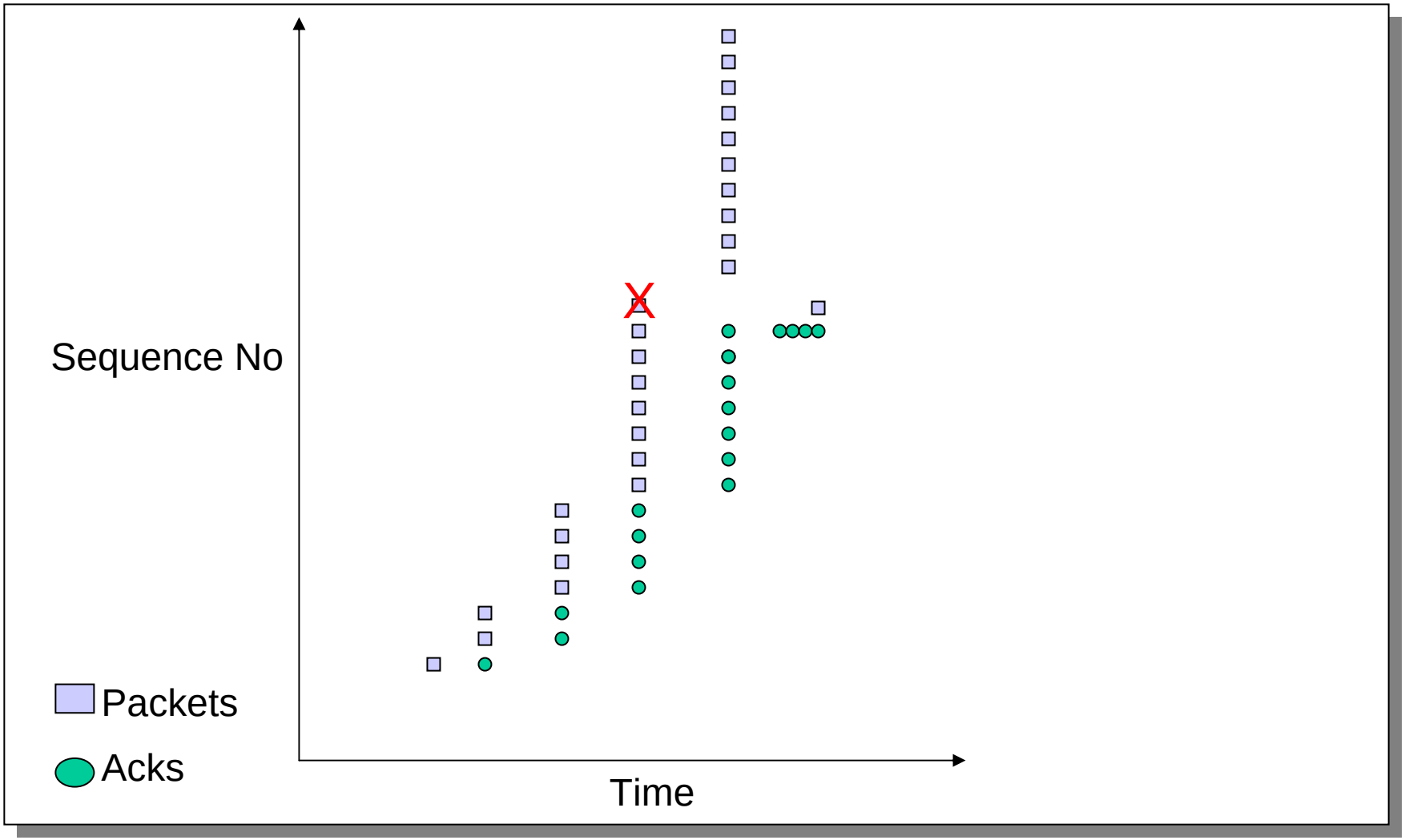
15-441 Computer Networks

ATM and MPLS

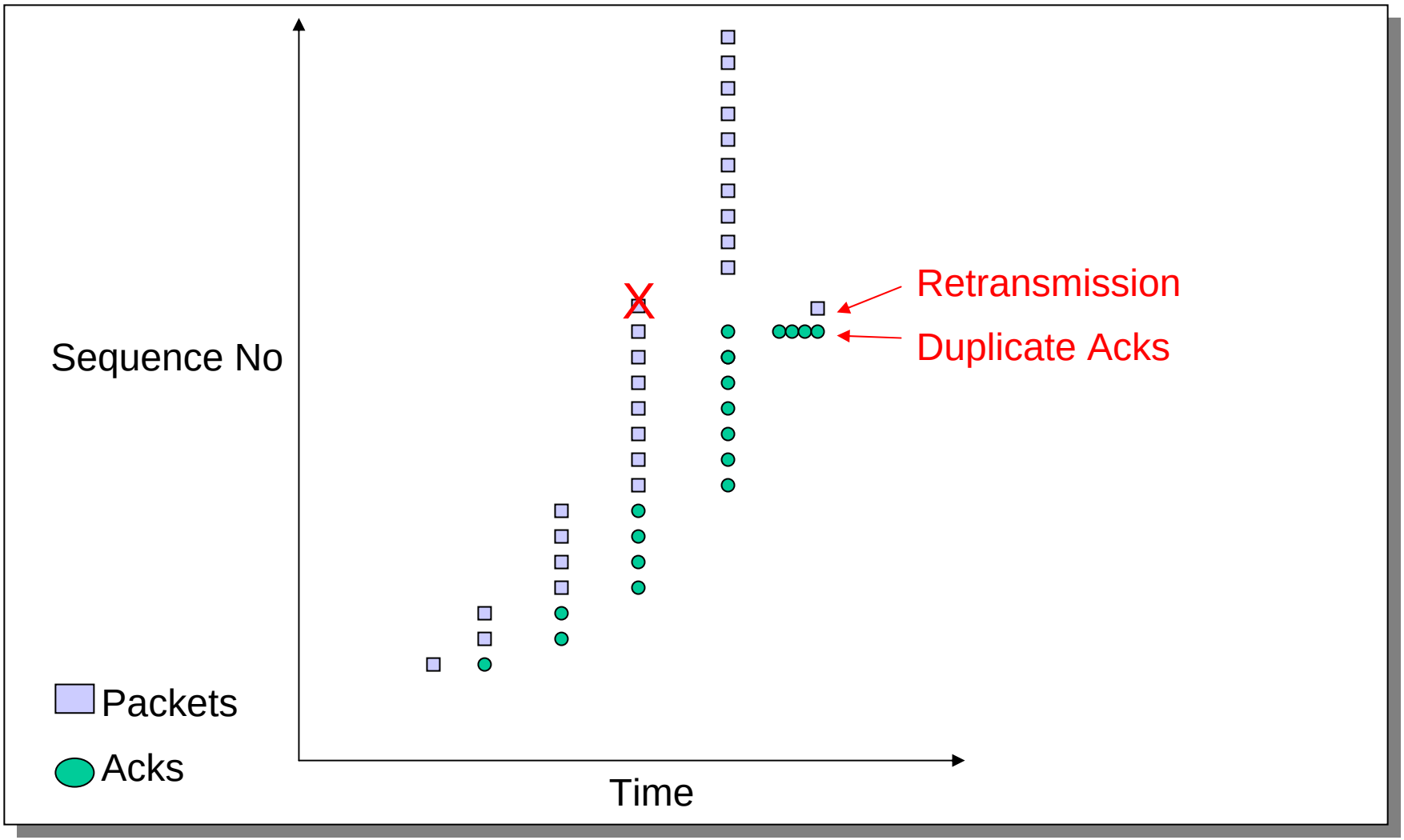
Professor Hui Zhang

hzhang@cs.cmu.edu

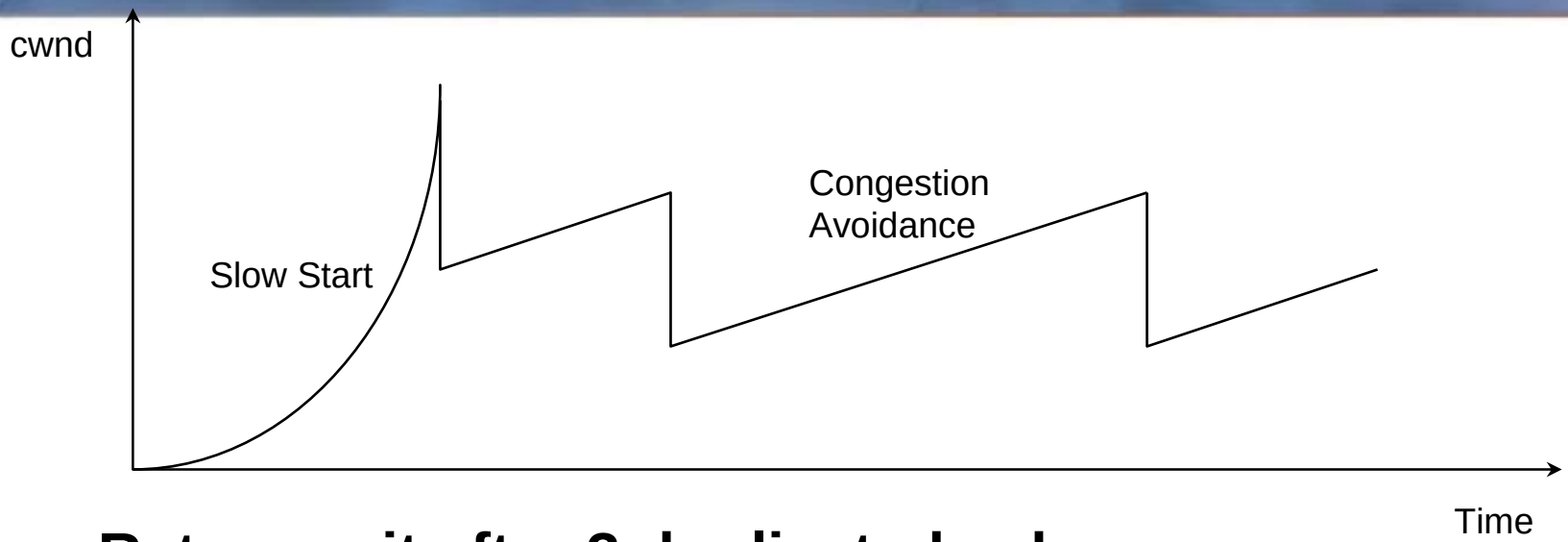
Fast Retransmit



Fast Retransmit

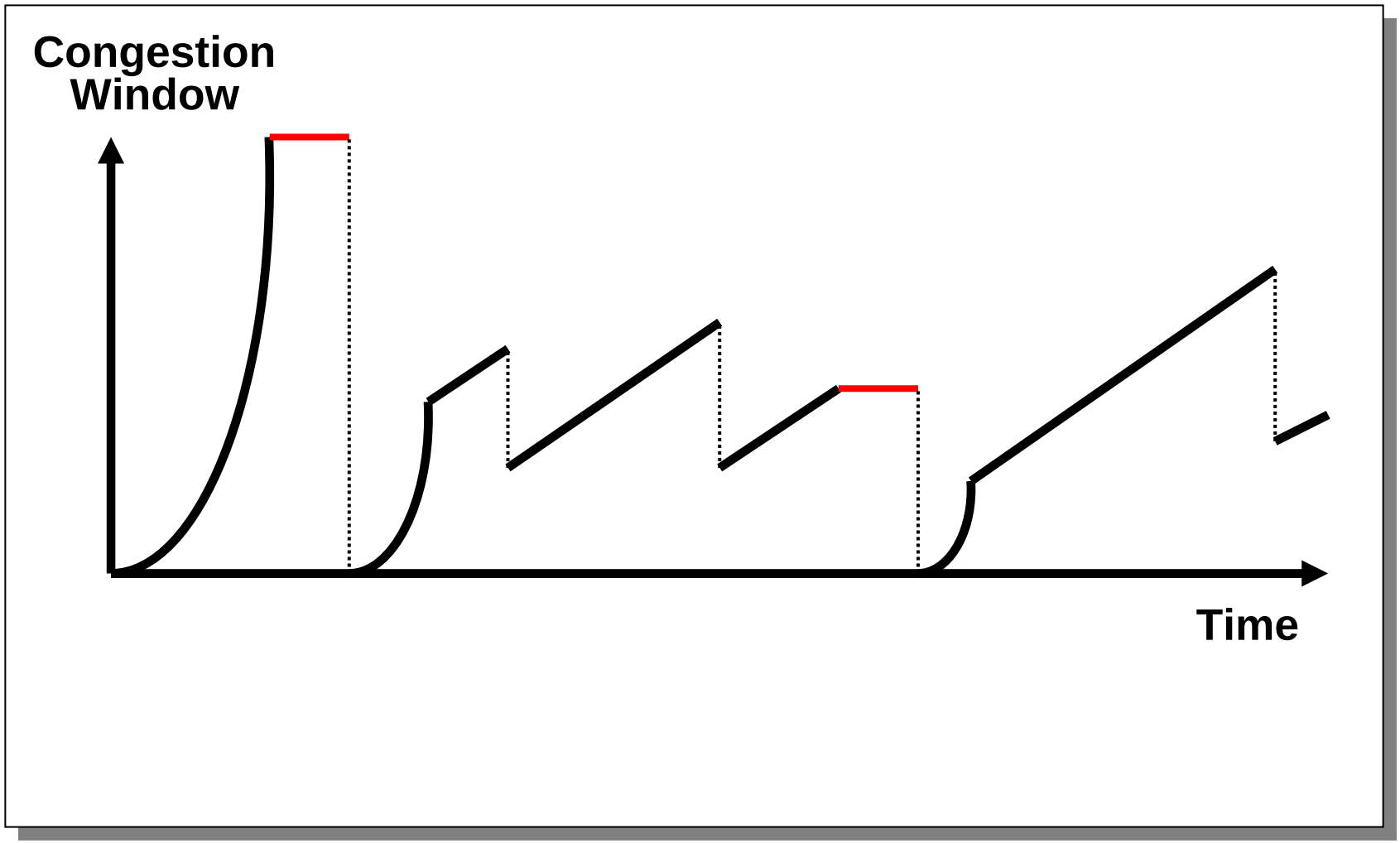


