Quality of Service in the Internet

TML,TKK, Helsinki, FINLAND

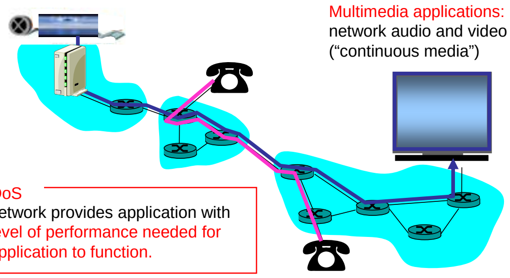
Chittaranjan Hota, PhD
Department of Computer Science and Information Systems
Birla Institute of Technology & Science, Pilani
Rajasthan, 333031, INDIA
E-mail: c_hota@bits-pilani.ac.in

(Few slides are adapted from Leon Garcia, Kuross Ross, and Tanenbaum)

Agenda

- What is QoS and it's Requirements
- Higher Layer Protocols for QoS Guarantee
- Mechanisms to achieve Quality of Service
- QoS Protocols and Models for the Internet
  - Integrated Services (IntServ)
  - Differentiated Services (DiffServ)
  - Multiprotocol Label Switching (MPLS)
- QoS in Mobile Networks
- Next Steps in Signaling (NSIS)

What is Quality of Service?



Multimedia applications: network audio and video ("continuous media")

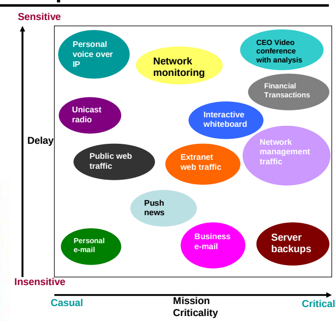
QoS
network provides application with level of performance needed for application to function.

Capability of a network to provide better service (high bandwidth, less delay, low jitter, and low loss probability) to a selected set of network traffic.

QoS Requirements

(With QoS)

(Without QoS)



Audio end-to-end delay : < 150 msec good, < 400 msec OK

Internet QoS

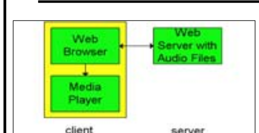
TCP/UDP/IP: "best-effort service"

- *no* guarantees on delay, loss

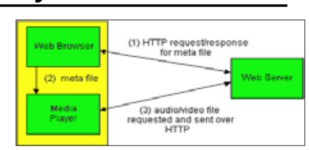
? ? ? ? ? ? ?
But you said multimedia apps require QoS and level of performance to be effective!
? ? ? ? ?

Today's Internet multimedia applications use application-level techniques to mitigate (as best possible) effects of delay, loss

