INTERNATIONAL COMPUTER SCIENCE INSTITUTE

CASE WESTERN RESERVE UNIVERSITY
EST. 1826

Life in the Fast Lane .... With a Model-T

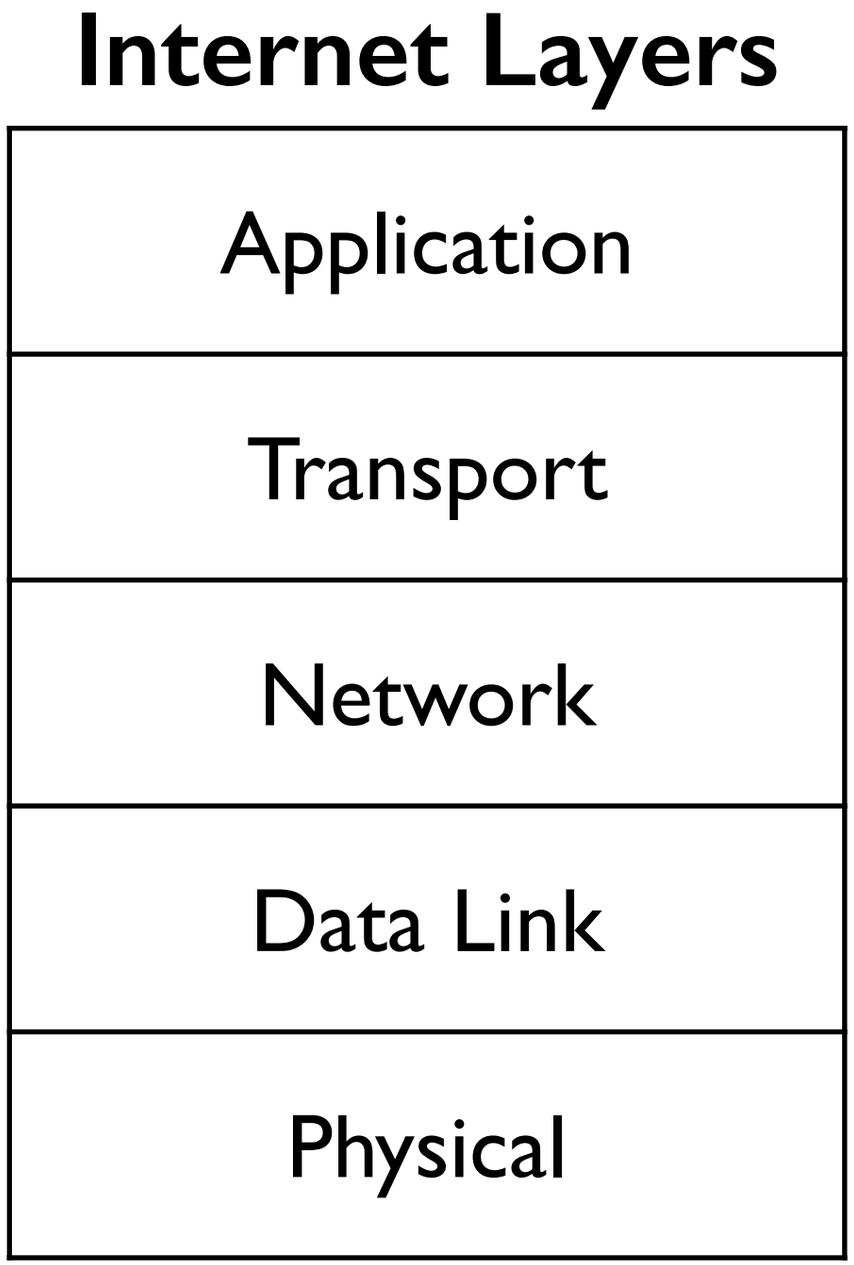
Why Last Mile Capacity is Just the Beginning

Mark Allman
ICSI / Case

Cyberinfrastructure Day
Case Western Reserve University
March 27, 2015

"With the radio blastin',
Goin' cruisin' just as fast as we can now"

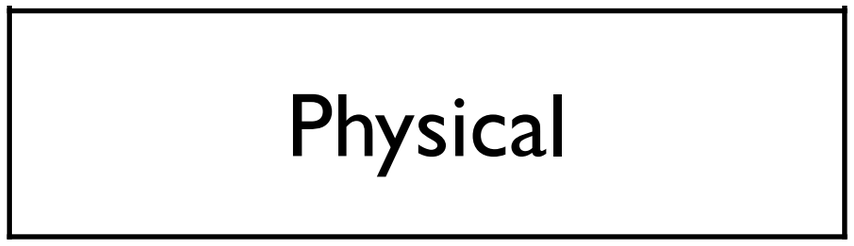
Thinking About Capacity



- This is how we theoretically and practically organize networks

Thinking About Capacity

Internet Layers



- This is what we consider when we discuss “capacity”
- Of course, layer 1 capacity is required

