Multiprotocol Label Switching

Multiprotocol Label Switching (MPLS) is a mechanism in high-performance telecommunications networks which directs and carries data from one network node to the next. MPLS makes it easy to create "virtual links" between distant nodes. It can encapsulate packets of various network protocols.

MPLS is a highly scalable, protocol agnostic, data-carrying mechanism. In an MPLS network, data packets are assigned labels. Packet-forwarding decisions are made solely on the contents of this label, without the need to examine the packet itself. This allows one to create end-to-end circuits across any type of transport medium, using any protocol. The primary benefit is to eliminate dependence on a particular Data Link Layer technology, such as ATM, frame relay, SONET or Ethernet, and eliminate the need for multiple Layer 2 networks to satisfy different types of traffic. MPLS belongs to the family of packet-switched networks.

MPLS operates at an OSI Model layer that is generally considered to lie between traditional definitions of Layer 2 (Data Link Layer) and Layer 3 (Network Layer), and thus is often referred to as a "Layer 2.5" protocol. It was designed to provide a unified data-carrying service for both circuit-based clients and packet-switching clients which provide a datagram service model. It can be used to carry many different kinds of traffic, including IP packets, as well as native ATM, SONET, and Ethernet frames.

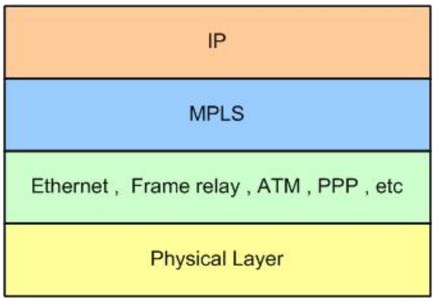
A number of different technologies were previously deployed with essentially identical goals, such as frame relay and ATM. MPLS technologies have evolved with the strengths and weaknesses of ATM in mind. Many network engineers agree that ATM should be replaced with a protocol that requires less overhead, while providing connection-oriented services for variable-length frames. MPLS is currently replacing some of these technologies in the marketplace. It is highly possible that MPLS will completely replace these technologies in the future, thus aligning these technologies with current and future technology needs.[1]

In particular, MPLS dispenses with the cell-switching and signaling-protocol baggage of ATM. MPLS recognizes that small ATM cells are not needed in the core of modern networks, since modern optical networks (as of 2008) are so fast (at 40 Gbit/s and beyond) that even full-length 1500 byte packets do not incur significant real-time queuing delays (the need to reduce such delays — *e.g.*, to support voice traffic — was the motivation for the cell nature of ATM).

At the same time, MPLS attempts to preserve the traffic engineering and out-of-band control that made frame relay and ATM attractive for deploying large-scale networks.

While the traffic management benefits of migrating to MPLS are quite valuable (better reliability, increased performance), there is a significant loss of visibility and access into the MPLS cloud for IT departments.[2]

History



MPLS Layer

MPLS was originally proposed by a group of engineers from Ipsilon Networks, but their "IP Switching" technology, which was defined only to work over ATM, did not achieve market dominance. Cisco Systems, Inc., introduced a related proposal, not restricted to ATM transmission, called "Tag Switching".[3] It was a Cisco proprietary proposal, and was renamed "Label Switching". It was handed over to the IETF for open standardization. The IETF work involved proposals from other vendors, and development of a consensus protocol that combined features from several vendors' work.

One original motivation was to allow the creation of simple high-speed switches, since for a significant length of time it was impossible to forward IP packets entirely in hardware. However, advances in VLSI have made such devices possible. Therefore the advantages of MPLS primarily revolve around the ability to support multiple service models and perform traffic management. MPLS also offers a robust recovery framework[4] that goes beyond the simple protection rings of synchronous optical networking (SONET/SDH).

How MPLS works

MPLS works by prefixing packets with an MPLS header, containing one or more "labels". This is called a label stack. Each label stack entry contains four fields:

- A 20-bit label value.
- a 3-bit *Traffic Class* field for QoS (quality of service) priority (experimental) and ECN (Explicit Congestion Notification).
- a 1-bit *bottom of stack* flag. If this is set, it signifies that the current label is the last in the stack.
- an 8-bit TTL (time to live) field.

These MPLS-labeled packets are switched after a label lookup/switch instead of a lookup into the IP table. As mentioned above, when MPLS was conceived, label lookup and label switching were faster than a routing table or RIB (Routing Information Base) lookup because they could take place directly within the switched fabric and not the CPU.

