COMP9332
Network Routing & Switching
Lecture 12 - Switching in IP Networks with MPLS

www.cse.unsw.edu.au/~cs9332
Lecture Overview

• Concept of switching with ATM switches
• Concept of label switching - MPLS
• Implementing MPLS on ATM switches
• Traffic engineering with MPLS
Forwarding vs Switching
Problems with forwarding

• Routing table lookup is slow
  – Routing tables are exploding
• Table lookup limits future growth of Internet
  – Cannot scale to ultra high-speed backbone
• Traffic engineering is difficult
Benefits of switching

• No longest prefix match
• Lookup is indexed and very fast
• Traffic engineering is easier
Switching with ATM Networks
Characteristics of ATM

- Packet-based technology like IP, but ...
- Unlike IP, it is *connection oriented*
  - IP is datagram based (connectionless)
- ATM packets (called cells) are *switched* inside ATM switches
ATM Protocol model

ATM is a 3-layer model

ATM terminal

ATM Adaptation Layer (AAL)
ATM Layer
Physical Layer

ATM switch

ATM Layer
Physical Layer

ATM terminal

ATM Adaptation Layer (AAL)
ATM Layer
Physical Layer
End devices such as routers use all three layers, while switches use only the bottom two layers.
ATM Cells

- 53-byte long
- 5-byte header
- 48-byte payload
- Slightly different format for User-Network and Network-Network communications
Advantages of Short, Fixed Size Cells

- Minimise delay for voice/video traffic
- Minimise delay variance (jitter)
- Easier to process (fast switching)
Without Short Cells

A voice packet waits behind a large data packet
With Short Cells

Voice packet to be transmitted after Data #1

• Voice packet can go immediately after data packet #1
• Waiting for voice is reduced significantly
ATM Interfaces

- User-Network Interface (UNI)
  - public and private
- Network-Network Interface (NNI)
Cell Formats

GFC: Generic flow control
VPI: Virtual path identifier
VCI: Virtual channel identifier

PT: Payload type
CLP: Cell loss priority
HEC: Header error control

 Payload data

UNI Cell

 Payload data

NNI Cell
<table>
<thead>
<tr>
<th>Virtual Channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>• ATM is connection oriented</td>
</tr>
<tr>
<td>• An ATM connection is called virtual channel/connection/circuit (VC)</td>
</tr>
<tr>
<td>• A VC must be established between two end points before data can be transmitted</td>
</tr>
<tr>
<td>• Tens of thousands of VCs per fiber</td>
</tr>
<tr>
<td>• Different VC, different bandwidth, QoS</td>
</tr>
<tr>
<td>• User data and signaling on different VCs</td>
</tr>
</tbody>
</table>
Virtual Paths

- Virtual Private Network over public ATM
  - many individual VCs between two sites
  - switches need to switch VCs individually even if they start and end at the same points
- VCs starting and ending at the same points can be bundled in virtual paths or VPs
- Switches may switch only VPs, not aware of individual VCs inside the VP
Virtual Paths & Virtual Channels

Physical Channel

VP

VC

VC

VC

VP

VC

VC

VC

VP

VC

VC

VC

CSE, UNSW
Virtual Path Connection (VPC)
Virtual Channel Connection (VCC)
VP and VC Labels

concept of hop-by-hop labeling

- VPs and VCs are labeled with integer identifiers
  - Virtual Path Identifier (VPI)
  - Virtual Channel Identifier (VCI)
- VPI/VCIs are assigned hop by hop
- Same end-to-end ATM connection may have different VPI/VCI in different hop
- Hop-by-hop label assignment allows reuse of the same number at different hops
- The network maps VPI/VCIs at adjacent hops to create an end-to-end virtual circuit
Virtual Channel Identifier (VCI)

- Each VC is identified by a unique VCI on a given link
- VCIs therefore have only local significance
- An end-to-end ATM connection traversing several links will be identified by different VCIs at different links
- VCI translation occurs at the link boundary
- VCIs are assigned to a connection during call set-up
Reserved Channels

- VCI 0-31 are reserved
  - Some of them have been pre-allocated to specific functions

<table>
<thead>
<tr>
<th>Reserved Channel</th>
<th>VCI</th>
<th>VPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle/unassigned cell</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F4 segment OAM</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>F4 end-to-end OAM</td>
<td>x</td>
<td>4</td>
</tr>
<tr>
<td><strong>Signaling</strong></td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>RM for VPs</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>ILMI</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LANE configuration direct</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>PNNI routing</td>
<td>0 or x</td>
<td>18</td>
</tr>
</tbody>
</table>
VP and VC Switch

