CSC4200/5200 – COMPUTER NETWORKING

Instructor: Susmit Shannigrahi

GLOBAL INTERNET AND TRANSPORT LAYER RECAP

sshannigrahi@tntech.edu

GTA: dereddick42@students.tntech.edu

Tennessee TECH

Apps (HTTP)

Data

Apps (HTTP)

Transport (TCP/UDP)

Segments

Transport (TCP/UDP)

Network (IP)

Packets

Network (IP)

Link (Ethernet)

Frames

Ethernet Interface

Ethernet Interface

Link (Ethernet)

Bits (1010001)

A Number of Logistical Updates

- **PA3 groups – let me know by tomorrow or I will create groups at random**
- All slides uploaded.
- Grading rubric for PA2 and PA3 has posted
- Homework 4 has posted, due in a week
- PA3 demo, December 6th, Tech Pride Room, 8AM -9:30AM
  - Make a poster, do a live demo, or screencast
- Final December 9th – Location/Time – Library 112, 1:00PM - 3:00PM Monday, December 9

Recap of Global Internet, TCP/IP

What is an Autonomous system? Why do we need ASes?

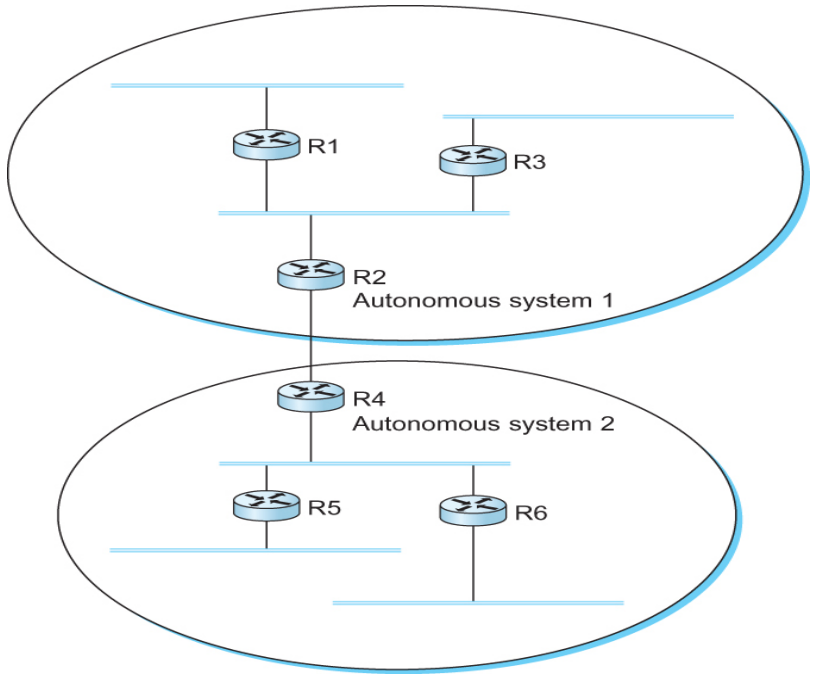
Autonomous systems (ASes)

- AS
  - A set of routers under a single technical administration
  - Uses IGP within the AS to route packets
  - Uses BGP between Ases to route packets

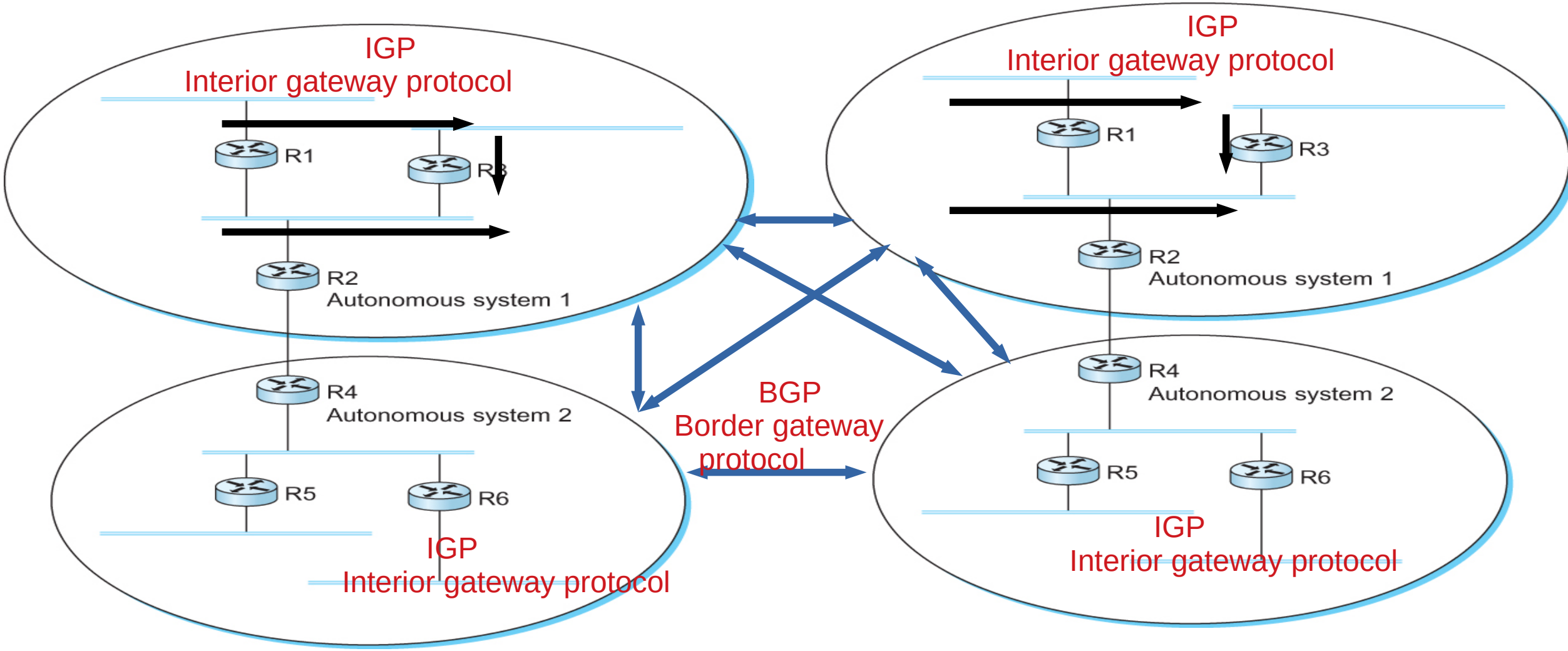
- What happens inside an AS stays within that AS!
  - That is, AS decides routing metrics internally

Recap of Global Internet, TCP/IP

Difference between Intra-AS and Inter-AS routing?



Interdomain Routing

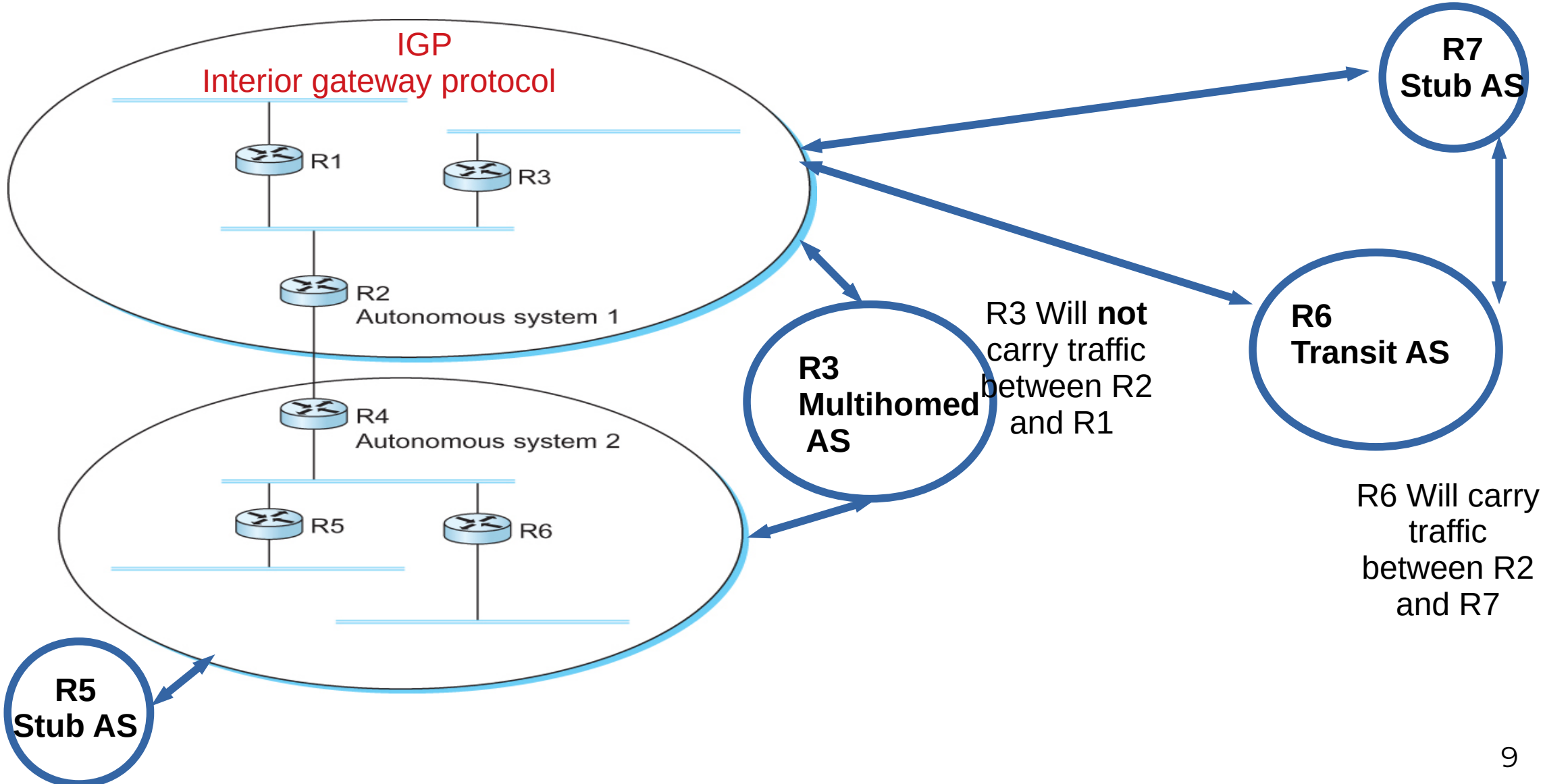


A network with four autonomous systems

Recap of Global Internet, TCP/IP

What are the three types of ASes in BGP?