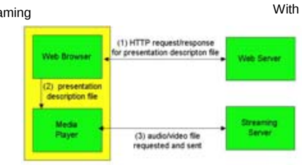
Application Layer Protocols



Without Streaming

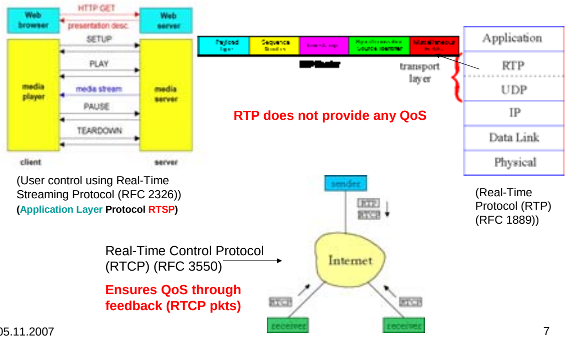


With Streaming

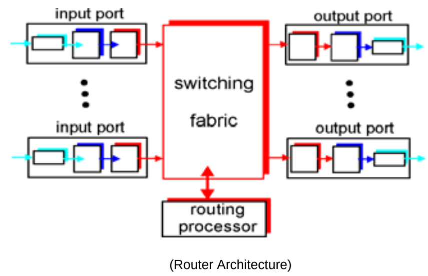


With Streaming

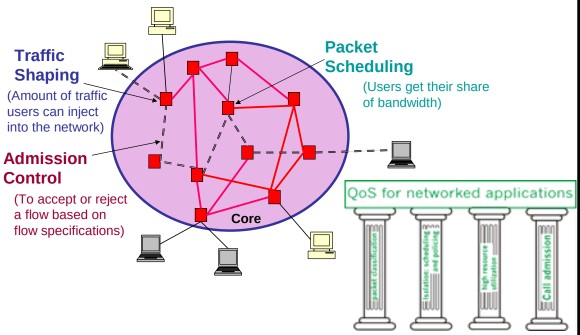
Higher Layer Protocols



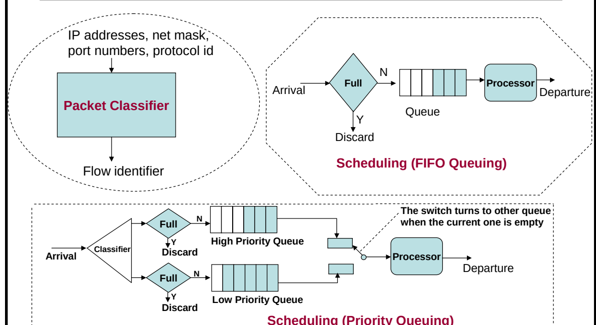
QoS at Network Layer



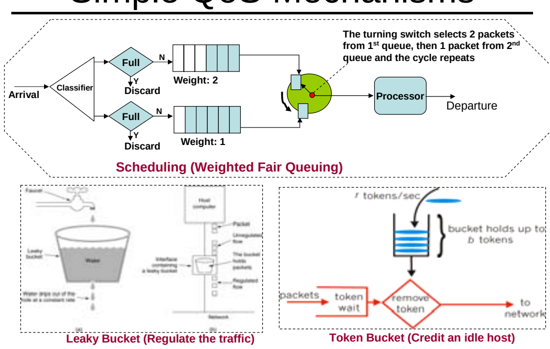
QoS Principles



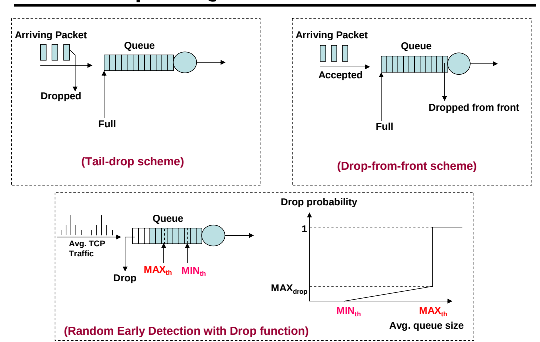
Simple QoS Mechanisms



Simple QoS Mechanisms



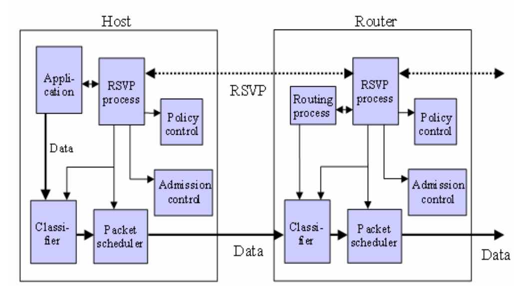
Simple QoS Mechanisms



QoS Architectures for Internet

- **Integrated Services (IntServ)**
  - **Flow Based QoS Model (Resources are available prior to establishing the session)**
  - **Session by session (end-to-end)**
  - **Uses RSVP (signaling protocol) to create a flow over a connectionless IP**
- **Differentiated Services (DiffServ)**
  - **Categorize traffic into different classes or priorities with high priority value assigned to real time traffic**
  - **Hop by hop (no assurance of end-to-end QoS)**
- **Multiprotocol Label Switching (MPLS)**
  - **Not primarily a QoS model, rather a Switching architecture**
  - **Ingress to the network decides a label according to FEC**

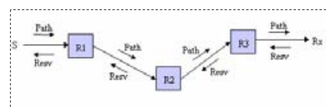
RSVP Architecture



RSVP Design Goals

- accommodate heterogeneous receivers (different bandwidth along paths)
- accommodate different applications with different resource requirements
- make multicast a first class service, with adaptation to multicast group membership
- leverage existing multicast/unicast routing, with adaptation to changes in underlying unicast, multicast routes
- control protocol overhead to grow (at worst) linear with # receivers

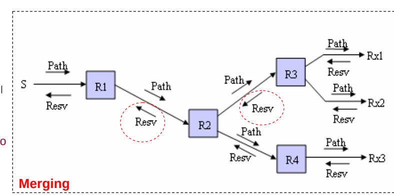
RSVP Messages



<Path Message> ::= <Common Header> [<INTEGRITY>] <SESSION> <RSVP_HOP> <TIME_VALUES> [<POLICY_DATA> ...] [<sender descriptor>] <sender descriptor> ::= <SENDER_TEMPLATE> <SENDER_TSPEC> [<ADSPEC>]

<Resv Message> ::= <Common Header> [<INTEGRITY>] <SESSION> <RSVP_HOP> <TIME_VALUES> [<RESV_CONFIRM>] [<SCOPE>] [<POLICY_DATA> ...] <STYLE> <flow descriptor list> <flow descriptor list> ::= <empty> | <flow descriptor list> <flow descriptor>

Reservation info needs to be refreshed: Soft State

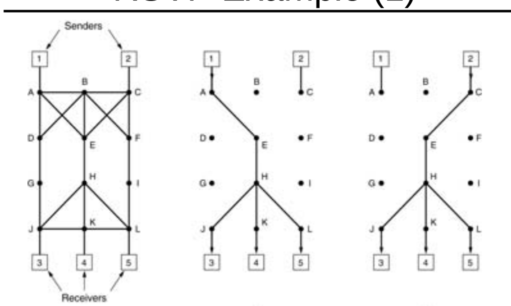


Merging

RSVP: Overview of Operation

- senders, receiver join a multicast group
  - done outside of RSVP
  - senders need not join group
- sender-to-network signaling
  - *path message:* make sender presence known to routers
  - path teardown: delete sender's path state from routers
- receiver-to-network signaling
  - *reservation message:* reserve resources from senders to receiver
  - reservation teardown: remove receiver reservations
- network-to-end-system signaling
  - path error, -reservation error

RSVP Example (1)

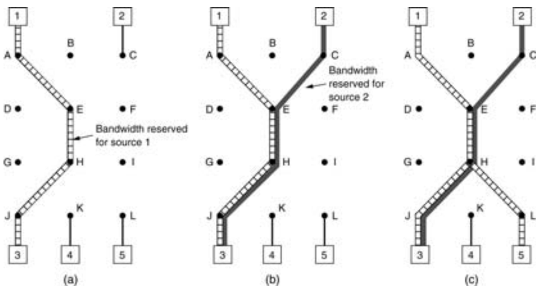


(a) (A network)

(b) (The multicast spanning tree for host 1)

(c) (The multicast spanning tree for host 2)

RSVP Example (1)



(Host 3 requests a channel to host 1)

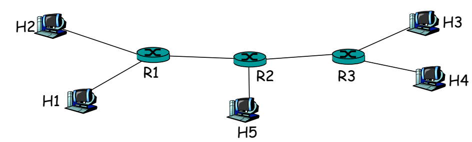
(Additionally, it requests a second channel, to host 2)

(Host 5 requests a channel to host 1)

05.11.2007 19

RSVP Example (2)

- (H1, H2, H3, H4, H5) : both senders and receivers
- multicast group m1
- no filtering: packets from any sender forwarded
- audio rate: *b*
- only one multicast routing tree possible