Thinking About Capacity

Internet Layers

- But, this is where capacity really matters

Thinking About Capacity

Internet Layers

Application
Transport
Network
Data Link
Physical

Wireless Networks

Internet Layers

Application
Transport
Network
Data Link
Physical

Residential Networks

Internet Layers

Application
Transport
Network
Data Link
Physical

Wired Enterprise Networks

Internet Layers

Application
Transport
Network
Data Link
Physical

Data Center Networks

Internet Layers

Application
Transport
Network
Data Link
Physical

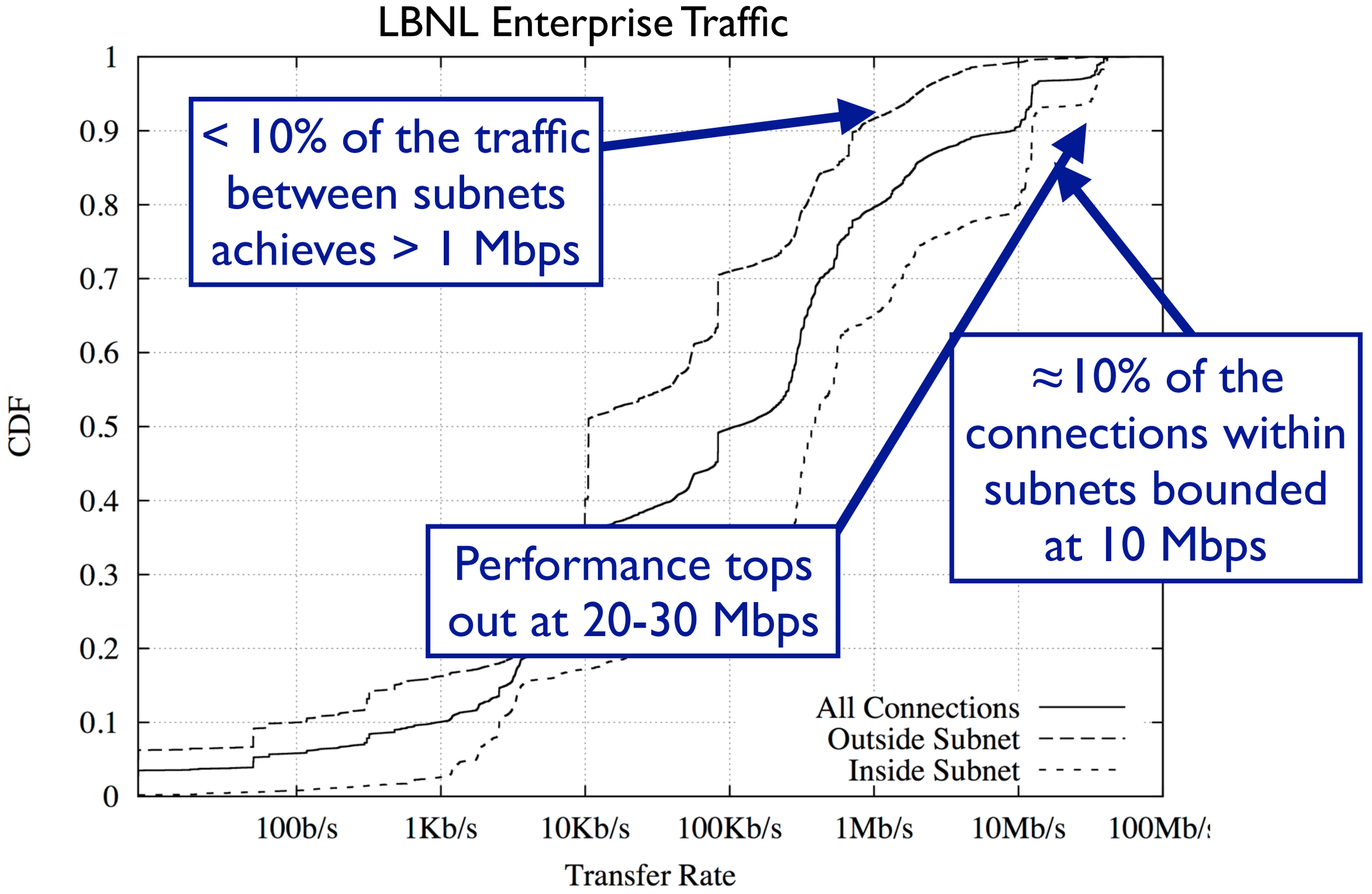
Highly Specialized Networks

Internet Layers

Application
Transport
Network
Data Link
Physical

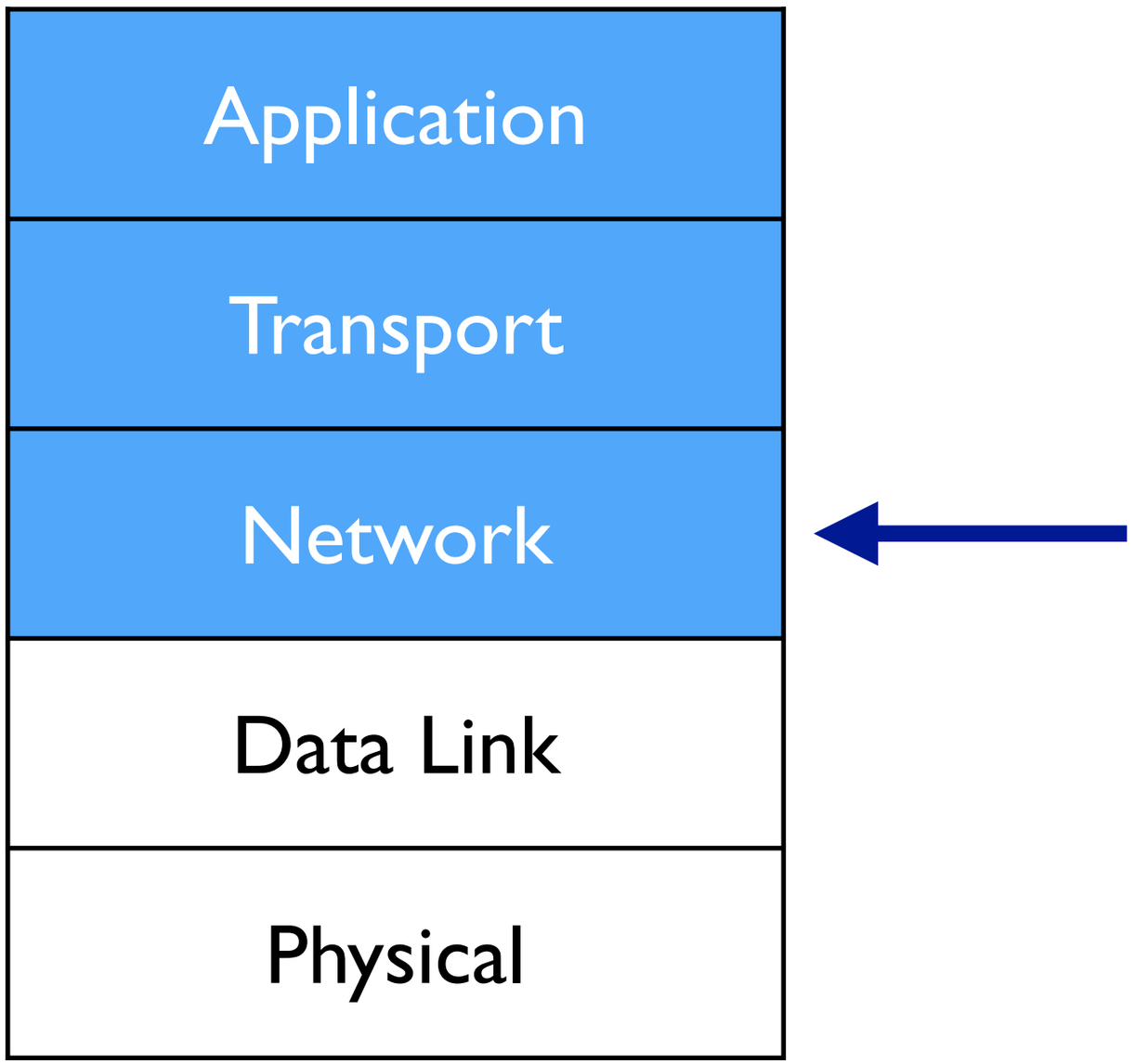
… but …

A First Look


LBNL Enterprise Traffic
< 10% of the traffic between subnets achieves > 1 Mbps
≈10% of the connections within subnets bounded at 10 Mbps
Performance tops out at 20-30 Mbps
All Connections
Outside Subnet
Inside Subnet
CDF
Transfer Rate
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
100b/s
1Kb/s
10Kb/s
100Kb/s
1Mb/s
10Mb/s
100Mb/s