BGP-4: Border Gateway Protocol



Recap of Global Internet, TCP/IP

Which of the following are the goals of BGP?

1) Loop free path
2) Optimal path
3) Multiple paths to the same destination
4) All paths to a destination

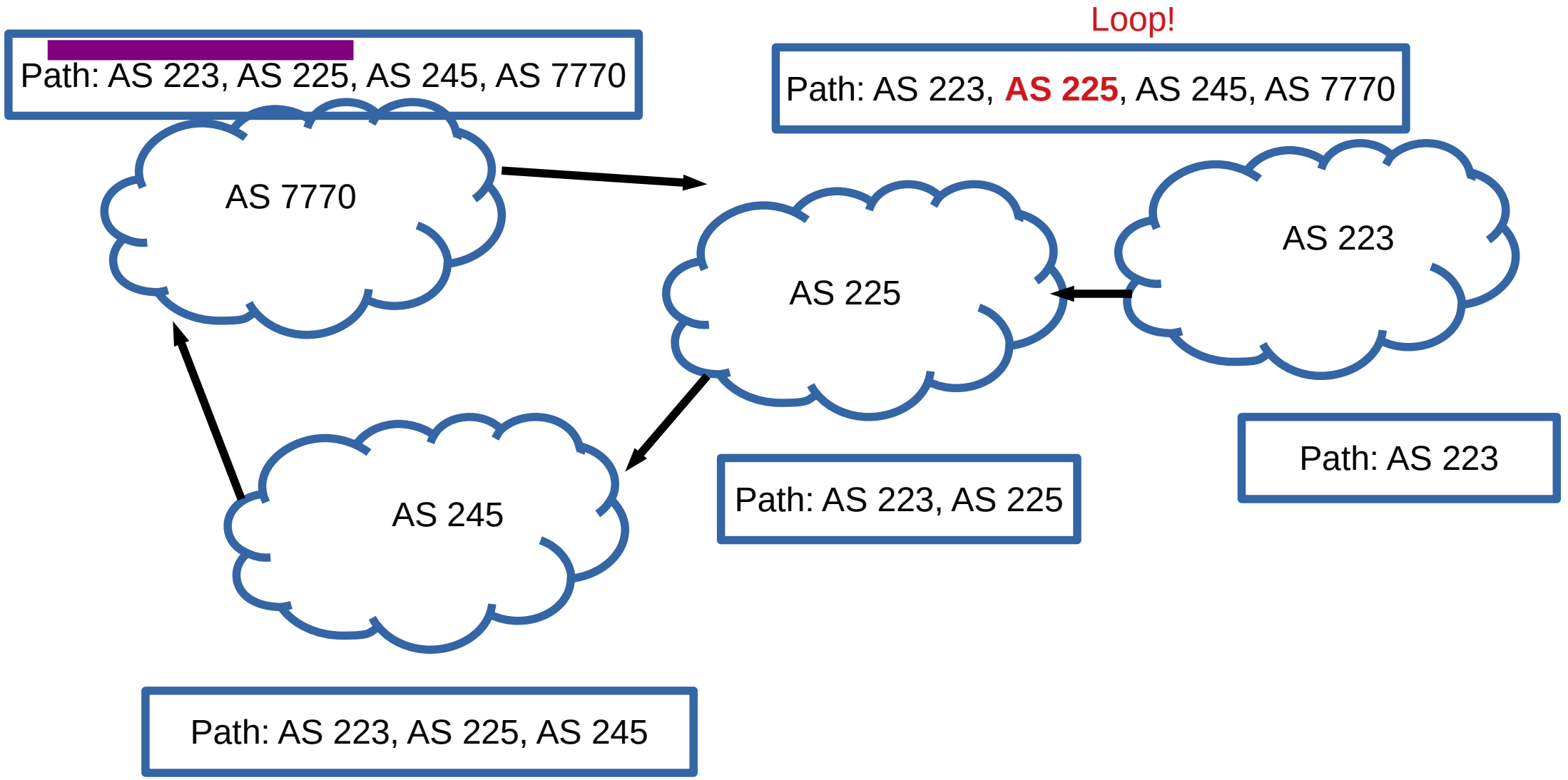
BGP - goals

- The goal of Inter-domain routing is to find **any path** to the intended destination that is **loop free**
  - **We are concerned with reachability than optimality**
  - Finding path anywhere close to optimal is considered to be a great achievement
- Why?

Recap of Global Internet, TCP/IP

What is the protocol used by BGP?
1)Path vector protocol
2)Distance vector protocol
3)Link state protocol
4) RIP
5) Flooding

BGP: Path vector protocol



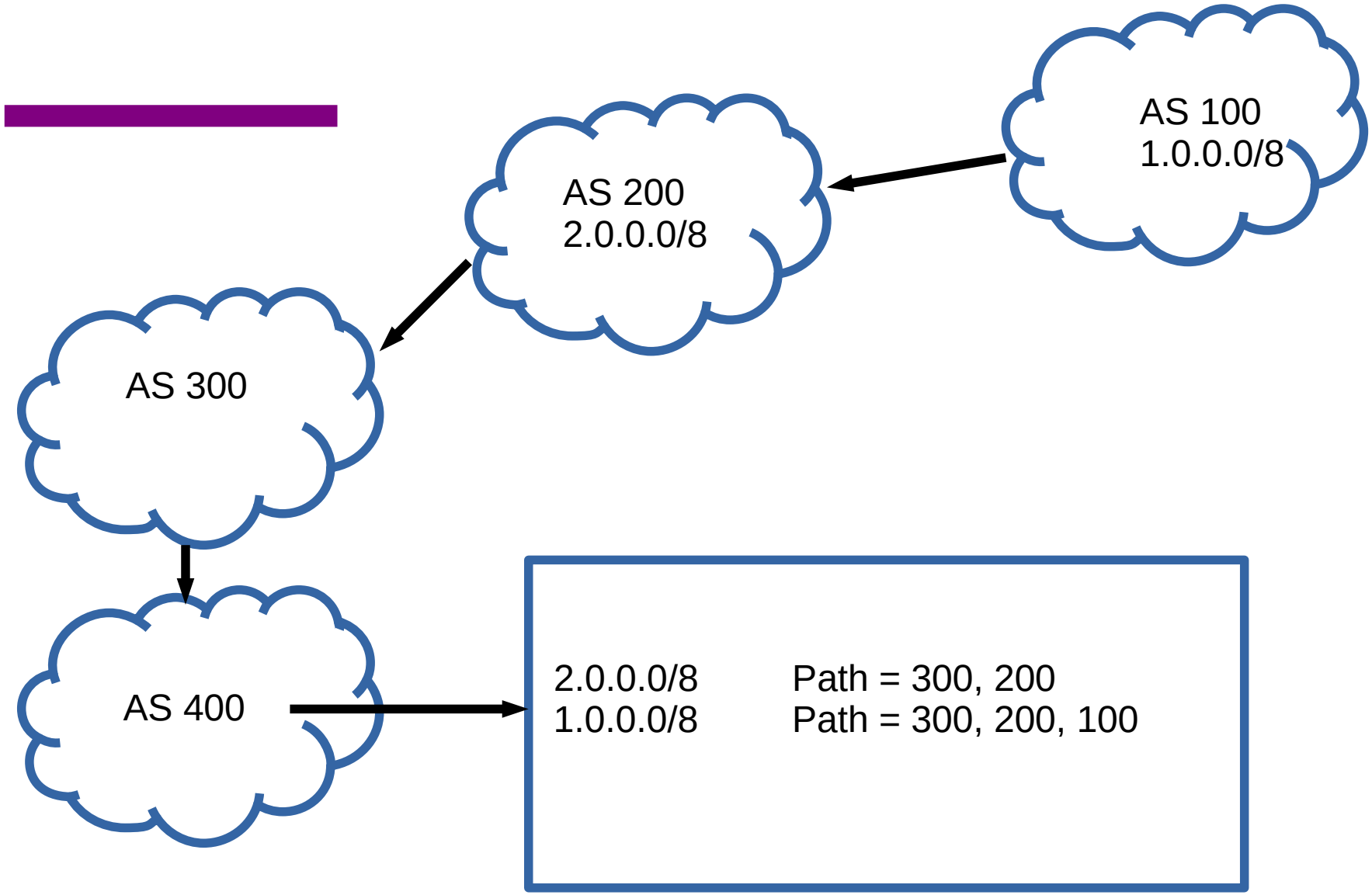
Recap of Global Internet, TCP/IP

At which layer is BGP?
1)Network
2)Transport
3)Physical
4)Application

Recap of Global Internet, TCP/IP

How does BGP announce a route to an AS?