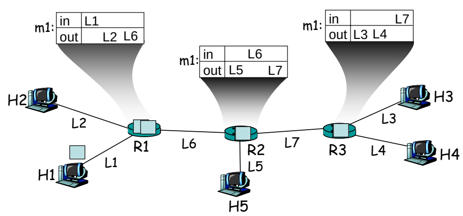


audio conference

05.11.2007 20

Building up Path State

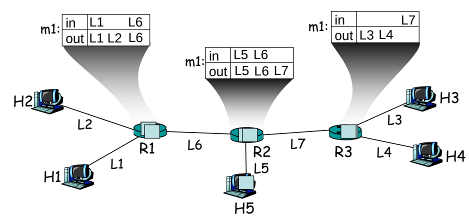
- H1, …, H5 all send path messages on *m1:* (address=*m1*, Tspec=*b*, filter-spec=no-filter, refresh=100)
- Suppose H1 sends first path message



05.11.2007 21

Building up Path State

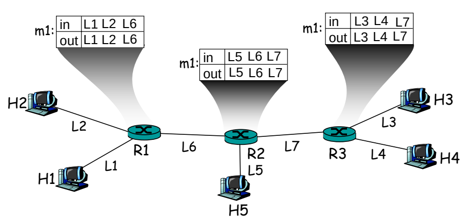
- next, H5 sends path message, creating more state in routers



05.11.2007 22

Building up Path State

- H2, H3, H5 send path msgs, completing path state tables

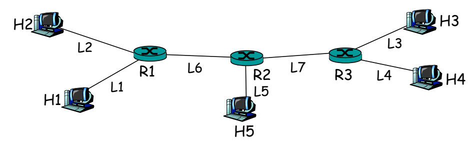


05.11.2007 23

Receiver Reservation

H1 wants to receive audio from all other senders

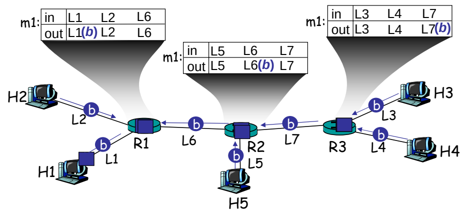
- H1 reservation msg flows uptree to sources
- H1 only reserves enough bandwidth for 1 audio stream
- reservation is of type “no filter” – any sender can use reserved bandwidth



05.11.2007 24

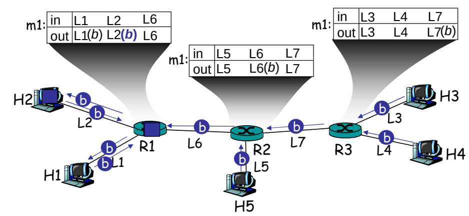
Receiver Reservation

- H1 reservation msgs flows uptree to sources
- routers, hosts reserve bandwidth b needed on downstream links towards H1

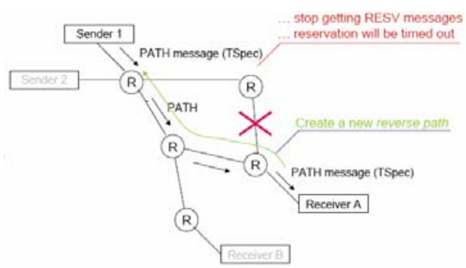


Receiver Reservation

- next, H2 makes no-filter reservation for bandwidth *b*
- H2 forwards to R1, R1 forwards to H1 and R2 (?)
- R2 takes no action, since *b* already reserved on L6

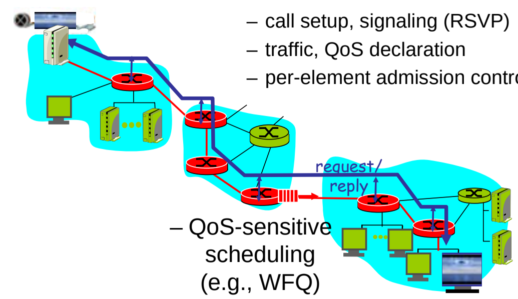


Link Failure



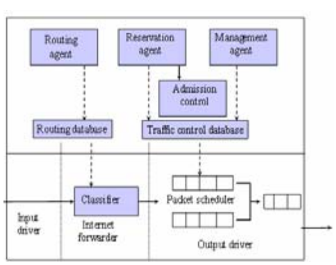
Integrated Services

- Resource reservation
  - call setup, signaling (RSVP)
  - traffic, QoS declaration
  - per-element admission control



  - QoS-sensitive scheduling (e.g., WFQ)

Router and Service Model



Router Model in Integrated Services IP

- Flow Descriptor
  - filterspec, flowspec
- filterspec is required for classifier
- flowspec(Tspec, Rspec)
  - Traffic behavior → Tspec; QoS → Rspec
- Guaranteed Service
  - Firm bound on end-to-end delay in a flow (real-time applications)
- Controlled-Load Service
  - Low delay, and low loss (adaptive applications)

IntServ Scalability

- RSVP Signaling Overhead
  - One PATH/RESV per flow for each refresh period
- Routers have to classify, police and queue each flow
- Admission control is also required
- State information stored in Routers
  - Flow identification (using IP address, port etc)
  - Previous hop identification
  - Reservation Status
  - Reserved Resources

DiffServ: Motivation and Design

- Complex processing is moved from core to edge
- Per flow service (IntServ) is replaced by per aggregate or per class service with an SLA with the provider. (to improve scalability)
- Label packets with a type field
  - e.g. a priority stamp
- Core uses the type field to manage QoS
- Defines an architecture and a set of forwarding behaviors
  - Up to the ISP to define an end-to-end service over this

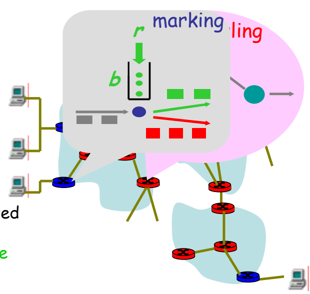
DiffServ Schema

- Source sends request message to first hop router
- First hop router sends request to Bandwidth Broker (BB) that replies with either accept or reject
- If the request is accepted, either the source or the first hop router will mark DSCP and will start sending packets
- Edge router checks compliance with the SLA and will do policing. It may drop or mark the packet with low priority to match the SLA
- Core routers will look into DSCP and decide the PHB

