Quality of Service and Traffic Engineering using Multiprotocol Label Switching

Raj Jain

Pro ciences

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

These slides are available on-line at:

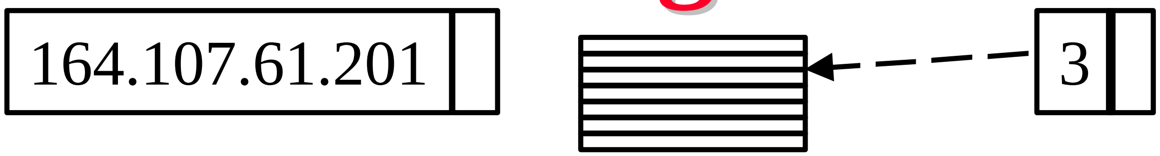
http://www.cis.ohio-state.edu/~jain/talks/mpls_te.htm

1. MPLS Overview
2. Traffic Engineering using MPLS
3. Our Simulation Results
4. Other QoS Approaches and their Interoperability with MPLS

Part 1: MPLS Overview

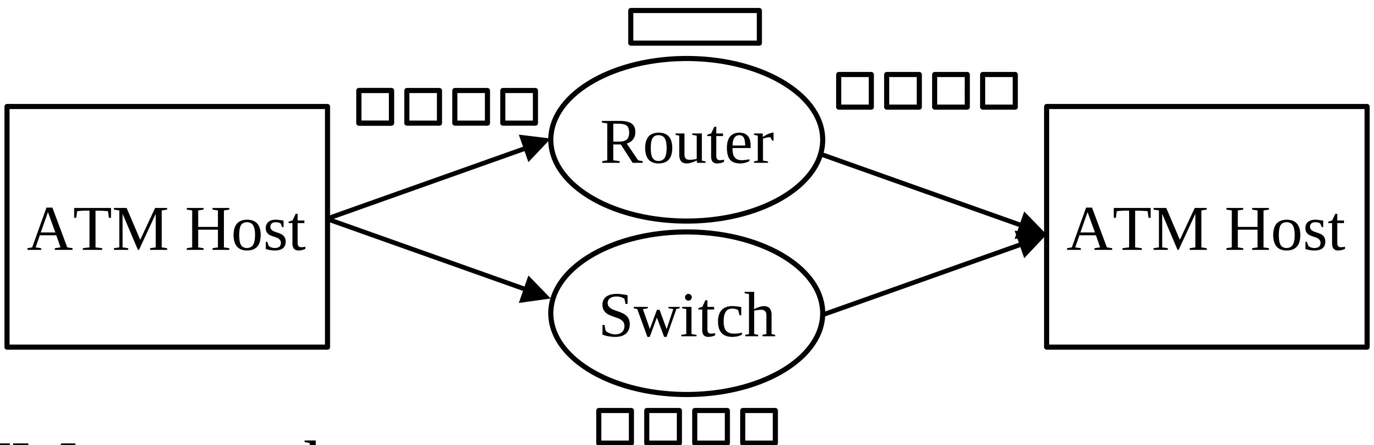
- Routing vs Switching
- Label Switching Concepts
- Label Stacks
- Label Distribution Protocol
- Independent vs Ordered Control

Routing vs Switching



- Routing: Based on address lookup. Max prefix match.
  ⇒ Search Operation
  ⇒ Complexity ≈ $O(\log_2 n)$
- Switching: Based on circuit numbers
  ⇒ Indexing operation
  ⇒ Complexity O(1)
  ⇒ Fast and Scalable for large networks and large address spaces
- These distinctions apply on all datalinks: ATM, Ethernet, SONET

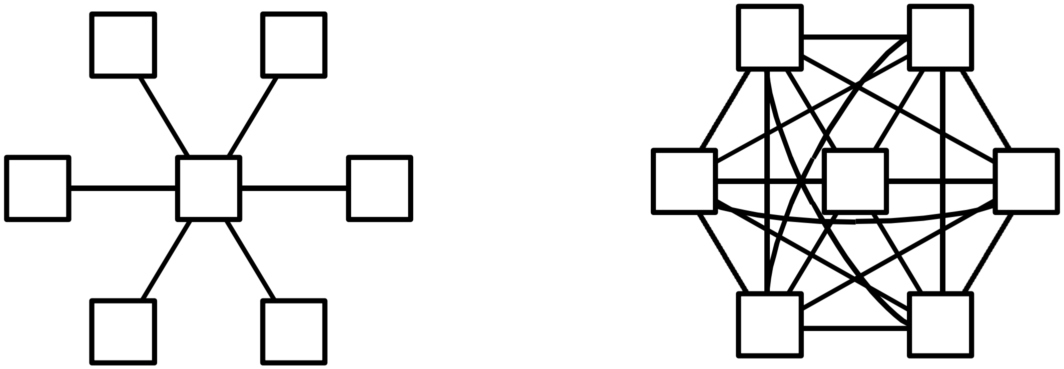
Routing vs Switching over ATM



On ATM networks:

- IP routers use IP addresses
  - ⇒ Reassemble IP datagrams from cells
- IP Switches use ATM Virtual circuit numbers
  - ⇒ Switch cells
  - ⇒ Do not need to reassemble IP datagrams
  - ⇒ Fast

High-Speed Backbone Alternatives

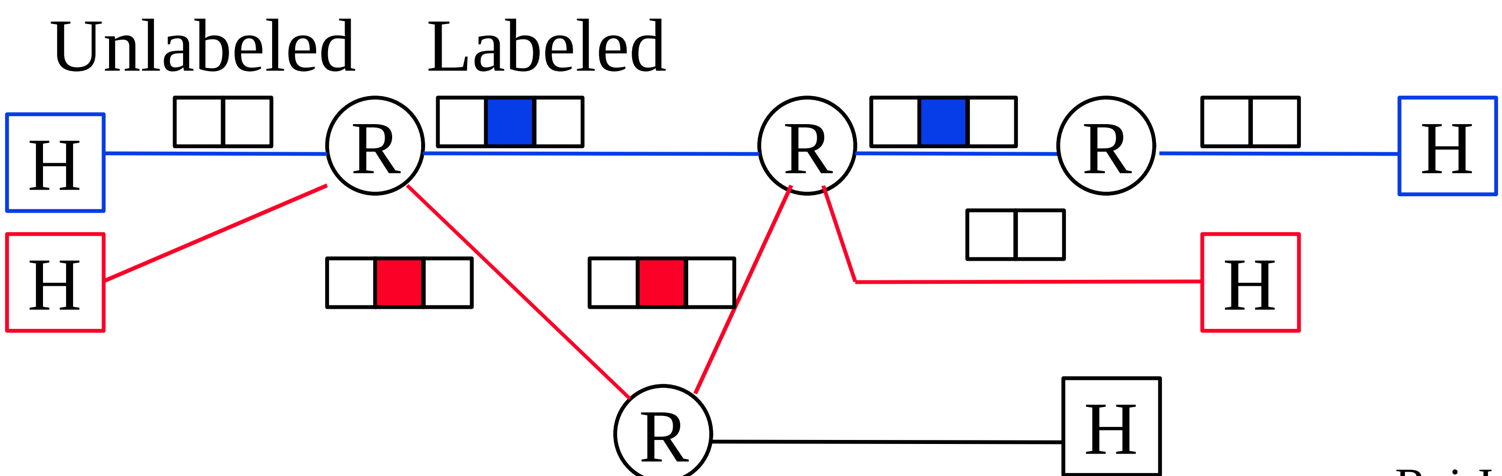


Physical Topology   Logical Topology

- High-speed (OC-3 and higher) ATM switches easily available. IP routers either not available or expensive.
- IP has no traffic engineering ⇒ Under/over-utilized links
- Logical ≠ Physical ⇒ ATM has n^2 scaling problem
- MPLS takes the best of both IP and ATM networks
- Works on both ATM and non-ATM networks ⇒ Easier management

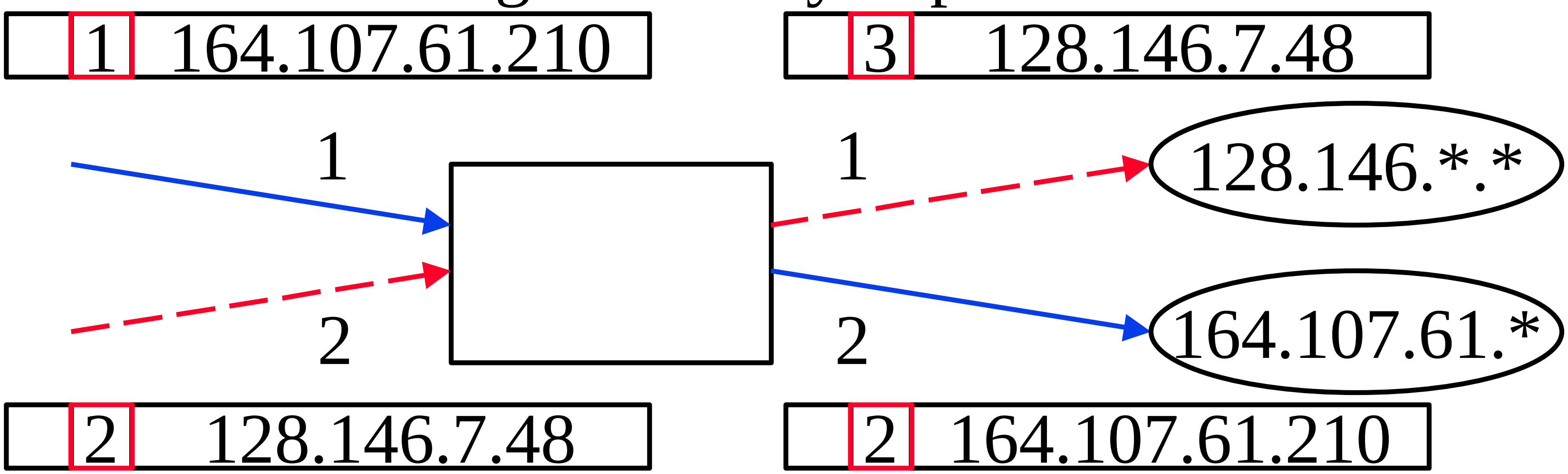
Label Switching

- Label = Circuit number = VC Id
- Ingress router/host puts a label. Exit router strips it off.
- Switches switch packets based on labels.
  Do not need to look inside ⇒ Fast.



Label Switching (Cont)

- Labels have local significance
- Labels are changed at every hop



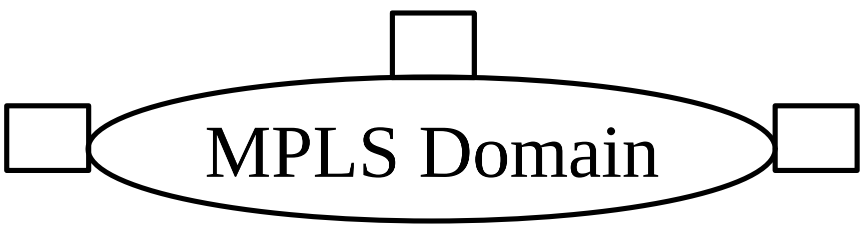
Input Port	Input Label	Adr Prefix	Output Port	Output Label
1	1	164.107.61.*	2	2
2	2	128.146.*.*	1	3

MPLS

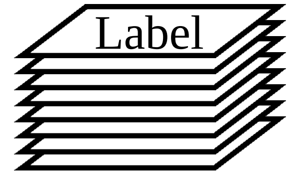
- Multiprotocol Label Switching
- IETF working group to develop switched IP forwarding
- Initially focused on IPv4 and IPv6. Technology extendible to other L3 protocols.
- Not specific to ATM. ATM or LANs.
- Not specific to a routing protocol (OSPF, RIP, ...)