- Two types of ATM switch
- VP switch does not look at VCIs, switching is based on VPIs only
- VCI does not change when passing through a VP switch; VPI may change
- VC switch looks at both VPI and VCI
- VCI (as well as VPI) may change when passing through a VC switch
VP Switch

VPI=5  VCI=15  VP Switch  VPI=7  VCI=15  VP Switch  VPI=9  VCI=15  VP Switch  VPI=6  VCI=15
VC Switch

VPI=5  VCI=15
VP Switch

VPI=7  VCI=15
VC Switch

VPI=9  VCI=24
VP Switch

VPI=6  VCI=24

CSE, UNSW
Permanent Virtual Circuit

No signalling required to setup a VC; the VC is pre-setup.
Permanent Virtual Circuit

- VPIs and VCIs for the end-to-end connection are preconfigured in hardware
- Semi-PVC: VPIs and VCIs are preconfigured, but can be erased and reconfigured later
- PVC is good for Virtual Leased Line for VPN
- PVC does not scale well for an ATM LAN/WAN with many stations: too many PVCs required
Switched Virtual Circuit

Signalling required to setup a VC
Switched Virtual Circuit

$X_1, X_2, Y_1, Y_2$ may take different values at successive call setups.
Switched Virtual Circuit

- Circuit on demand
- Scalable solution to ATM connectivity
- Different VPIs and VCIs at successive call setups for the same source-destination.
Label Switching with MPLS
Introduction to Label Switching

• Router has two main elements
  – Forwarding and routing
• Label switching is an alternative to IP forwarding
• MPLS (multiprotocol label switching) combines benefits of virtual circuit with flexibility and robustness of datagram forwarding
• An MPLS-capable router uses traditional IP routing, but replaces IP forwarding with label switching
Why label switching

• Improves IP forwarding speed
• Seamless integration of IP and ATM
• Traffic engineering
Terminologies

- **flow**: a single instance of an application to application flow of data

- **forwarding equivalence class**: a group of IP packets which are forwarded along same path (or same treatment)

- **label**: a short fixed length physically contiguous identifier which is used to identify a FEC (local significance)

- **label Switched Router**: an MPLS node which is capable of forwarding L3 packets
How Does MPLS Work
MPLS Devices

MPLS edge node

MPLS ingress node

MPLS domain

LSR: label switching router
Conventional IP Routing

• packet travels from one router to the next
  – an independent forwarding decision made at each hop

• At each hop packets assigned to a
  Forwarding Equivalence Class (FEC)
  – packets considered to be in same FEC if the routing
    table contains some prefix X such that X is the longest
    match for each packet’s destination address
MPLS Approach

• assignments of packets to FEC done once
  – as the packet enters the network
• FEC to which packet is assigned is encoded with a label
• packet is forwarded to next hop with label
  – no further analysis of packet header at next hop
Label encoding

• MPLS uses 32 bits for label encoding
  – 20 bits for actual label
  – 2 power 20 (2^20) labels possible

<table>
<thead>
<tr>
<th>label</th>
<th>exp</th>
<th>s</th>
<th>ttl</th>
</tr>
</thead>
</table>

Label: Label Value, 20 bits (0-16 reserved)
Exp.: Experimental, 3 bits (earlier Class of Service)
S: Bottom of Stack, 1 bit (1 = last entry in label stack)
TTL: 8 bit Time to Live
Reserved Labels

• Some labels are reserved, for example
• 1 = Router Alert Label
  ⇒ Looked by the router software. Packet forwarded based on the next label in stack
  Similar to “Router Alert Option” in IP Packets
• 3 = Implicit Null Label
  Used only for label distribution
  Should not appear in any label in data packets
TTL Handling

• At the beginning TTL = TTL from IP header
• At every hop: TTL = TTL - 1
• Drop the packet if TTL=0
• At the exit: TTL in IP Header = TTL from label
• In some cases, entire MPLS domain may be considered one hop, e.g., ATM
MPLS Encapsulation

- For ATM, VPI/VCI fields in the ATM cell header carries MPLS labels
  - No separate encapsulation is necessary
- For non-ATM (PPP or Ethernet), MPLS uses a “shim” layer between link layer and network layer headers
  - *shim* layer may consist of sequence of labels (concept of label stack)
  - MPLS is layer 2.5!
Label Processing

• *label* used as index into a table which specifies next hop and a new label
  – longest match calculation eliminated at subsequent hops
• old *label* replaced with new *label* and packet forwarded to next hop
• label may also contain class of service (COS)
  – for scheduling/ discarding packets etc
  – again saves header processing
Label Distribution

- Who assigns labels for communication between A and B?
  - A, B, or someone else?
  - Downstream, upstream, ...