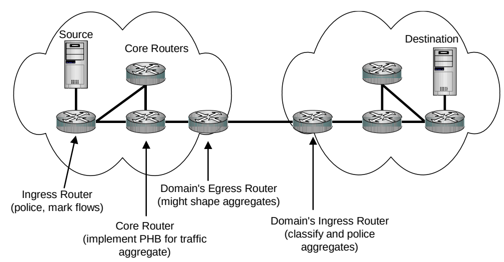
DiffServ Architecture

Edge router:
- per-flow traffic management
- marks packets as in-profile and out-profile

Core router:
- per class traffic management
- buffering and scheduling based on marking at edge
- preference given to in-profile packets

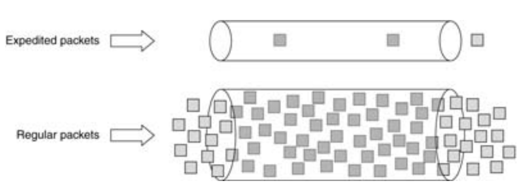


Multi-Domain Example



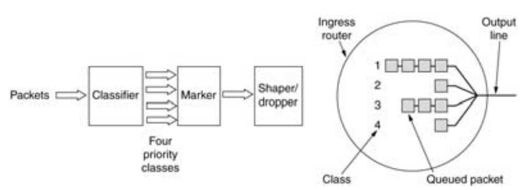
Expedited Forwarding

- Expedited packets experience a traffic-free network (low loss, low latency, low jitter, and assured bandwidth (premium service)
- EF PHB (101110)

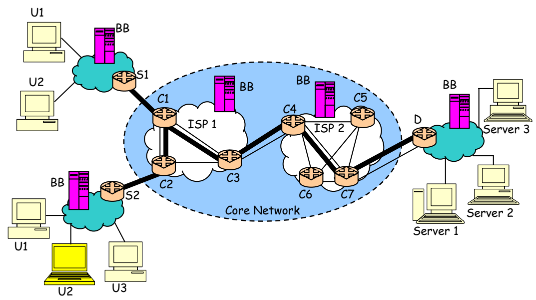


Assured Forwarding

- A possible implementation of the data flow for assured forwarding is shown below.
- AF PHB delivers the packet with high assurance as long as its' class does not exceed the traffic profile of the node.

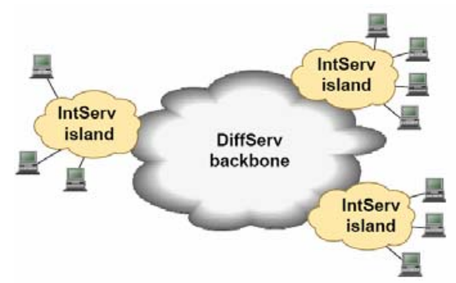


Bandwidth Brokers



05.11.2007 37

Integrated Solution



05.11.2007 38

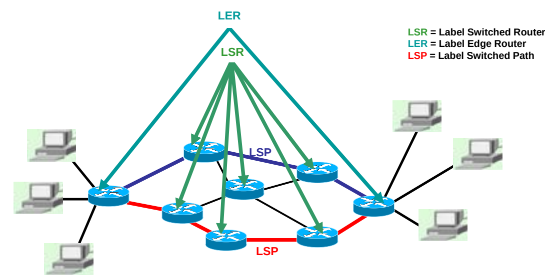
Multiprotocol Label Switching

- **MPLS is a traffic engineering tool whereby we allocate specific path and network resources to specific types of traffic ensuring QoS**
- **Supports Multiple protocols like IPv4, IPv6, IPX, AppleTalk at the network layer, and Ethernet, Token Ring, FDDI, ATM, Frame Relay, PPP at the link layer**
- **Independent of layer 2 and layer 3**
- **Data transmission occurs on Label Switched Paths (LSP)**
- **Labels are distributed using Label Distribution Protocol (LDP), or RSVP, or piggybacked on BGP and OSPF**
- **FEC (Forward Equivalence Class) is a representation of group of packets that share the same requirements for their transport**
- **Assignment of FEC to a packet is done once only as it enters into the network**

05.11.2007 39

Model for MPLS Network

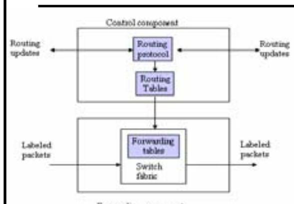
- Convergence of connection oriented forwarding techniques and Internet's routing protocols



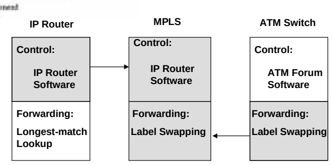
Route at edge and Switch at core

05.11.2007 40

Separate forwarding and control

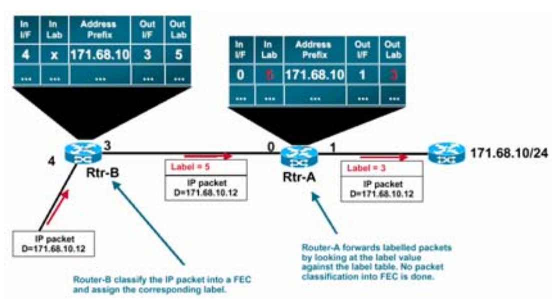


- **Not a longest prefix match like IP, MPLS does exact match of a label, hence faster routing decisions**



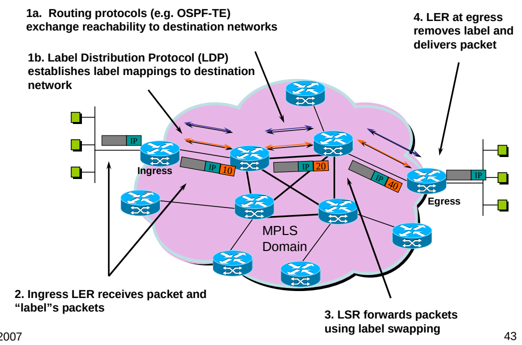
05.11.2007 41

MPLS Forwarding



05.11.2007 42

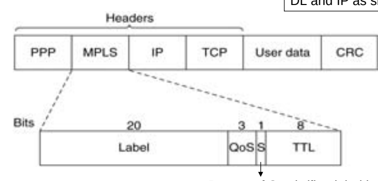
MPLS Operation



MPLS Labels

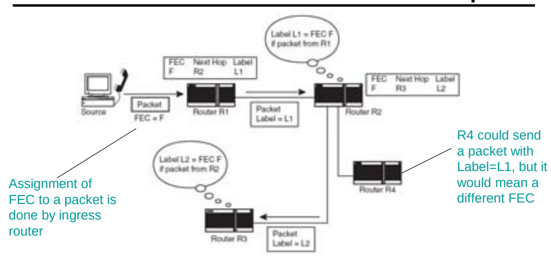
- Label assignment decisions are based on forwarding criteria like
  - Destination unicast routing
  - Traffic engineering
  - Multicast
  - Virtual Private Network
  - Quality of Service