BGP Attributes - AS_PATH

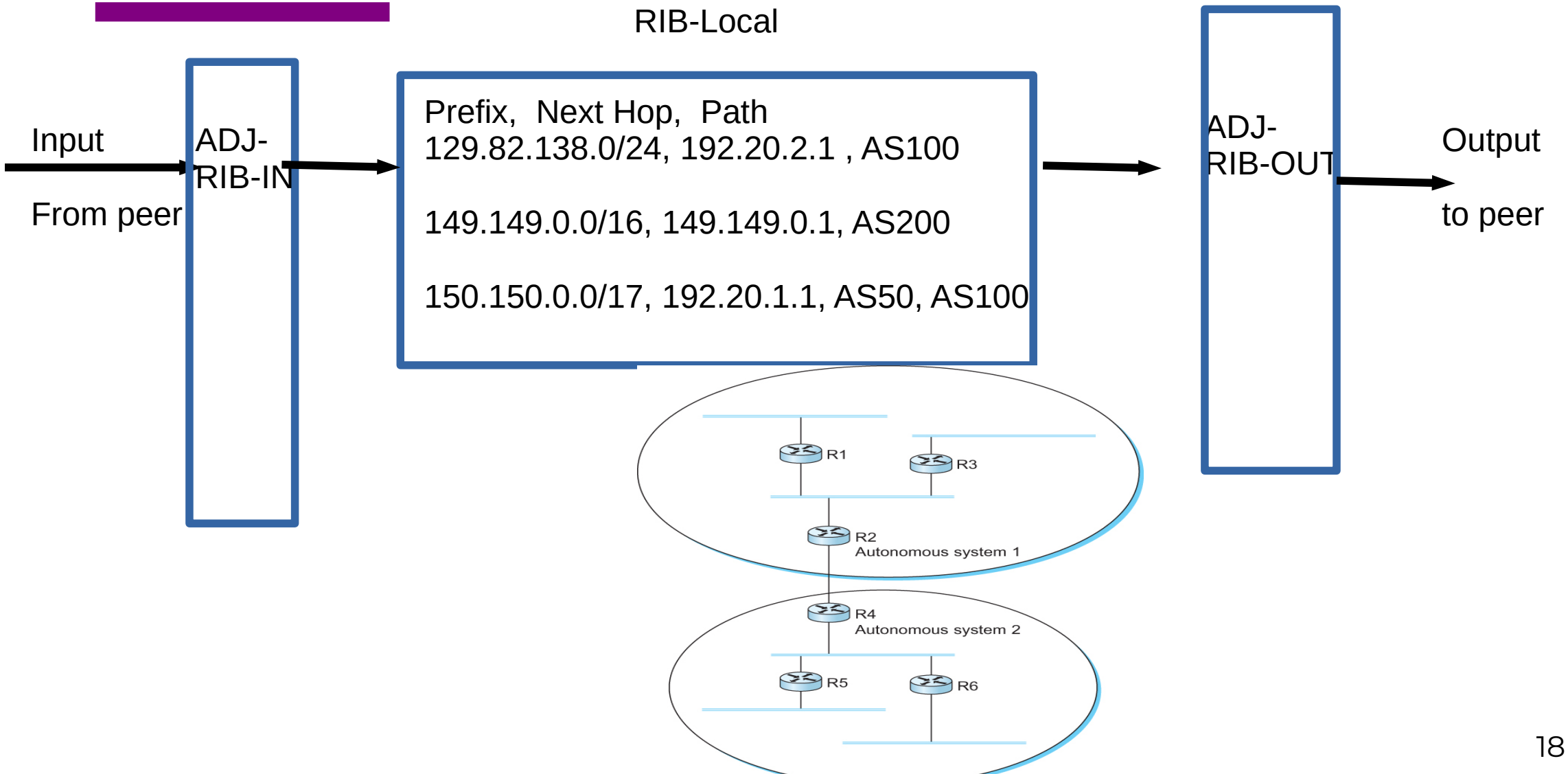


Recap of Global Internet, TCP/IP

How does BGP announce a route to an AS?

Recap of Global Internet, TCP/IP

What are the three tables and what do they do?

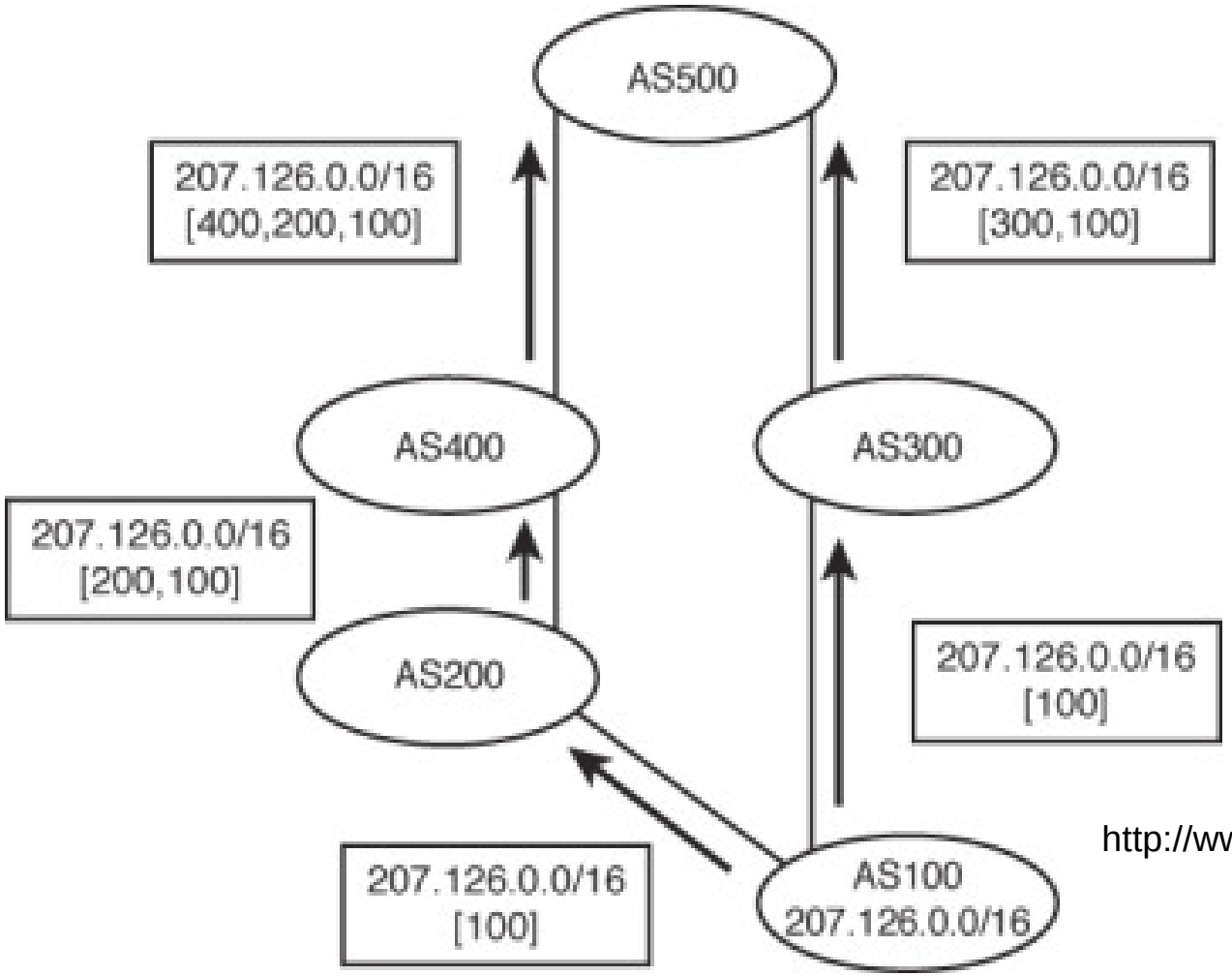


Recap of Global Internet, TCP/IP

How does BGP enforce policies?

Attributes -
Local Preference, MED, Path Length

BGP Attribute – AS PATH



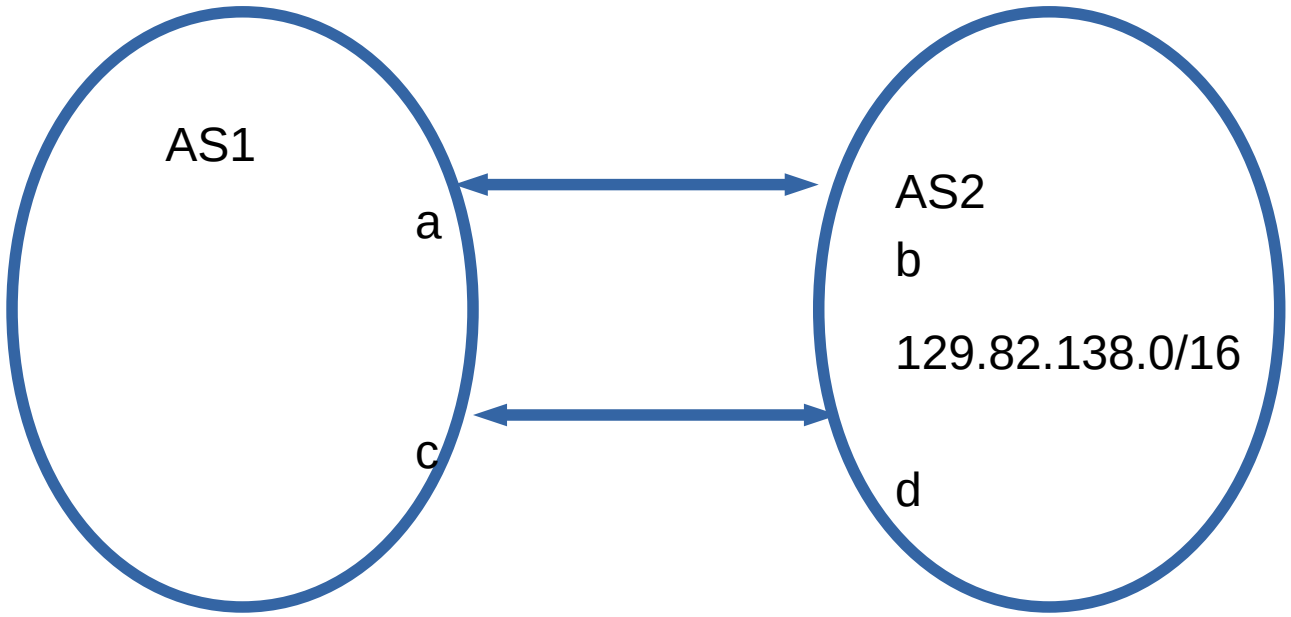
Each hop adds ASN to the path
-Only externally

http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2

Recap of Global Internet, TCP/IP

Local Preference vs MED

BGP Attribute – Local Preference



How do you load balance between two links using BGP?

