocol

Although an LSP can carry any data, there are occasions when knowledge of the layer 3 protocol can be useful to an intermediate or egress LSR.

If an intermediate LSR is unable to deliver a packet (e.g. because of a resource failure) it can return an error packet specific to the layer 3 protocol (such as ICMP for IP packets) to notify the sender of the problem. For this to work, the LSR that detects the error must know the layer 3 protocol in use.

Also, at an egress, it may help the LSR to forward data packets if the layer 3 protocol is known.

RSVP identifies a single payload protocol during LSP setup, but there is no scope within the protocol for CR-LDP to do this. Even RSVP is unable to help when more than one protocol is routed to a particular LSP.

Recent discussions led by Metaswitch in the MPLS Working Group have considered options for identifying the payload protocol in CR-LDP, and for marking the payload packets so that their protocol can be easily determined.

5.14 QoS and Diff-Serv

CR-LDP and RSVP have different approaches to Quality of Service (QoS) parameters.

The RSVP Tspec object carried on Path messages describes the data that will flow rather than the QoS that is required from the connection. Various RFCs and Internet drafts describe how to map from different QoS requirements to the Tspec (for example, RFC 2210 - The Use of RSVP with IETF Integrated Services).

The CR-LDP specification is more explicit about how the information carried on a LABEL_REQUEST message is mapped for QoS.

Support for Diff-Serv (IP Differentiated Services) is addressed by an Internet draft (draft-ietf-mpls-diff-ext), which defines extensions to LDP, RSVP and CR-LDP. If implemented, this draft extends the full function of Diff-Serv to an MPLS network.

5.15 Provision of VPNs

Virtual Private Networks (VPNs) are an important feature of the service provided by ISPs to their customers. VPNs allow physically private networks to be extended to encompass remote sites by connecting them through the Internet.

A customer in these circumstances expects to be able to preserve their IP addresses (which might not be globally unique) and to have the security of their data guaranteed. MPLS can provide an excellent solution as described in RFC 2547.

Both CR-LDP and RSVP are suitable MPLS signaling protocols for VPNs over MPLS.

5.16 Voice over IP and Voice over MPLS

Voice over IP (VoIP) is an exciting development in Internet technology. The concept of a single infrastructure for voice and data, providing faster, cheaper and value-added services, is very attractive.

MPLS is set to be a major component in VoIP networks, offering connection-oriented paths with resource reservation through the connectionless Internet.

Voice over MPLS (VoMPLS) is the term given to the transfer of voice traffic over an MPLS network. This could involve establishing LSP Tunnels to act as trunks for multiple calls, or setting up LSPs for the duration of individual calls. Alternatively, VoMPLS could mean sending voice samples as labeled MPLS packets without including IP headers.

Whichever approach is used, both CR-LDP and RSVP are suitable MPLS signaling protocols.

5.17 MIB Management

Traffic Engineered LSPs can be managed at their ingress and inspected at their egress through the MPLS Traffic Engineering MIB. This MIB is currently in an early stage which slightly favors CR-LDP, but new drafts will be produced that fully support RSVP and CR-LDP.

5.18 Acceptance/Availability

There is currently no clear “winner” between RSVP and CR-LDP in terms of market acceptance.

Although generic RSVP has been available for a number of years from a variety of equipment vendors, and in that sense is an established network protocol, the changes required to a generic RSVP stack to add support for Labels RSVP are non-trivial, and hence Labels RSVP is in many respects a new protocol.

CR-LDP is based on ideas that have been implemented in proprietary networks for as long as ten years, but as an IETF protocol it is very new and somewhat unproven.

Manufacturers are currently hedging their bets, favoring one of the two protocols, but planning to offer both in the long run. It is often suggested that Nortel Networks Corp. and Nokia Corp. favor CR-LDP, while Cisco Systems Inc. and Juniper Networks Inc. favor RSVP.

The ITU has a Study Group (SG13) investigating general aspects of network architectures, interfaces, performance and interworking. As yet they have not devoted much energy to MPLS, although the current preference within the submissions that they have received is for CR-LDP.

5.19 Interoperability

There are two interoperability issues to be addressed. Do two implementations support a compatible set of options, and do they interpret the specifications in the same way?

The option sets are functions of the flexibility of the protocol. RSVP has more implementation options than CR-LDP and so is perhaps at more risk. However, the protocol is specified to allow interworking between implementations that support different function sets. An IETF MPLS draft (draft-loa-mpls-cap-set) provides a list of capability sets to allow implementations to identify the functions that they provide.

Interoperability testing is clearly the only way to prove that two implementations interwork correctly. Interoperability forums are being set up in many places including

- the University of New Hampshire InterOperability Labs (UNH IOL)
- George Mason University in Washington DC with the support of UUNET
- EANTC in Berlin
- NetWorld & Interop events.