It's A Long Way To The Top …

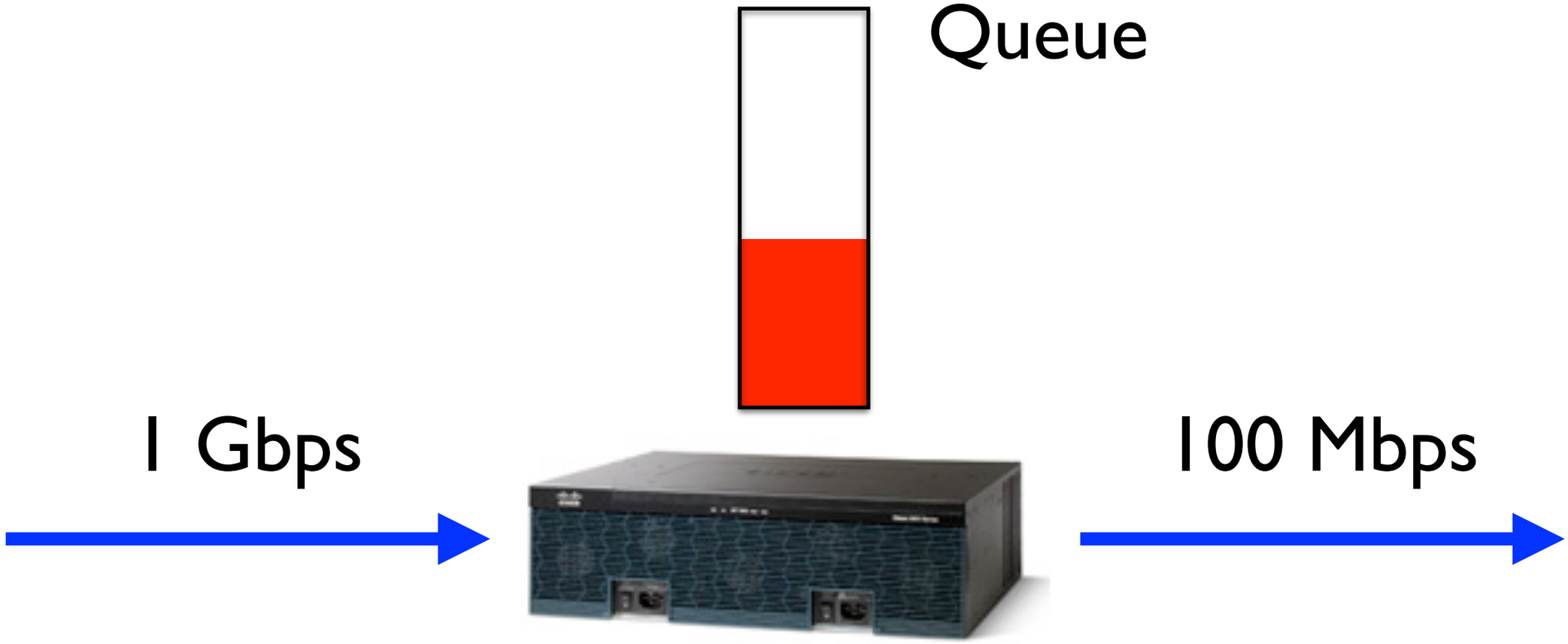
Internet Layers



Network Layer

- Fundamentally about moving packets from one host to another
- But, also, naturally must deal with *impedance mis-matches*