Fast Retransmit and Fast Recovery

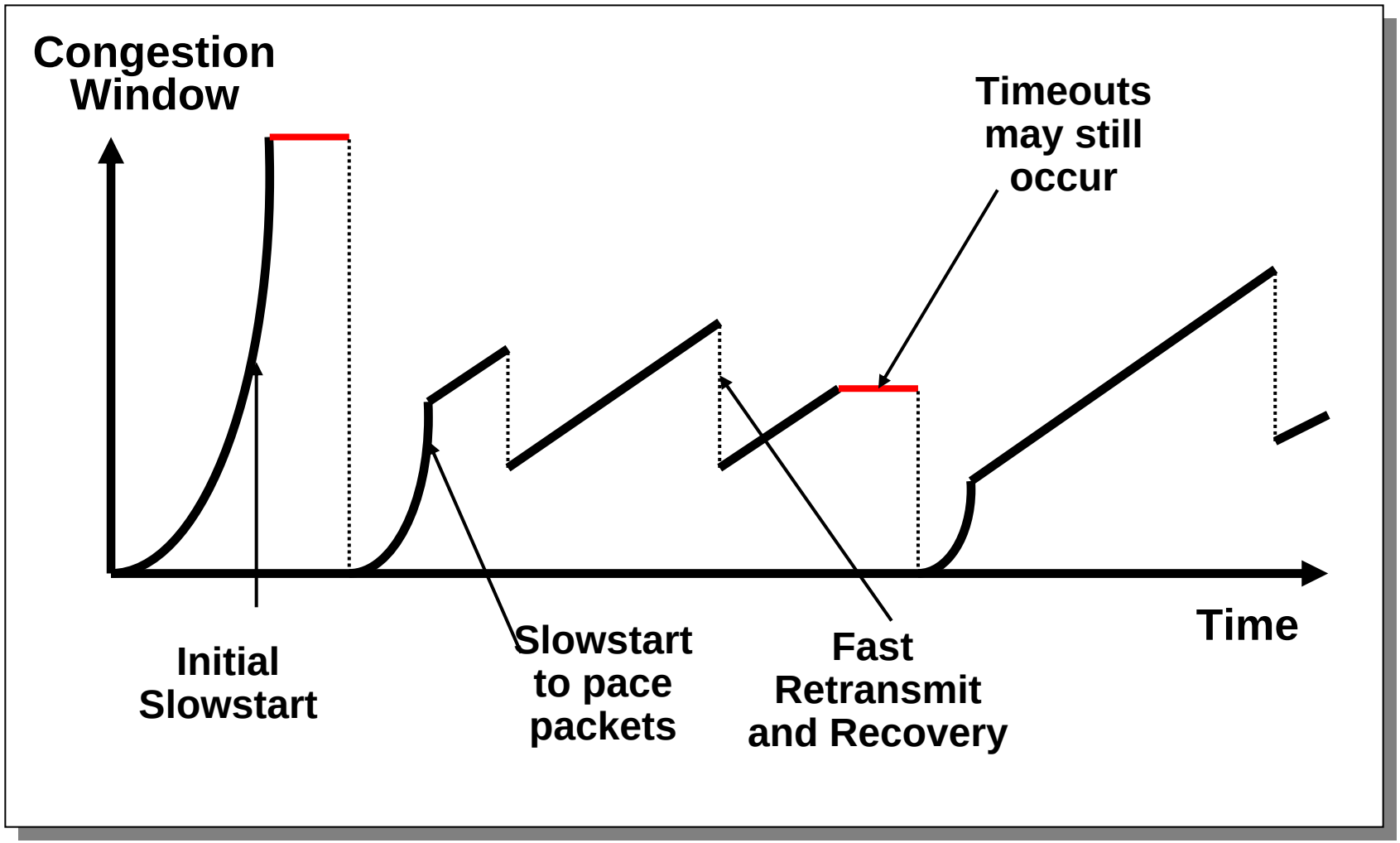


- **Retransmit after 3 duplicated acks**
  - prevent expensive timeouts
- **No need to slow start again**
- **At steady state, cwnd oscillates around the optimal window size.**