All of these forums will include work on RSVP and CR-LDP.

Participation in interoperability events is a clear requirement for all MPLS software vendors. Testing for hardware vendors will be a combination of involvement in interoperability events, in-house testing with competitors' equipment, and collaborative work with other vendors. Hardware vendors have a right to expect the support of their software suppliers during interoperability testing.

5.20 Interoperation with Other Label Distribution Methods

LDP is another label distribution protocol specified by the IETF. It is used to set up basic (unconstrained), end-to-end LSPs using "hop-by-hop" routing. It is also used to request and distribute labels for multiple or single hops, a feature which is useful in conjunction with topology-driven LSP setup.

Since CR-LDP is built as an extension to LDP it is easier for a CR-LDP implementation also to support the features of LDP. RSVP is entirely different and, although RSVP also supports hop-by-hop LSP setup, a second protocol stack must be implemented to support the features of LDP.

6. Summary

CR-LDP and Labels RSVP are both good technical solutions for setting up and managing Traffic Engineered LSPs. Early versions of both protocols had some functional omissions, but these are being fixed by subsequent Internet drafts so that the level of function provided by each protocol is similar.

Some key differences in the structure of the protocols and the underlying transport mean that the support that the protocols can provide will never converge completely. These differences and the differences in speed and scope of deployment will be the main factors that influence vendors when they are selecting a protocol.

The choice between RSVP and CR-LDP should be guided by the function of the target system. What LSP setup model will be used? How stable are the LSPs – do they represent permanent trunks or short-duration calls? How large is the network and how complex is it? Is this a stand-alone network or must the components interwork with other hardware and other networks?

A final consideration must be the robustness of the hardware solution. What level of fault tolerance is required? How important is high availability?

Two informational Internet drafts may help guide the choice of protocol.

- Applicability Statement for Extensions to RSVP for LSP-Tunnels
- Applicability Statement for CR-LDP

7. Glossary

AS: Autonomous System. A part of the network under a single administration and usually running a single routing protocol for internal routing.

BGP: Border Gateway Protocol. The Exterior Gateway Protocol used for distributing routes over the Internet backbone.

CR-LDP: Constraint-based Routed Label Distribution Protocol. Extensions to LDP to set up Traffic Engineered LSPs, as defined in the Internet Draft "Constraint-based LSP Setup using LDP".

DLCI: Data Link Circuit Identifier. The labels used in Frame Relay that are equivalent to MPLS labels.

EGP: Exterior Gateway Protocol. Any routing protocol used for distributing routes between Autonomous Systems. Also the name of the first such protocol, now superseded by BGP.

ER: Explicit Route. A route specified during setup and not determined by the routing protocol at each hop across the network.

IGP: Interior Gateway Protocol. Any routing protocol used for distributing routes within a single Autonomous System.

Labels RSVP: Extensions to RSVP to set up Traffic Engineered LSPs.

LDP: Label Distribution Protocol. A protocol defined by the IETF for distributing labels to set up MPLS LSPs.

LSP: Label Switched Path. A data forwarding path determined by labels attached to each data packet where the data is forwarded at each hop according to the value of the labels.

LSP Tunnel: A Traffic Engineered LSP capable of carrying multiple data flows.

LSR: Label Switching Router. A component of an MPLS network that forwards data based on the labels associated with each data packet.

MPLS: MultiProtocol Label Switching. A standardized technology that provides connection-oriented switching based on IP routing protocols and labeling of data packets.

OSPF: Open Shortest Path First. A common routing protocol that provides IGP function.

RSVP: Resource ReSerVation Protocol (RFC 2205). A setup protocol designed to reserve resources in an Integrated Services Internet.

VoIP: Voice over IP. The process of carrying voice over an IP network.

VoMPLS: Voice over MPLS. The process of carrying voice traffic over MPLS LSPs with or without using IP.

VPI/VCI: Virtual Path Identifier / Virtual Channel Identifier. The labels used in ATM layer 2 networks that are equivalent to MPLS labels.

VPN: Virtual Private Network. A private network provided by securely sharing resources with a wider, common network.

8. References

The most recent versions of Internet drafts and RFCs for MPLS can be found listed at the MPLS Working Group's home page at http://www.ietf.org/html.charters/mpls-charter.html

Internet drafts and RFCs relevant to generic RSVP can be found listed at the RSVP Working Group's home page at http://www.ietf.org/html.charters/rsvp-charter.html

The following drafts are of particular relevance.