Network Layer

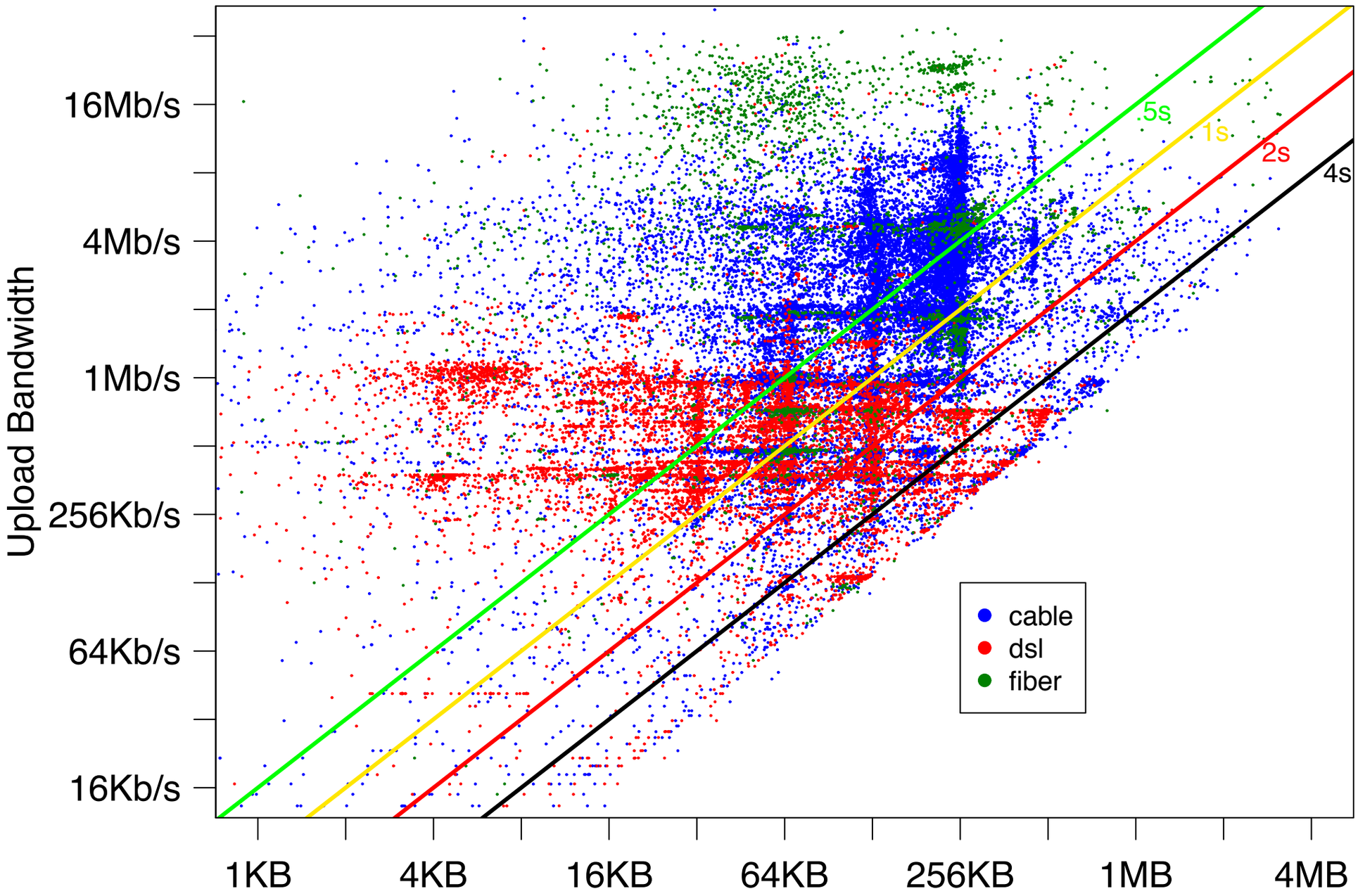


Bufferbloat

- Claim: over buffering in networking gear is a ***big problem***
  - *"dark storm clouds surround[ing] us" (Gettys)*