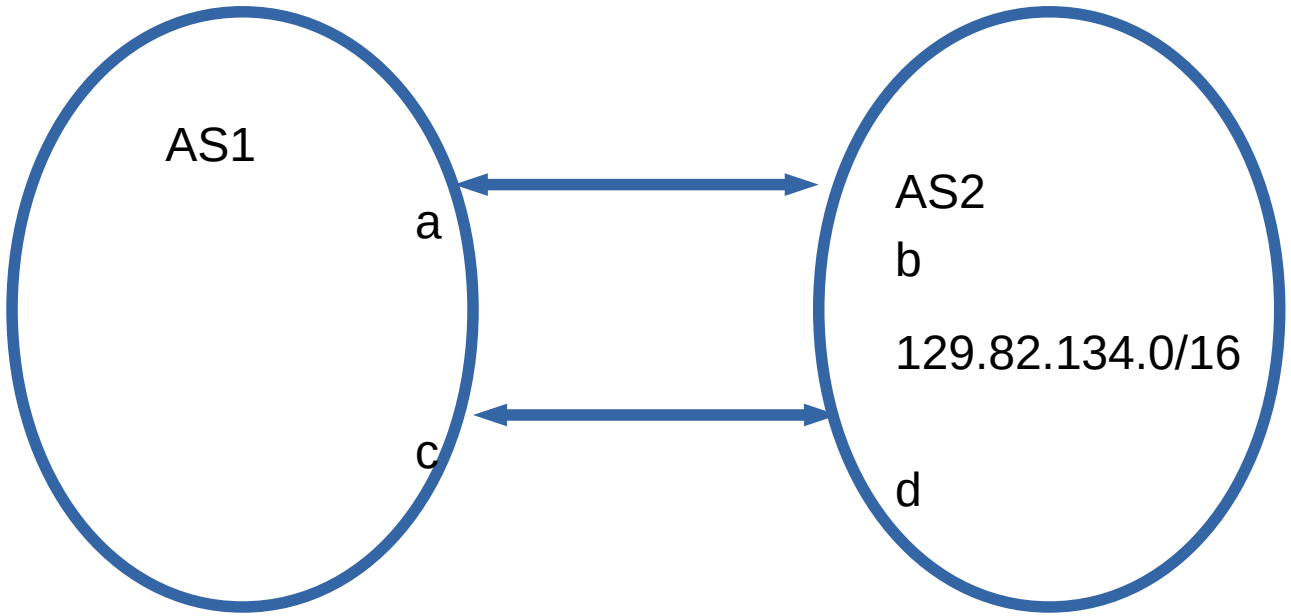
At A:

129.82.138.0/17 → 10
129.82.138.128/17 → 5

At C:
129.82.138.0/17 → 5
129.82.138.128/17 → 10

http://www.ciscopress.com/articles/article.asp?p=2738462&seqNum=2

BGP Attribute – MED (Multi exit discriminator



AS1 and AS2 has two paths between them

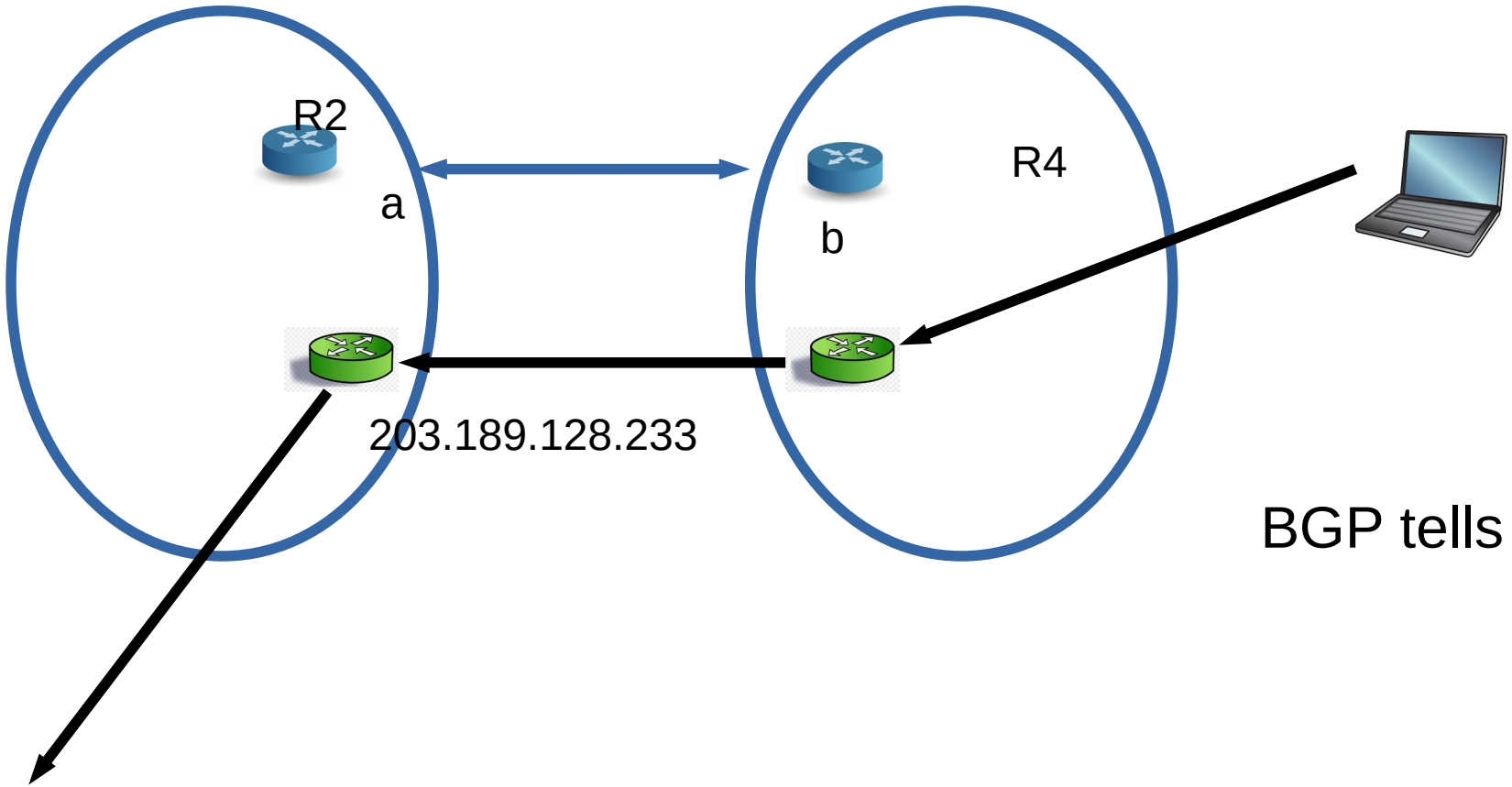
AS1 tells AS2 it's MED for influencing AS2's path selection

Lower cost wins

How does BGP relate to IP?

Next hop | Announcing AS| Target Prefix| Path

203.189.128.233 | 23673 | 149.149.0.0/16 | 23673 1299



BGP tells you which IP router to use

BGP Decision process

At ADJ-RIB-IN calculate degree of preference until one route for each destination remains!!

- select route with highest LOCAL-PREF
- Select route with shortest AS-PATH
- Select route with lowest MED
- Select route with smallest NEXT-HOP cost
- Select route learned from E-BGP peer with lowest ID
- Select route learned from I-BGP peer with lowest ID

- Install selected route in LOC-RIB

- Update ADJ-RIB-OUT, notify peers
  - You can only send what is in LOC-RIB (or a subset of it)

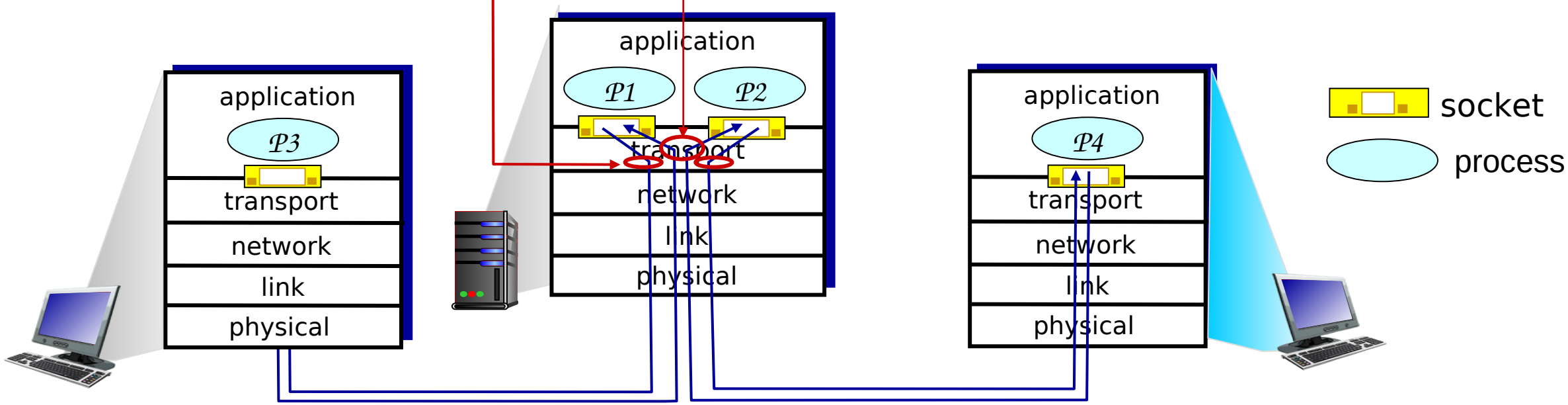
Transport vs. network layer

- network layer: logical communication between hosts
- transport layer: logical communication between processes
  - relies on, enhances, network layer services

Multiplexing/demultiplexing

multiplexing at sender:
handle data from multiple sockets, add transport header (later used for demultiplexing)

demultiplexing at receiver:
use header info to deliver received segments to correct socket