MPLS Terminology

- Label = Short fixed length,
  physically contiguous, locally significant
- Label Switching Router (LSR): Routers that use labels
- Forwarding Equivalence Class (FEC):
  Same Path + treatment $\Rightarrow$ Same Label
- MPLS Domain: Contiguous set of MPLS nodes in one Administrative domain
- MPLS edge node = Egress or ingress node
- Label distribution protocol $\cong$ Routing protocols

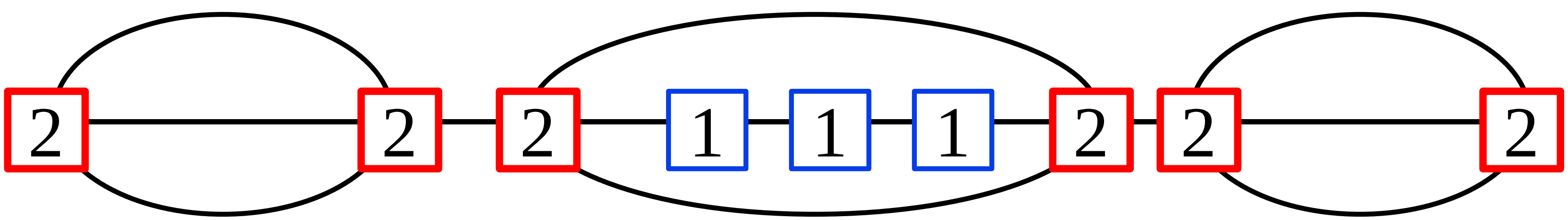


Label Stacks



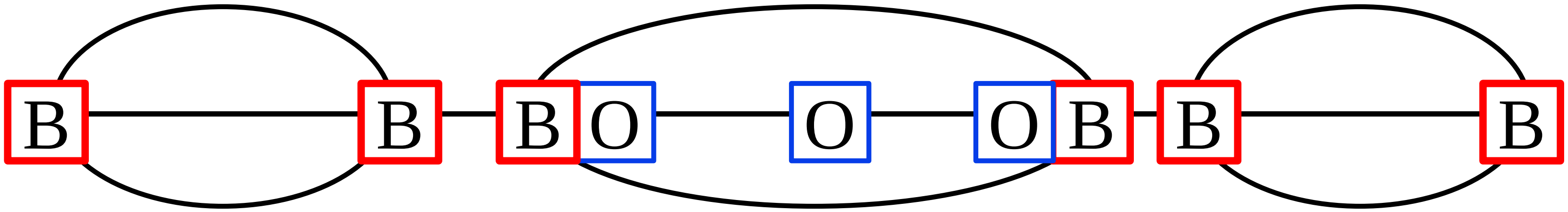
- A MPLS packet may have multiple labels
- Labels are pushed/popped
  as they enter/leave MPLS domain
- Stack allows hierarchy of MPLS domains
- Bottom label may indicate protocol (0=IPv4, 2=IPv6)

L2 Header	Label 1	Label 2	...	Label n	

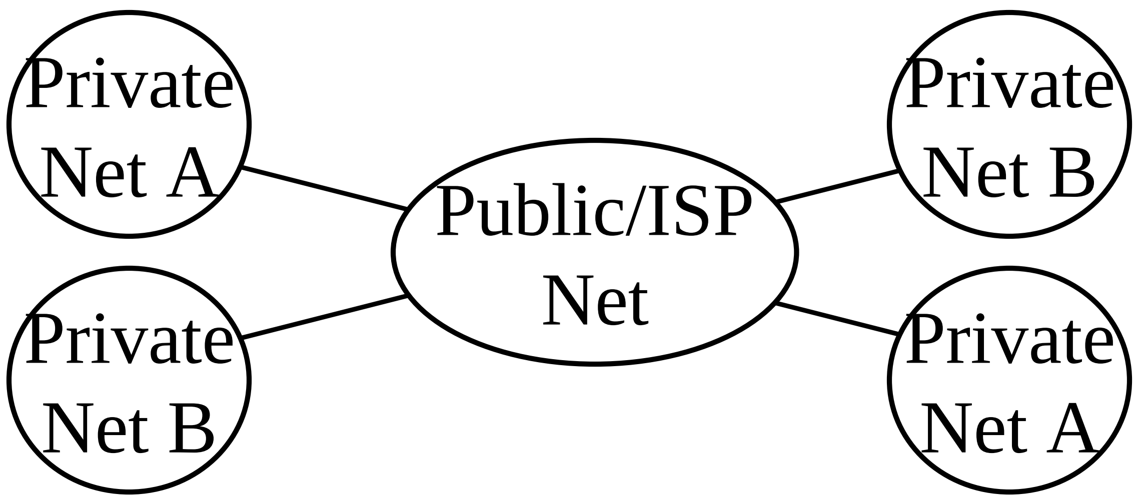


Label Stack Examples

1. BGP/OSPF Routing Hierarchy

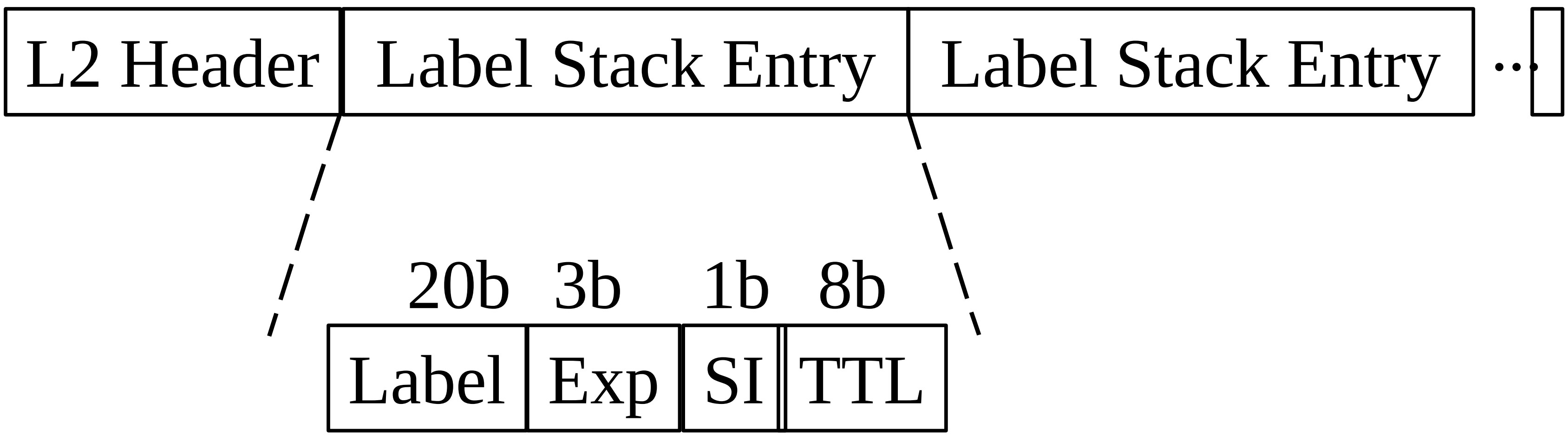


2. VPN: Top label used in public network.
   Net A and B can use the same private addresses.



Label Stack Entry Format

- Labels = Explicit or implicit L2 header
- TTL = Time to live
- Exp = Experimental
- SI = Stack indicator, 1⇒ Bottom of Stack



Label Assignment

- Unsolicited: Topology driven ⇒ Routing protocols exchange labels with routing information.
  Many existing routing protocols are being extended: BGP, OSPF
- On-Demand:
  ⇒ Label assigned when requested,
  e.g., when a packet arrives ⇒ latency
- A new Label Distribution Protocol called LDP is being defined.
- RSVP is being extended to allow label request and response

Label Distribution Protocol

- LDP peers: LSRs that exchange LDP messages. Using an LDP session.
- LDP messages:
  - Session establishment/termination messages
  - Discovery messages to announce LSRs (Hello)
  - Advertisement msgs to create/delete/change label
  - Notification messages for errors and advice
- Discovery messages are UDP based. All others TCP.
- Hello messages are sent on UDP port 646.
- Session establishment messages sent on TCP port 646.
- No multicast, multipath, or QoS in the first version.

LDP Messages

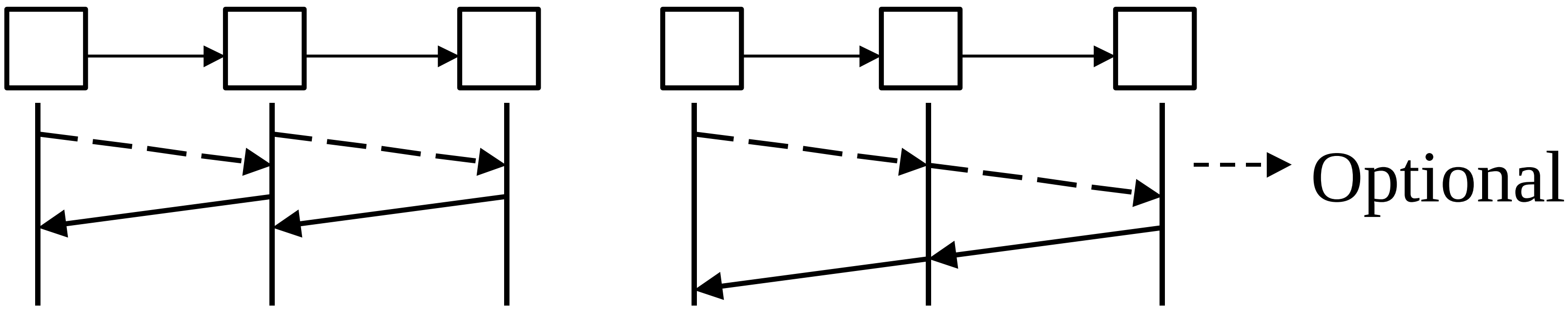
- Hello
- Initialization
- Label Request
- Label Mapping (Label Response)
- Label Withdraw (No longer recognized by downstream)
- Label Release (No longer needed by upstream)
- Label Abort Request
- KeepAlive
- Notification
- Address (advertise interface addresses)
- Address Withdraw
- Vendor-Private
- Experimental

LDP TLVs

- FEC (Wild card, prefix, or host address)
- Address List
- Hop Count
- Path Vector
- Generic Label
- ATM Label
- Frame Relay Label
- Status
- Extended Status
- Returned PDU
- Returned Message
- Common Hello parameters

Independent vs Ordered Control

- Independent: Each router issues Labels for FECs. May cause loops.
- Ordered: A router issues labels for an FEC only if it is the egress router or if it has received a label from the next hop ⇒ Use LSP only after it is fully setup
- Use ordered LSP control if you need QoS for LSP
- LSRs can use either method.

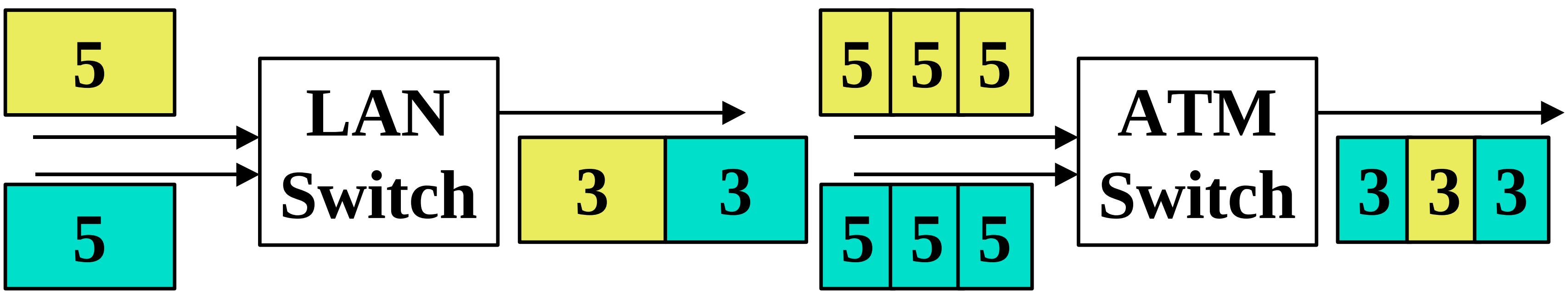


MPLS Over ATM

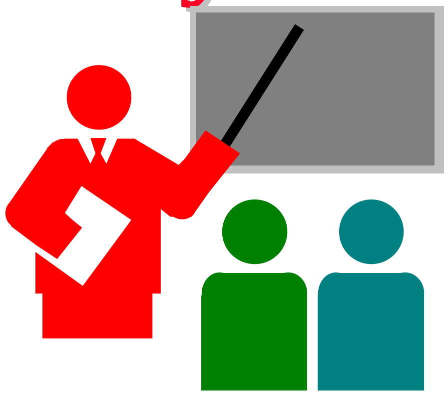
- With MPLS software, ATM switches can act as LSRs.
- VPI/VCI fields are used for labels.
- No Stack bit ⇒ Maximum two possible levels of hierarchy: VCI, VPI
  All ATM switches should use the same encoding.
- No TTL field ⇒ Hops between ingress and egress can be computed during LSP setup.
  Ingress router drops if TTL < hops to egress
- ATM LSRs need to participate in network layer routing protocols (OSPF, BGP)
- VPI/VCI space may be segmented for label switching and normal ATM switching