A Label could be embedded in the header of the DL layer like ATM (VPI/VCI) and FR (DLCI) or could be between DL and IP as shown below:



Bottom of Stack (first label in stack)

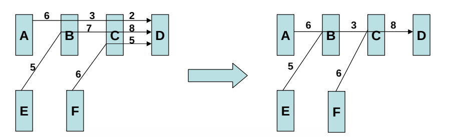
Label and FEC Relationship



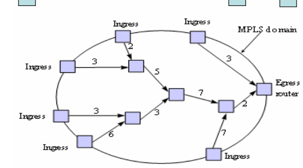
- FEC (Forwarding Equivalence Class): Assigned on the basis of IP addresses, port numbers or TOS bits.
- FEC could be associated with all the flows destined to an egress LSR.

Label Merging

- Label Switched Path (LSP): A unidirectional connection through multiple LSRs.

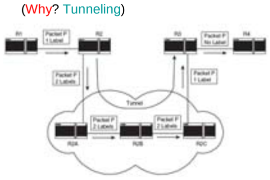


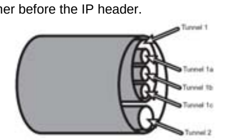
Multi-point to Single point tree routed at Egress router



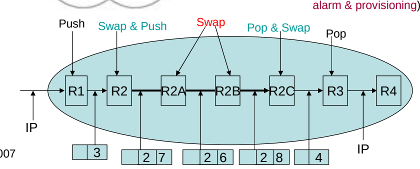
LSP Hierarchy

- A Packet can have several labels one after the other before the IP header. (Why? Tunneling)

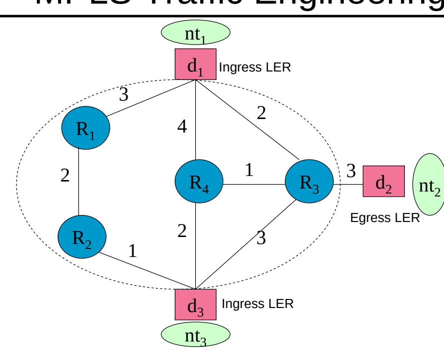




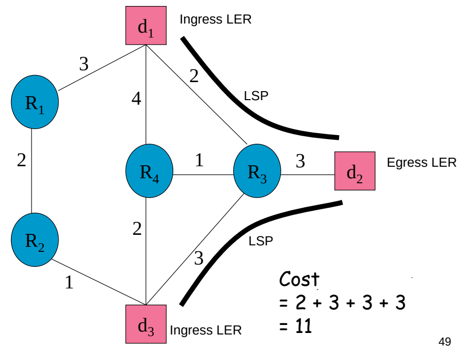
(Multiple Levels of Nesting)
(Tunnel 1 may be for the Enterprise with 1a for VoIP data, 1b for billing, and 1c for alarm & provisioning)



MPLS Traffic Engineering

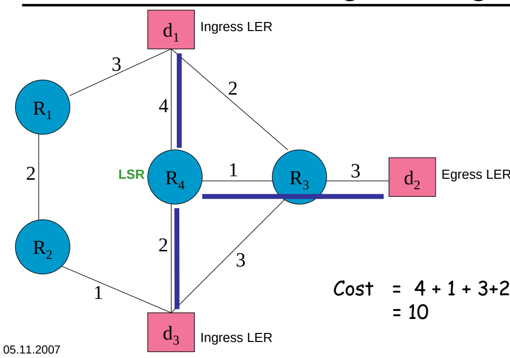


MPLS Traffic Engineering



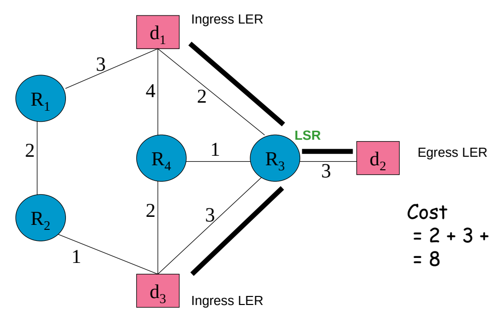
Cost
= 2 + 3 + 3 + 3
= 11

MPLS Traffic Engineering



Cost = 4 + 1 + 3+2
= 10

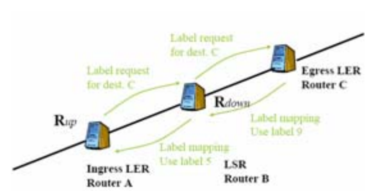
MPLS Traffic Engineering



Cost
= 2 + 3 + 3
= 8

Label Distribution

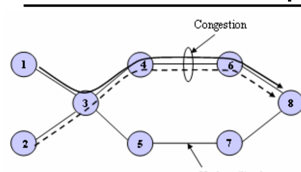
- Establishes and Maintains a LSP that includes establishment of Label/FEC bindings between LSRs in the LSP.
- A downstream LSR can directly distribute Label/FEC (unsolicited downstream).
- An upstream LSR requests a downstream for Label/FEC (downstream on demand).
- Protocols like LDP, RSVP-TE are used to distribute Labels in the LSP