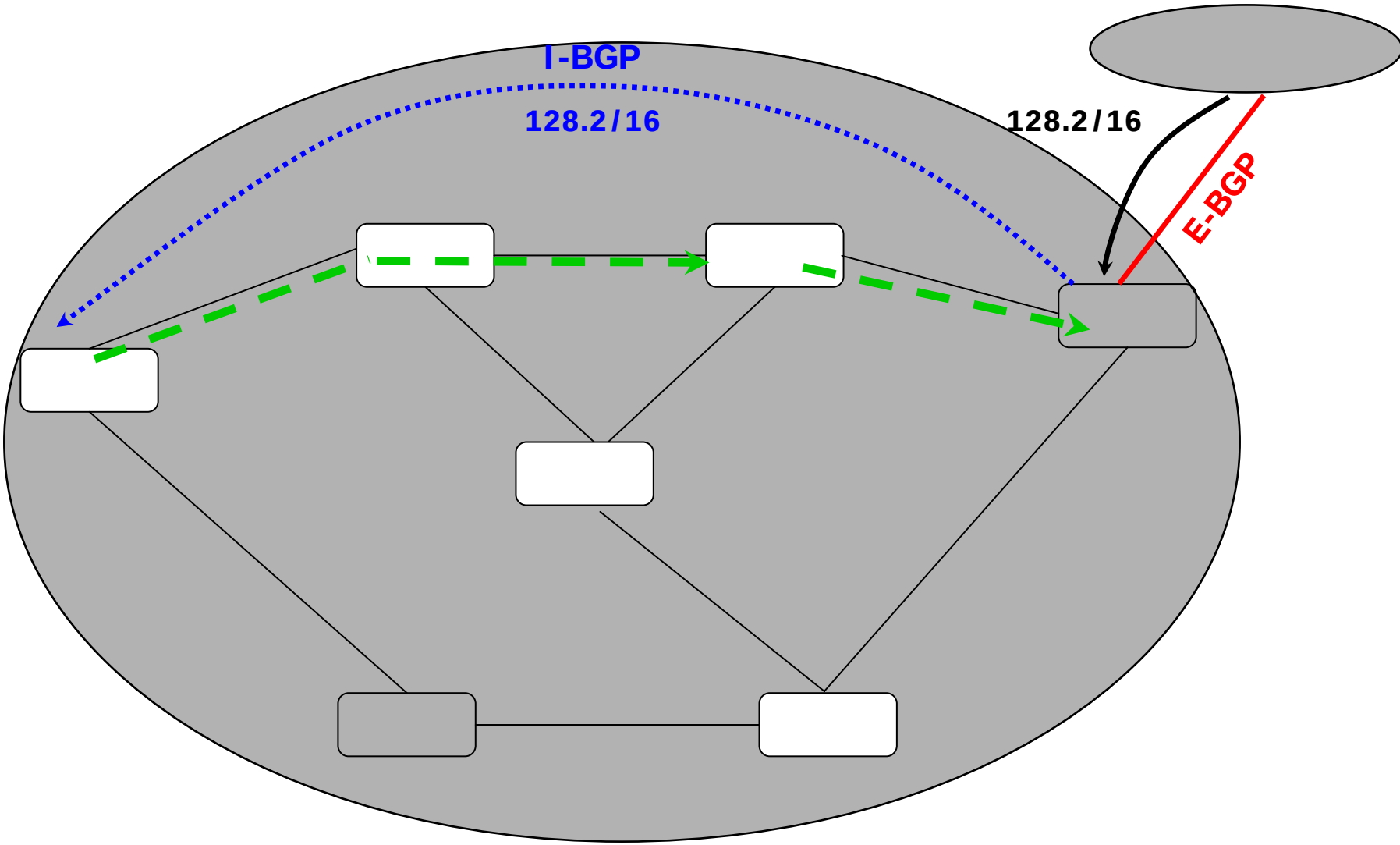
Review: TCP Congestion Control



Review: TCP Congestion Control



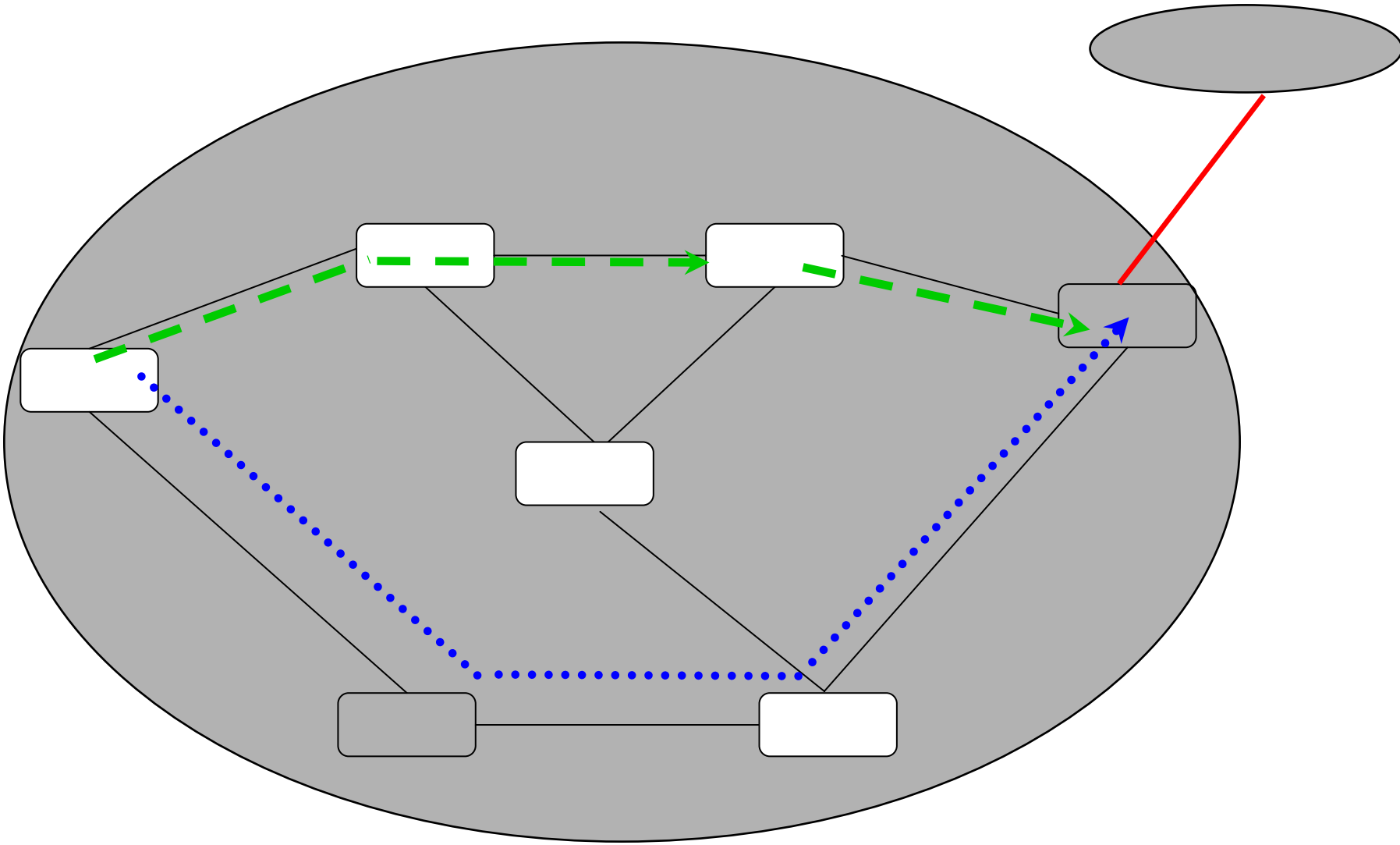
Datagram Routing



Inter-Domain Routing Review

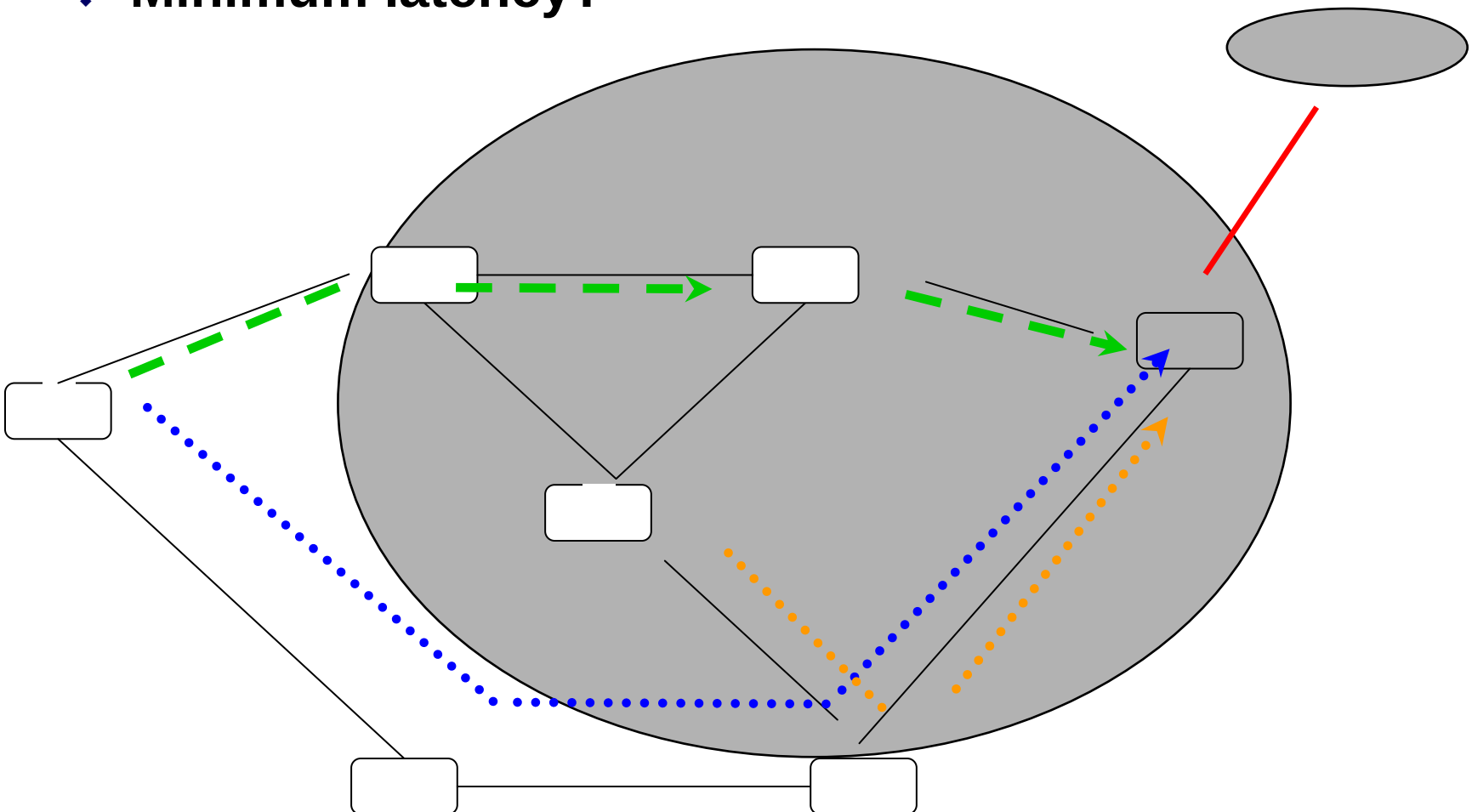
- **Each border router receives multiple E-BGP and I-BGP feeds, it needs to select, among many choices, one AS path for each prefix**
  - Consistency requirement: no AS level loop
- **Each internal router receives multiple I-BGP feeds, it needs to select, among many choices, one egress border router for each prefix**
  - Consistency requirement: different internal nodes need to pick consistent egress nodes for each prefix

Traffic Engineering?