Stream Merging

- Required for egress based labels. Helpful for mpt-to-pt streams.
- In ATM/AAL5, cells of frames on the same VC cannot be intermingled ⇒ VCs cannot be merged.
- VC-merge: Store all cells of a frame and forward together ⇒ Need more buffering. Delay.
- VP Merge: VPI = Labels, VCI = source



Summary of Part 1: MPLS



- MPLS combines the best of ATM and IP. Works on all media: ATM and non-ATM.
- Label is similar to circuit number or VC Id.
- Label stacks allow hierarchy of MPLS domains.
- Common routing protocols and RSVP are being extended to include label exchange.
- LDP allows independent or ordered control

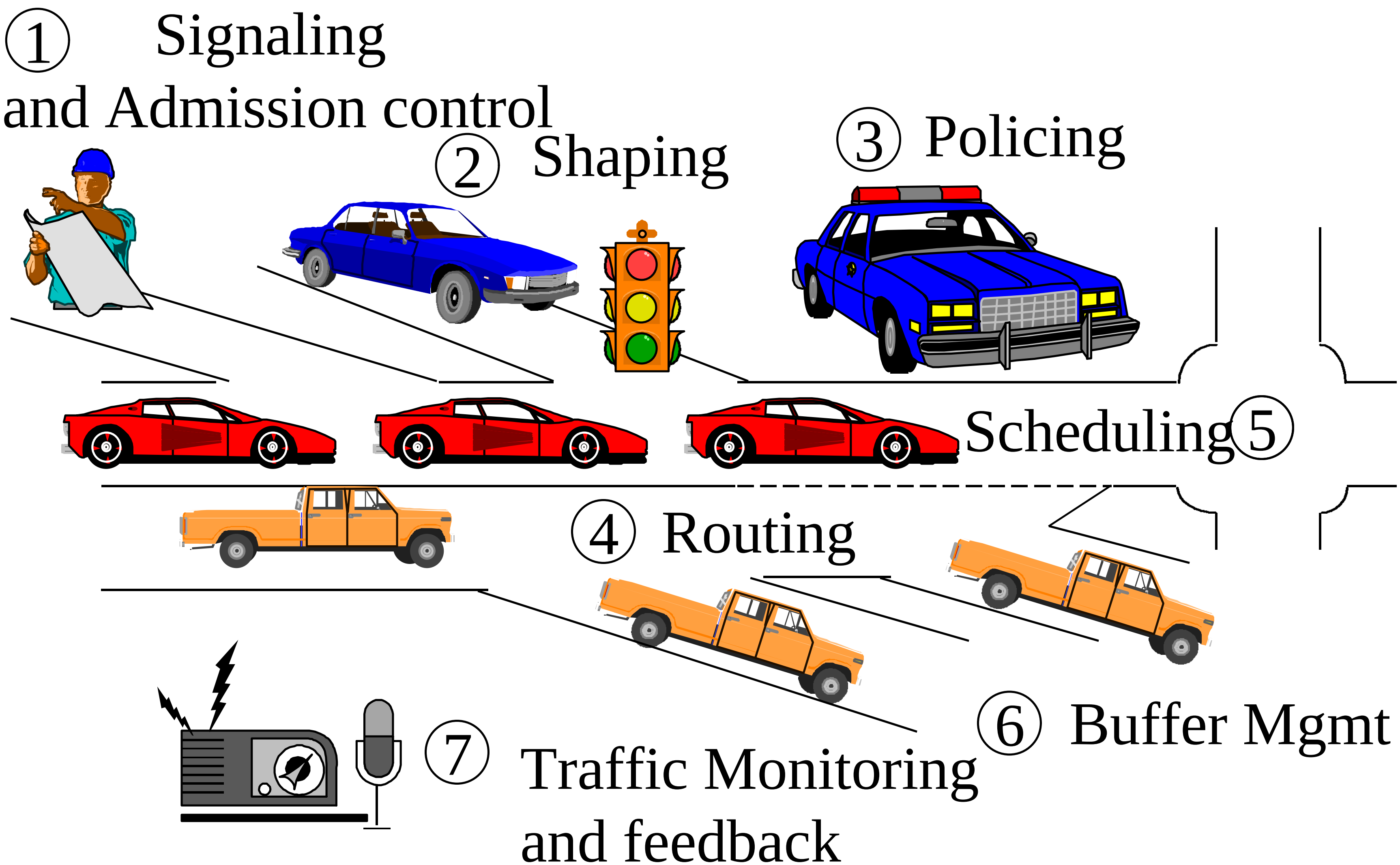
Part 2: Traffic Engineering

- Objectives and Mechanisms
- Traffic Trunks
- CR-LDP
- Explicit Route
- Priority and Preemption
- Traffic Engineering Extensions to OSPF and IS-IS

Traffic Engineering Objectives

- User's Performance Optimization
  $\Rightarrow$ Maximum throughput, Min delay, min loss, min delay variation
- Efficient resource allocation for the provider
  $\Rightarrow$ Efficient Utilization of all links
  $\Rightarrow$ Load Balancing on parallel paths
  $\Rightarrow$ Minimize buffer utilization
  - Current routing protocols (e.g., RIP and OSPF) find the shortest path (may be over-utilized).
- QoS Guarantee: Selecting paths that can meet QoS
- Enforce Service Level agreements
- Enforce policies: Constraint based routing $\supseteq$ QoSR

Traffic Engineering Components



Traffic Engineering Components

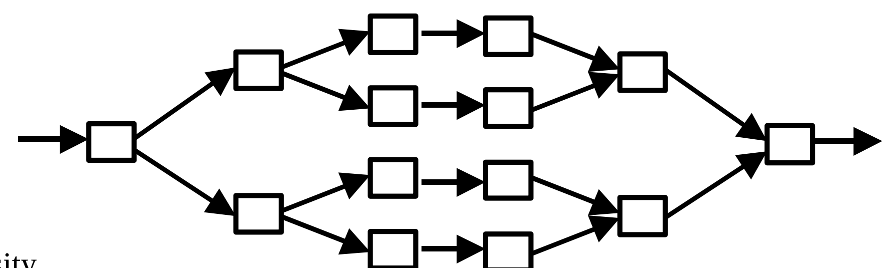
1. Signaling: Tell the network about traffic and QoS.
   Admission Control: Network may deny the request.
2. Shaping: Smoothen the bursts
3. Policing: Ensure that users are following rules
4. Routing: Path Selection, Request Prioritization, Preemption, Re-optimization/Pinning, Fault Recovery
5. Scheduling: Weight, Prioritization, Preemption
6. Buffer Management: Drop Thresholds, Drop Priority
7. Feedback: Implicit, Explicit

Accounting/Billing

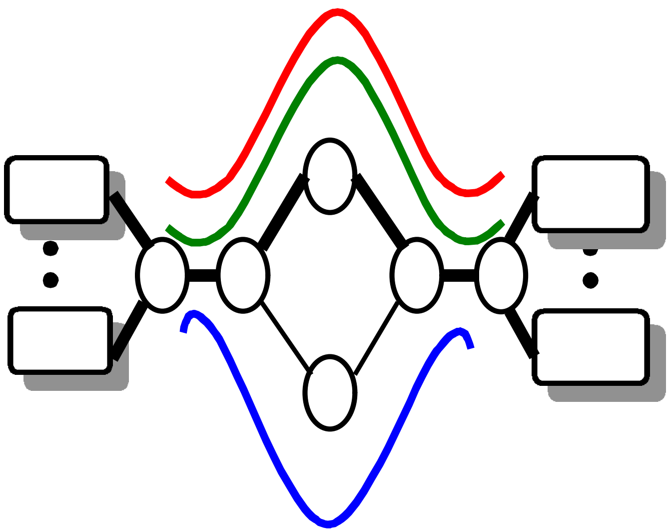
Performance Monitoring/Capacity Planning

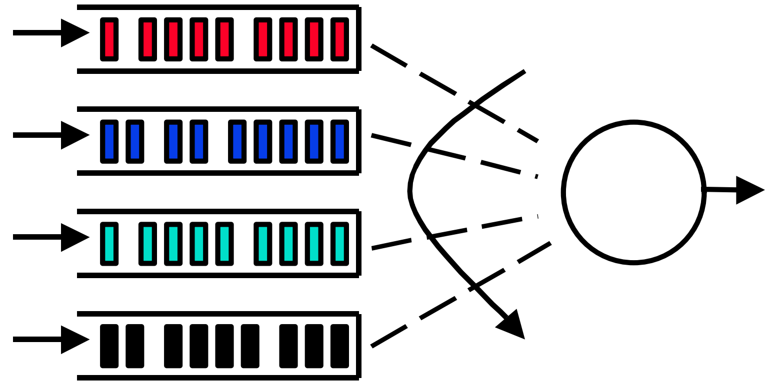
MPLS Mechanisms for TE

- Signaling, Admission Control, Routing
- Explicit routing of LSPs
- Constrained based routing of LSPs
  Allows both Traffic constraints and Resource Constraints (Resource Attributes)
- Hierarchical division of the problem (Label Stacks)
- Traffic trunks allow aggregation and disaggregation (Shortest path routing allows only aggregation)



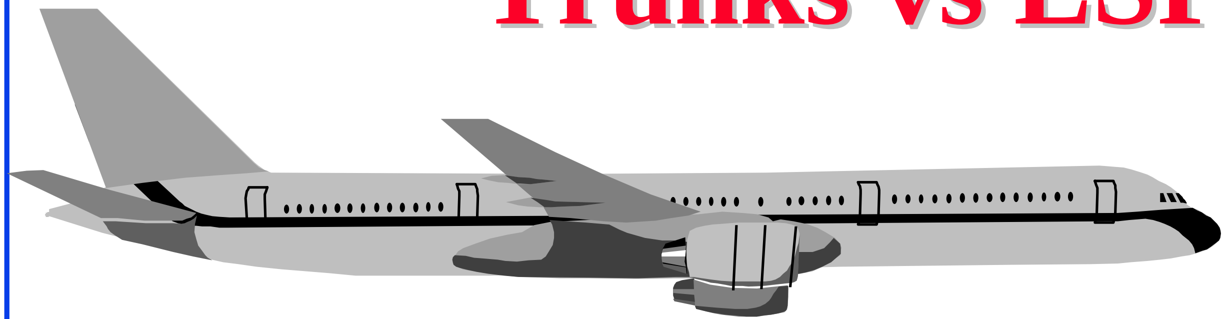
Traffic Trunks



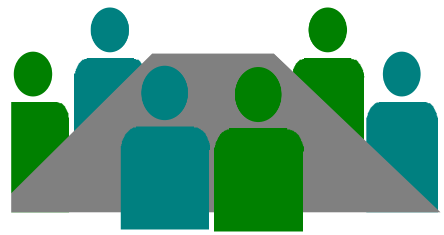


- Trunk: Aggregation of flows of same class on same LSP
- Trunks are routable
  ⇒ LSP through which trunk passes can be changed
- Class ⇒ Queue, LSP ⇒ Next hop
  Class can be coded in Exp or Label field. Assume Exp.

