A Short Introduction to: Generalized Multiprotocol Label Switching (GMPLS)
The Need for a Signaling Plane

- Internet traffic is growing exponentially.
- Networks have to be scalable (easily upgradeable).
- Routing (forwarding) needs to be made easier and faster (much more data to be relayed).
- IP routing cannot easily specify routes through the network.
- The same network has to provide for different services (voice, data).
- Different services may need to have guaranteed services.
- Traditional IP does not provide easily for the above, a signaling plane is needed to interwork with IP.
IP over ATM

- ATM was the first attempt to provide with a “signaling layer” for IP.
- Yet, in the evolution of ATM, ATM (connection oriented, VPI, VCI) moved far away from the traditional IP (connection less, IP address) architectural model (resource allocation, multicasting).
- ATM was “the next thing,” native ATM applications were started to be developed. Yet soon everybody realized that the major function of ATM was the forwarding of IP packets (ATM came way too late).
- IP over ATM became a serious issue to tackle and turned out to be extremely complex (lots of different protocols).
ATM and IP were developed disregarding each other’s architecture and behavior.

The mapping of IP packets into ATM cells is inefficient.

The price of ATM equipment never dropped and bit-rates remained relatively low.

ATM did not solve the routing problem but introduced additional problems.
Toshiba’s Cell Switching Router

- The first attempt to control an ATM switch with IP routing (e.g., OSPF) and signaling (e.g., RSVP) messages.
- Toshiba claimed that such approach would make all other ATM signaling unnecessary.
- The IETF did not find this problem to be a viable problem at that time (1994) and thus they did not form a working group for standardization.
IP Switching

- Ipsilon (a startup company) has announced IP Switching in 1996.
- IP Switching maps forwarding to routing.
- IP switches are needed not ATM switches to efficiently work with IP.
- ATM signaling is not needed, IP needs to be added a new signaling plane (ATM over IP is way too complex).
- IP switching received considerable attention, and started the decay of the ATM era.
Cisco’s Tag Switching

- A few months after Ipsilon’s IP Switching Cisco announced another approach to label switching – Tag Switching.

- Forwarding tables are not implicitly set-up by IP flows anymore.

- Tag switching bid farewell to ATM technology (yet it could work on ATM).

- Cisco announced that it will pursue a standardization with IETF on Tag Switching, this effort kicked off the MPLS working group.
IBM’s ARIS

- Shortly after Tag and IP switching, IBM has announced Aggregate Route-based IP Switching (ARIS).
- Similar to Tag switching (separate control plane defined).
- Most of the ideas from ARIS were incorporated into MPLS.
The IETF MPLS WG.

- The “Birds of a Feather” session for the IP-Tag-ARI switching standardization effort at IETF in 1996 was the most well attended session in IETF history.

- Some people argued, saying that faster ATM switches and IP router would make this problem irrelevant, yet

- The IETF MPLS Working Group was established in April 1997.

- The name “multiprotocol” has been decided on because they did not want to name it either IP or Tag or ARI switching but show that it will be a combined effort of all of those.
MPLS – Concepts
Forwarding and Control Components
MPLS

- MPLS can manage (packet-switched) traffic in IP core networks no matter whether they relay voice or data.
- MPLS offers a control plane to IP with which resources can be allocated (reserved) to flows, and flows can have a pre-determined route over the network.
An MPLS Sample

- LSR - Label Switched Router
Label Switching (and Swapping)

- A small - fixed label is “inserted in front of the packet.”
- Forwarding (determining the output port) in LSRs is based on the label and the input port (optional) by label switching forwarding tables (table is indexed by the label for fast lookup).
- Before sending the packet off, an LSR will update the label of the outgoing packet (label swapping).
Label Switched Paths and Forwarding Equivalence Classes

- The path packets follow is determined by the label of the ingress LSR.
- Such paths are called Label Switched Paths (LSP).
- A set of packets that should be labeled with the same label value on entry to the MPLS network, and that will therefore follow the same LSP, is known as a Forwarding Equivalence Class (FEC).
Carrying Labels

- Labels can be carried in the link layer headers of existing DLL framings, e.g., the VPI, VCI fields of ATM, or the DLCI field of frame relay.