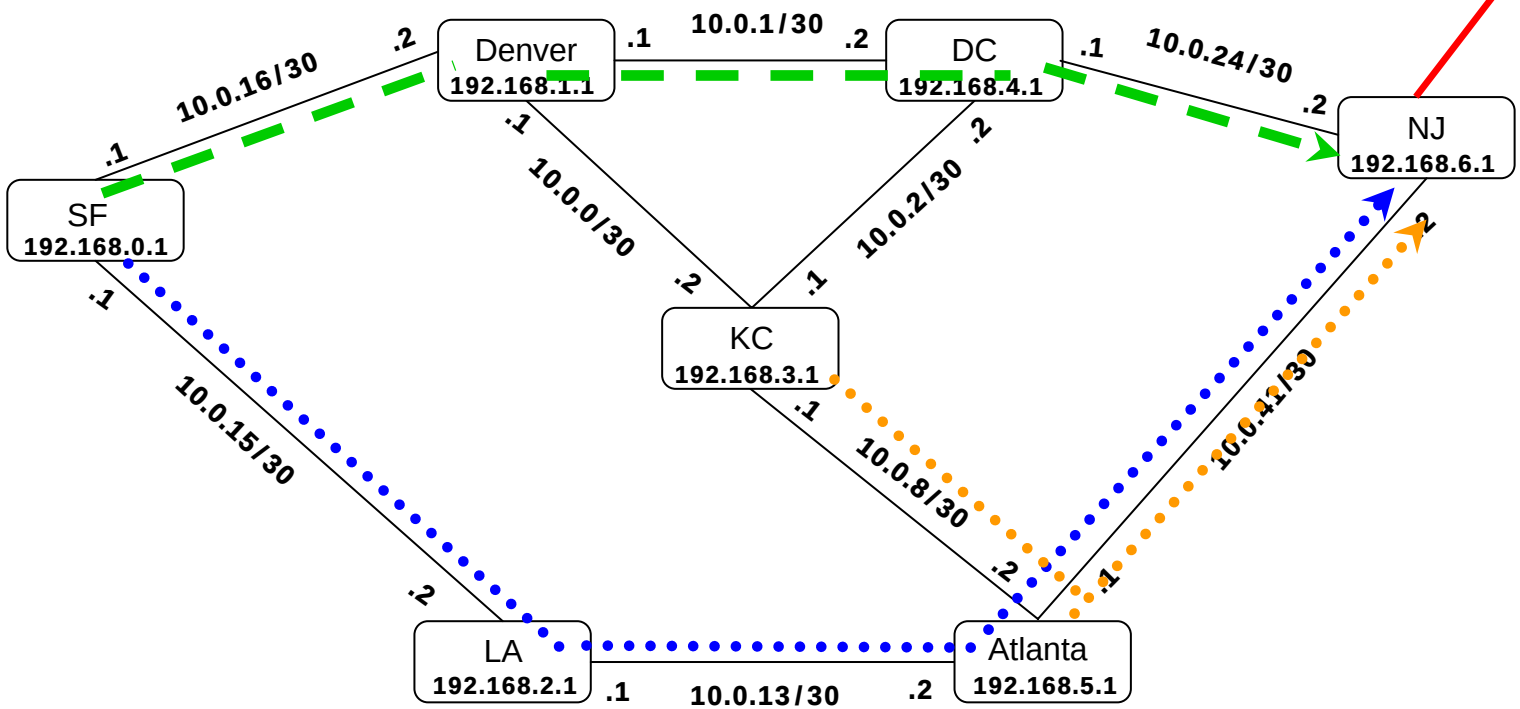
Quality of Service ?

- Guaranteed 40Mbps throughput for blue traffic?
- Minimum latency?

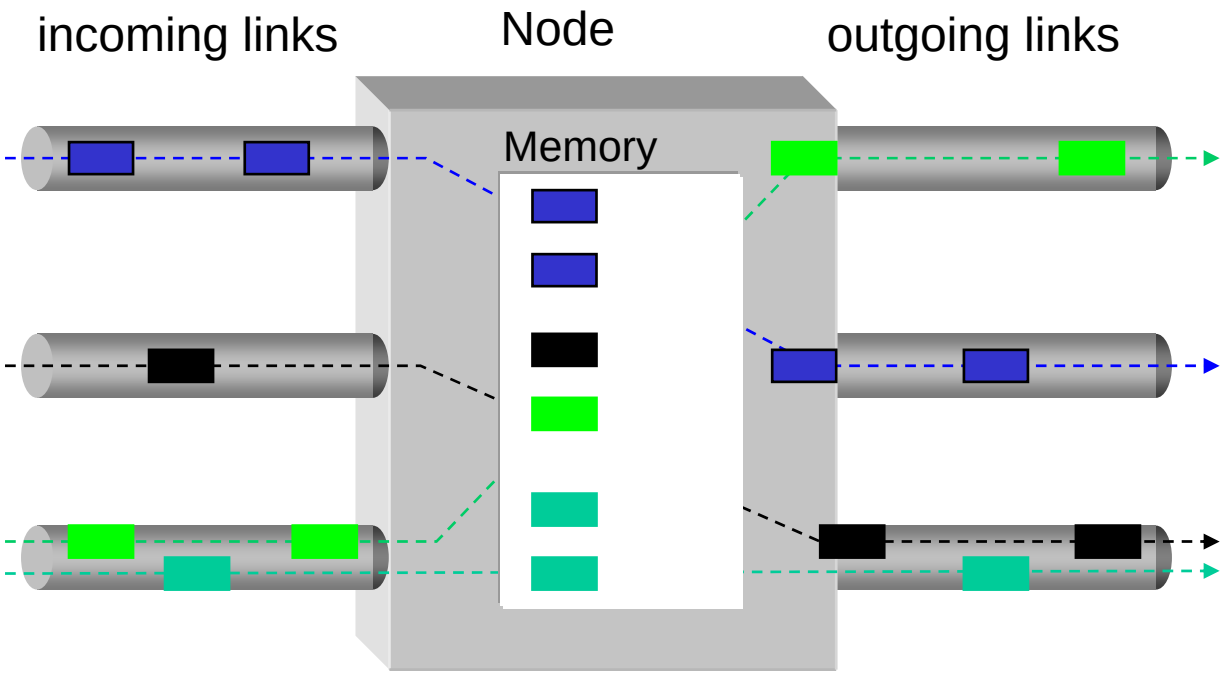


Quality of Service?

- Guaranteed 40Mbps throughput for blue traffic?
- Minimum latency?



High Speed IP Router?



- **Dest IP address lookup for every packet**
- **Switching and buffer management for variable size packet**

Asynchronous Transfer Mode: ATM

- ITU standard for high-speed **(155Mbps to 622 Mbps and higher)** ***Broadband Integrated Service Digital Network*** **architecture**
  - Most of the work done in 90's
- **Goal:** ***integrated, end-end transport of carry voice, video, data***
  - meeting timing/QoS requirements of voice, video (versus Internet best-effort model)
  - "next generation" telephony: technical roots in telephone world

Four Ideas

- **Flexible and efficient → packet switching**
- **Support voice, video, data**
  - → needs end-to-end Quality of Service
  - → connection-oriented network
  - → virtual circuit network
- **Scale to high performance switch → fixed size packet**
- **Support low jitter voice → small size packet**

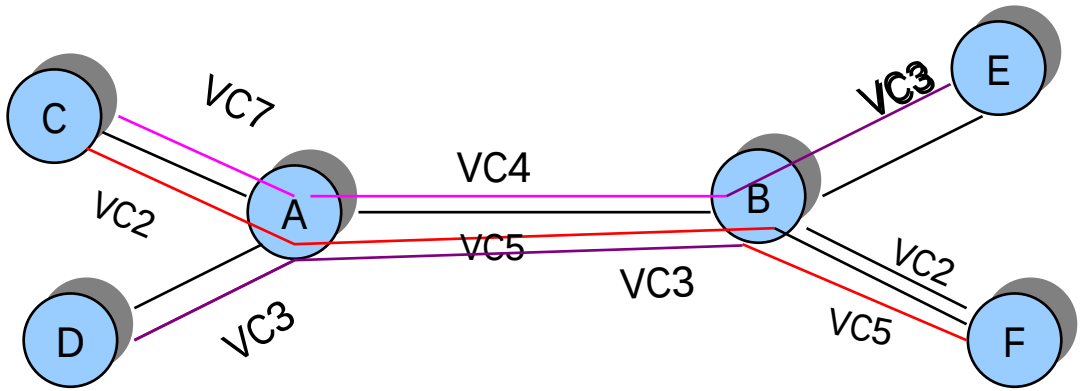
Virtual Circuit Concept

- **Logical Connection**
- **Connection is first established using signaling protocol**
  - Route from the source to the destination is chosen
  - The same route is used for all packets of the connection
- **No routing decision for every cell**

Virtual Circuit Concepts

- **No dedicated capacity**
  - Packet switching to enable statistical multiplexing
- **Each packet contains enough information for node (switch) to forward it towards the destination**

Virtual Circuit Switching: Label Swapping



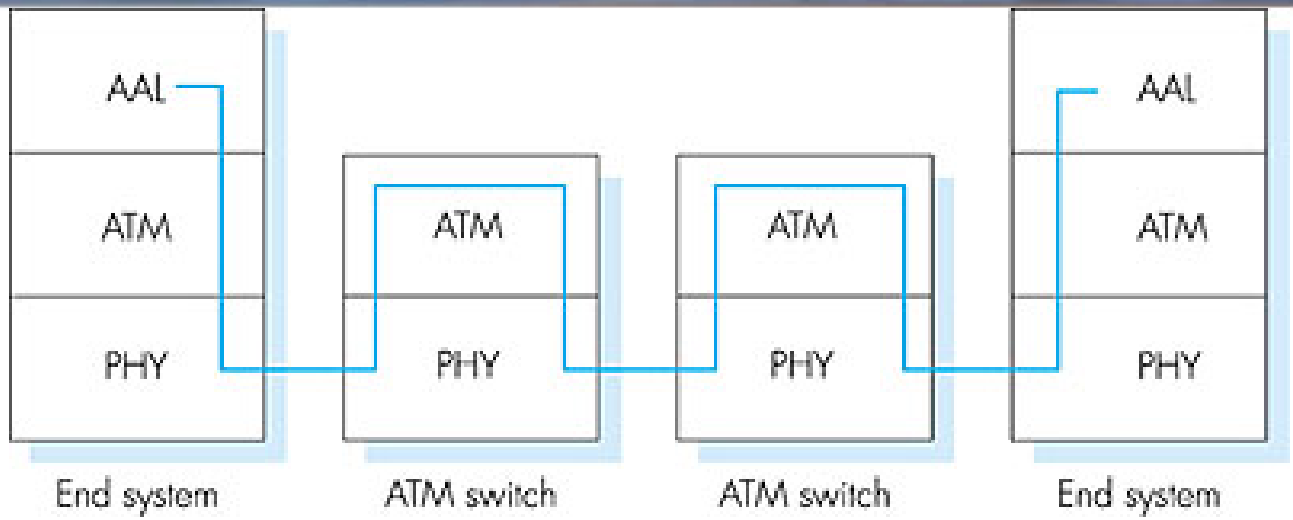
IN LINK	IN VC	OUT LINK	OUT VC
CA	**7**	**AB**	**4**
CA	**2**	**AB**	**5**
DA	**3**	**AB**	**3**