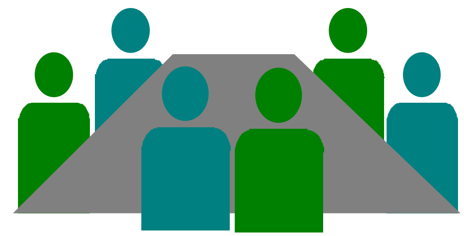
Trunks vs LSPs



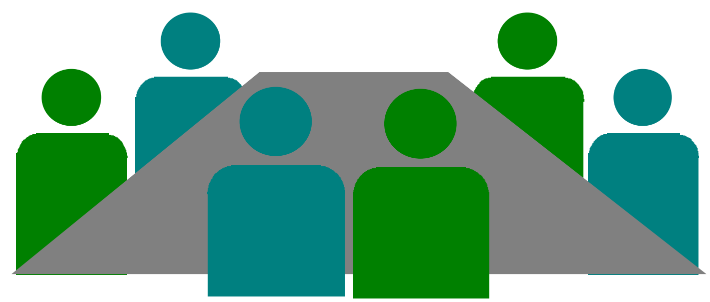
Tour Group



First Class



Business Class



Coach Class

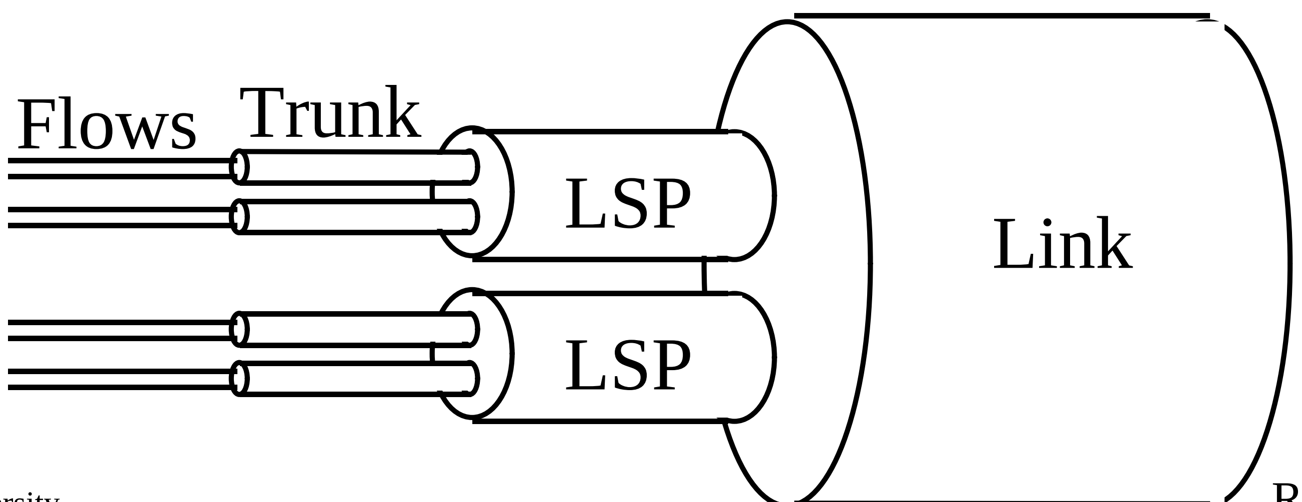
Flights = LSP

Tour Groups = Trunks

Flows, Trunks, LSPs, and Links

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

DL	Label	Exp	SI	TTL	IP	TCP	



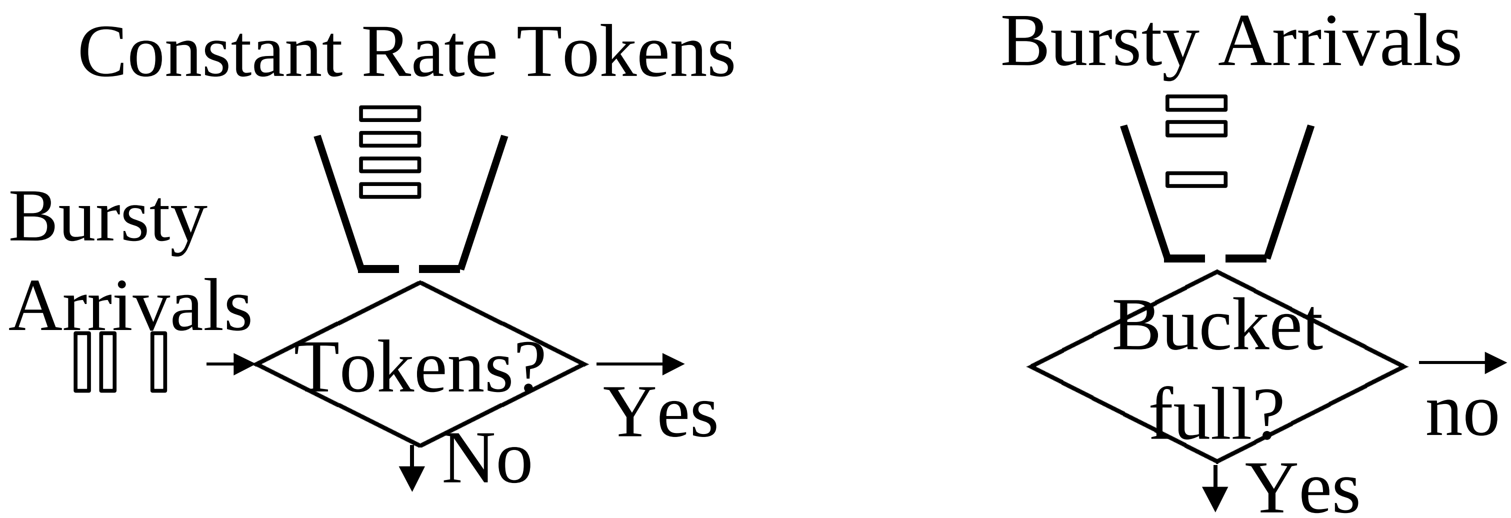
Traffic Trunks

- Each traffic trunk can have a set of associated characteristics, e.g., priority, preemption, policing
- Some trunks may preempt other trunks. A trunk can be preemptor, non-preemptor, preemptable, or non-preemptable.
- 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.
- Multiple LSPs can be used in parallel to the same egress.

Trunk Attributes

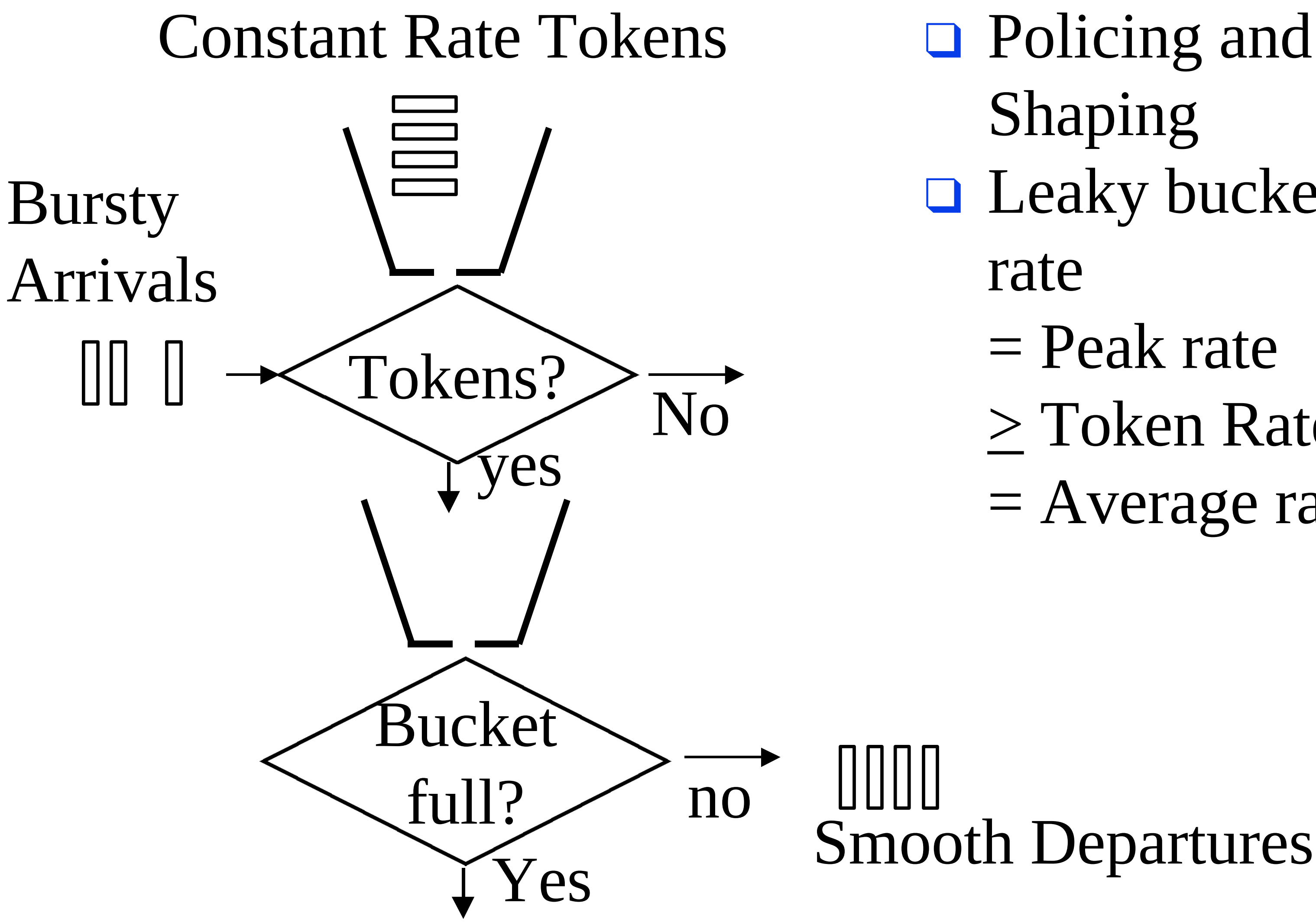
- **Signaling**: Routing Protocols, RSVP, CR-LDP
- **Admission Control**: Network may deny the request.
- **Policing**: Token Bucket
- Shaping: Smoothen the bursts
- **Routing: Path Selection, Request Prioritization, Preemption, Re-optimization/Pinning, Fault Recovery**
- Scheduling: Class Weight, Prioritization, Preemption
- Buffer Management: Class drop thresholds/priority
- Feedback: Implicit, Explicit (ICMP being discussed)
- Accounting/Billing
- Performance Monitoring/Capacity Planning

Token Bucket vs Leaky Bucket



- Both designed for controlling average rate.
- Token bucket sends less than b+rt. Used by IETF.
  Leaky bucket sends less than rt. Used by ATM.
- On bursty arrivals after a long idle:
  - Token bucket results in bursty departures
  - Leaky bucket results in smooth departures

Token and Leaky Bucket



- Policing and Shaping
- Leaky bucket rate
  = Peak rate
  ≥ Token Rate
  = Average rate

Traffic Granularity

- Same label ⇒ Same port quadruples (source/destination address, IP protocol, source/destination port)
- Same QoS + Port quadruples
- Same host pair (Source/destination address)
- Same network pairs (Source/destination address prefixes)
- Same destination network
- Same Egress router

Traffic Granularity (Cont)

- Same BGP next hop AS
- Same BGP destination AS
- Same Shared multicast tree (*,G)
- Same Source specific multicast tree (S,G)

CR-LDP

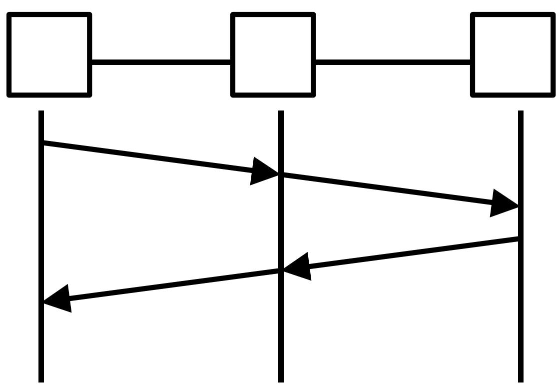
- Extension of LDP for constraint-based routing (CR)
- New Features:
  - Traffic parameters
  - Explicit Routing
  - Preemption of existing route. Based on holding priority of existing route and setup priority of new route
  - Route pinning: To prevent path changes

CR-LDP (Cont)

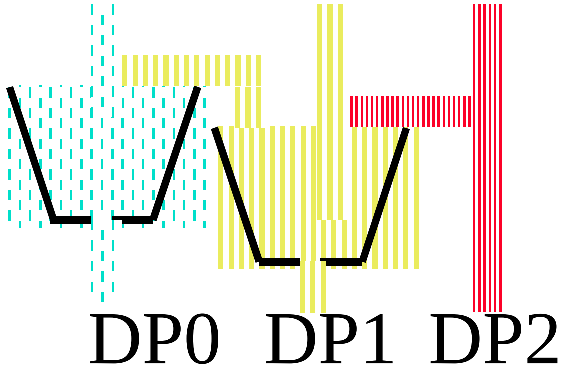
- No new messages
- Enhanced Messages: Label request, Label Mapping, Notification
- New TLVs: Explicit Route, Explicit Route Hop, Traffic, Route Pinning, Resource Class, Pre-emption, LSP Id
- Enhanced TLVs: FEC (CRLSP)
- Each setup (label request) message has a unique connection ID (LSPID)

CR-LSP Setup

- New CR-TLV ⇒ Use "downstream on demand" label advertisement with ordered control
- Similar to ATM connection setup message.
- Egress router indicates the negotiated values in the response (label mapping message)
- Other LSRs return the response towards the ingress and reserve.