The entry and exit points of an MPLS network are called label edge routers (LER), which, respectively, **push** an MPLS label onto an incoming packet and **pop** it off the outgoing packet. Routers that perform routing based only on the label are called label switch routers (LSR). In some applications, the packet presented to the LER already may have a label, so that the new LER pushes a second label onto the packet. For more information see penultimate hop popping.

Labels are distributed between LERs and LSRs using the “Label Distribution Protocol” (LDP).[5] Label Switch Routers in an MPLS network regularly exchange label and reachability information with each other using standardized procedures in order to build a complete picture of the network they can then use to forward packets. *Label Switch Paths (LSPs)* are established by the network operator for a variety of purposes, such as to create network-based IP virtual private networks or to route traffic along specified paths through the network. In many respects, LSPs are not different from PVCs in ATM or Frame Relay networks, except that they are not dependent on a particular Layer 2 technology.

In the specific context of an MPLS-based virtual private network (VPN), LSRs that function as ingress and/or egress routers to the VPN are often called PE (Provider Edge) routers. Devices that function only as transit routers are similarly called P (Provider) routers. See RFC 2547.[6] The job of a P router is significantly easier than that of a PE router, so they can be less complex and may be more dependable because of this.

When an unlabeled packet enters the ingress router and needs to be passed on to an MPLS tunnel, the router first determines the forwarding equivalence class (FEC) the packet should be in, and then inserts one or more labels in the packet's newly-created MPLS header. The packet is then passed on to the next hop router for this tunnel.

When a labeled packet is received by an MPLS router, the topmost label is examined. Based on the contents of the label a *swap*, *push* (*impose*) or *pop* (*dispose*) operation can be performed on the packet's label stack. Routers can have prebuilt lookup tables that tell them which kind of operation to do based on the topmost label of the incoming packet so they can process the packet very quickly.

In a *swap* operation the label is swapped with a new label, and the packet is forwarded along the path associated with the new label.

In a *push* operation a new label is pushed on top of the existing label, effectively "encapsulating" the packet in another layer of MPLS. This allows hierarchical routing of MPLS packets. Notably, this is used by MPLS VPNs.

In a *pop* operation the label is removed from the packet, which may reveal an inner label below. This process is called "decapsulation". If the popped label was the last on the label stack, the packet "leaves" the MPLS tunnel. This is usually done by the egress router, but see Penultimate Hop Popping (PHP) below.

During these operations, the contents of the packet below the MPLS Label stack are not examined. Indeed transit routers typically need only to examine the topmost label on the stack. The forwarding of the packet is done based on the contents of the labels, which allows "protocol-independent packet forwarding" that does not need to look at a protocol-dependent routing table and avoids the expensive IP longest prefix match at each hop.

At the egress router, when the last label has been popped, only the payload remains. This can be an IP packet, or any of a number of other kinds of payload packet. The egress router must therefore have routing information for the packet's payload, since it must forward it without the help of label lookup tables. An MPLS transit router has no such requirement.

In some special cases, the last label can also be popped off at the penultimate hop (the hop before the egress router). This is called Penultimate Hop Popping (PHP). This may be interesting in cases where the egress router has lots of packets leaving MPLS tunnels, and thus spends inordinate amounts of CPU time on this. By using PHP, transit routers connected directly to this egress router effectively offload it, by popping the last label themselves.

MPLS can make use of existing ATM network infrastructure, as its labeled flows can be mapped to ATM virtual circuit identifiers, and vice versa.

Installing and removing MPLS paths

There are two standardized protocols for managing MPLS paths: CR-LDP (Constraint-based Routing Label Distribution Protocol) and RSVP-TE, an extension of the Resource Reservation Protocol (RSVP) for traffic engineering.[7] As of February 2003 the IETF MPLS working group deprecated CR-LDP and decided to focus purely on RSVP-TE.[8] Furthermore, there exist extensions of the BGP protocol that can be used to manage an MPLS path.[9] [10] [11]

An MPLS header does not identify the type of data carried inside the MPLS path. If one wants to carry two different types of traffic between the same two routers, with different treatment by the core routers for each type, one has to establish a separate MPLS path for each type of traffic.