Table at Node A

Signaling Protocol

- **Signaling protocol establishes/tears down virtual circuit**
  - Signaling message are routed
  - Signaling protocol fills the forwarding table
- **Parameters used for establishing Virtual Circuits**
  - Source and destination Addresses
  - Traffic Characteristics
  - QoS Parameters
  - Others?
- **Parameters can be stored in forwarding table to help forwarding decision**

ATM architecture



- Adaptation layer: only at edge of ATM network
  - data segmentation/reassembly
  - roughly analagous to Internet transport layer
- **ATM layer: "network" layer**
  - cell switching, routing
- **Physical layer**

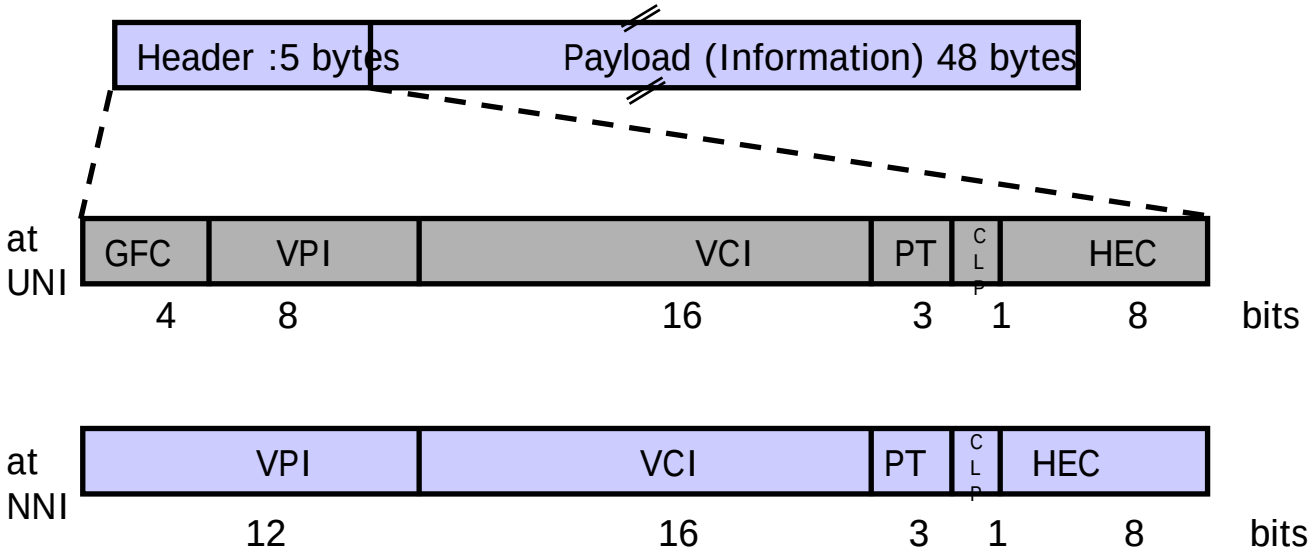
ATM Layer: Virtual Circuits

- **VC transport: cells carried on VC from source to dest**
  - call setup, teardown for each call *before* data can flow
  - each packet carries VC identifier (not destination ID)
  - *every* switch on source-dest path maintain “state” for each passing connection
  - link,switch resources (bandwidth, buffers) may be *allocated* to VC: to get circuit-like performance
- **Permanent VCs (PVCs)**
  - long lasting connections
- **Switched VCs (SVC):**
  - dynamically set up on per-call basis

Cell Size: 32 bytes or 64 bytes?

- **Cell size of 32 and 64 bytes:**
  - 64 bytes cells have better transmission efficiency
  - 32 bytes cells have small delay
  - both sizes are integer power of 2
- **Europe wanted 32 bytes size, US and Japan wanted 64 bytes size**
- **Compromise: 48 bytes**

ATM Cell Format



GFC	**:**	**Generic Flow Control**	**VPI**	**:**	**Virtual Path Identifier**
VCI	**:**	**Virtual Circuit Identifier**	**PT**	**:**	**Payload Type**
CLP	**:**	**Cell Loss Priority**	**HEC**	**:**	**Header error Check**
UNI	**:**	**User Network Interface**	**NNI**	**:**	**Network-Network Interface**

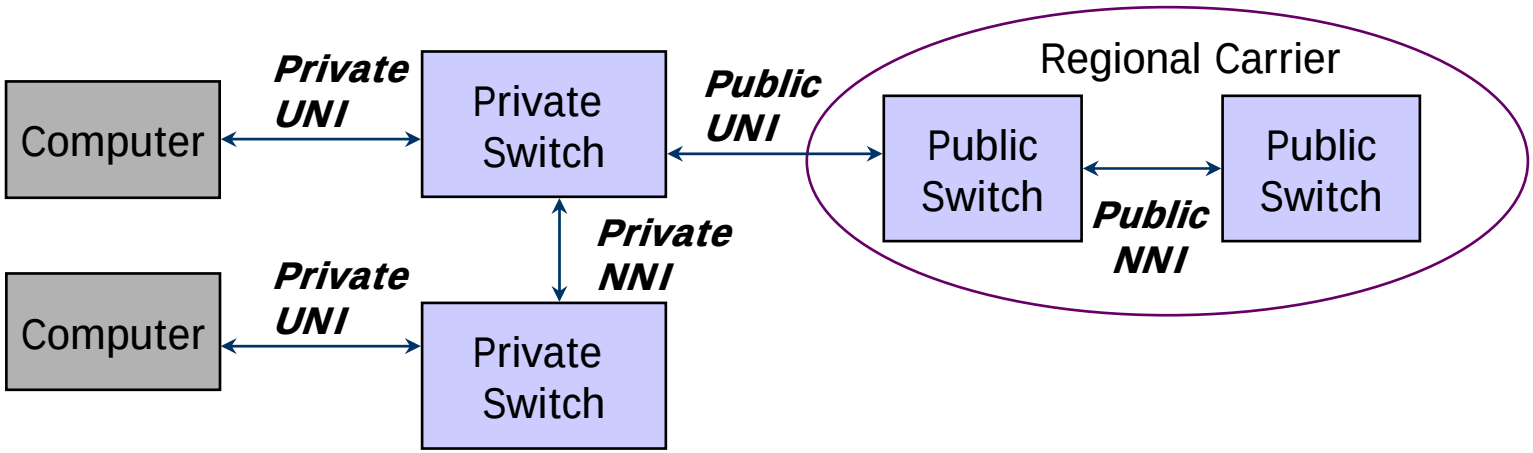
Questions

- How many Virtual Circuits can one ATM switch support?
- What is the purpose of Virtual Path?

Protocol vs. Service

- **Service – says what a layer does**
  - Ethernet: unreliable subnet unicast/multicast/broadcast datagram service
  - IP: unreliable end-to-end unicast datagram service
  - TCP: reliable end-to-end bi-directional byte stream service
  - Guaranteed bandwidth/latency unicast service
- **Protocol – says how is the service implemented**
  - a set of rules and formats that govern the communication between two peers
    - Packet format, how to interpret packet fields
    - State machine of protocol messages

User Network Interface (UNI) vs. Network Network Interface (NNI)



ATM Layer Service

Service: transport cells across ATM network

- **analagous to IP network layer**
- **very different services than IP network layer**

Network Architecture	Service Model	Guarantees ?				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	best effort	none	no	no	no	no (inferred via loss)
ATM	CBR	constant rate	yes	yes	yes	no congestion
ATM	VBR	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

Issues of IP over ATM

- Fragmentation/Reassembly
- How to do ARP?
- When to set up an ATM VC?
- How to represent IP level topology?

ATM Adaptation Layer (AAL) 5

- AAL5: low overhead **AAL used to carry IP datagrams**
  - 4 byte cyclic redundancy check
  - PAD ensures payload multiple of 48bytes
  - large AAL5 data unit to be fragmented into 48-byte ATM cells

payload	PAD	Rev	Len	CRC
0-65535	0-47	2	2	4

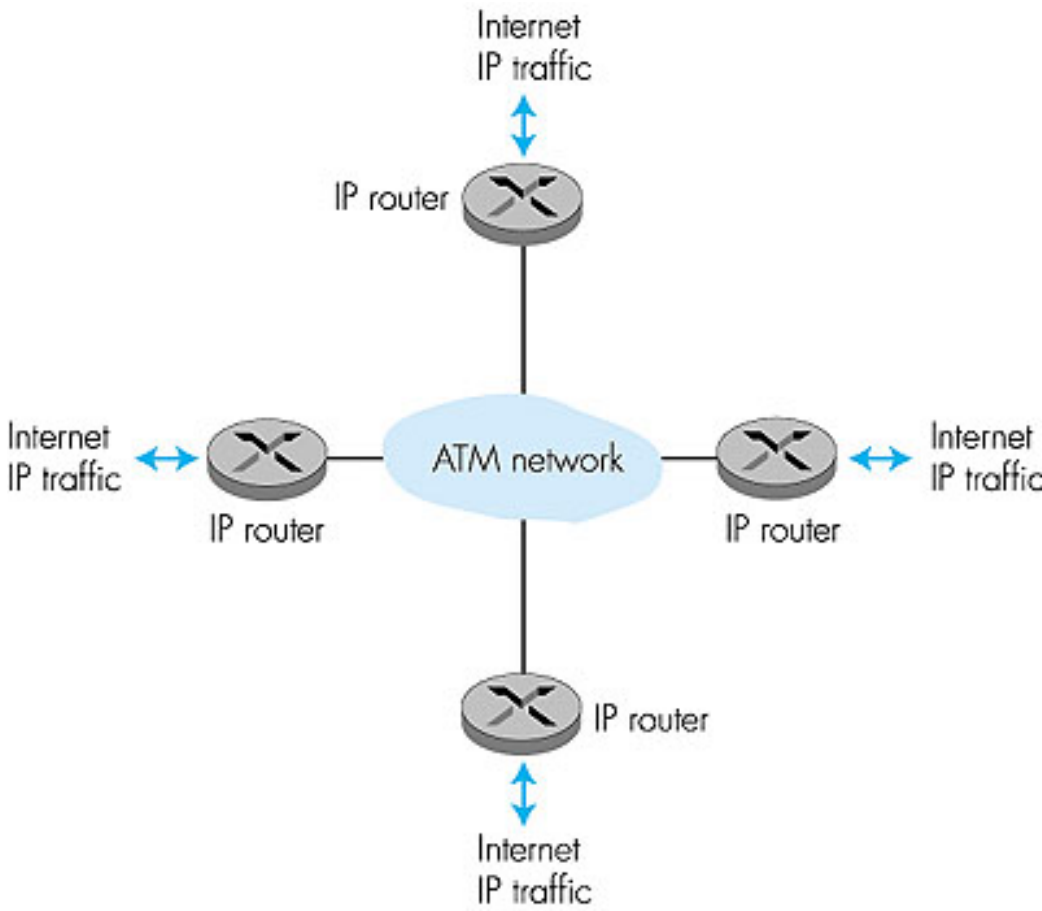
ATM: network or link layer?