- Case: anecdotal or mis-understood

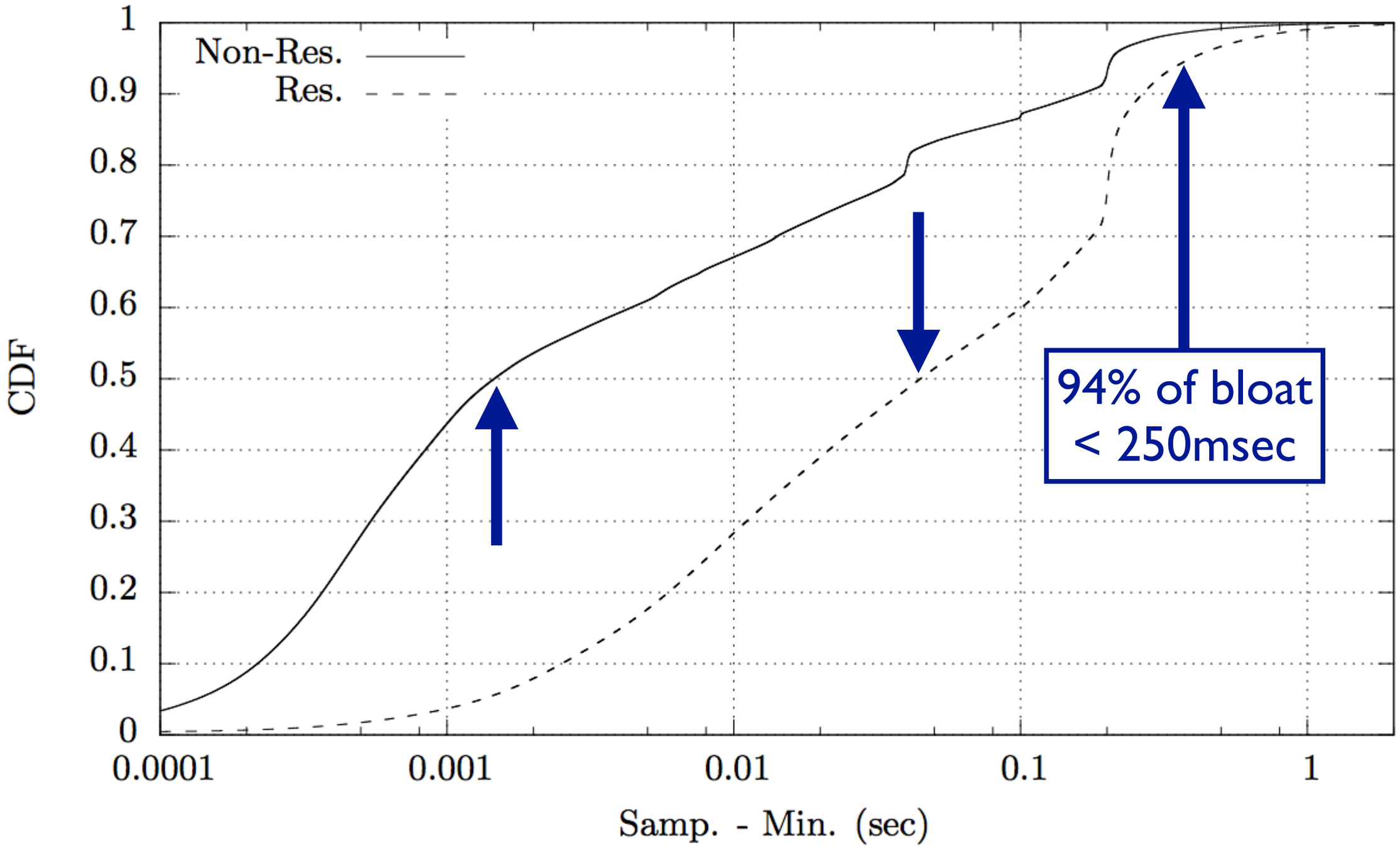
Bufferbloat Can Happen …



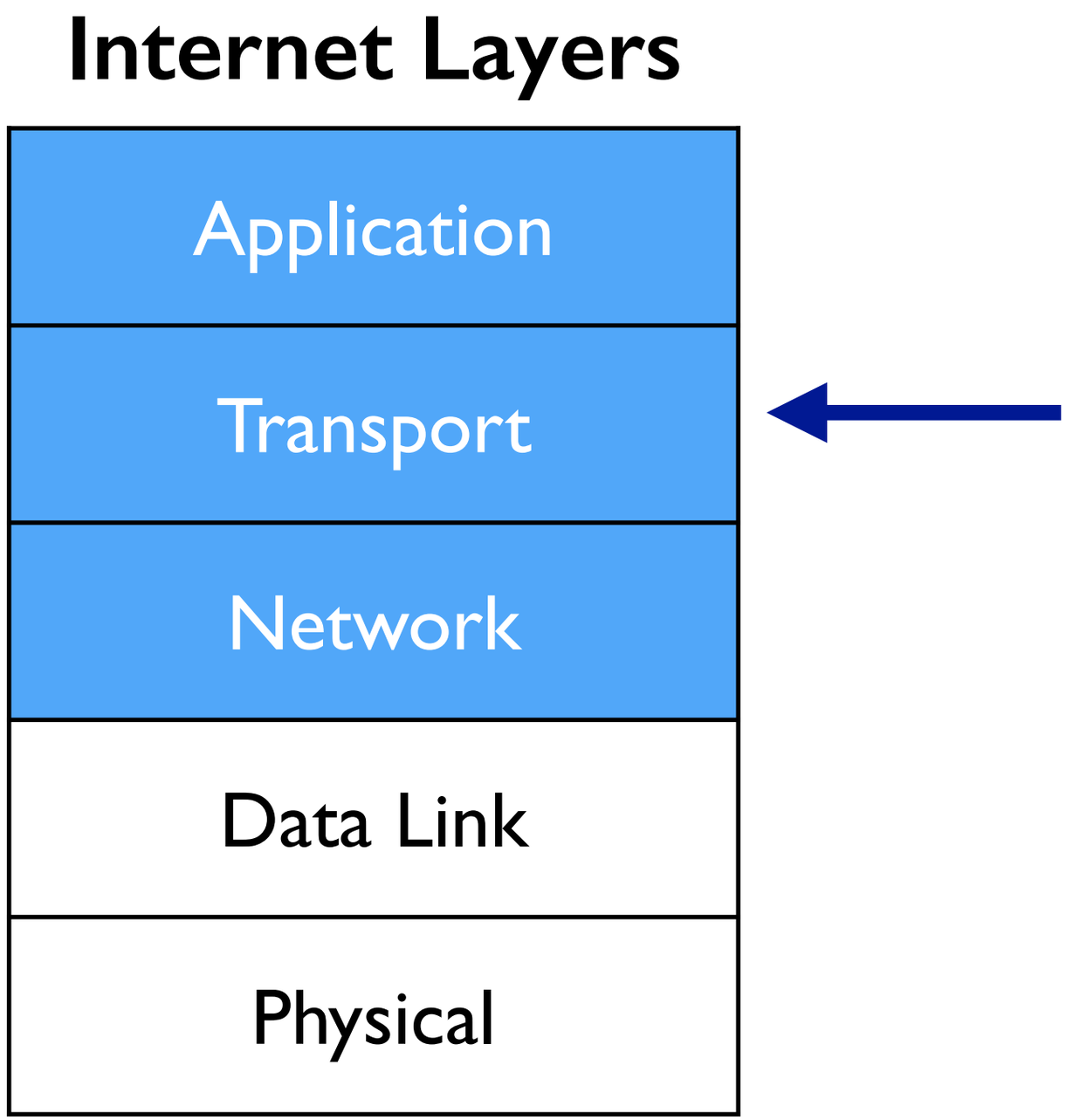
Inferred Buffer Capacity

Kreibich, et.al., IMC 2010