- Shim labels: the shim label is inserted between the link layer header (if any) and the network layer header. This can be (and is) used over Ethernet, FDDI, point-to-point links, etc.
# Multiprotocol Above and Below

## Network layer protocols
- IPv6
- IPv4
- IPX
- AppleTalk

## Link layer protocols
- Ethernet
- FDDI
- ATM
- Frame Relay
- Point-to-point

<table>
<thead>
<tr>
<th>IPv6</th>
<th>IPv4</th>
<th>IPX</th>
<th>AppleTalk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Label switching</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethernet</td>
<td>FDDI</td>
<td>ATM</td>
<td>Frame Relay</td>
</tr>
</tbody>
</table>

*Forwarding Component*
Control Component

- Responsible for:
  - Distributing routing information among LSRs (protocols)
  - Algorithms to convert the distributed information into forwarding tables and decision making processes.

- Similarity between the routing subcomponent and IP routing subcomponent, in fact the routing subcomponent includes IP routing protocols (e.g., OSPF, BGP, etc.), yet routing information dissemination is not enough for label switching.

- Additional procedures are needed for:
  - Bindings between FECs and labels
  - Dissemination of such bindings to other LSRs
  - Constructing the forwarding tables based on routing information and label mappings.
Control Component

- Network Layer Routing Protocols (e.g., OSPF, BGP).
- Protocols and Algorithms to create binding between FECs and labels.
- Procedures to disseminate label binding information.

Label switching forwarding table (next hop mapping)
Label Distribution Protocols

- MPLS does not define what kind of protocol should be used for label distribution.
- Protocols such as LDP can be used for the sole purpose of label exchange.
- Protocols such as RSVP and CR-LDP may be used to tie label distribution with traffic engineering, where QoS provisioning and bandwidth availability can also be considered.
Tunnels

Since packets are transmitted between the two LSRs that terminate the LSP, the intermediate LSRs do not have to investigate the network layer header thus creating tunnels (like VPIs in ATM), thus a tunnel carries packet between ingress and egress LSRs opaquely.
Label Stacks

- If several LSPs are routed over the same route using a large number of LSRs, they can be added another label (stacked) for these LSRs and thus the number of labels (thus forwarding tables and control information) can be reduced (like VCI in ATM).

- This is a very important feature of MPLS enabling MPLS to scale with the network (hierarchy).

- Label popping is the process of eliminating the top label from the label stack (and thus returning to a finer granularity).

- Label stacks are also essential to support Virtual Private Networks (VPN) over public IP infrastructures.
The optical layer is connection oriented (circuit switched), Lightpaths are easy to be established.

Lightpaths can be seen as LSPs between ingress and egress OXCs.

Multiprotocol Lambda Switching (MP\(\lambda\)S) was defined as a control plane for optical networks.

MPLS and MP\(\lambda\)S were then unified and called Generalized MPLS – GMPLS.
Generalized Multiprotocol Label Switching (GMPLS)
GMPLS

- A label can be associated to anything sufficient to carry a traffic flow.
- A label could correspond to a TDM slot on a wavelength (timeslot label and/or SONET/SDH labels), a wavelength (wavelength label), a band (waveband label), or an entire fiber (whole fiber label).
- Note, that all the labels have a one-to-one mapping with bandwidth. This is due to the inherited (train like) behavior of the optical domain.
- GMPLS labels are not fixed (32 bit) labels anymore but arbitrary byte length. Thus LDPs have to be updated to handle this situation.
Requesting Generalized Labels

The basic label agreeing protocols are unchanged from MPLS. They can be upstream or downstream binding. Downstream binding:

1. Upstream LSR send request to downstream LSR (using RSVP, or CR-LDP). This request also contains information on requested bandwidth.
2. The downstream LSR allocates a label value to the request.
3. The downstream LSR sends a response to the upstream to communicate the selected value.
Generalized Labels

- Switches have to advertise what kind of labels they are capable to switch.
- Currently supported label types (relevant for optical networks) are:
  - ANSI PDH
  - ETSI PDH
  - SDH/SONET (what timeslots, what speed, what concatenation, etc.)
  - Lambda (which wavelength)
  - Fiber (which fiber)
Constraining Label Choice

- Let us assume that LSRs want to set up wavelength based LSPs (Lightpaths).
- Since labels are assigned by the downstream nodes, these nodes have to select a wavelength that is available throughout the entire path. Thus advice has to be given to them on how to select (constrain) their pick of the label (or wavelength).
- This can be done by either label sets or explicit label control.
Label Sets

- In the downstream direction (where the setup request propagates), nodes also transmit their available wavelength pool.