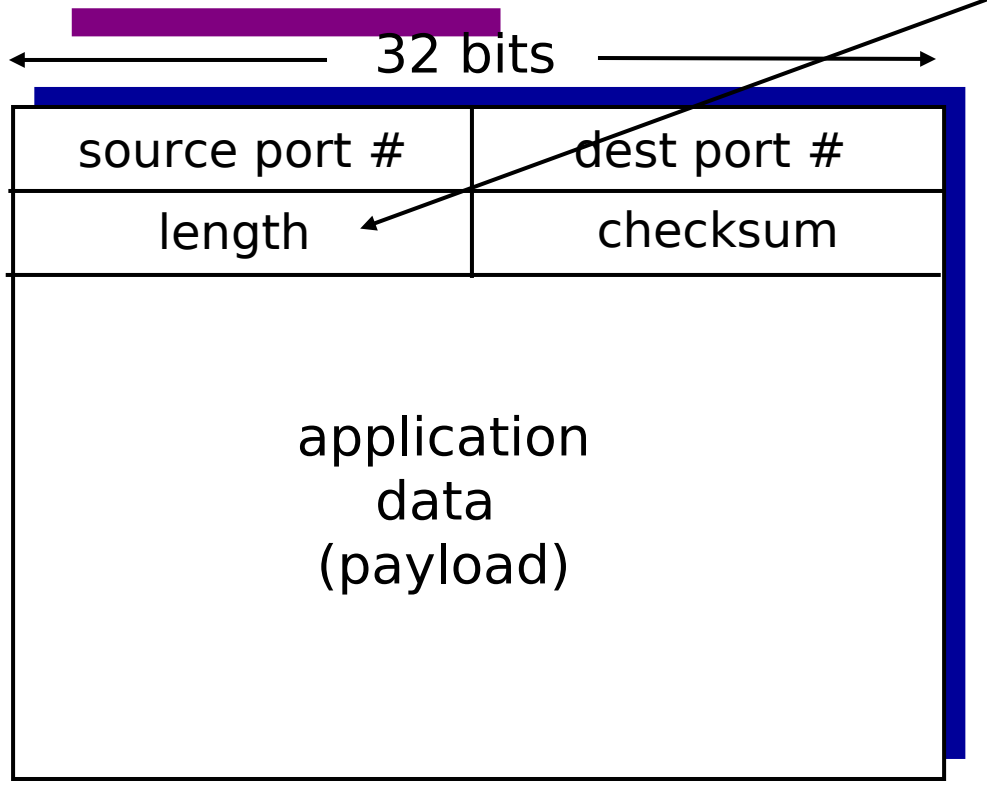


Transport protocols

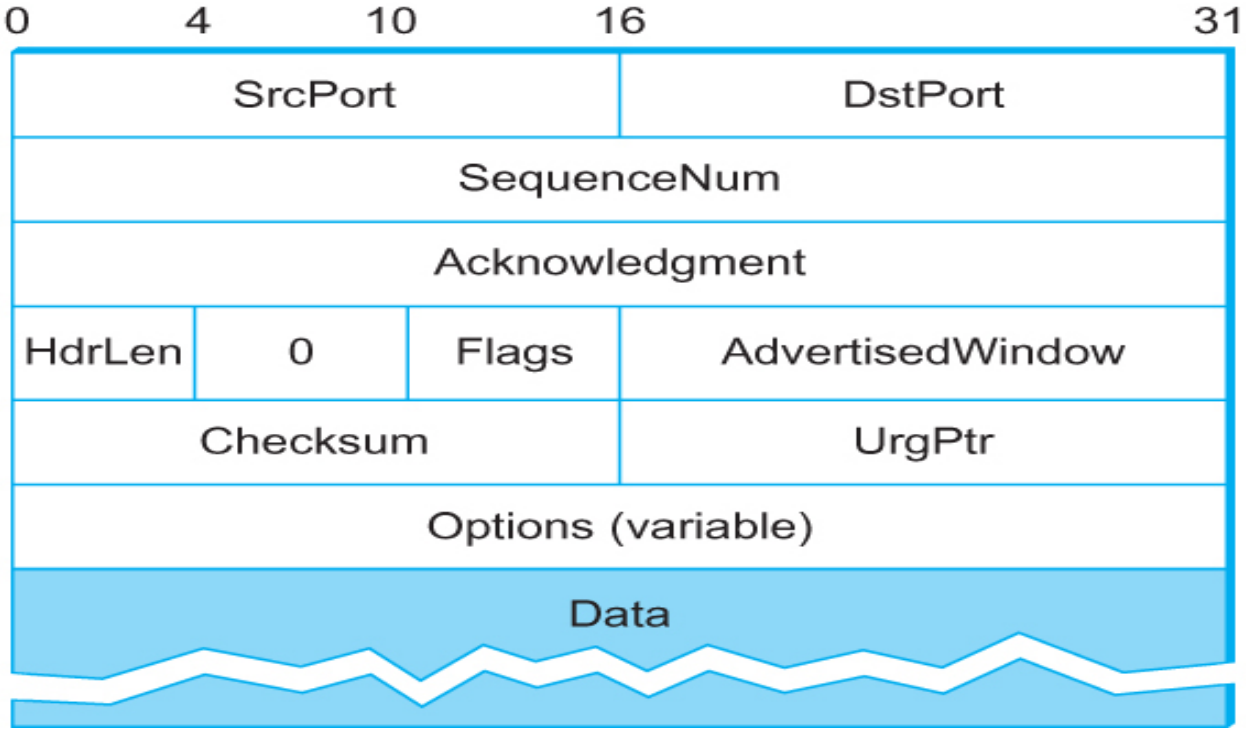
What are the two types of Transport protocols?

If you had to write your own transport protocol, what functionality would you implement?

TCP vs UDP



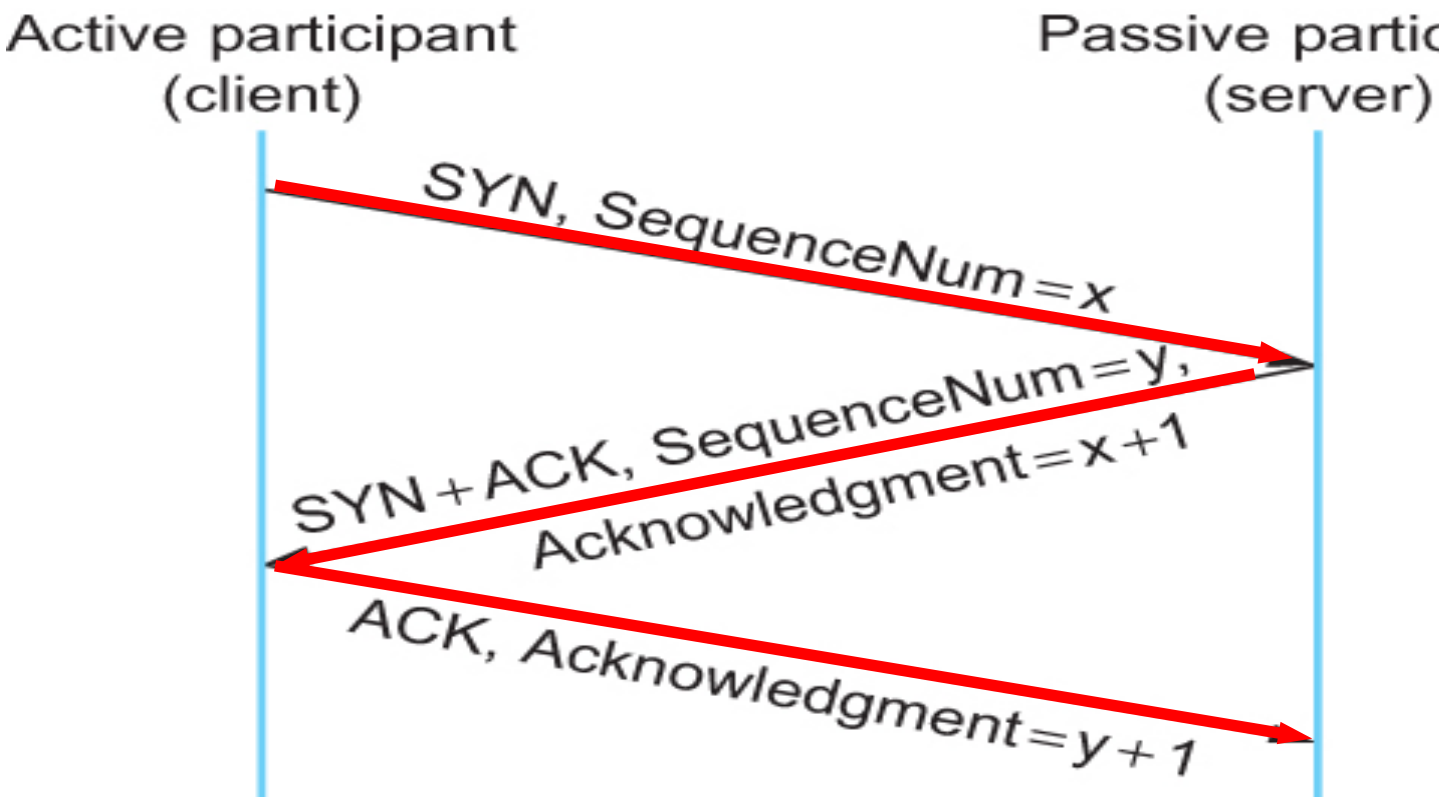
UDP segment format



TCP vs UDP

- Which is connection oriented and which is connection less?
- If there is no connection, how does UDP know where to return packets?

TCP Three-way Handshake



Timeline for three-way handshake algorithm

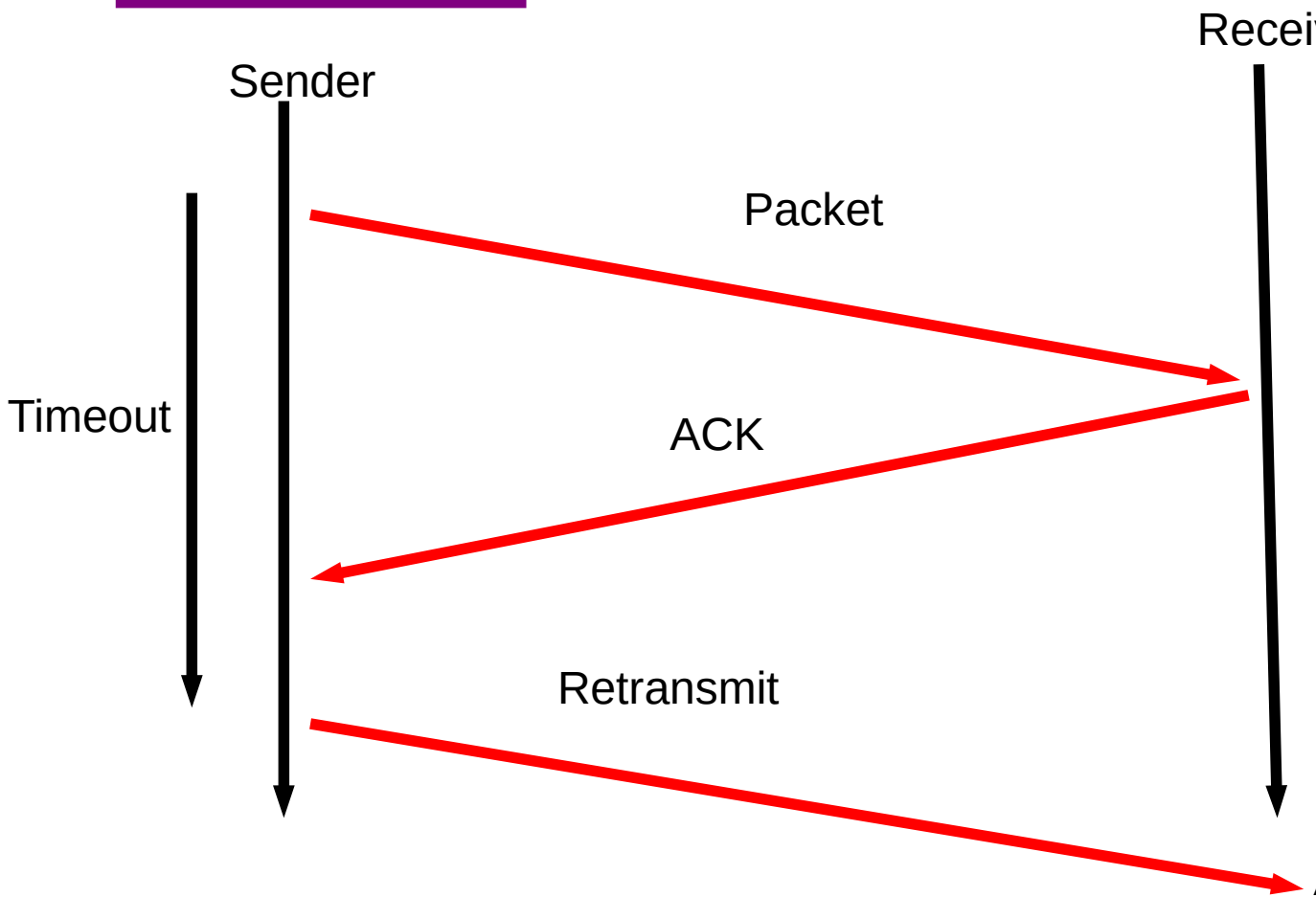
The idea is to tell each other The ISNs

SYN → Client tells server that it wants to open a connection, Client's ISN = x

SYN+ ACK → Server tells Client → Okay → Server's ISN = y, ACK = CLSeq + 1

Why increment by 1?