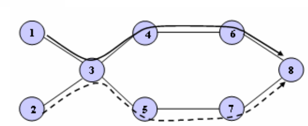
Label Distribution

- **Label Distribution Protocol (LDP) [RFC 3036]**
  - **An LSR sends HELLO messages over UDP periodically to its' neighbors to discover LDP peers (routing protocol tells about peers)**
  - **Upon discovery, it establishes a TCP connection to its peer**
  - **Two peers then may negotiate Session parameters (label distribution option, valid label ranges, and valid timers)**
  - **They may then exchange LDP messages over the session (label request, label mapping, label withdraw etc)**
- **RSVP-TE (Resource Reservation Protocol-Traffic Extension) [RFC 3209]**
  - **Path message includes a label request object, and Resv message contains a label object**
  - **Follows a downstream-on-demand model to distribute labels**
  - **Path message could contain an Explicit Route Object (ERO) to specify list of nodes**
  - **Priorities can be assigned to LSPs, where a higher one can preempt a lower one**

RSVP-TE Explicit Routing



(Hop-by-Hop Routing)

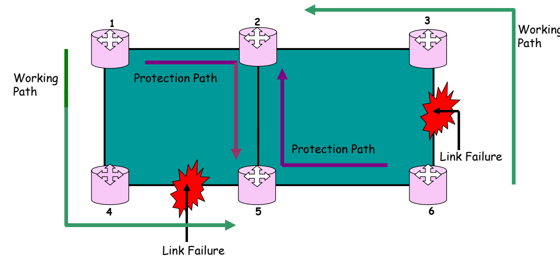


(Explicit Routing)

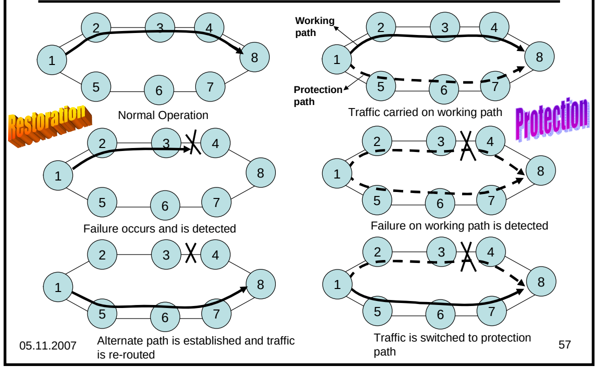
MPLS Survivability

- Survivability is the capability of a network to maintain existing services in the face of failures
- Dynamic routing restores the traffic (upon a failure) based on the convergence time of the protocol
- For a packet network carrying mission critical or high priority data (like MPLS network), we may need specific fast restoration or protection mechanisms

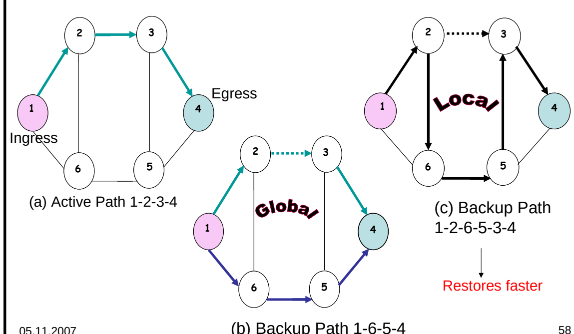
Working Path & Protection Path



Approaches to Survivability

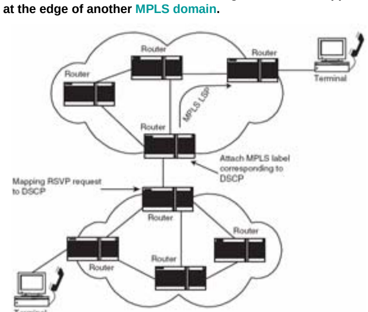


Local and Global Restoration



IntServ, DiffServ and MPLS

- **An RSVP request (say guaranteed service) from one domain could be mapped to an appropriate DiffServ PHB at another domain that again could be mapped to a possible MPLS FEC at the edge of another MPLS domain.**

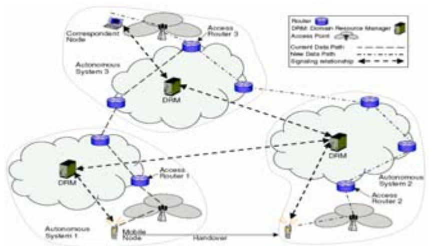


QoS for Mobile Networks

- Problems:
  - Current IP QoS Signaling is not mobility aware (RSVP, DiffServ etc).
  - QoS breaks in new packet path.
  - Resources may not be available for the new path.
  - Handoff latency.
  - Different QoS mechanisms.
- Objectives:
  - Minimize handoff latency.
  - Release any old QoS state after handoff as early as possible.
  - Trigger QoS Signaling as soon as handoff starts.
  - Deal with multiple QoS mechanisms deployed.

A Mobile Environment

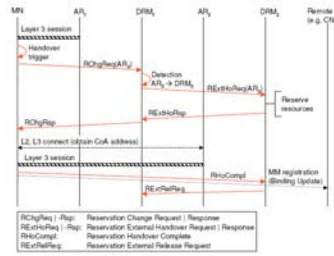
[Ref: 9]



- Domain Resource Manager (DRM) controls QoS for one domain
  - Maintains up-to-date model of resource usage
  - Admission control for reservations
- Supports heterogeneous QoS provisioning
  - per-flow reservations, aggregate reservations (DiffServ) and overprovisioning

Anticipated Inter-Domain Handover

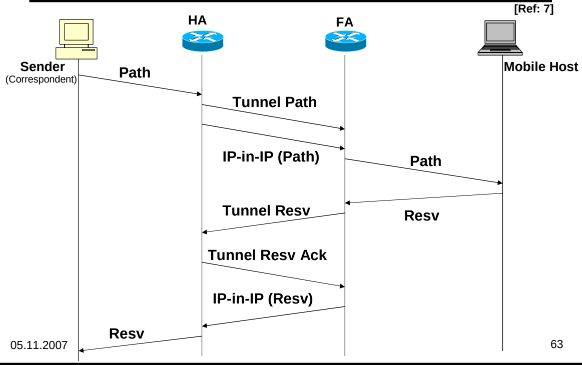
[Ref: 9]