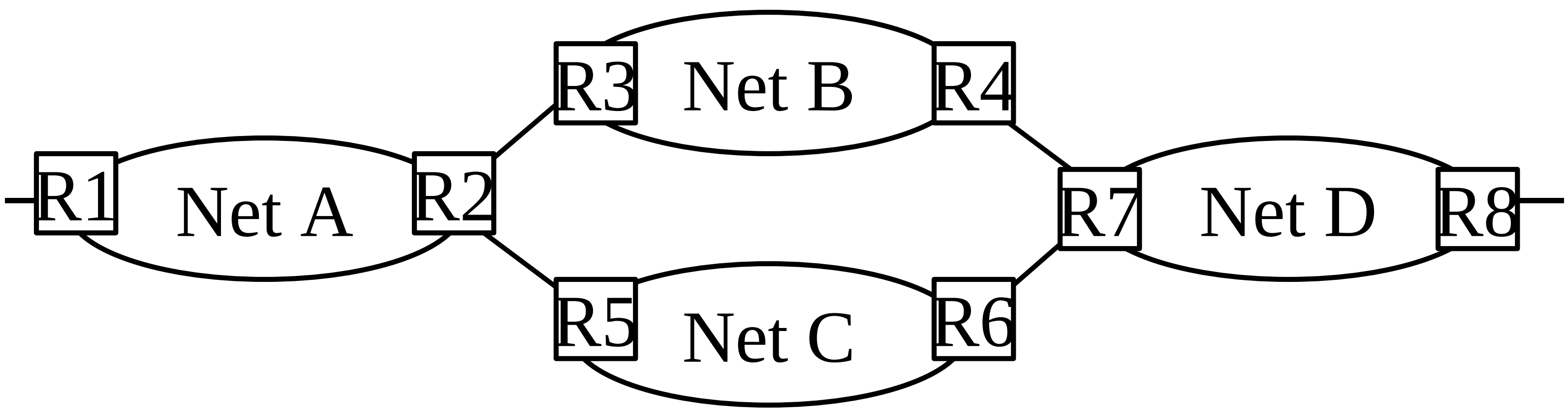
Traffic Parameters



- Single-rate dual-token-bucket
- Tokens generated at "Committed Data Rate" (CDR). Tokens go to 1st bucket, if full go to the 2nd bucket
- Peak, committed data rate, committed burst size, excess burst size (Dual-bucket single rate)
- Negotiation Allowed
- Color Aware ⇒ Use incoming drop precedence (DP) Color unaware ⇒ Ignore incoming drop precedence

Explicit Route

- Explicit route specified as a list of Explicit Route Hops (group of nodes)
- Hops can include IPv4 prefix, IPv6 prefix, MPLS tunnels or Autonomous systems
- Example: R1-R2-Net B-R7-R8



Explicit Route (Cont)

- All or a subset may be traversed
- The list is specified by edge router based on imperfect info (Strict/loose)
  - Strict $\Rightarrow$ Path must include only nodes from the previous and this abstract node
  - Loose $\Rightarrow$ path between two nodes may include other nodes
- Managed like ATM PNNI Designated Transit Lists (DTLs)

Path Selection

- Manual/Administrative
- Dynamically computed
- Explicitly specified: Partially/fully, strict/loose, Mandatory/non-mandatory, Single/Set
- Non-Mandatory
  $\Rightarrow$ Use any available path if specified not available
- Set $\Rightarrow$ Preference ordered list
- Resource class affinity

Resource Attributes

- Capacity
- Overbooking Factor: Maximum Allocation Multiplier
- Class: Allows policy enforcement
- Class Examples: secure/non-secure, transit/local-only
- A resource can be member of multiple classes

Resource Class Affinity

- Each resource has a class
- Affinity = Desirability
- Binary Affinity: 0 $\Rightarrow$ Must Exclude,
  1 $\Rightarrow$ Must Include, Not-specified $\Rightarrow$ Don't care
- <Class, affinity> pair can be used to implement policies

Adaptivity and Resilience

- Stability: Route pinning
- Resource availability is dynamic
- Trunks can live for long time
- Adaptivity: Re-optimization when availability changes
- Resilience: Reroute if path breaks
- Adaptivity $\Rightarrow$ Resilience. Resilience $\not\Rightarrow$ Adaptivity
- Idea: Adaptivity is not binary $\Rightarrow$ Rerouting period

Priority and Preemption

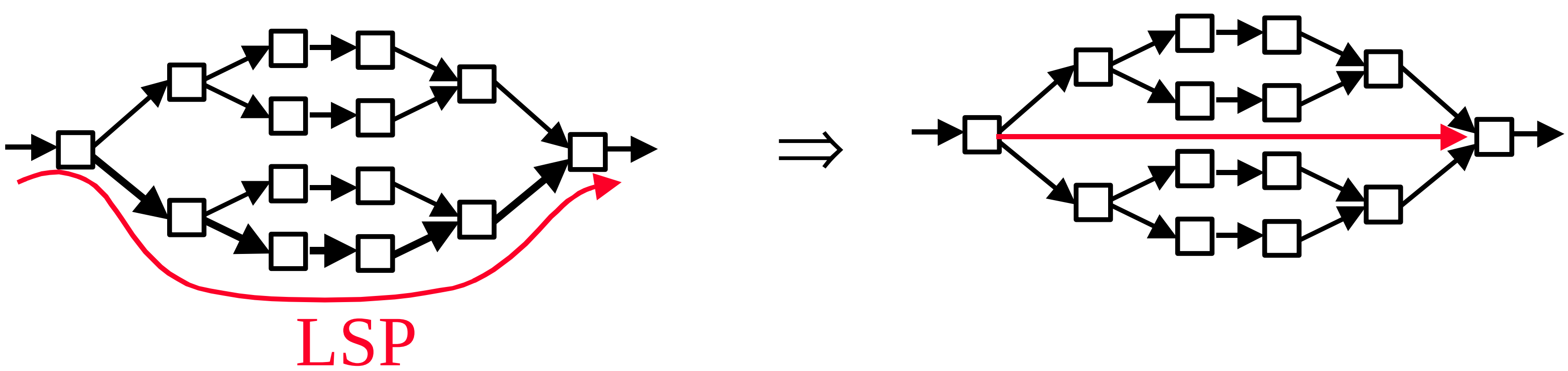
- Preemptor-enabled: Can preempt other trunks
- Non-Preemptor: Can't preempt other trunks
- Preemptable: Can be preempted by other trunks
- Non-Preemptable: Can't be preempted by other trunks
- These attributes and priority are used to decide preemption

Traffic Engineering Extensions to OSPF

- Add to Link State Advertisements:
- TE Metric: May be different from standard OSPF link metric
- Maximum bandwidth
- Maximum Reservable Bandwidth:
  May be more than maximum bandwidth
- Unreserved Bandwidth
- Resource Class/color
- Ref: draft-katz-yeung-ospf-traffic-00.txt

TE Extensions to OSPF (Cont)

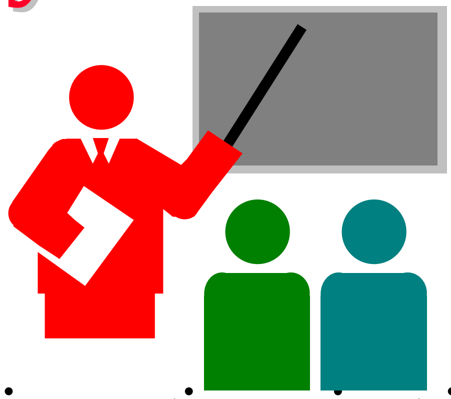
- Link Delay and Link Loss rate also proposed in draft-wimer-ospf-traffic-00.txt
- In path calculations, TE tunnels are used as links to tunnel egress



Traffic Engineering Extensions to IS-IS

- Add to Link State Protocol Data Units:
- TE Metric
- Maximum bandwidth
- Maximum Reservable Bandwidth: May be more than maximum bandwidth
- Unreserved Bandwidth
- Resource Class/color
- Ref: draft-ietf-isis-traffic-01.txt

Summary of Part 2: Traffic Engg

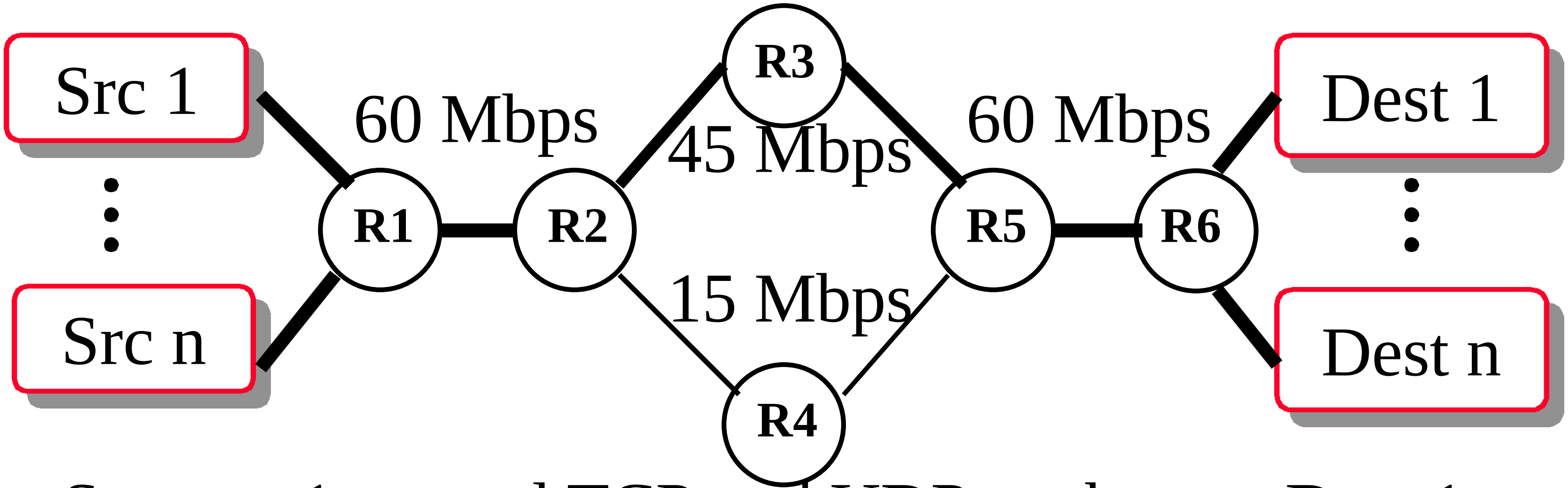


- Goal of traffic engineering is to optimize performance for users and providers and ensure QoS
- MPLS traffic trunks are like ATM VCs that can be routed based on explicit route or policies
- CR-LDP allows explicit routing, constraint-based routing, traffic parameters, and QoS
- OSPF and IS-IS is being modified for traffic engg

A Simulation Analysis of Traffic Engineering

- Simulation Model
- Four Simulation Scenarios
  - Case 1: No Trunks, No MPLS
  - Case 2: Two trunks w UDP + TCP Mixed
  - Case 3: Three Trunks w Isolated TCP, UDP
  - Case 4: Non End-to-End Trunks
- Future Work