How long should the sender wait?



For TCP, timeout is a function of RTT

Keep a running estimate of RTTs by watching the ACKs

EstimatedRTT = (1 – α) • EstimatedRTT
+ α • SampleRTT

Alpha is generally 0.125

Timeout = 2* EstimatedRTT

Sliding Window Revisited

Sending Side

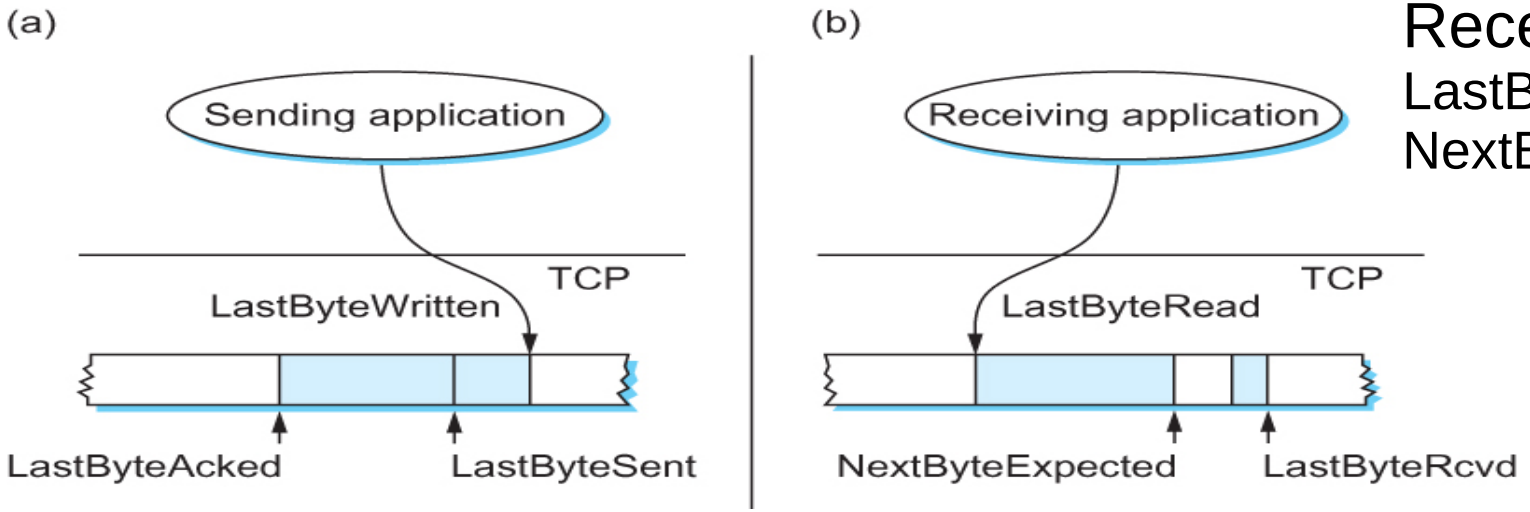
LastByteAcked ≤ LastByteSent

LastByteSent ≤ LastByteWritten

Receiving Side

LastByteRead < NextByteExpected

NextByteExpected ≤ LastByteRcvd + 1



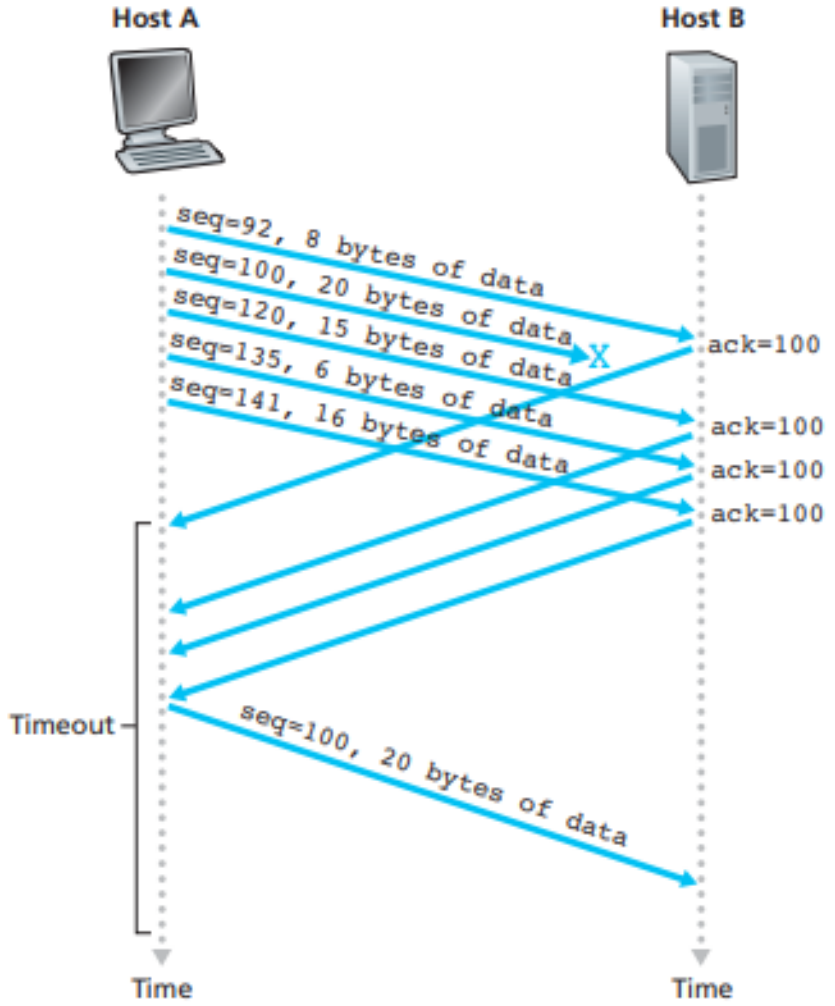
Relationship between TCP send buffer (a) and receive buffer (b).