Vision: **end-to-end transport: "ATM from desktop to desktop"**

- ATM *is* a network technology

Reality: **used to connect IP backbone routers**

- "IP over ATM"
- ATM as switched link layer, connecting IP routers



Where is ATM Today?

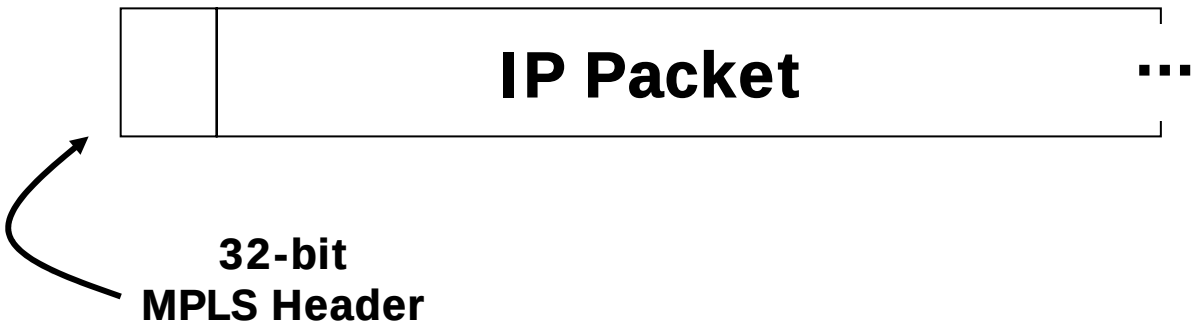
- DSL (Digital Subscriber Loop)
- Multi-service switching
- Interconnection of IP routers

MPLS

- **Multi-Protocol Label Switching**
- **Bringing virtual circuit concept into IP**
- **Driven by multiple forces**
  - QoS
  - traffic engineering
  - High performance forwarding
  - VPN

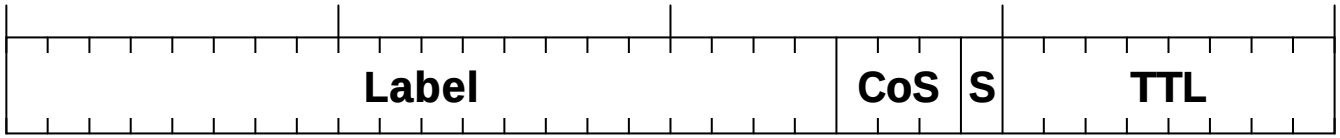
MPLS Header

- **IP packet is encapsulated in MPLS header and sent down LSP**



- **IP packet is restored at end of LSP by egress router**
  - TTL is adjusted also

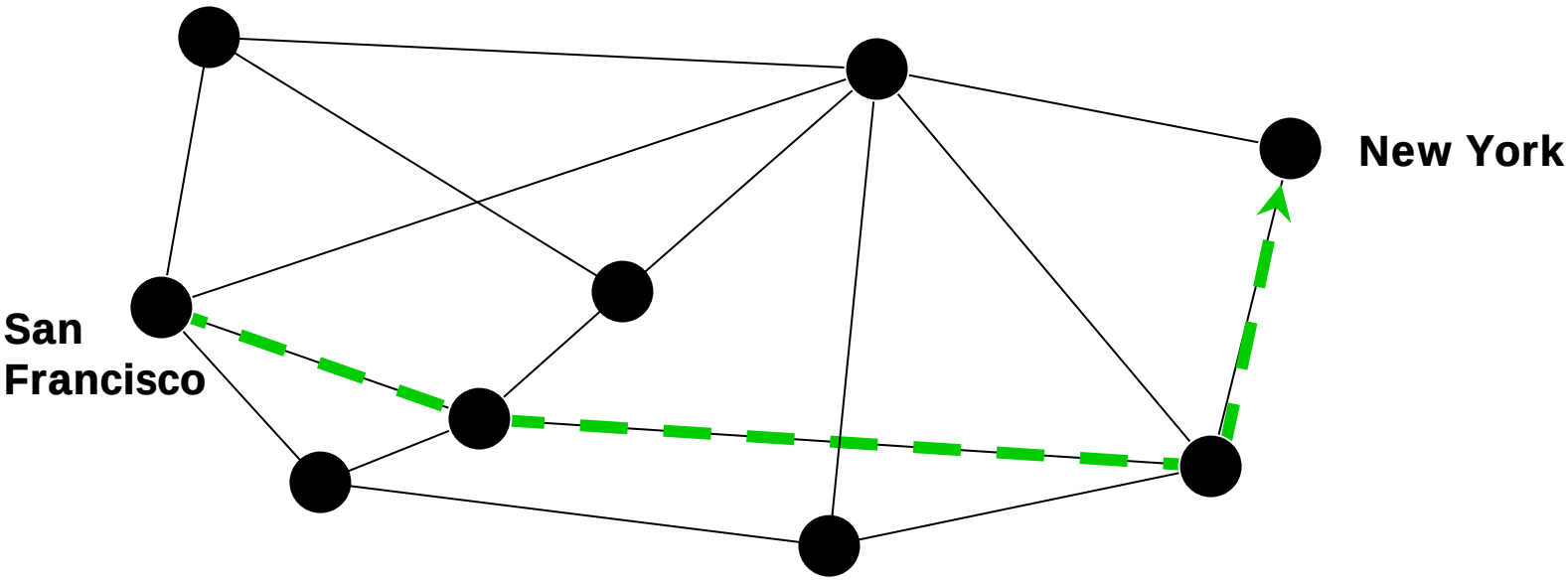
MPLS Header



- **Label**
- **Class of service**
- **Stacking bit**
- **Time to live**
  - Decrement at each LSR, or
  - Pass through unchanged