Simulation Model

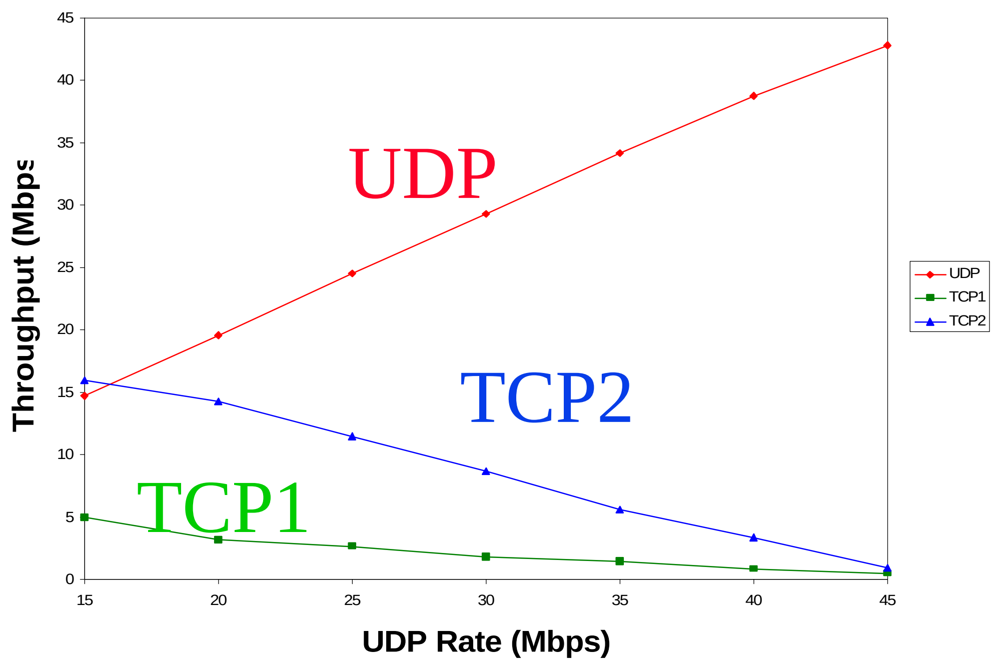


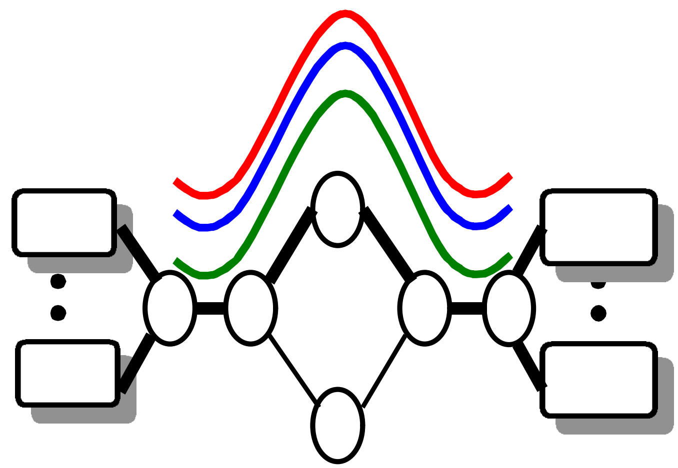
- Sources 1..n send TCP and UDP packets to Dest 1..n
- R2-R3-R5 is a high bandwidth (45 Mbps) path.
- R2-R4-R5 is a low bandwidth (15 Mbps) path.
- All links have 5ms delay
- TCP1 MSS = 512 B, TCP2 MSS = 1024 B, UDP MSS = 210B

Simulation Scenarios

1. Normal IP with Best Effort routing
2. Two trunks using Label Switched Paths
   - Trunk 1: R1-R2-R3-R5-R6
     - TCP and UDP sources are multiplexed over this trunk
   - Trunk 2: R1-R2-R4-R5-R6
     - Only TCP sources over this trunk
3. Three trunks using Label Switched Paths
   - All three flows are isolated.
4. Non End-to-end trunks.

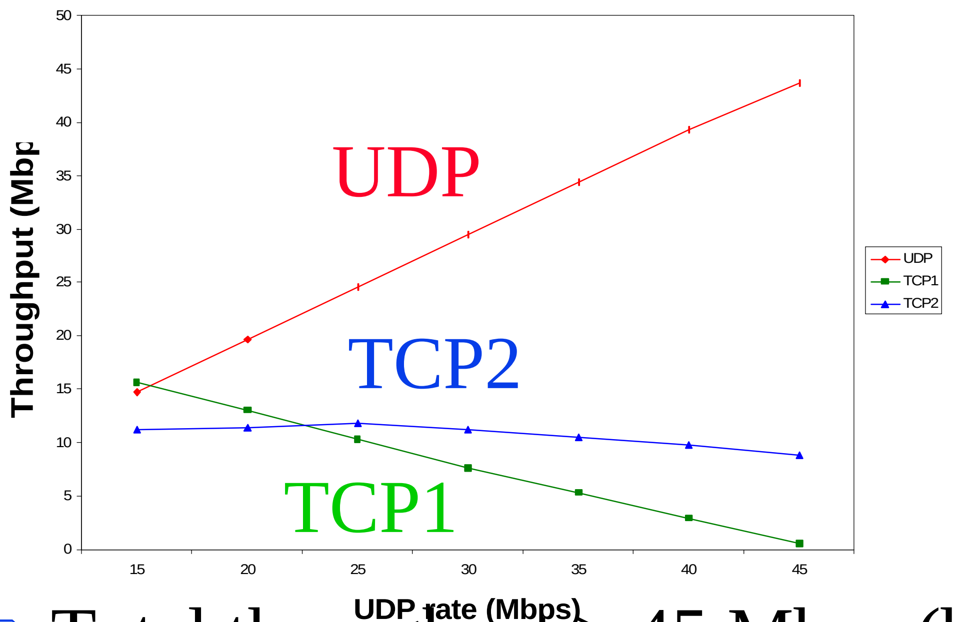
Case 1: No Trunks, No MPLS

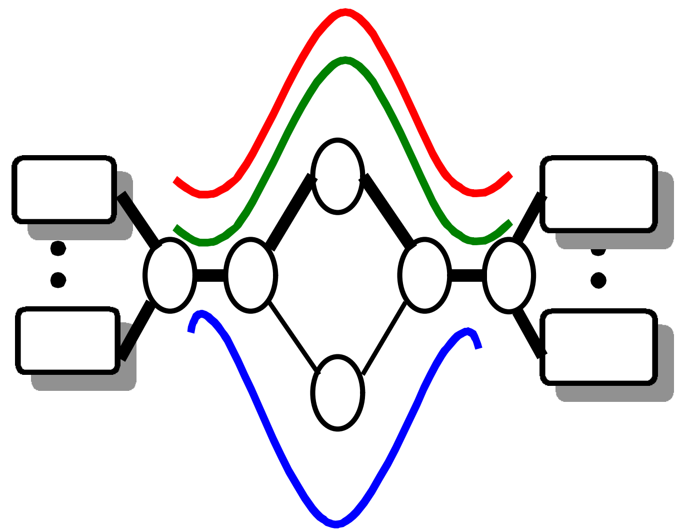




- 15 Mbps path not used at all
- TCP suffers as UDP increases its rate
- Unfairness among TCP flows

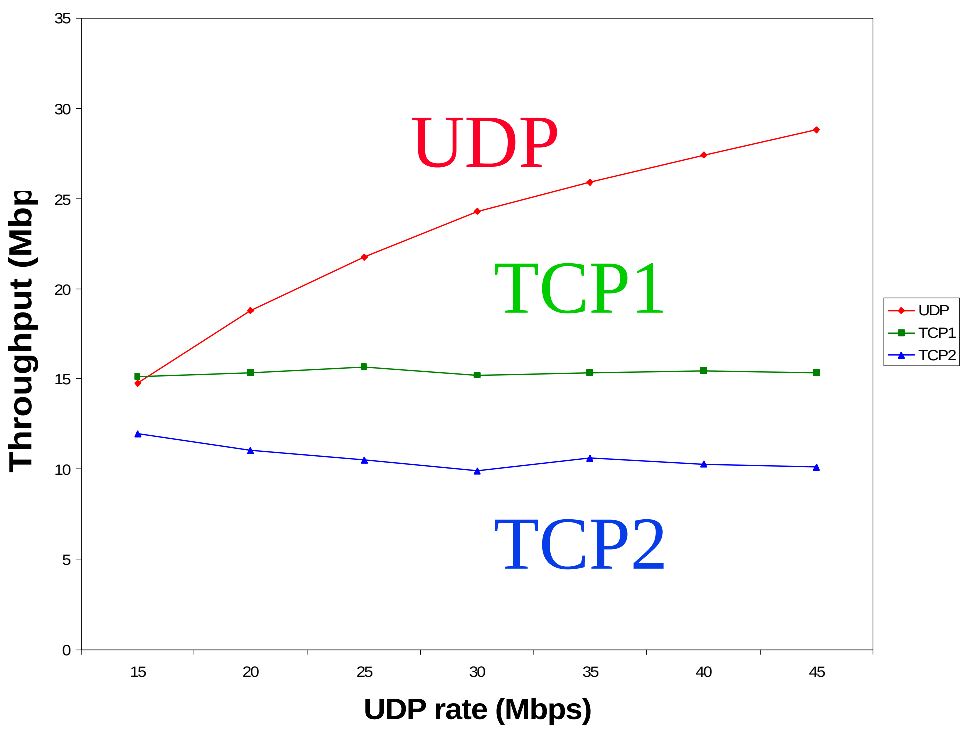
Two trunks w UDP + TCP Mixed

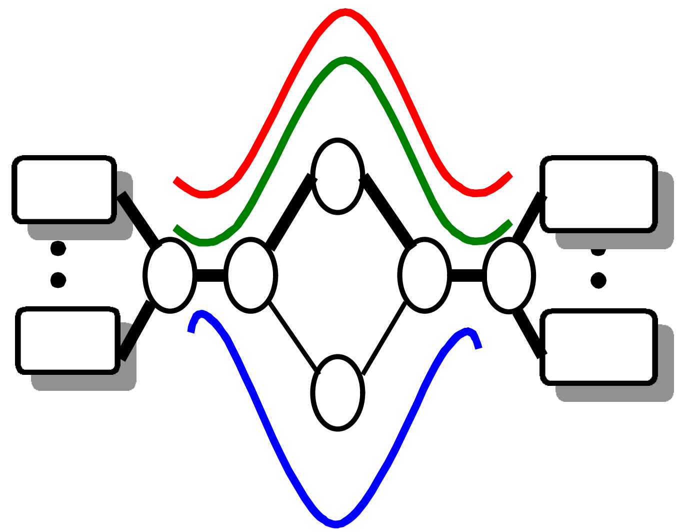




- Total throughput > 45 Mbps (both paths used)
- TCP flows sharing the trunk with UDP suffer
- TCP flow not sharing with UDP do not suffer

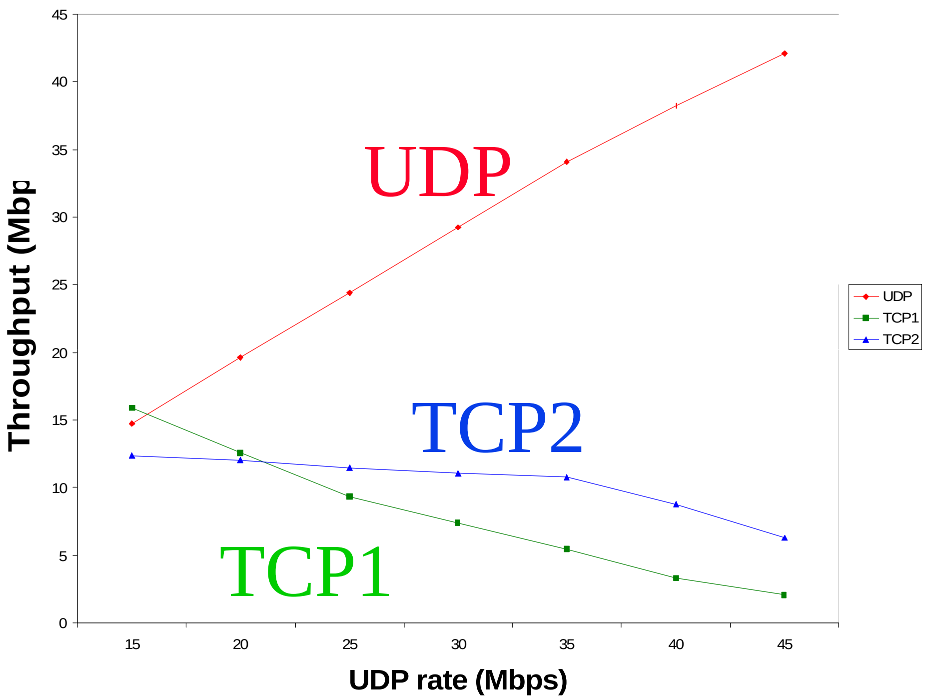
3 Trunks w Isolated TCP, UDP

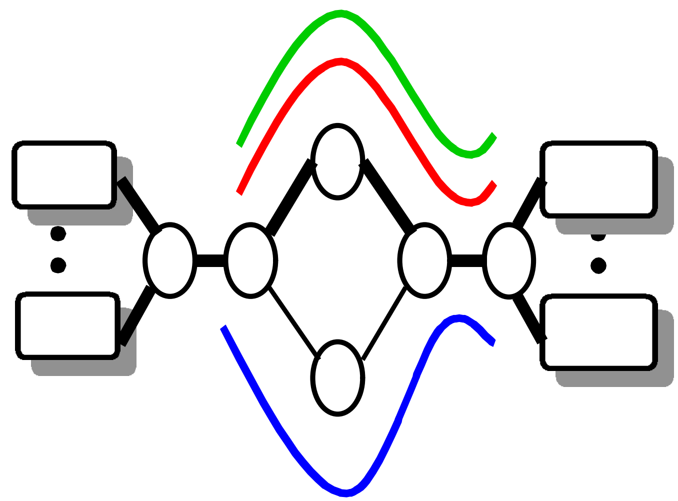




- TCP flows are not affected by UDP and achieve a fairly constant throughput

Non End-to-End Trunks



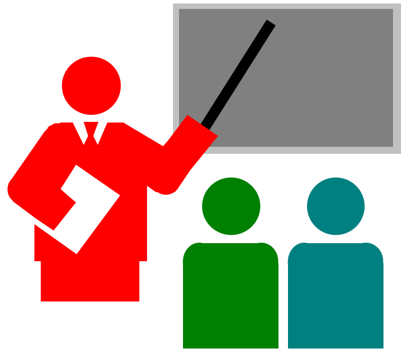


- TCP flows are affected by UDP in the shared path

Future Work

- Other Traffic Scenarios:
  - Aggregate flows: TCP+UDP
  - Short duration TCP connections
  - Bursty (Web) traffic
- Queue Service Policies: WFQ, WF2Q, WF2Q+
- Packet drop policies: RED, Tail drop
- Round Trip Time
- TCP parameters: MSS, window size, etc.
- DiffServ vs MPLS, DiffServ+MPLS