TCP Fast Retransmission

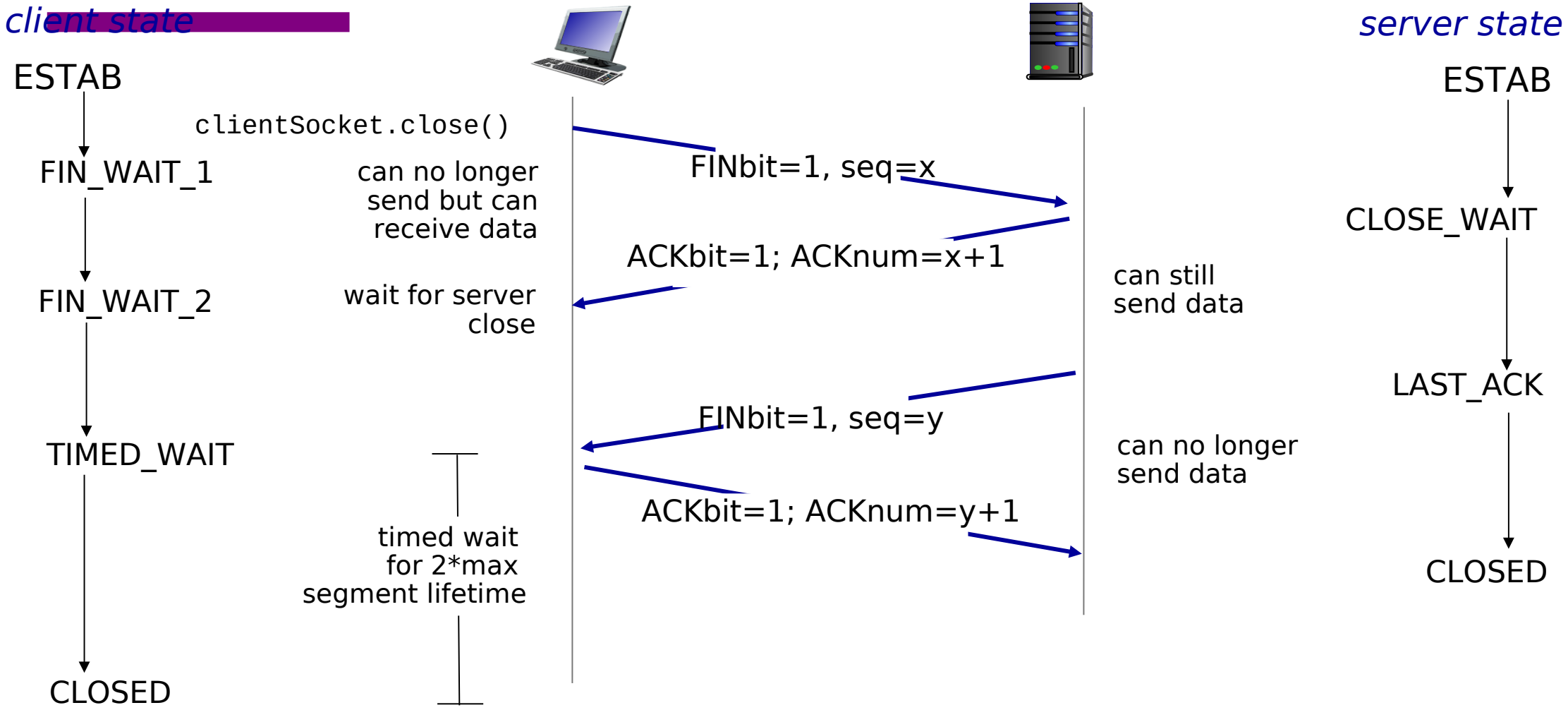
Timeouts are wasteful

Triple duplicate ACKs

Retransmits before timeout



TCP: closing a connection

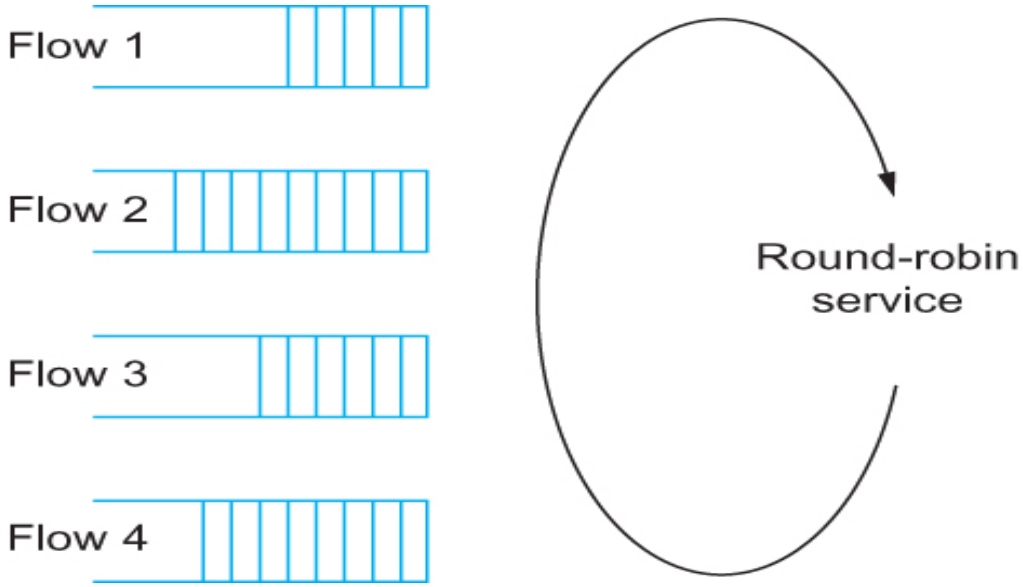


TCP Flow Control vs Congestion Control

- Flow control - between sender and receiver
- Congestion control – In the network!
- FC – Save the receiver
- CC – Save the network! ← Literally (Congestive collapse 1986→ 1987/1988)

Queuing Disciplines

- Fair Queuing



Round-robin service of four flows at a router

Metrics: Throughput vs Delay

High throughput –

- Throughput: measured performance of a system –E.g., number of bits/second of data that get through
- Low delay –
- Delay: time required to deliver a packet or message –E.g., number of ms to deliver a packet •
- These two metrics are sometimes at odds –
  - More packets = more queuing

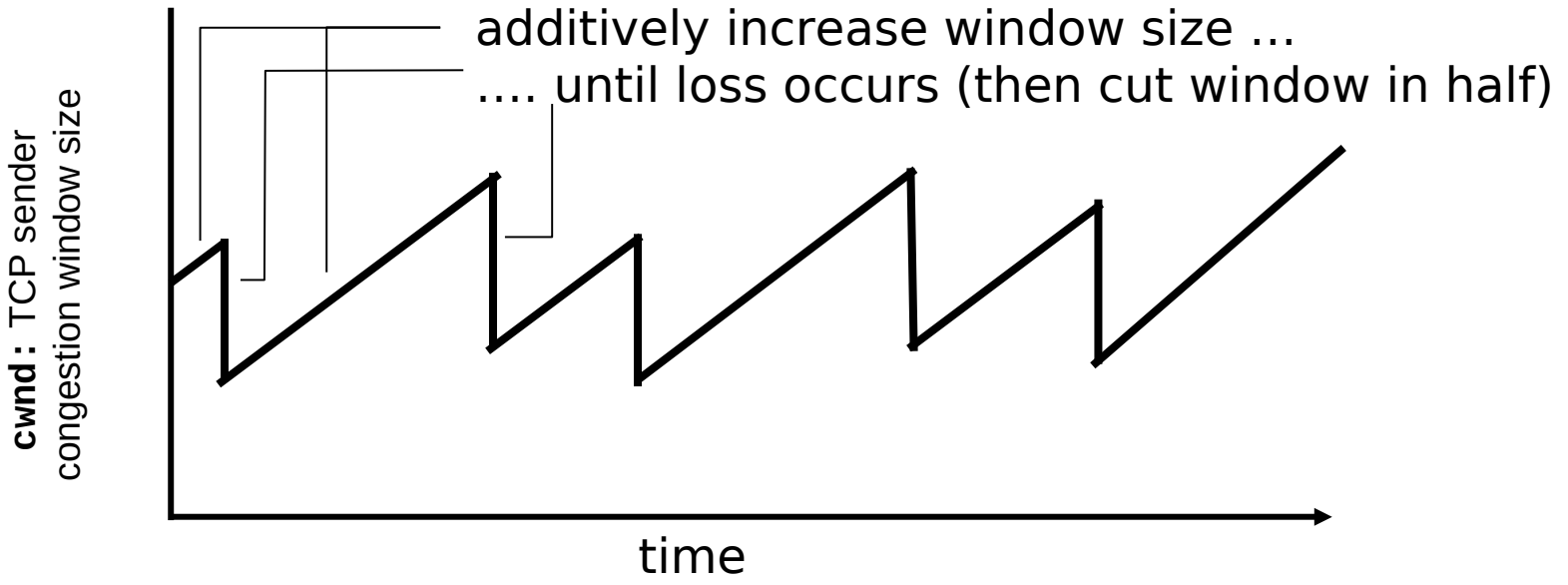
Min Max Fair queuing

- Assume ***n*** clients
- Channel capacity **C**
- Give **c/n** to each client
  - If C1 does not want c/n
  - Divide the excess capacity equally among others
  - So everyone else gets c/n + (c/n – c1)/n-1
  - Repeat for C2 and others

TCP Congestion Control

What is the basic idea?

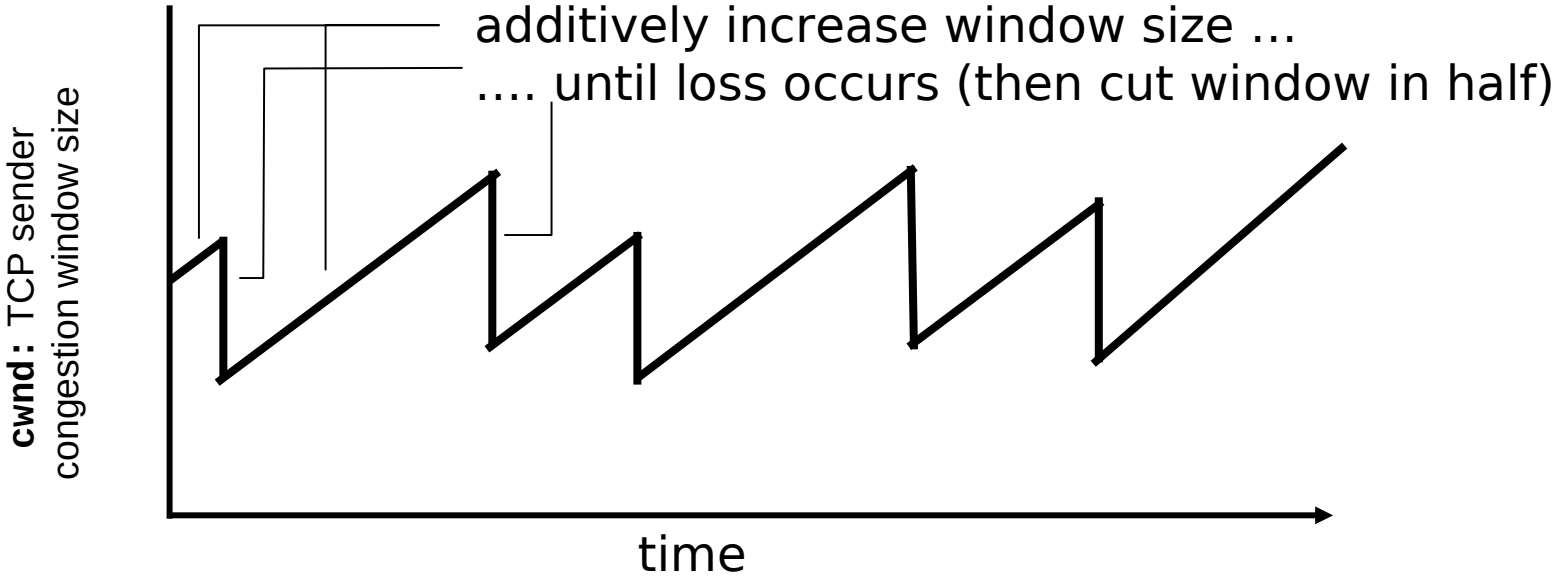
AIMD saw tooth behavior: probing for bandwidth



Congestion control - How much to increase and decrease?