… But Doesn't That Much



It's A Long Way To The Top …



Transport Layer

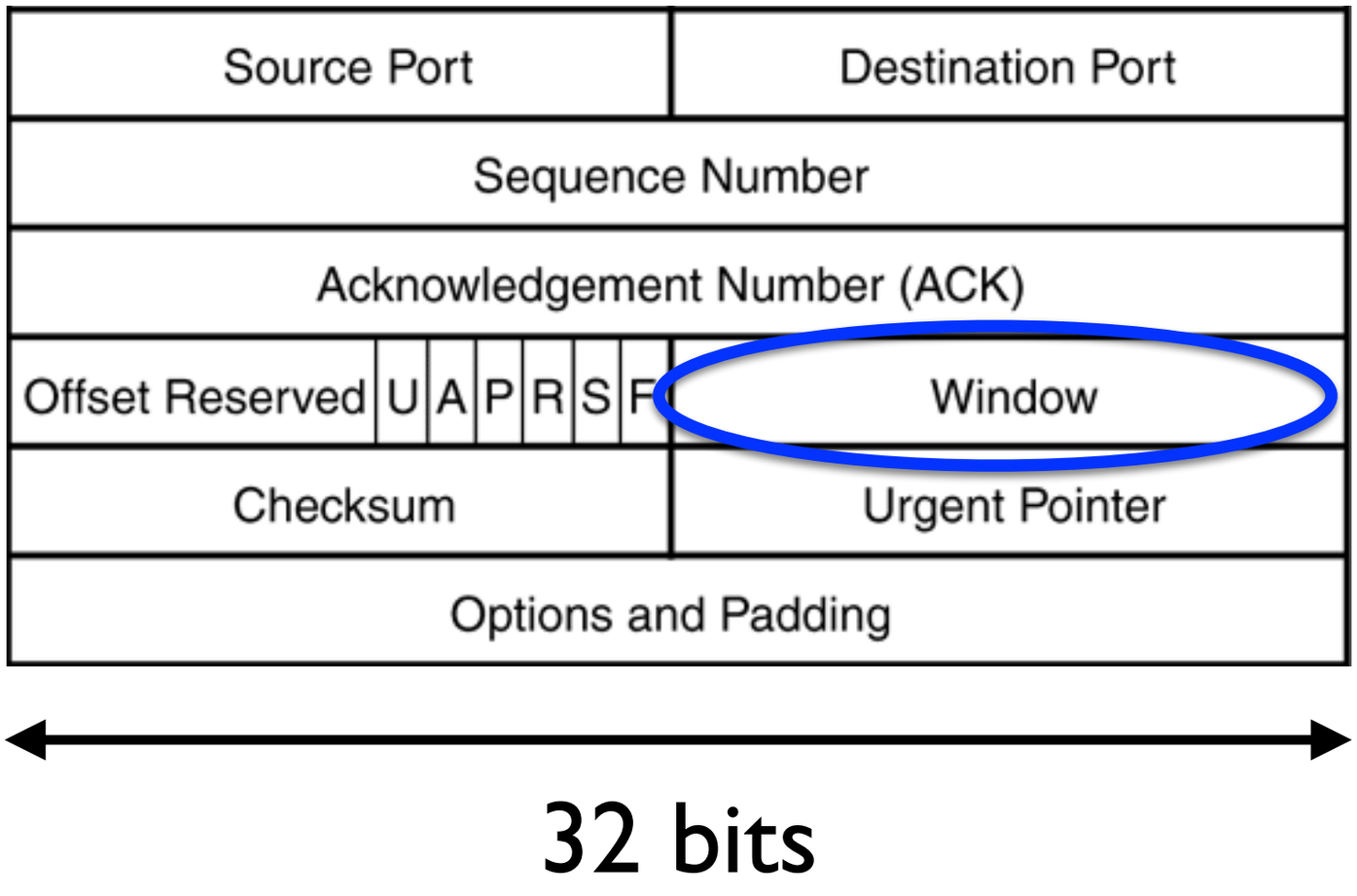
- TCP imposes two kinds of control on traffic
  - flow control
  - congestion control