RFC 2205	Resource ReSerVation Protocol (RSVP)
draft-ietf-mpls-rsvp-lsp-tunnel	Extensions to RSVP for LSP Tunnels
draft-ietf-rsvp-refresh-reduct	RSVP Refresh Reduction Extensions
draft-ietf-mpls-cr-ldp	Constraint-Based LSP Setup Using LDP
draft-ietf-mpls-crlsp-modify	LSP Modification Using CR-LDP
draft-ietf-mpls-ldp	LDP specification
draft-ietf-mpls-te-mib	MPLS Traffic Engineering Management Information Base Using SMIv2
RFC 2207	RSVP Extensions for IPSEC Data Flows
RFC 2210	The Use of RSVP with IETF Integrated Services
draft-ietf-mpls-diff-ext	MPLS Support of Differentiated Services
RFC 2547	BGP/MPLS VPNs
draft-ietf-mpls-rsvp-tunnel-applicability	Applicability Statement for Extensions to RSVP for LSP-Tunnels
draft-ietf-mpls-crldp-applic	Applicability Statement for CR-LDP
draft-loa-mpls-cap-set	MPLS Capability Set

9. About Metaswitch

Metaswitch is a privately owned technology company based in London, UK. We have US offices in Alameda, CA, Reston, VA, and Boxborough, MA.

Our Network Protocols Division is the leading developer and supplier of (G)MPLS, OSPF(-TE), ISIS(-TE), BGP, VPN, RIP, PIM, IGMP, MLD, ATM, MGCP, Megaco, SCTP, SIP, VoIP Conferencing, Messaging, Directory and SNA portable products. Customers include Alcatel, Cisco, Fujitsu, Hewlett-Packard, Hitachi, IBM Corp., Microsoft, Nortel and Sun.

Our company culture focuses on building software of consistently high quality, developed and supported by engineers who are with Metaswitch for the long term.

- Founded in 1981, we have over 450 employees, of whom 280 are engineers. The average length of service of engineers at Metaswitch is 8 years, and the annual attrition rate is 3%.
- Throughout this period, Metaswitch has been consistently profitable with profits exceeding 15% of revenue. 2007-2008 revenues were $118m with $22m profit.
- Over 90% of revenue is generated from exports and 80% is from customers in the US (so we are very used to working with American companies).
- The company is privately held by top-tier investment firms Francisco Partners and Sequoia Capital, as well as the Employee Benefit Trust (EBT). As part of this ownership structure, Metaswitch distributes a share of profit to all employees, equitably rewarding them for their contribution and encouraging long-term commitment.
- As a private company with an emphasis on long-term stability, we are not driven by the short-term requirements of quarterly profit statements. This means that we can concentrate on providing software as we would like – that is, developing high quality implementations of complex technologies.

The DC-MPLS product family provides OEMs with a flexible source code solution with the same high quality architecture and support for which Data Connection's other communications software products are renowned. It runs within Data Connection's existing high performance portable execution environment (the N-BASE). This provides extensive scalability and flexibility by enabling distribution of protocol components across a wide range of hardware configurations from DSPs to line cards to specialized signaling processors. It has fault tolerance designed in from the start, providing hot swap on a component by component basis on failure or upgrade of hardware or software.

DC-MPLS is suitable for use in a wide range of IP switching and routing devices including Label Switch Routers (LSRs) and Label Edge Routers (LERs). Support is provided for a range of label distribution methods including Resource ReSerVation Protocol (RSVP), Constraint-based Routed Label Distribution Protocol (CR-LDP) and Label Distribution Protocol (LDP). The rich feature set gives DC-MPLS the performance, scalability and reliability required for the most demanding MPLS applications.

DC-MPLS integrates seamlessly with Data Connection's other converged network software products, and uses the same proven N-BASE communications execution environment. The N-BASE has been ported to a large number of operating systems including VxWorks, pSOS, Chorus, Nucleus, Solaris, HP-UX and Windows NT, and has been used on all common processors including x86, i960, Motorola 860, Sparc, IDT and MIPS. Proprietary OSs and chipsets can be supported with minimal effort.

All of the Metaswitch protocol implementations are built with scalability, distribution across multiple processors and fault tolerance architected in from the beginning. We have developed extremely consistent development processes that result in on-time delivery of highly robust and efficient software. This is backed up by an exceptionally responsive and expert support service, staffed by engineers with direct experience in developing the protocol solutions.

About the authors

Paul Brittain is a senior networking architect with Data Connection.

Adrian Farrel was originally Architect and Development Manager for the DC-MPLS product family, and contributed to the GMPLS drafts in the IETF. He is now a Member of Technical Staff with Movaz Networks Inc.

Metaswitch Networks
100 Church Street
Enfield
EN2 6BQ
England
+44 20 8366 1177
http://www.metaswitch.com