- Where is the control for the entire path? A, B, ingress or egress LSR?

- Separate protocol or existing route distribution mechanisms?
  - IETF’s Label Distribution Protocol (LDP)
    - based on Cisco’s TDP and IBM’s ARIS protocol
Example Network

Table at R2

<table>
<thead>
<tr>
<th>Index of table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 162.25.8 1</td>
</tr>
<tr>
<td>8 192.35.10 0</td>
</tr>
</tbody>
</table>

Table at R4

| 162.25.8 1 |
| 192.35.10 1 |
Label Binding

Example shows control packets only
Label allocation

<table>
<thead>
<tr>
<th>In label</th>
<th>out label</th>
<th>Addr prefix</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>7</td>
<td>162.25.8</td>
<td>1</td>
</tr>
<tr>
<td>?</td>
<td>8</td>
<td>192.35.10</td>
<td>1</td>
</tr>
</tbody>
</table>

Table at R2

<table>
<thead>
<tr>
<th>In label</th>
<th>out label</th>
<th>Addr prefix</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>9</td>
<td>162.25.8</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>192.35.10</td>
<td>0</td>
</tr>
</tbody>
</table>
Label Switching

Example shows data packets
Hierarchical Routing (1)

BGP: Border Gateway Protocol (inter-domain)
IGP: Interior Gateway Protocol (intra-domain)

IGP Routers (R3, R4) in transit domain A maintain all the routes provided by Interdomain routing
- necessary to forward transit traffic to/from domains B & C
Hierarchical Routing (2)

**BGP entries**

<table>
<thead>
<tr>
<th></th>
<th>in label</th>
<th>out label</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>....</td>
<td>4</td>
<td>R2</td>
</tr>
<tr>
<td>R2</td>
<td>4</td>
<td>3</td>
<td>R5</td>
</tr>
<tr>
<td>R5</td>
<td>3</td>
<td>7</td>
<td>R6</td>
</tr>
<tr>
<td>R6</td>
<td>7</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

**IGP entries**

<table>
<thead>
<tr>
<th></th>
<th>in label</th>
<th>out label</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>?</td>
<td>9</td>
<td>R3</td>
</tr>
<tr>
<td>R3</td>
<td>9</td>
<td>11</td>
<td>R4</td>
</tr>
<tr>
<td>R4</td>
<td>11</td>
<td>14</td>
<td>R5</td>
</tr>
<tr>
<td>R5</td>
<td>14</td>
<td>?</td>
<td>R5</td>
</tr>
</tbody>
</table>
Stack of labels

- Packets carry several labels organised as a label stack
- packet forwarded from one domain to other contains one label
- packet forwarded through a transit domain contains two labels
- LSRs use label from top of stack
  - labels pushed (ingress) and popped (egress) at domain boundaries
push/pop on stack

push

BGP entries

<table>
<thead>
<tr>
<th></th>
<th>in label</th>
<th>out label</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>....</td>
<td>4</td>
<td>R2</td>
</tr>
<tr>
<td>R2</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>R5</td>
<td>3</td>
<td>7</td>
<td>R6</td>
</tr>
<tr>
<td>R6</td>
<td>7</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

pop

IGP entries

<table>
<thead>
<tr>
<th></th>
<th>in label</th>
<th>out label</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>?</td>
<td>9</td>
<td>R3</td>
</tr>
<tr>
<td>R3</td>
<td>9</td>
<td>11</td>
<td>R4</td>
</tr>
<tr>
<td>R4</td>
<td>11</td>
<td>14</td>
<td>R5</td>
</tr>
<tr>
<td>R5</td>
<td>14</td>
<td>?</td>
<td>R5</td>
</tr>
</tbody>
</table>

swap label 4 to 3
R5 not directly connected
Find direct conn. and label
push label 9 and send to R3

next hop R5 (self)
pop off the top label of the stack
swap this label 3 to 7 (BGP)