Summary of Part 3: TE Analysis



- Total network throughput improves significantly with proper traffic engineering
- Congestion-unresponsive flows affect congestion-responsive flows
  - Separate trunks for different types of flows
- Trunks should be end-to-end
  - Trunk + No Trunk = No Trunk

Part 4: Other QoS Approaches and MPLS Interoperability

- ATM
- Integrated Services/RSVP
- Differentiated Services
- IEEE 802.1D

ATM Service Categories

- **CBR**: Throughput, delay, delay variation
- **rt-VBR**: Throughput, delay, delay variation
- **nrt-VBR**: Throughput
- **UBR**: No Guarantees
- **GFR**: Minimum Throughput
- **ABR**: Minimum Throughput. Very low loss. Feedback.
- ATM also has QoS-based routing (PNNI)

ATM QoS: Issues

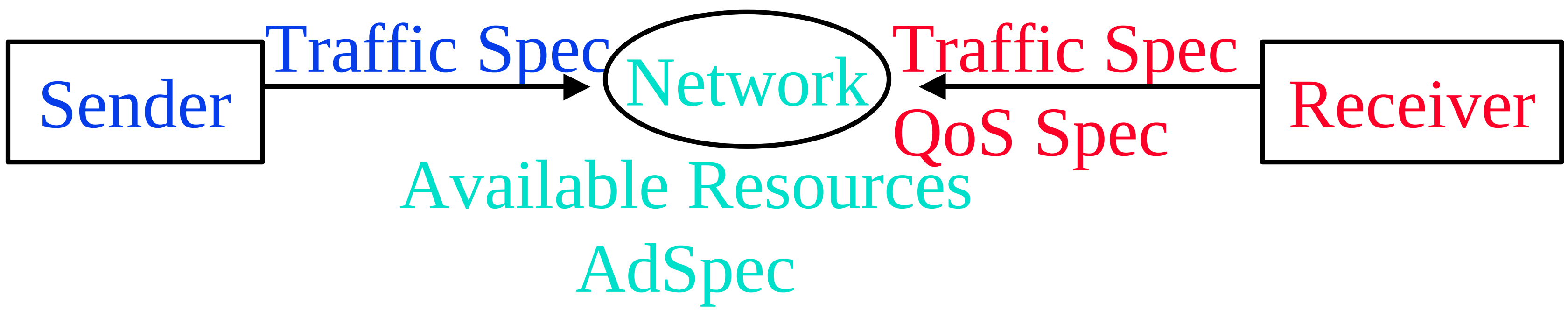
- Can't easily aggregate QoS: VP = Σ VCs
- Can't easily specify QoS: What is the CDV required for a movie?
- Signaling too complex ⇒ Need Lightweight Signaling
- Need Heterogeneous Point-to-Multipoint: Variegated VCs
- Need QoS Renegotiation
- Need Group Address
- Need priority or weight among VCs to map DiffServ and 802.1D
- MPLS also has many of these problems.

Integrated Services

- Best Effort Service: Like UBR.
- Controlled-Load Service: Performance as good as in an unloaded datagram network. No quantitative assurances. Like nrt-VBR or UBR w MCR
- Guaranteed Service: rt-VBR
  - Firm bound on data throughput and delay.
  - Delay jitter or average delay not guaranteed or minimized.
  - Every element along the path must provide delay bound.
  - Is not always implementable, e.g., Shared Ethernet.
  - Like CBR or rt-VBR

RSVP

- Resource ReSerVation Protocol
- Internet signaling protocol
- Carries resource reservation requests through the network including traffic specs, QoS specs, network resource availability
- Sets up reservations at each hop



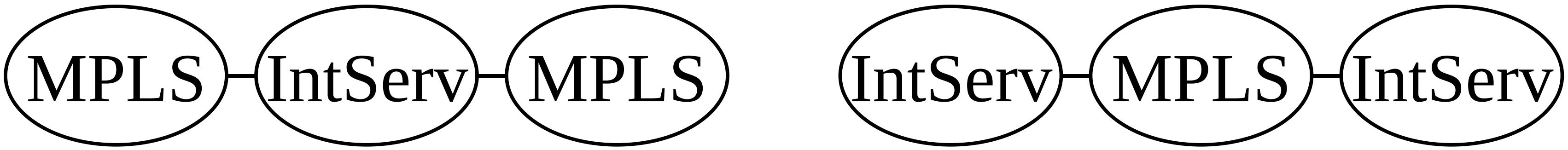
Problems with IntServ/RSVP

- Complexity in routers: packet classification, scheduling
- Per-Flow State: O(n) ⇒ Not scalable with # of flows. Number of flows in the backbone may be large.
  ⇒ Suitable for small private networks
- Need a concept of "Virtual Paths" or aggregated flow groups for the backbone
- Need policy controls: Who can make reservations? Support for accounting and security.
  ⇒ RSVP admission policy (rap) working group.

Problems (Cont)

- Receiver Based:
  Need sender control/notifications in some cases.
  Which receiver pays for shared part of the tree?
- Soft State: Need route/path pinning (stability).
  Limit number of changes during a session.
- RSVP does not have negotiation and backtracking
- Throughput and delay guarantees require support of lower layers. Shared Ethernet $\Rightarrow$ IP can't do GS or CLS. Need switched full-duplex LANs.
- MPLS solves many of these problems.

MPLS-IntServ Interoperability



- IntServ is more complex and will be less widely implemented.
- MPLS over IntServ: Not a realistic scenario.
- IntServ over MPLS:
    - MPLS can provide controlled service, guaranteed service and best effort services without the need for classification at each hop.

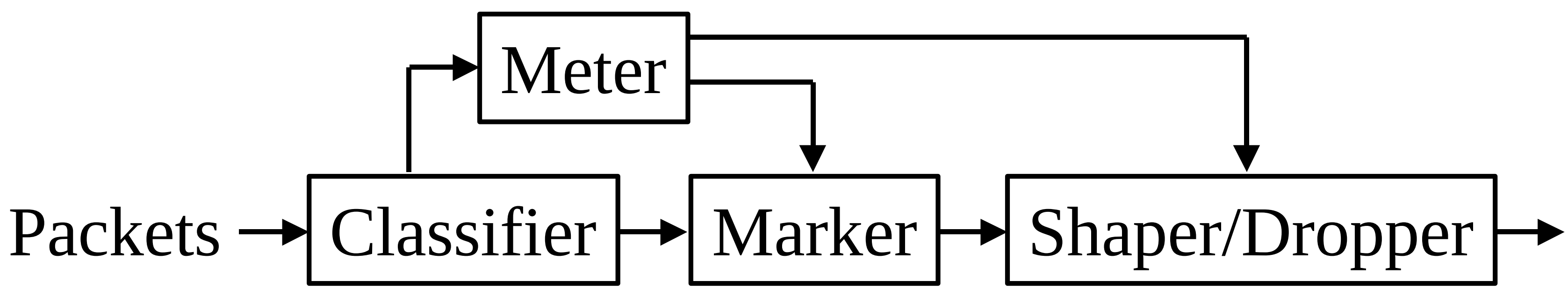
Differentiated Services

Ver	Hdr Len	Precedence	ToS	Unused	Tot Len
4b	4b	3b	4b	1b	16b

- IPv4: 3-bit precedence + 4-bit ToS
- OSPF and integrated IS-IS can compute paths for each ToS
- Many vendors use IP precedence bits but the service varies ⇒ Need a standard ⇒ Differentiated Services
- DS working group formed February 1998
- Charter: Define ds byte (IPv4 ToS field)
- Mail Archive: http://www-nrg.ee.lbl.gov/diff-serv-arch/

DiffServ Concepts

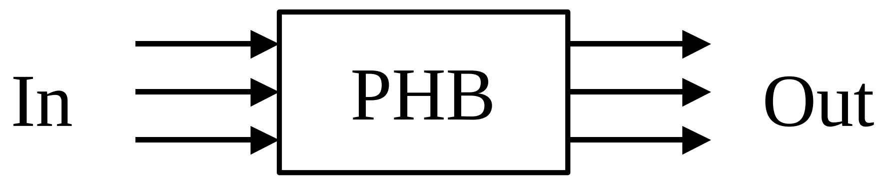
- Micro-flow = A single application-to-application flow
- Traffic Conditioners: Meters (token bucket), Markers (tag), Shapers (delay), Droppers (drop)
- Behavior Aggregate (BA) Classifier:
  Based on DS byte only
- Multi-field (MF) Classifiers:
  Based on IP addresses, ports, DS-byte, etc..



Diff-Serv Concepts (Cont)

- Service: Offered by the protocol layer
  - Application: Mail, FTP, WWW, Video,...
  - Transport: Delivery, Express Delivery,...
    Best effort, controlled load, guaranteed service
  - DS group will not develop services
    They will standardize "Per-Hop Behaviors"

Per-hop Behaviors

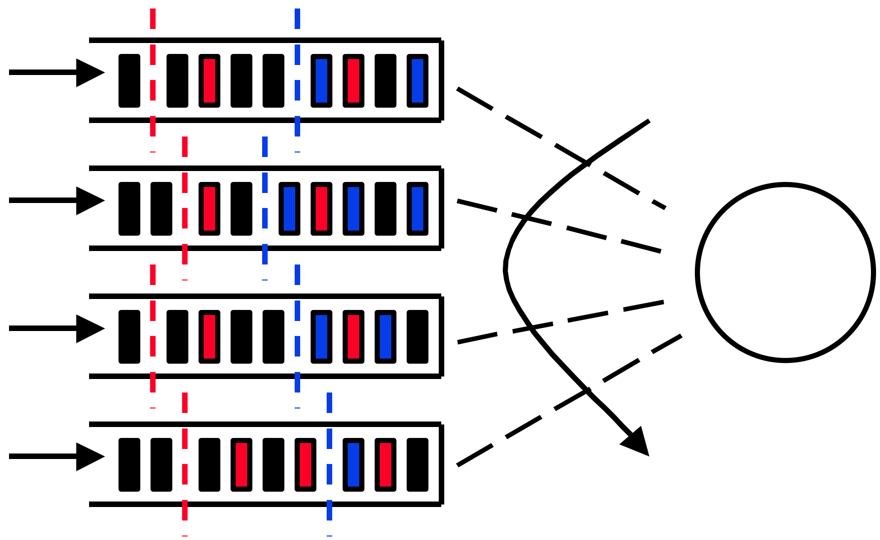


- Externally Observable Forwarding Behavior
- x% of link bandwidth
- Minimum x% and fair share of excess bandwidth
- Priority relative to other PHBs
- PHB Groups: Related PHBs. PHBs in the group share common constraints, e.g., loss priority, relative delay

Expedited Forwarding

- Also known as "Premium Service"
- Virtual leased line
- Similar to CBR
- Guaranteed minimum service rate
- Policed: Arrival rate < Minimum Service Rate
- Not affected by other data PHBs
  $\Rightarrow$ Highest data priority (if priority queueing)
- Code point: 101 110

Assured Forwarding



- PHB Group
- Four Classes: No particular ordering
- Three drop preference per class

Assured Forwarding (Cont)

- DS nodes SHOULD implement all 4 classes and MUST accept all 3 drop preferences. Can implement 2 drop preferences.
- Similar to nrt-VBR/ABR/GFR
- Code Points:

Drop Prec.	Class 1	Class 2	Class 3	Class 4
Low	010 000	011 000	100 000	101 000
Medium	010 010	011 010	100 010	101 010
High	010 100	011 100	100 100	101 100

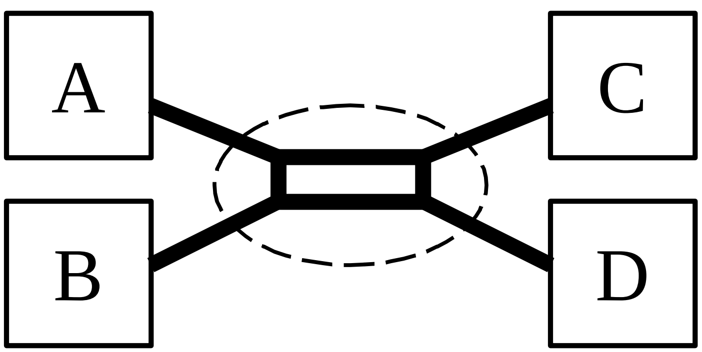
- Avoids 11x000 (used for network control)

Problems with DiffServ

- per-hop ⇒ Need at every hop
  One non-DiffServ hop can spoil all QoS
  This applies to almost all QoS approaches.
- End-to-end ≠ Σ per-Hop
  Designing end-to-end services with weighted guarantees at individual hops is difficult.
  Only EF will work.
- Designed for static Service Level Agreements (SLAs)
  Both the network topology and traffic are highly dynamic.
- Multicast ⇒ Difficult to provision
  Dynamic multicast membership ⇒ Dynamic SLAs?