MPLS and IP

MPLS cannot be compared to IP as a separate entity because it works in conjunction with IP and IP's IGP routing protocols. MPLS gives IP networks simple traffic engineering, the ability to transport Layer 3 (IP) VPNs with overlapping address spaces, and support for Layer 2 pseudowires (with Any Transport Over MPLS, or ATOM - see Martini draft). Routers with programmable CPUs and without LSP can be

- explicitly configured hop by hop,
- dynamically routed by the Constrained Shortest Path First (CSPF) algorithm, or
- configured as a loose route that avoids a particular IP or that is partly explicit and partly dynamic.

In a pure IP network, the shortest path to a destination is chosen even when it becomes more congested. Meanwhile, in an IP network with MPLS Traffic Engineering CSPF routing, constraints such as the RSVP bandwidth of the traversed links can also be considered, such that the shortest path with available bandwidth will be chosen. MPLS Traffic Engineering relies upon the use of TE extensions to OSPF or IS-IS and RSVP. Besides the constraint of RSVP bandwidth, users can also define their own constraints by specifying link attributes and special requirements for tunnels to route (or not to route) over links with certain attributes.[12]

MPLS local protection (Fast Reroute)

In the event of a network element failure when recovery mechanisms are employed at the IP layer, restoration may take several seconds which is unacceptable for real-time applications such as VoIP.[13] [14] [15] In contrast, MPLS local protection meets the requirements of real-time applications with recovery times comparable to those of SONET rings of less than 50 ms.[13] [15] [16]

MPLS and Multicast

Multicast was for the most part an after-thought in MPLS design. It was introduced by point-to-multipoint RSVP-TE.[17] It was driven by Service Provider requirements to transport broadband video over MPLS. Since the inception of RFC 4875 there has been tremendous surge in interest and deployment of MPLS multicast and this has led to several new developments both in the IETF and in shipping products.

Comparison of MPLS versus Frame Relay

Frame relay aimed to make more efficient use of existing physical resources, which allow for the underprovisioning of data services by telecommunications companies (telcos) to their customers, as clients were unlikely to be utilizing a data service 100 percent of the time. In more recent years, frame relay has acquired a bad reputation in some markets because of excessive bandwidth overbooking by these telcos.

Telcos often sell frame relay to businesses looking for a cheaper alternative to dedicated lines; its use in different geographic areas depended greatly on governmental and telecommunication companies' policies.

AT&T is currently (as of June 2007) the largest frame relay service provider in the United States, with local networks in 22 states, plus national and international networks. This number is expected to change between 2007 and 2009 when most of these frame relay contracts expire. Many customers are likely to migrate from frame relay to MPLS over IP or Ethernet within the next two years, which in many cases will reduce costs and improve manageability and performance of their wide area networks.[18]

Comparison of MPLS versus ATM

While the underlying protocols and technologies are different, both MPLS and ATM provide a connection-oriented service for transporting data across computer networks. In both technologies, connections are signaled between endpoints, connection state is maintained at each node in the path, and encapsulation techniques are used to carry data across the connection. Excluding differences in the signaling protocols (RSVP/LDP for MPLS and PNNI:Private Network-to-Network Interface for ATM) there still remain significant differences in the behavior of the technologies.

The most significant difference is in the transport and encapsulation methods. MPLS is able to work with variable length packets while ATM transports fixed-length (53 byte) cells. Packets must be segmented, transported and re-assembled over an ATM network using an adaptation layer, which adds significant complexity and overhead to the data stream. MPLS, on the other hand, simply adds a label to the head of each packet and transmits it on the network.

Differences exist, as well, in the nature of the connections. An MPLS connection (LSP) is unidirectional—allowing data to flow in only one direction between two endpoints. Establishing two-way communications between endpoints requires a pair of LSPs to be established. Because 2 LSPs are required for connectivity, data flowing in the forward direction may use a different path from data flowing in the reverse direction. ATM point-to-point connections (virtual circuits), on the other hand, are bidirectional, allowing data to flow in both directions over the same path (only SVC ATM connections are bidirectional; PVC ATM connections are unidirectional).

Both ATM and MPLS support tunneling of connections inside connections. MPLS uses label stacking to accomplish this while ATM uses *virtual paths*. MPLS can stack multiple labels to form tunnels within tunnels. The ATM virtual path indicator (VPI) and virtual circuit indicator (VCI) are both carried together in the cell header, limiting ATM to a single level of tunnelling.