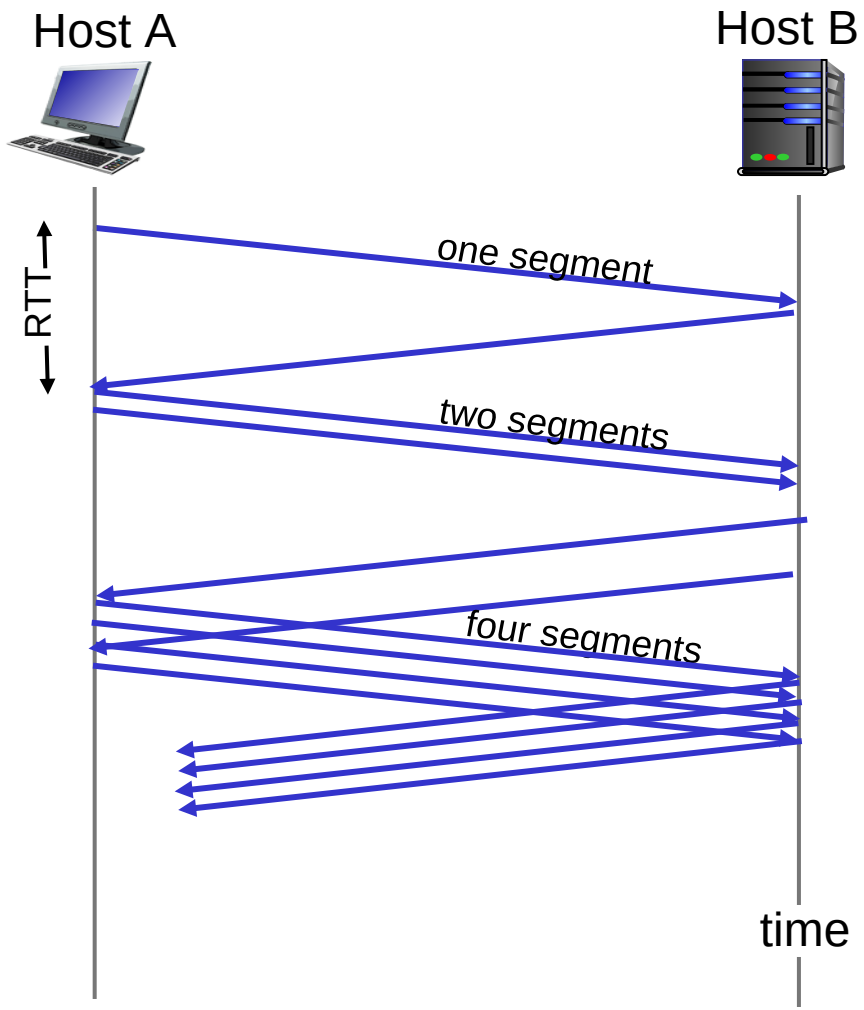
- *approach:* sender increases transmission rate (window size), probing for usable bandwidth, until loss occurs
  - *additive increase:* increase **cwnd** by 1 MSS every RTT until loss detected
  - *multiplicative decrease*: cut **cwnd** in half after loss

AIMD saw tooth behavior: probing for bandwidth

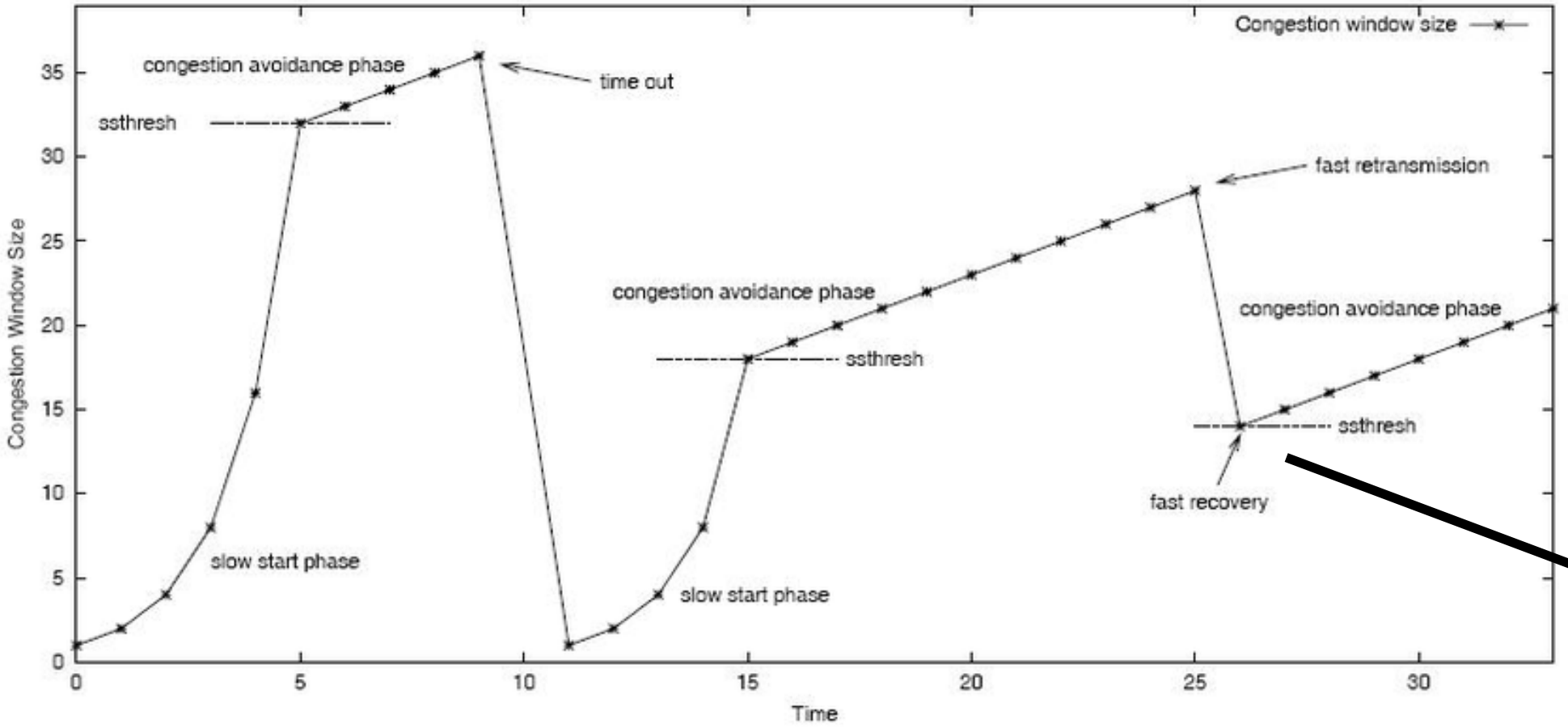


TCP Slow Start

Why not use a large number of segments?



TCP Slow Start and congestion avoidance



How to set ssthresh?

Initially – Randomly high

Later – adjusted as congestion happens

Why cut in half? Why not Start from the previous value?

https://www.researchgate.net/figure/3-TCP-slow-start-phase-and-congestion-avoidance-phase_fig3_225731524

TCP Throughput

TCP average throughput as a function of window size and RTT?
Ignore slow start, assume long TCP flow

Let W be the window size

Throughput = W/RTT
After loss, throughput = W/2*RTT
Average throughput = 0.75W/RTT

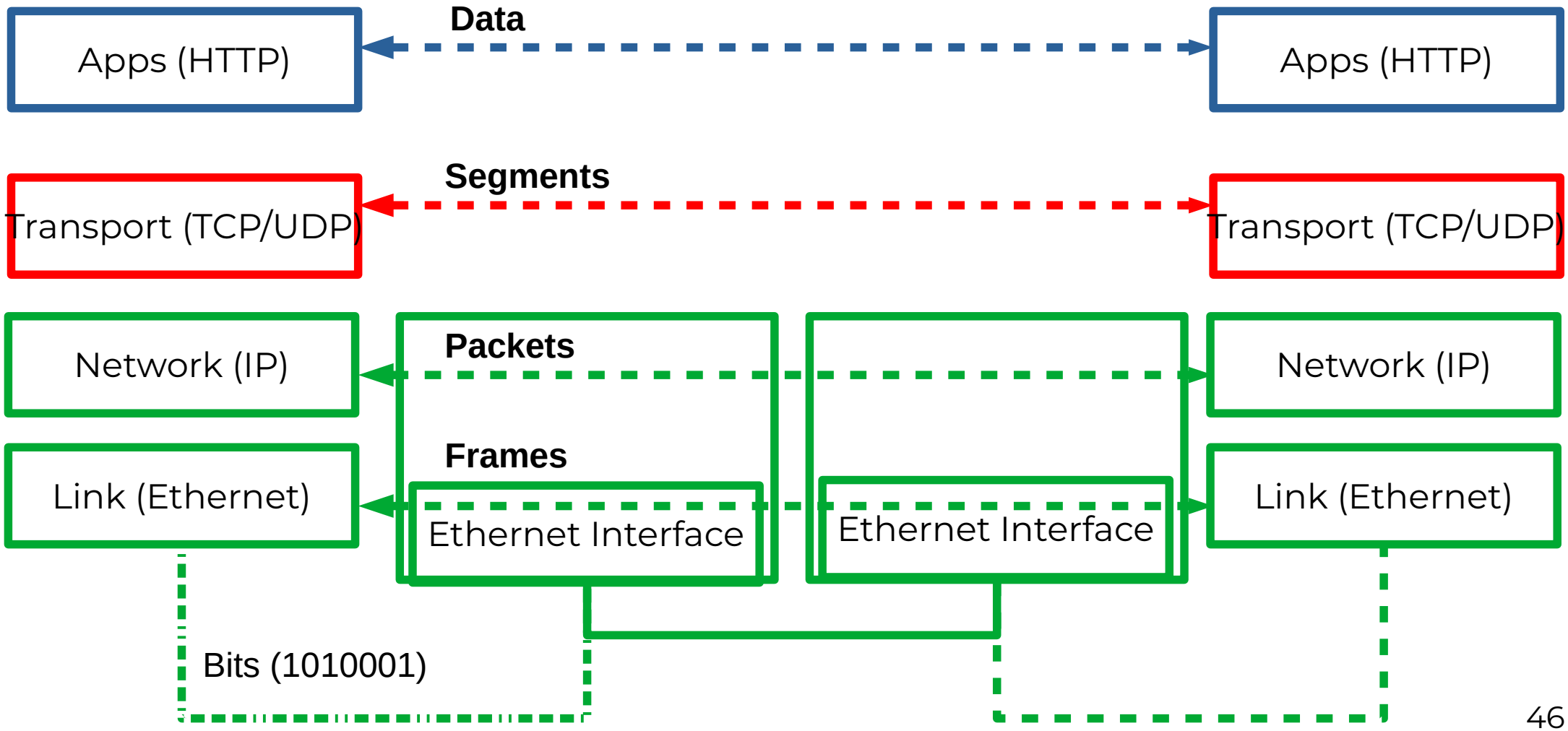
Throughput = (1.22*MSS)*(RTT/sqrt(Loss)) ← Magic formula

What is the loss rate to maximize 100Gbps pipe with 9000 bytes segments and 100ms RTT? Hint – must be very very low

Next steps: TCP Delay

We will briefly touch on TCP delay in the next class

Start with applications!


Apps (HTTP)
Data
Apps (HTTP)
Transport (TCP/UDP)
Segments
Transport (TCP/UDP)
Network (IP)
Packets
Network (IP)
Link (Ethernet)
Frames
Ethernet Interface
Ethernet Interface
Link (Ethernet)
Bits (1010001)