CSE, UNSW
IP-ATM Integration using MPLS
MPLS over ATM

- ATM switches performing label switching are called ATM-LSRs
- *labels* need to be encoded in VCI/VPI fields
- ATM forwarding hardware remains unchanged
  - Only software upgrade is needed to convert an ATM switch to ATM-LSR
- Same concept can be extended to optical switches (later in the course)
MPLS over ATM

- Actual label is encoded in VPI/VCI.
- No TTL decrement
- VCI = 0 through 32 should not be used in labels
Cell Interleave Problem

R3 needs to reassemble cells into packets, but can’t distinguish which cell belongs to which packet.

<table>
<thead>
<tr>
<th>in label</th>
<th>out label</th>
<th>addr prefix</th>
<th>interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>5</td>
<td>162.25</td>
<td>1</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>...</td>
<td>..</td>
</tr>
</tbody>
</table>
Solution 1
Request Separate labels

get a label for 162.25

R1

get a label for 162.25

R2

ATM

162.25

request separate labels for R1 and R2 from R3

R3

Predefined VPI/VCI used for Label binding

<table>
<thead>
<tr>
<th>in interface</th>
<th>in label</th>
<th>out label</th>
<th>addr prefix</th>
<th>interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>5</td>
<td>162.25</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>4</td>
<td>162.25</td>
<td>1</td>
</tr>
</tbody>
</table>
Solution 2
VC - Merge

Delay cells from one packet (AAL5 EOF marker)
MPLS on ATM: Issues

- VCI field is sufficient for one level tagging. VPI may be used for the 2nd level.
- LSR switches need to participate in network layer routing protocols (OSPF, BGP).
- Multiple labels per destination may be used to avoid cell interleave problem.
- VPI/VCI space may be segmented for label switching and normal ATM switching.
What Do ATM LSRs Bring to Us

- Now we know how to convert an ATM switch to an ATM LSR
- What do we gain by converting existing ATM switches to ATM LSRs?
- Two main benefits (see picture next slide)
  - Reduced number of neighbours
  - Full view of the network topology
Existing Systems
[Overlay Approach]

R1 has 5 neighbours

ATM switches are NOT converted to LSR
ATM-LSR Approach

[Direct Peering]

R1 has only 1 neighbour!

ATM switches are converted to LSRs
Traffic Engineering with MPLS
Traffic Engineering Using MPLS

- Traffic Engineering
  - Performance Optimization
  - Efficient resource allocation
  - Constrained routing / Load balancing
  - Maximum throughput, Min delay, min loss
    \[ \Rightarrow \] Quality of service
  - Meet policy requirements of operators
Shortest-Path Routing
Constrained Routed LSP

- Hard to do non-shortest path routing with connectionless forwarding
- Constrained Routed label-switched paths allows selective non-shortest path routing
  - load balancing across path possible
  - Signaling protocols for such path establishment being developed by IETF
    - M-RSVP is based on Resource Reservation Protocol RSVP (soft-state model)
    - CR-LDP defines hard-state signaling protocols
CR-LSP Example

MPLS Backbone

A
R1

R2
R3
R4
R6

B

C
Traffic Trunk

- In MPLS networks: “Traffic Trunks” = SVCs
- A set of traffic parameters can be specified to determine the “Forwarding Equivalence Class (FEC)” or set of packets assigned to that trunk
- Features of trunks
  - Multiple trunks can be used in parallel to the same egress.
  - Traffic trunks are routable entities like VCs
  - Each traffic trunk can have a set of associated characteristics, e.g., priority, preemption, policing, overbooking
Traffic Trunks Features

- Traffic Trunk Features
  - Trunk paths are setup based on policies or specified resource availability.
  - A traffic trunk can have alternate sets of paths in case of failure of the main path. Trunks can be rerouted.
  - Some trunks may preempt other trunks. A trunk can be preemptor, non-preemptor, preemptable, or non-preemptable.
  - Each trunk can have its own overbooking rate
Flows, Trunks, LSPs, and Links

- Label Switched Path (LSP): All packets with the same label
- Trunk: Same Label+Exp
- Flow: Same MPLS+IP+TCP headers
Summary

• Simplified forwarding based on exact match of fixed length label
• Separation of routing and forwarding in IP networks
• Facilitates the integration of ATM and IP
• Traffic Engineering Benefits