The biggest single advantage that MPLS has over ATM is that it was designed from the start to be complementary to IP. Modern routers are able to support both MPLS and IP natively across a common interface allowing network operators great flexibility in network design and operation. ATM's incompatibilities with IP require complex adaptation, making it comparatively less suitable for today's predominantly IP networks.

MPLS deployment

MPLS is currently in use in IP-only networks and is standardized by the IETF in RFC 3031. It is deployed to connect as few as two facilities to very large deployments. For example, in the retail sector, it is not uncommon to see deployments of 2000 to 5000 locations to communicate transaction data to a headquarters data center.

In practice, MPLS is mainly used to forward IP datagrams and Ethernet traffic. Major applications of MPLS are telecommunications traffic engineering and MPLS VPN.

Competitors to MPLS

MPLS can exist in both IPv4 environment (IPv4 routing protocols) and IPv6 environment (IPv6 routing protocols). The major goal of MPLS development - the increase of routing speed - is no longer relevant because of the usage of ASIC, TCAM and CAM-based switching. Therefore the major usage of MPLS is to implement limited traffic engineering and Layer 3/Layer 2 “service provider type” VPNs over existing IPv4 networks. The main competitors to MPLS are Provider Backbone Bridges (PBB), and MPLS-TP that also provide services such as service provider Layer 2 and Layer 3 VPNs. L2TPv3 has been suggested as a competitor, but has not reached any wider success.

IEEE 1355 is a completely unrelated technology that does something similar in hardware.

IPv6 references: Grossetete, Patrick, IPv6 over MPLS, Cisco Systems 2001; Juniper Networks IPv6 and Infranets White Paper; Juniper Networks DoD's Research and Engineering Community White Paper.

Access to MPLS networks

MPLS supports a range of access technologies, including T1, ATM, frame relay and DSL

See also

- IPv4
- IPv6
- VPLS, virtual private LAN service over MPLS.

Books

- "Deploying IP and MPLS QoS for Multiservice Networks: Theory and Practice" by John Evans, Clarence Filsfils (Morgan Kaufmann, 2007, ISBN 0-12-370549-5)
- Rick Gallaher's MPLS Training Guide (ISBN 1932266003)

External links

- MPLS Working Group [19], IETF.

References

[1] Applied Data Communications (A Business-Oriented Approach) James E. Goldman & Phillip T. Rawles, 2004 (ISBN 0-471-34640-3)

[2] Routers Hold key to MPLS Measurement (http://www.networkperformancedaily.com/2006/11/routers_hold_the_key_to_mpls_n_1.html)

[3] Y. Rekhter et al., Tag switching architecture overview (http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=650179), Proc. IEEE 82 (December, 1997), 1973–1983.

[4] Framework for Multi-Protocol Label Switching (MPLS)-based Recovery (http://tools.ietf.org/html/rfc3469), RFC 3469, V. Sharma & F. Hellstrand, February 2003

[5] Thomas, B; Gray, E (January). "RFC 3037: LDP Applicability" (http://www.ietf.org/rfc/rfc3037.txt). . Retrieved 2007-09-21

[6] "RFC 2547 - BGP/MPLS VPNs" (http://tools.ietf.org/html/rfc2547). tools.ietf.org. . Retrieved 2008-06-11.

[7] *RFC 3209 — RSVP-TE: Extensions to RSVP for LSP Tunnels* (http://tools.ietf.org/html/rfc3209), IETF, December 2001,

[8] RFC 3468

[9] starting with RFC 2547

[10] Carrying Label Information in BGP-4 (http://www.ietf.org/rfc/rfc3107.txt), RFC 3107, Rekhter Y and Rosen E, May 2001.

[11] Graceful Restart Mechanism for BGP with MPLS (http://www.ietf.org/rfc/rfc4781.txt), RFC4781, Yakov Rekhter and Rahul Aggarwal, January 2007.

[12] de Ghein, Luc. *MPLS Fundamentals*. pp. 249–326.

[13] Aslam et al. (2005-02-02). *NPP: A Facility Based Computation Framework for Restoration Routing Using Aggregate Link Usage Information* (http://cat.inist.fr/?aModele=afficheN&cpsidt=16546516). QoS-IP 2005 : quality of service in multiservice IP network. . Retrieved 2006-10-27.

[14] Raza et al.. *Online routing of bandwidth guaranteed paths with local restoration using optimized aggregate usage information* (http://ieeexplore.ieee.org/iel5/9996/32109/01494347.pdf). IEEE-ICC 2005. . Retrieved 2006-10-27.