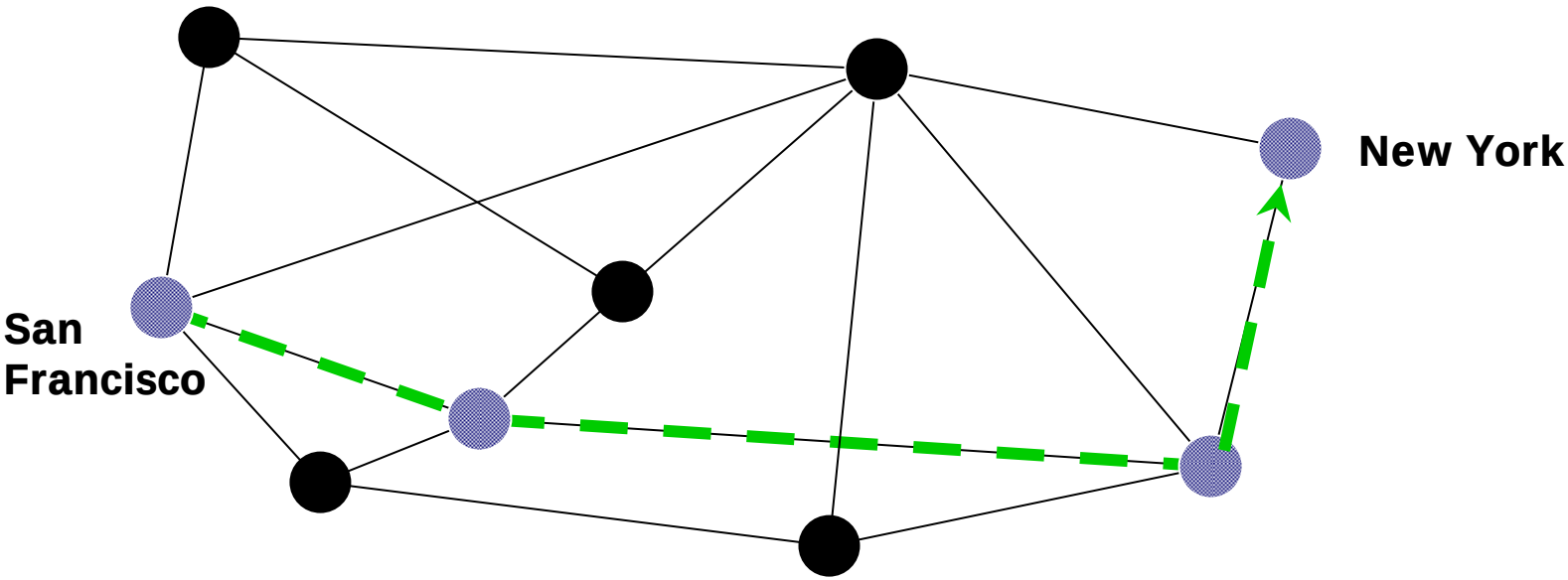
MPLS Vocabulary

- **Label-switched path (LSP)**
  - Simplex path through interior network



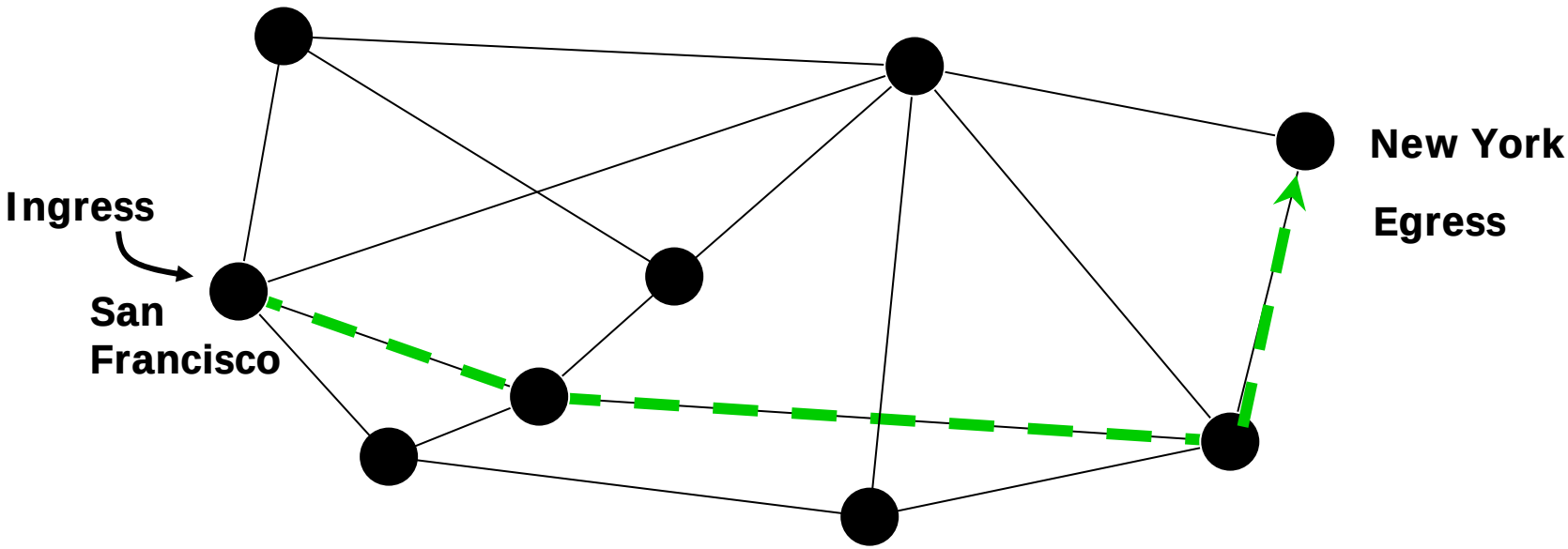
MPLS Vocabulary

- **Label-switching router (LSR) performs**
  - MPLS packet forwarding
  - LSP setup



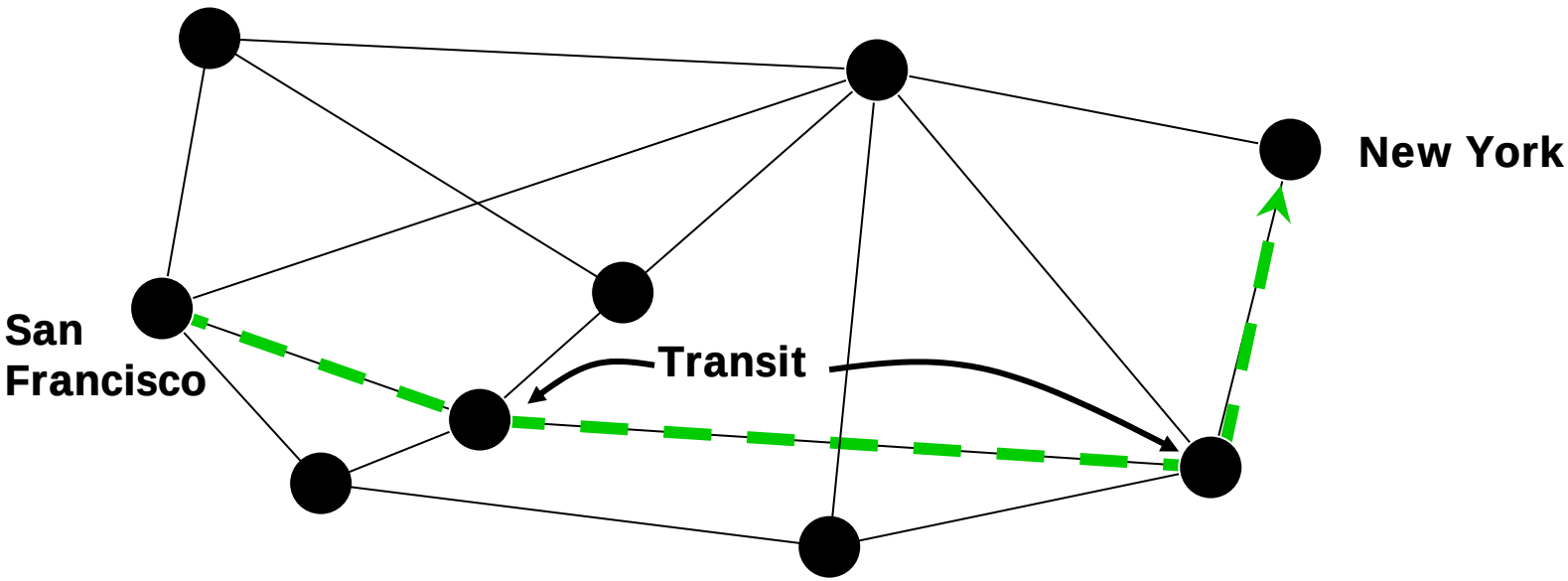
MPLS Vocabulary

- **Label Edge Router (LER)**
  - Ingress and egress node of LSP



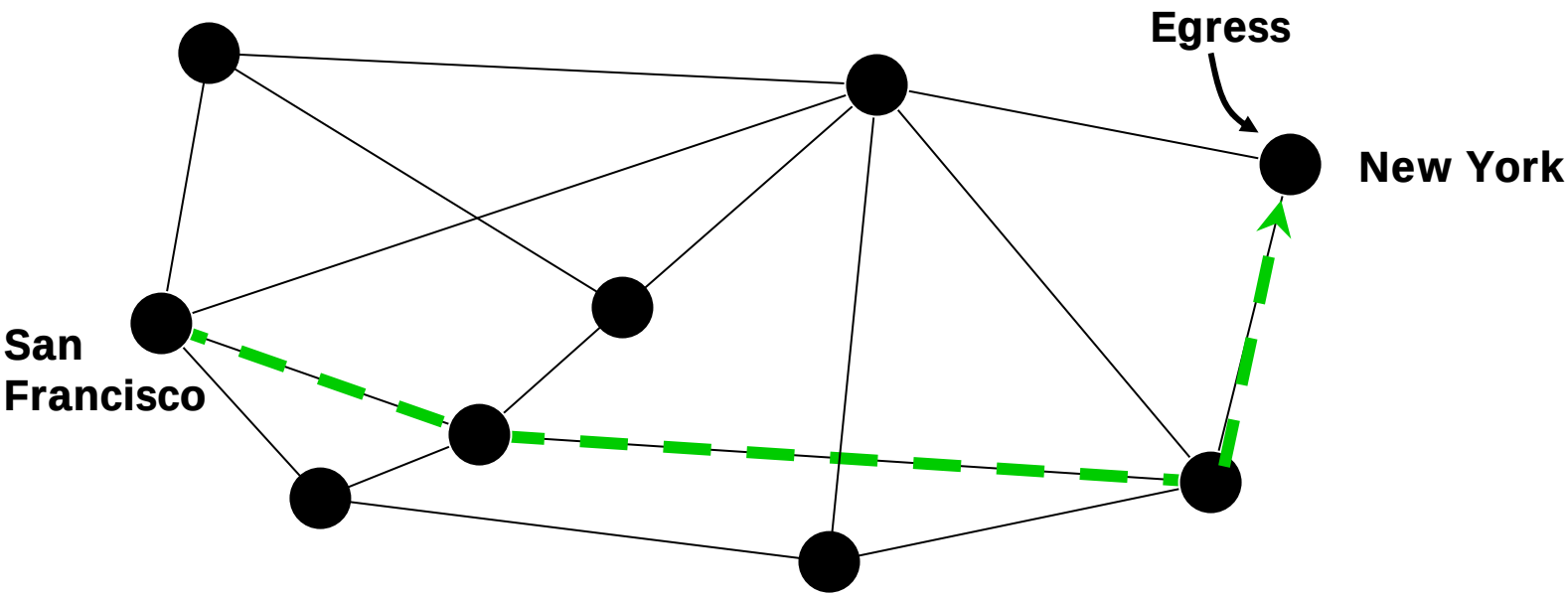
MPLS Vocabulary

- **Transit router**
  - Zero or more transit routers
  - Swaps MPLS label
  - Sends traffic to next hop in LSP

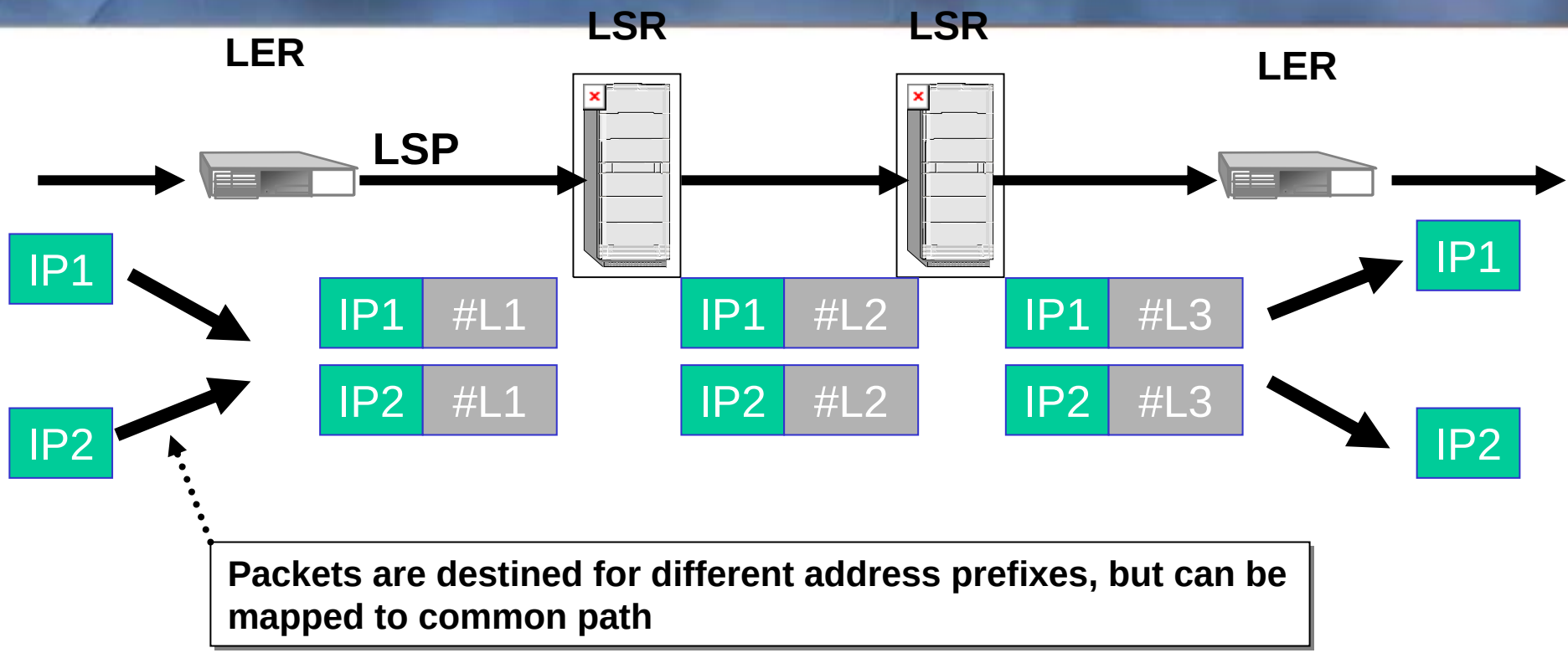


MPLS Vocabulary

- **Egress router**
  - Packets exit LSP here
  - Also called “tail-end” router
  - Downstream from other routers