DiffServ Problems (Cont)

- DiffServ is unidirectional ⇒ No receiver control
- Modified DS field ⇒ Theft and Denial of service. Ingress node should ensure.
- How to ensure resource availability inside the network?
- QoS is for the aggregate not per-destination. Multi-campus enterprises need inter-campus QoS.



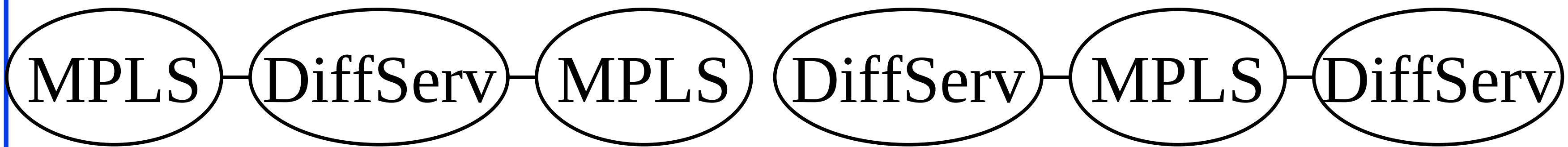
DiffServ Problems (Cont)

- QoS is for the aggregate not micro-flows.
  Not intended/useful for end users. Only ISPs.
  - Large number of short flows are better handled by aggregates.
  - Long flows (voice and video sessions) need per-flow guarantees.
  - High-bandwidth flows (1 Mbps video) need per-flow guarantees.
- All IETF approaches are open loop control $\Rightarrow$ Drop
  Closed loop control $\Rightarrow$ Wait at source
  Data prefers waiting $\Rightarrow$ Feedback

DiffServ Problems (Cont)

- Guarantees ⇒ Stability of paths
  ⇒ Connections (hard or soft)
  Need route pinning or connections.

MPLS-DiffServ Interoperability



- MPLS is borrowing the best of DiffServ and can be end-to-end.
- MPLS over DiffServ:
  No end-to-end guarantees $\Rightarrow$ Not useful
- DiffServ over MPLS:
  - DS byte can be encoded in CR-LDP label requests and responses.

IEEE 802.1D Model

Dest Addr	Src Addr	Tag Prot ID	Pri	CFI	VLAN ID

802.1Q header

Prot Type	Payload	FCS

CFI = Canonical Format Indicator (Source Routing)

- **Up to eight priorities:** Strict.

  1 Background

  2 Spare

  0 Best Effort

  3 Excellent Effort

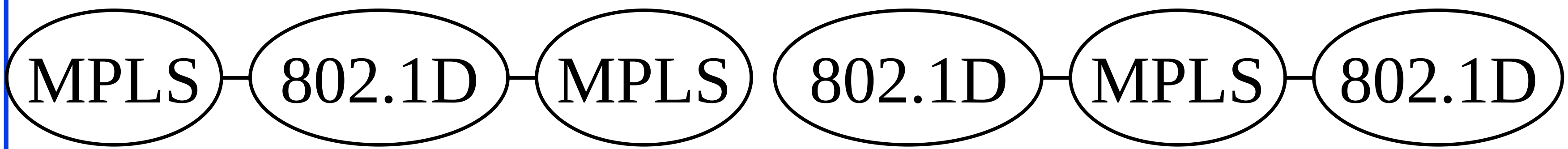
  4 Control load

  5 Video (Less than 100 ms latency and jitter)

  6 Voice (Less than 10 ms latency and jitter)

  7 Network Control

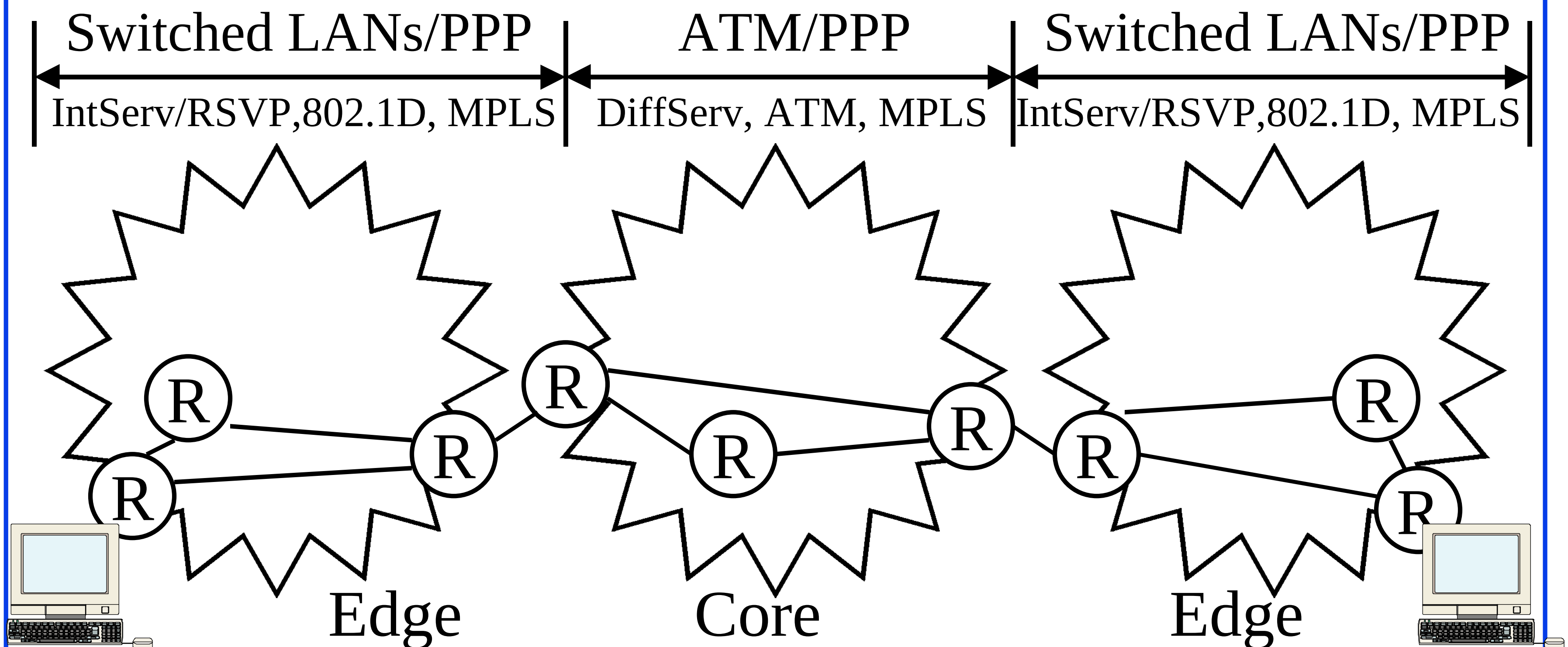
MPLS-802.1D Interoperability



- MPLS over 802.1D: Priority among packets at the same node. Lower priority traffic from other nodes can get through.
- 802.1D Traffic over MPLS:
  - Packet priority can be encoded in Exp field, label
  - Trunk priority can be encoded in CR-LDP label requests and responses.

End-to-end View

- ATM/PPP backbone, Switched LANs/PPP in Stub
- IntServ/RSVP, 802.1D, MPLS in Stub networks
- DiffServ, ATM, MPLS in the core



QoS Debate Issues

- Massive Bandwidth vs Managed Bandwidth
- Per-Flow vs Aggregate
- Source-Controlled vs Receiver Controlled
- Soft State vs Hard State
- Path based vs Access based
- Quantitative vs Qualitative
- Absolute vs Relative
- End-to-end vs Per-hop
- Static vs Feedback-based
- One-way multicast vs n-way multicast
- Homogeneous multicast vs heterogeneous multicast
- Single vs multiple bottlenecks: Scheduling

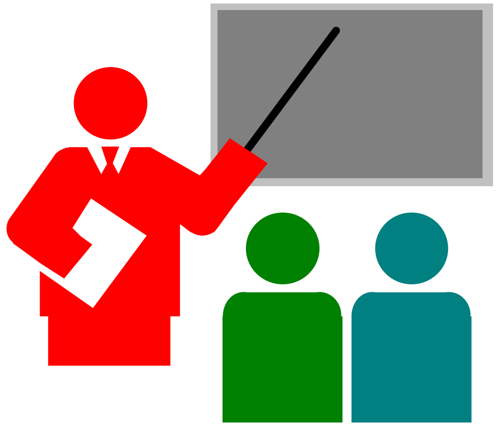
Comparison of QoS Approaches

Issue	ATM	IntServ	DiffServ	MPLS	IEEE 802.3D
Massive Bandwidth vs Managed Bandwidth	Managed	Managed	Massive	Managed	Massive
Per-Flow vs Aggregate	Both	Per-flow	Aggregate	Both	Aggregate
Source-Controlled vs Receiver Controlled	Unicast Source, Multicast both	Receiver	Ingress	Both	Source
Soft State vs Hard State	Hard	Soft	None	Hard	Hard
Path based vs Access based	Path	Path	Access	Path	Access
Quantitative vs Qualitative	Quantitative	Quantitative+Qualitative	Mostly qualitative	Both	Qualitative
Absolute vs Relative	Absolute	Absolute	Mostly Relative	Absolute plus relative	Relative

Comparison (Cont)

Issue	ATM	IntServ	DiffServ	MPLS	IEEE 802.3D
End-to-end vs Per-hop	e-e	e-e	Per-hop	e-e	Per-hop
Static vs Feedback-based	Both	Static	Static	Static	Static
One-way multicast vs n-way multicast	Only one-way				
Homogeneous multicast vs heterogeneous multicast	Homogeneous	Heterogeneous	N/A	Homogeneous	N/A
Single vs multiple bottlenecks: Scheduling	Multiple bottleneck	Multiple		Multiple	

Summary of Part 4



- MPLS is taking the best features of ATM, IntServ, DiffServ, and 802.1D QoS approaches
  $\Rightarrow$ MPLS is most promising
- MPLS provides a superset of functionality of many of these other technologies
- Features $\Rightarrow$ Complexity
  Complexity has to be controlled.

References

- References on MPOA, MPLS, and IP Switching, http://www.cis.ohio-state.edu/~jain/refs/ipsw_ref.htm
- Quality of Service using Traffic Engineering over MPLS: An Analysis, http://www.cis.ohio-state.edu/~jain/papers/mpls-te-anal.htm
- IP Switching, http://www.cis.ohio-state.edu/~jain/cis788-97/ip_switching/index.htm
- References on QoS over IP, http://www.cis.ohio-state.edu/~jain/refs/ipqs_ref.htm
- IP Switching: Issues and Alternatives, http://www.cis.ohio-state.edu/~jain/talks/ipsw.htm

References (Cont)

- Quality of Service in IP Networks, http://www.cis.ohio-state.edu/~jain/talks/ipqos.htm
- Requirements for Traffic Engineering over MPLS, draft-ietf-mpls-traffic-eng-01.txt
- Constraint-based LSP Setup using LDP, draft-ietf-mpls-cr-ldp-01.txt
- Optimizing Routing Software for Reliable Internet Growth, http://www.juniper.net/techcenter/techpapers/optimizing-routing-sw.fm.html

References (Cont)

- Cisco - Multiprotocol Label Switching, http://www.cisco.com/warp/public/784/packet/apr99/6.html

Thank You!