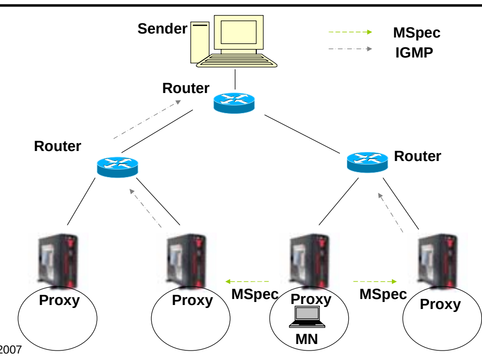
- Signaling for new resources before hand-over
  - Request can be sent over old access router to new DRM
  - Resources can be reserved in advance
- Not possible with on-path signaling approaches!
  - Current IETF approaches (RSVP, NSIS) not sufficient

Mobile RSVP

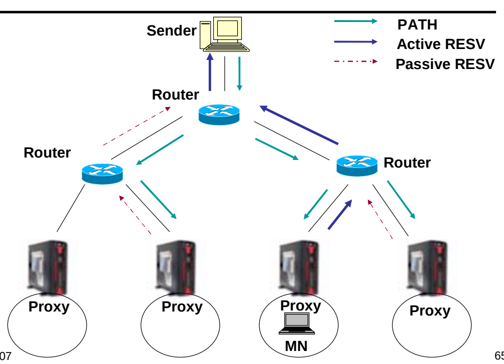
[Ref: 7]



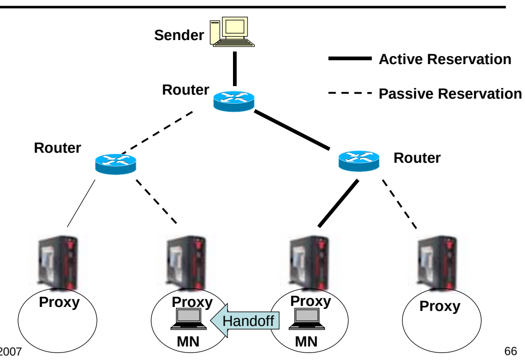
MRSVP Multicast



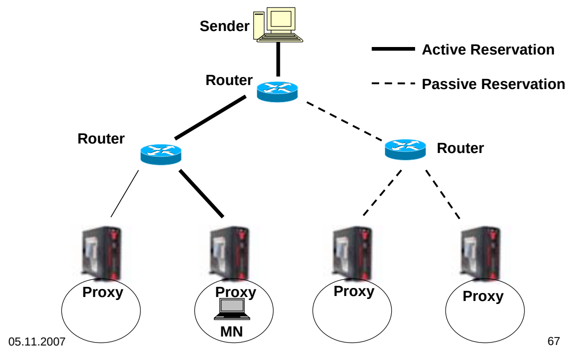
MRSVP Path and Reservation



MRSVP - Handoff



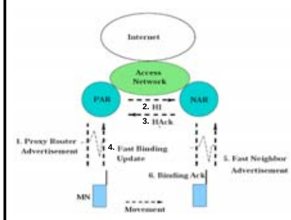
MRSVP – After Handoff



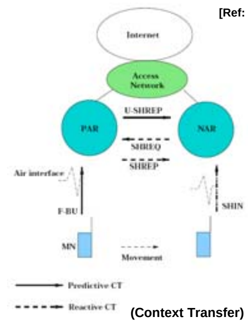
05.11.2007 67

QoS through Context Transfers

[Ref: 8]



(Fast Handover Signaling)



(Context Transfer)

05.11.2007 68

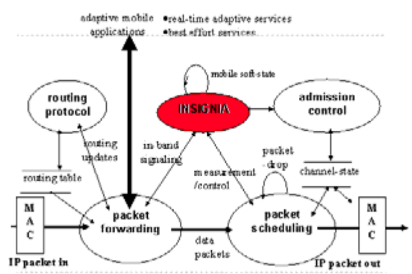
QoS in Mobile Ad Hoc Networks

- All mobile nodes with limited battery life, and wireless connections
- Frequent topology changes leads to rerouting
- High traffic load and mobility degrades service quality
- Hard QoS is difficult
- INSIGNIA uses Adaptive approach (Fast reservation, Fast restoration, QoS reporting, and Adaptation according to network conditions)

05.11.2007 69

INSIGNIA Framework

[Ref: 10]

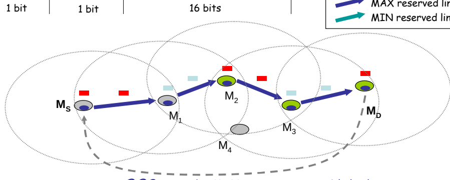


05.11.2007 70

Reservation Set-up

SERVICE MODE	PAYLOAD INDICATOR	BANDWIDTH INDICATOR	BANDWIDTH REQUEST	
RES/BE	BQ/EQ	BW_IND	MAX	MIN
1 bit	1 bit	1 bit	16 bits	

Legend
- RES/BQ packet
- RES/EQ packet
- BE packet
- MAX reserved link
- MIN reserved link



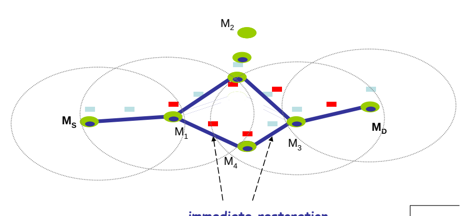
Packets Received at Destination Mobile Node

RES	BQ	MAX	Max_BW	Min_BW
RES	EQ	MAX	Max_BW	Min_BW

05.11.2007 71

[Source: Seoung-Bum Lee's presentation about INSIGNIA]

Re-routing / Restoration

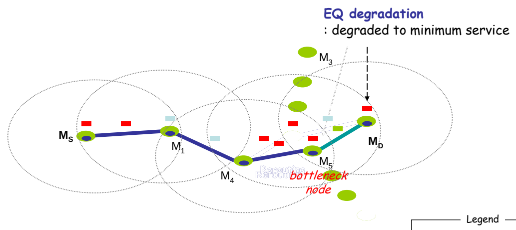


Legend
- RES/BQ packet
- RES/EQ packet
- BE packet
- MAX reserved link
- MIN reserved link