- This pool is recalculated at each intermediate node, and labels are removed (or added) from the pool if the current LSR (OXC) cannot support it on the outgoing link.

- The egress router will receive the available pool on the link and select (and reverse propagate) the binding for the LSP. (Contention can arise between parallel requests for labels).
Explicit Label Control

- The ingress LSR can “insist” on a label for the path.
- Paths can be computed using labels for each hop (exact opposite than before – like source routing).
- Request fails if the predefined label is not available at a downstream node.
Out of Band Signaling

- With original MPLS, signaling information traveled on the same interfaces and routes the data did.
- In Optical networks there is a control networks (data communication network) for OXCs to exchange small bandwidth information, thus the above will not be true anymore. The following new issues have to be dealt with:

1. LSP setup messages will have to propagate between physically adjacent OXCs, containing the next hop IP addresses and the interfaces numbers.
2. Signaling messages have to have guaranteed (low latency) delivery.
3. Signaling messages have to be able to indicate as to which interface they refer to.
Extending Route Calculation

- Since the topologies for the signaling and data networks are different, first the next IP on the data link has to be calculated and then (using this as the final address) the next hop on the signaling interface has to be calculated.

- Thus protocols such as OSPF has to be enhanced to carry topology information for both the data as well as the signaling network.
Signaling Message Encapsulation

- In RSVP (CR-LDP does not have the same problem) signaling, each message is addressed to the egress LSR (not to the next hop). LSRs on the path will intercept these messages (by detecting a flag that is set).

- Problem: In case of separate signaling and data networks, the signaling data may find shortcuts to jump over intermediate nodes, or a transit node that should not be involved intercepts the message and attempts to process it.

- Solutions:
  - The signaling message can be addressed to the next hop (change in existing protocols).
  - Double encapsulated with two IP headers, one for the final and one for the next hop (using the flag) – better.
Switch Programming Latency

- With regular RSVP, switches are programmed as the signaling message propagates backwards. As soon as this reverse signaling message reaches the ingress node again, the LSP is set up.

- Currently OXCs are slow to reconfigure (transients after turning mirrors, etc.). This needs to be accounted for!

- The latency can be reduced by the introduction of suggested labels: each LSR selects a label (incorporating it into the path message) which it thinks will be suitable and programs its XC assuming that its prediction was correct. If the returning signal (Resv) supports that choice, the signaling message can be passed on right away otherwise the switch has to reconfigure itself (no lose situation).
RSVP (not an issue with CR-LDP) requires regular refreshes to keep LSPs alive (the shorter the refresh period the more robust the process but the more soft state overhead). The overhead may be too large for an optical LSP, which may be long lived.

This overhead can be reduced by the introduction of message ids: flow control, where each message has to be acknowledged by a two-way handshake.
Bidirectional LSPs

- LSPs often need to be bidirectional.
- In MPLS, this was done by the setup of two unidirectional LSPs, requiring two signaling processes, not ensuring that both LSPs are routed along the same route and requiring the higher layers to coordinate on the virtual binding of the two LSPs.
- GMPLS extends MPLS with bidirectional LSPs:
  1. The Path message also carries an upstream label, if this label is not acceptable by the downstream node, it will immediately reject the Path message (PathErr) and may return an acceptable label set.
  2. If the upstream label is acceptable then the downstream OXC will continue with the above process (depending on whether it has label modification capabilities—e.g., wavelength conversion).
  3. By the time the Path message reaches the egress, the reverse path is set up. The forward path will be set up by the Resv messages as before.
Bidirectional LSPs

- Using the bidirectional path setup, the egress node can assume the availability of the reverse link before it actually has been established.

- This can be avoided by either:
  1. Each node while processing the Path message will start configuring the reverse path and pass on the path message only if it is sure that it can do it (egress has to wait some time too).
  2. By the egress node waiting for a ResvConf message coming from the ingress after it has received the Resv message (three way handshake) – slower but more robust.
Asymmetric Bidirectional LSPs

- In highly utilized networks it may happen that although bidirectional LSPs are not available, two diversely routed unidirectional LSPs could be established (this is called an asymmetric bidirectional LSP).

- Yet, the tying of the two endpoints have to be made transparent and should be mapped to a single virtual entity (not yet standardized).