- Both provide fundamental limits on performance via sliding windows

Sliding Window Protocols

$$Thput = \frac{Window}{RTT}$$

Flow Control

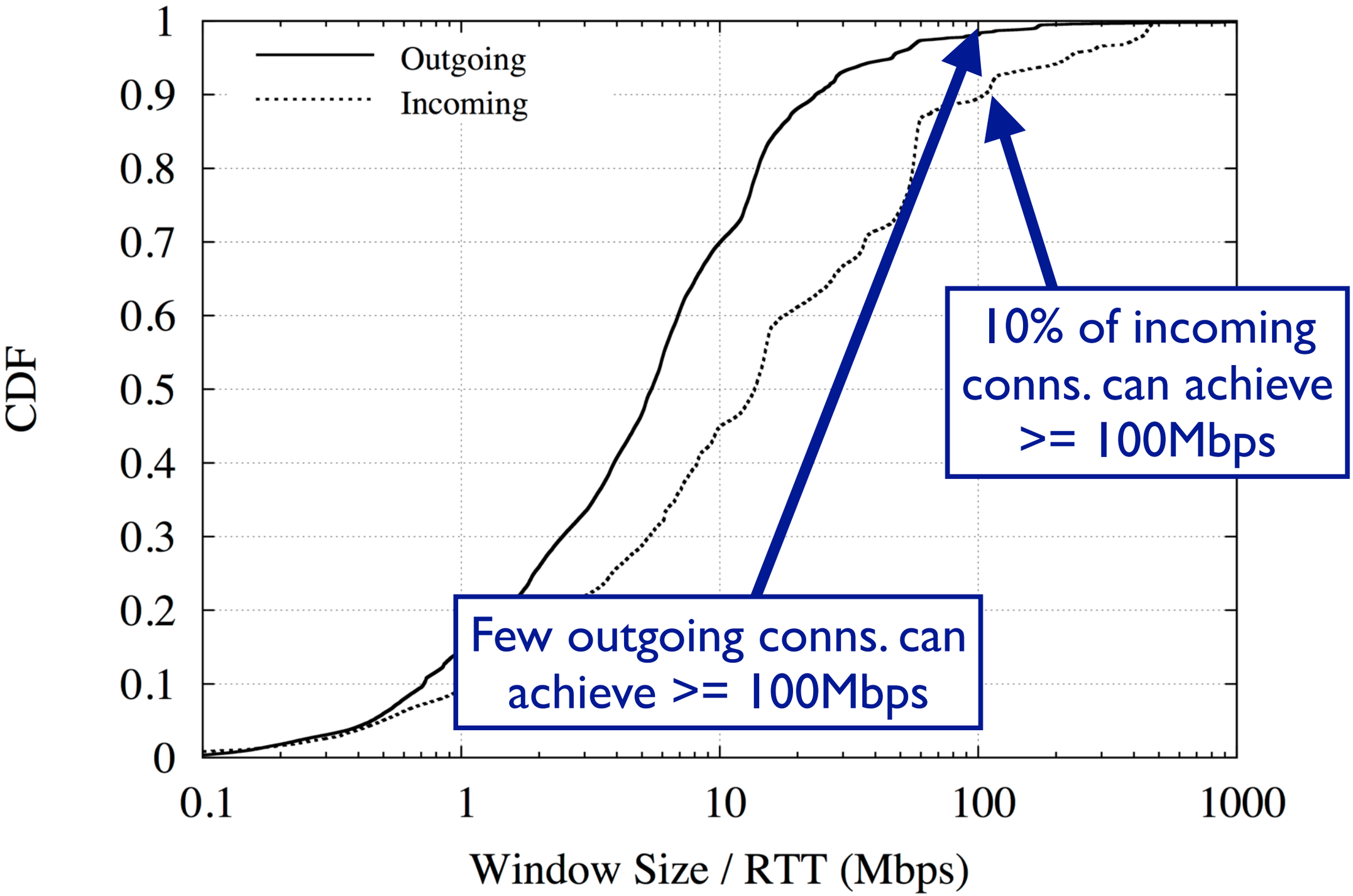


x 2^{14} = 1 GB max

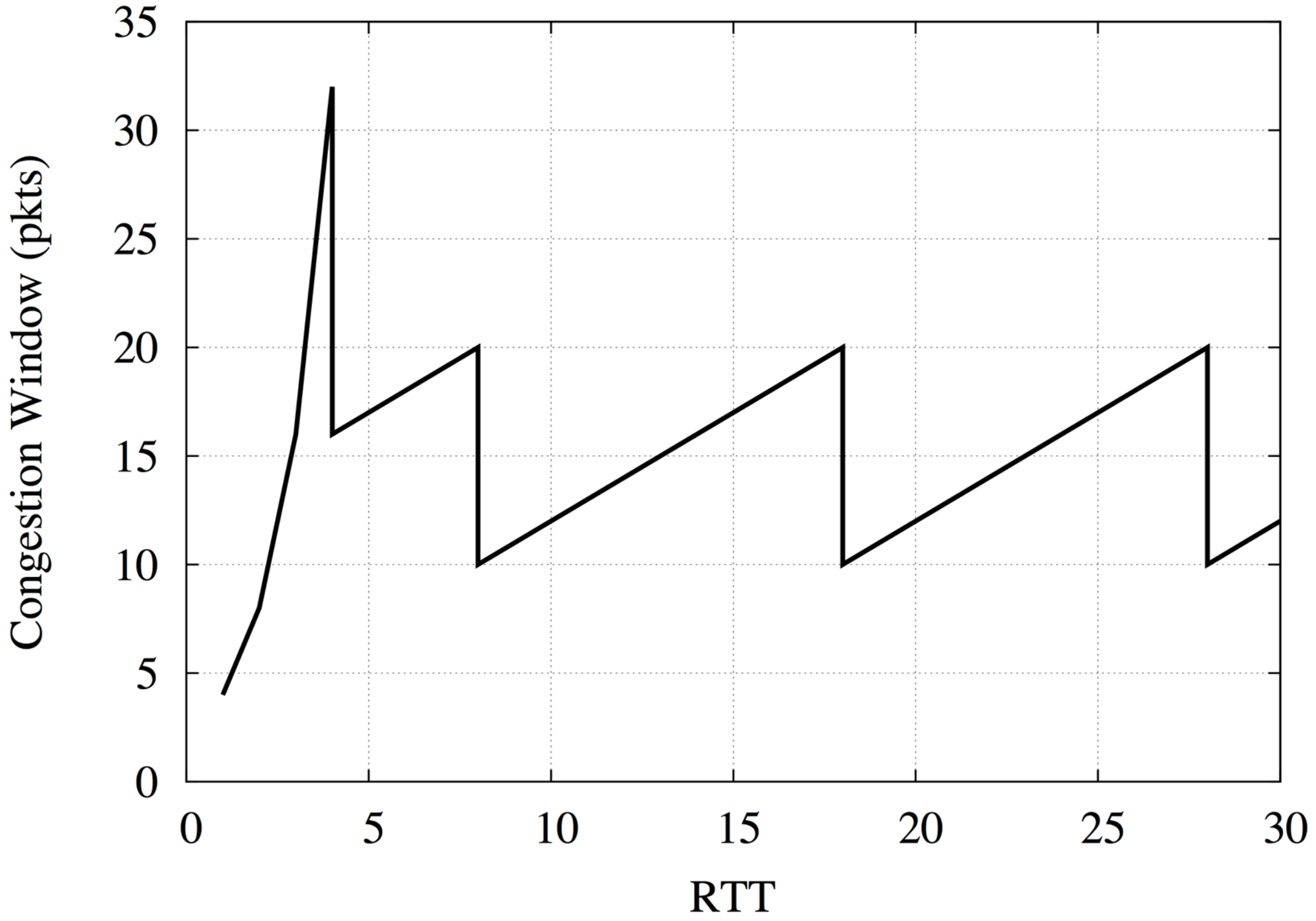
Flow Control

Window	RTT	Thput
1 MB	1 msec	7.8 Gbps
1 MB	10 msec	800 Mbps
1 MB	100 msec	80 Mbps
64 KB	1 msec	500 Mbps
1 GB	1 msec	8 Tbps
640 KB	500 msec	10 Mbps