05.11.2007 72

[Source: Seoung-Bum Lee's presentation about INSIGNIA]

Re-routing / Degradation



Packets Received at Destination Mobile Node

RES	BQ	MIN	Max_BW	Min_BW
BE	EQ	-	Max_BW	Min_BW

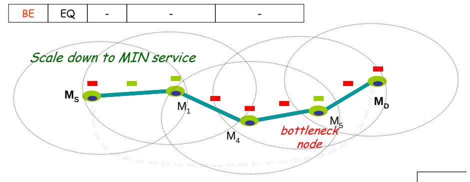
Legend
- RES/BQ packet
- RES/EQ packet
- BE packet
- MAX reserved link
- MIN reserved link

[Source: Seoung-Bum Lee's presentation about INSIGNIA]

Adaptation : Scale Down

Packets sent at Source Mobile Node after "Scaling Down" to MINIMUM service

RES	BQ	MAX	Max_BW	Min_BW
BE	EQ	-	-	-



Pkts Received at Destination after "Scaling Down to MINIMUM service

RES	BQ	MIN	Max_BW	Min_BW
BE	EQ	-	-	-

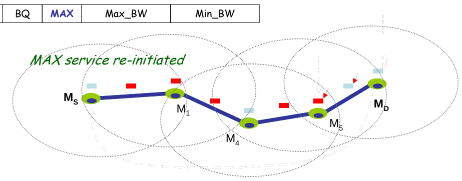
Legend
- RES/BQ packet
- RES/EQ packet
- BE packet
- MAX reserved link
- MIN reserved link

[Source: Seoung-Bum Lee's presentation about INSIGNIA]

Adaptation : Scale Up

Packets sent by Source Mobile Node in MIN service

RES	BQ	MAX	Max_BW	Min_BW



Legend
- RES/BQ packet
- RES/EQ packet
- BE packet
- MAX reserved link
- MIN reserved link

Pkts Received at Destination in MIN service

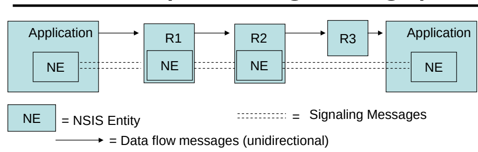
RES	BQ	MAX	Max_BW	Min_BW

[Source: Seoung-Bum Lee's presentation about INSIGNIA]

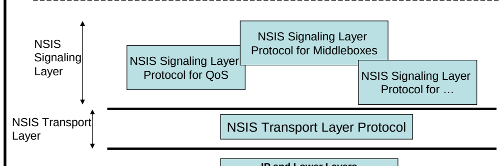
Next Steps in Signaling (NSIS)

- RSVP not widely used for resource reservation
  - but is used for MPLS path setup
  - design heavily biased by multicast needs
  - marginal and after-the-fact security
  - limited support for IP mobility
- Thus, IETF NSIS working group is developing new frameworks for general state management protocol
  - Protocols for signaling information about a data flow along it's path in the network
  - Envisioned to support various signaling applications
  - Resource Reservation
  - NAT and Firewall control (by examining the flow identifier)
  - Traffic and QoS Measurement
  - Security and AAA issues
  - Interaction with other protocols (IP Routing, Mobility, Load Sharing)

Next Steps in Signaling (NSIS)



(Signaling and Data Flow in NSIS)



(NSIS Protocol Components)

References

1. Andrew S. Tanenbaum, Computer Networks, Fourth Edition, Pearson Education, 2006.
2. James F. Kurose, and Keith W. Ross: Computer Networking: A Top-Down Approach Featuring the Internet, Third Edition, Pearson Education, 2006.
3. Alberto Leon-Garcia and Indra Widjaja, Communication Networks: Fundamental Concepts and Key Architectures, Second Edition, Tata McGraw-Hill, 2005.
4. IP QoS Architectures and Protocols, Packet Broadband Network Handbook, Digital Engineering Library, McGraw Hill, 2004.
5. Congestion Control and Quality of Service, Data Communication and Networking, Digital Engineering Library, McGraw Hill, 2006.
6. Manner Jukka, Lopez A, Mihailovi A, Velayos H, Hepworth E, and Y Khouaja, Evaluation of Mobility and QoS Interaction, Computer Networks Volume 38, Issue 2, 5 Feb 2002, pp. 137-163.
7. Anup Kumar Talukdar, B. R. Badrinath, and Arup Acharya, MRSVP: A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts, Wireless Networks, 7, 5–19, 2001.
8. Rajeev Koodli, and Charles E. Perkins, Fast Handovers and Context Transfers in Mobile Networks, ACM SIGCOMM Computer Communications Review, Special Issue on Wireless Extensions to Internet, 2001.
9. J. Hillebrand, C.Prehofer, R. Bless, M. Zitterbart, Quality-of-Service Signaling for Next-Generation IP-based Mobile Networks, IEEE Communications Magazine, June 2004.
10. Seoung-Bum Lee, G. Ahn, X. Zhang and A. T. Campbell, INSIGNIA: An IP-Based Quality of Service Framework For Mobile Ad Hoc Networks, Journal of Parallel and Distributed Computing, 2000.
11. Chittaranjan Hota, Sanjay Jha, G Raghurama, Distributed Dynamic Resource Management in IP VPNs to Guarantee Quality of Service, IEEE ICON 2004, Singapore.
12. RFC 2205: Resource Reservation Protocol, Braden, Zhang et al.
13. RFC 3031: Multiprotocol Label Switching (MPLS), Rosen, Viswanathan and Callon.
14. RFC 2475: An Architecture for Differentiated Services, S. Blake, D. Black et al.
15. RFC 4080: Next Steps in Signaling (NSIS): Framework, R. Hancock, G. Karagiannis et al., 2005.
16. RFC 2326: Real Time Streaming Protocol (RTSP), H. Schulzrinne et al., 1998

Thank you!

Questions