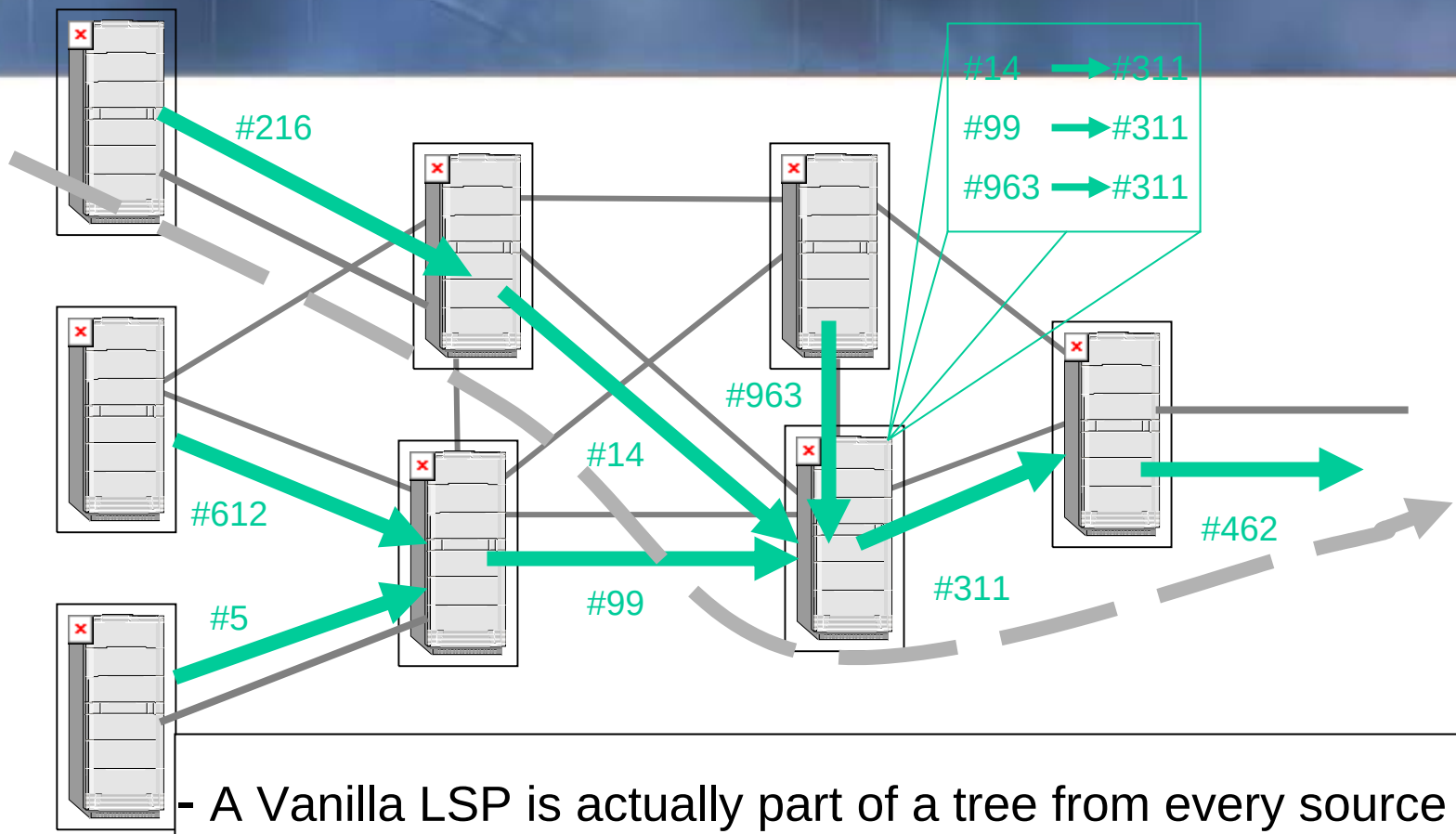


Forwarding Equivalence Classes



- FEC = “A subset of packets that are all treated the same way by a router”
- The concept of FECs provides for a great deal of flexibility and scalability
- In conventional routing, a packet is assigned to a FEC at each hop (i.e. L3 look-up), in MPLS it is only done once at the network ingress.

LABEL SWITCHED PATH (vanilla)



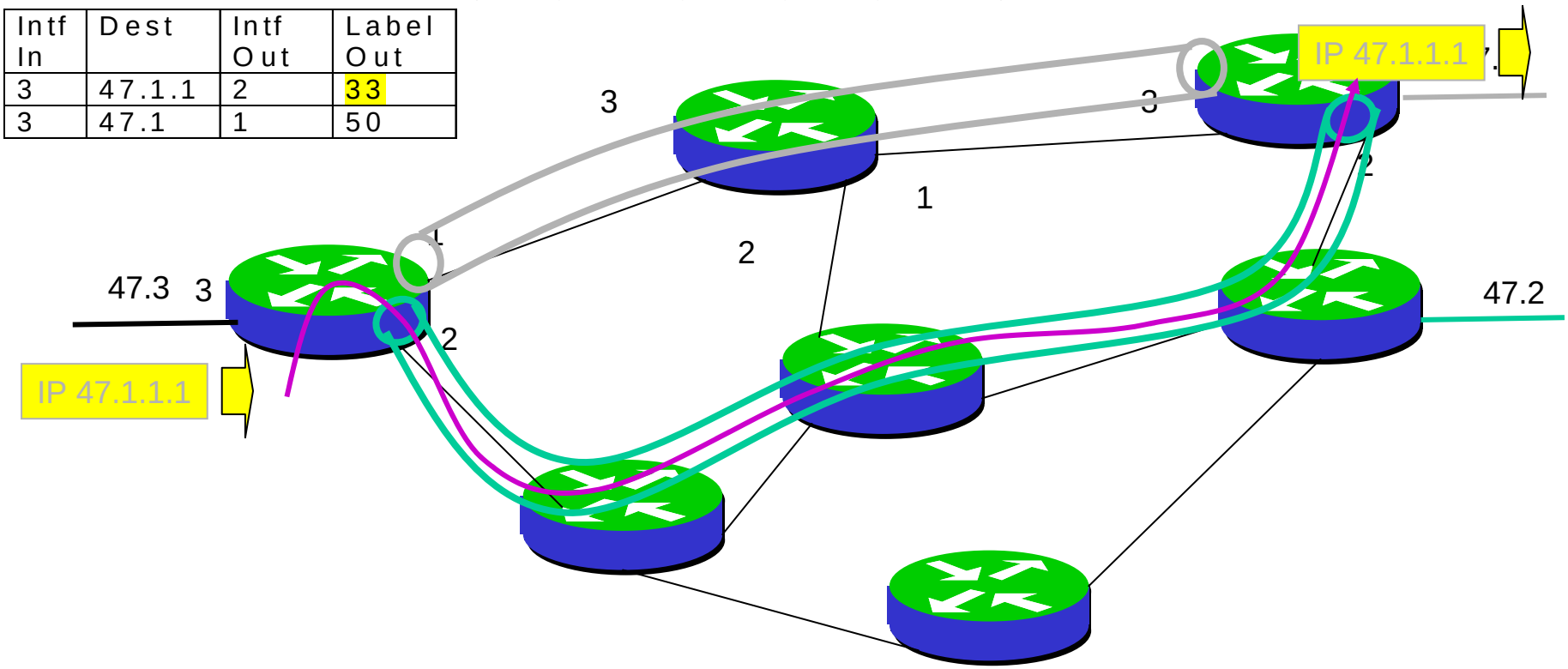
- A Vanilla LSP is actually part of a tree from every source to that destination (unidirectional).
- Control protocol (e.g. LDP) builds that tree using existing IP forwarding tables to route the control messages.

EXPLICITLY ROUTED LSP ER-LSP

Intf In	Label In	Dest	Intf Out	Label Out
3	50	47.1	1	40

Intf In	Label In	Dest	Intf Out
3	40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1.1	2	33
3	47.1	1	50



MPLS just another ATM?

- **Same IP addressing and routing protocols**
  - ATM: new addressing, new routing protocols
- **Same IP packet with another header**
  - No fragmentation/reassembly
- **Merging of LSP more flexible**
- **Label stacking allow more flexible tunneling**

What is a “LABEL”?

A property that uniquely identifies a flow on a logical or physical interface

There are three basic types of LABEL:

StatMux : inband bits identify the flow.
TimeDivMux : position in time identifies the flow.
FreqDivMux : frequency identifies the flow.

Labels may be module/network wide unique or more commonly interface wide unique.

A label by any other name ……

- **There are many examples of label substitution protocols already in existence**

  - **ATM - label is called VPI/VCI and travels with cell**
  - **Frame Relay - label is called a DLCI and travels with frame**
  - **TDM - label is called a timeslot its implied, like a lane**
  - **X25 - a label is an LCN**
  - **Proprietary PORS, TAG etc..**
  - **Frequency substitution where label is a light frequency via DWDM, OXC etc.**

Protocol Comparison

	Forwarding	Control Protocols
Ethernet	Dest MAC address Exact match	Learning Spanning tree
IP	Dest IP address Longest prefix match	Routing protocol
TDM	Time slot, exact match Time Slot Exchange (TSE)	E2E signaling protocol Routing protocol
ATM	Label, exact match Label swapping	E2E signaling protocol Routing protocol
MPLS	Label, Dest IP Address	Flexible signaling Routing Protocol