[15] Li Li et al.. *Routing bandwidth guaranteed paths with local restoration in label switched networks* (http://ieeexplore.ieee.org/iel5/49/30289/01391048.pdf?tp=&arnumber=1391048&isnumber=30289). IEEE Journal on Selected Areas in Communications. . Retrieved 2006-10-27.

[16] Kodialam et al.. *Dynamic Routing of Locally Restorable Bandwidth Guaranteed Tunnels using Aggregated Link Usage Information* (http://ieeexplore.ieee.org/iel5/7321/19793/00916720.pdf). IEEE Infocom. pp. 376–385. 2001. . Retrieved 2006-10-27.

[17] RFC 4461: Extensions to Resource Reservation Protocol — Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs), R. Aggarwal, D. Papadimitriou, S. Yasukawa, Eds.

[18] "AT&T — Frame Relay and IP-Enabled Frame Relay Service (Product Advisor)" (http://www.researchandmarkets.com/reportinfo.asp?report_id=451078), Research and Markets, June 2007.

[19] http://www.ietf.org/html.charters/mpls-charter.html

Article Sources and Contributors

Multiprotocol Label Switching *Source*: http://en.wikipedia.org/w/index.php?oldid=380671480 *Contributors*: 7x10, ALM scientist, ARTamb, Adamantios, Akr7577, Aldie, Alvestrand, Amacanally, Amoliyer80, Andrew.whittle777, Andypdavis, Apelleti, Aphstein, Ari.takanen, Arved, Athaenara, Ayengar, Azgs, Aznxsinn, Baeksu, Bdw4, BertK, Bertra g, Biot, Bluezy, Bobblewik, Boffob, BorgHunter, Brodes18, Bseirup, Can't sleep, clown will eat me, Carlwp, Cburnett, Celada, Charles Matthews, Choij, Cleduc, Comsign, Conversion script, Cwacker1, Cwolfsheep, Cybercobra, DKEdwards, Dave6, Derek Ross, Disdero, Dman727, Dmktg, Dnas, Dogcow, Dosman, Dougher, DougsTech, Drugonot, Dwandelt, DylanW, Echoray, Edcolins, ElKevbo, EnOreg, Entheta, Faizullah, Falcon8765, Falcorian, FreedomFireCom, Frumphammer, Fuzheado, Gadfium, Gaius Cornelius, Grin, Guoc, Guy Harris, Hankwang, Hcberkowitz, Hom sepanta, Hu12, Humble226, Iceberg3k, Indon, Inkling, IoannisK, Isnow, J.delanoy, JTN, Jadams76, JeffHolton, Jim.henderson, Jnc, JoanneB, Josephsukhbir, Juansempere, Kaz219, Kubanczyk, Legoktm, Leujohn, Lightmouse, LimoWreck, Loftenter, Ludovic.ferre, Mam711, Mange01, Mani1, Manscher, Marudubshinki, Mattisse, Mauls, MeekMark, Merovingian, Michael Hardy, Mindmatrix, Minesweeper, Miralin, Monkeyman, Moorlock, Mulad, Mzajac, Naveen dev83, Nealcardwell, Neale Monks, Niteowlneils, Nubiatech, Nunocordeiro, Oheckmann, Ohnoitsjamie, Oli Filth, Ortolan88, PAStheLoD, Paul, Pensees, Peroba, Petri Krohn, Pgallert, Phatom87, Phisches, Pinkadelica, PropellerHead, R. S. Shaw, Ranshe, Rchandra, Red123abc, RichardNeill, Rickgall1, Rjwilmsi, Robert Brockway, Rohitthakral, Ronz, Rossami, Ruwanindika, Sbeyer, Schallee, Sf internet tech, Sgarson, Shantavira, Spearhead, Splatg, Sridev, Srleffler, Stevehughes, Sudeoo, Ta bu shi da yu, TastyPoutine, Tckma, Tedmund, The Anome, Thesauce101, Thingg, Three3wall, TimBentley, Tmaufer, Tohd8BohaithuGh1, Tsarbanis, Ttwaring, Twilsonb, Visiting1, Wk muriithi, Wrs1864, Wynnfarm, Xlaran, Zoicon5, 429 anonymous edits

Image Sources, Licenses and Contributors

File:Mpls rotocolstack.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:Mpls_rotocolstack.jpg *License*: Public Domain *Contributors*: Ruwanindika, 1 anonymous edits

License