CCZ Flow Control

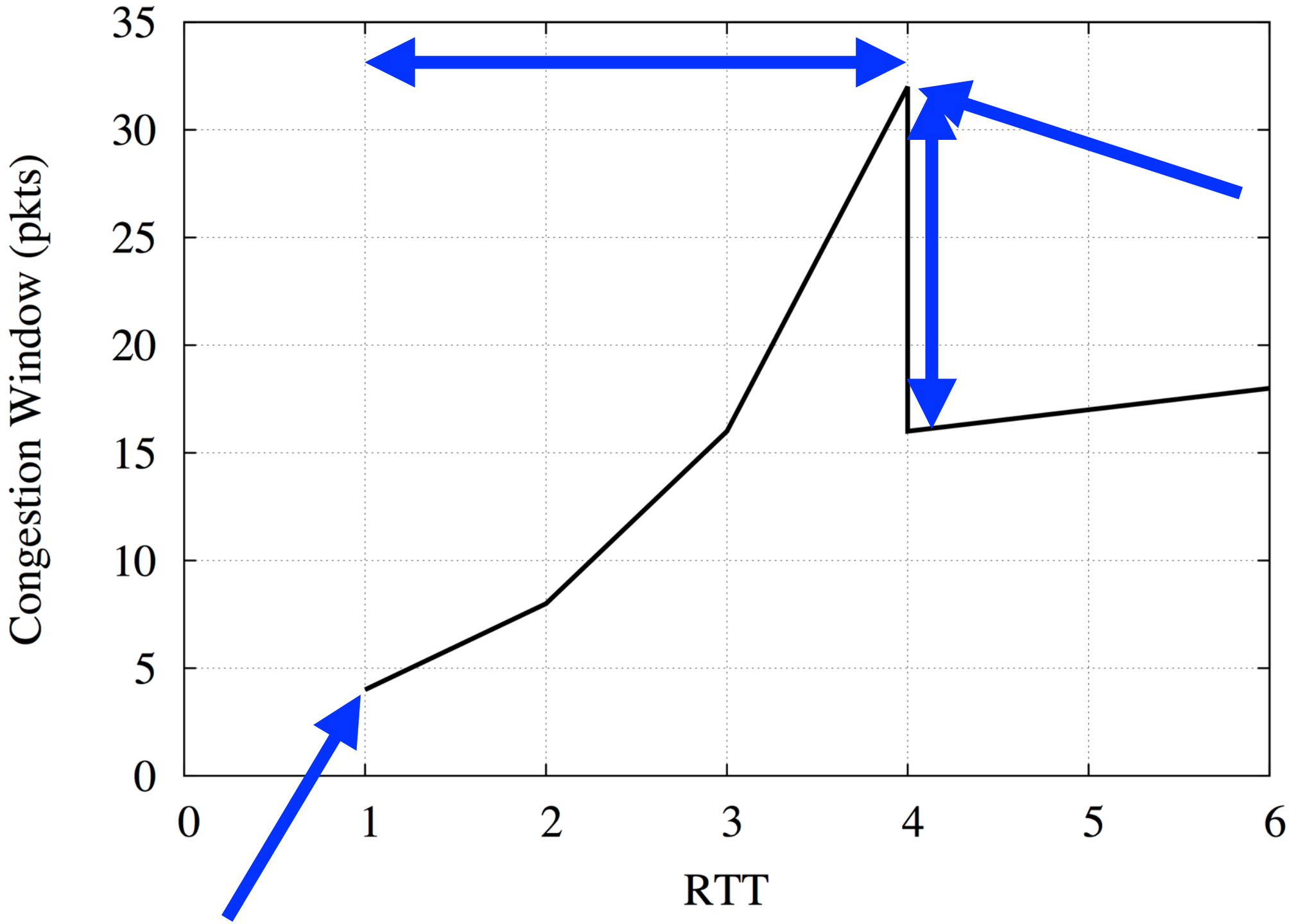


Congestion Control



RFC 5681

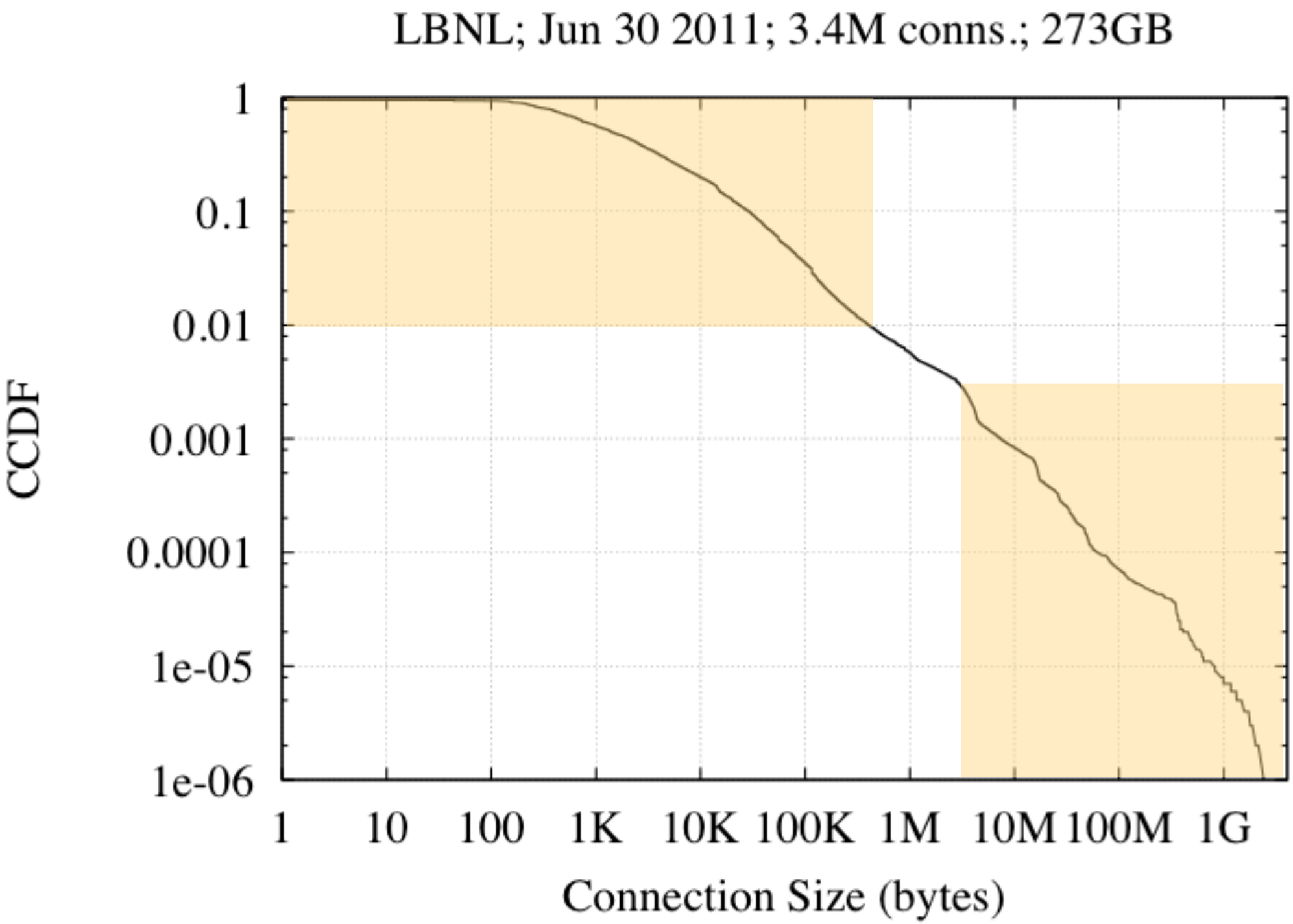
Congestion Control



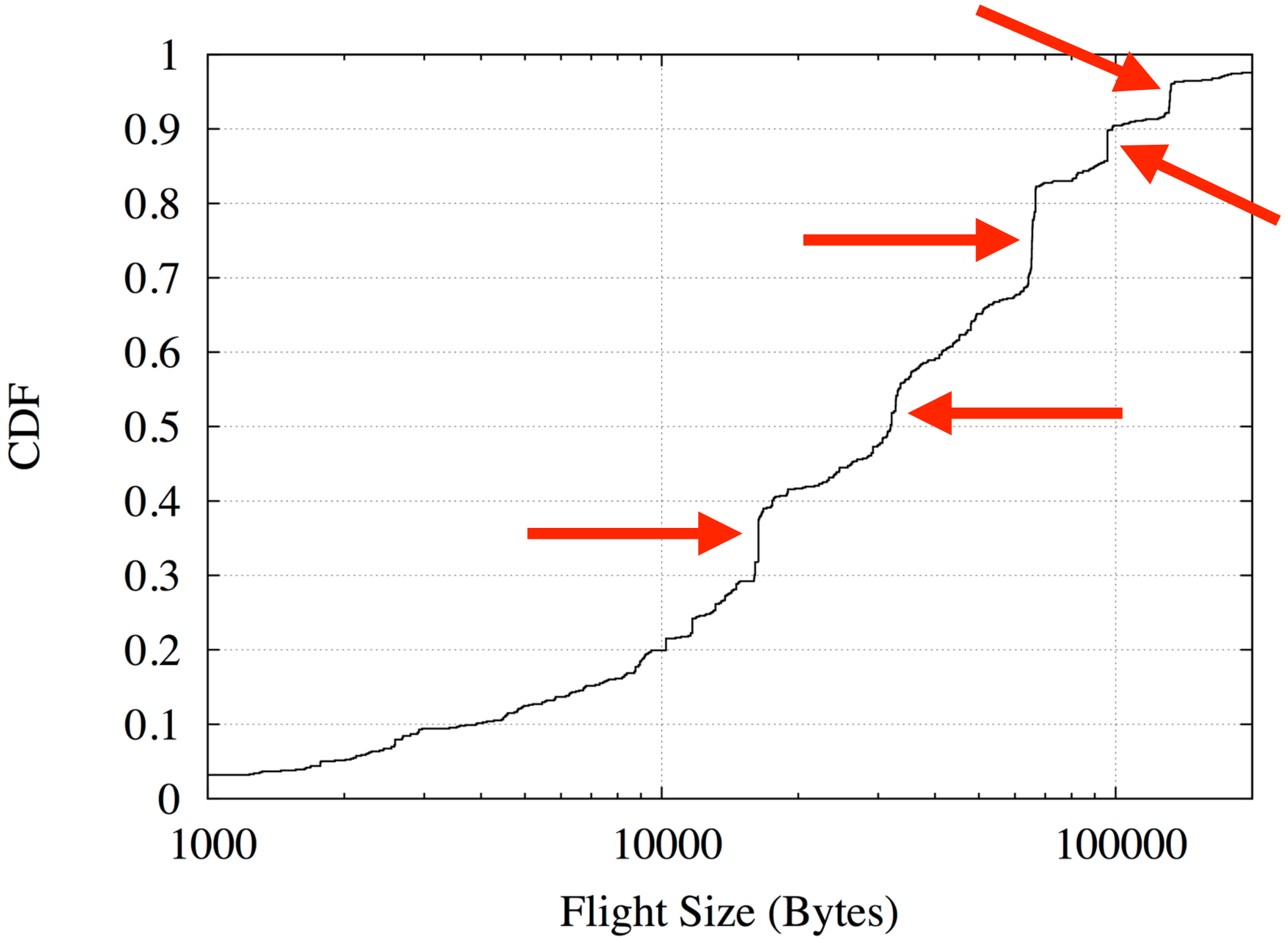
Congestion Control

- Takes *data* to build *speed*
- To reach a congestion window of X bytes requires sending ≈2X data bytes

Connection Sizes



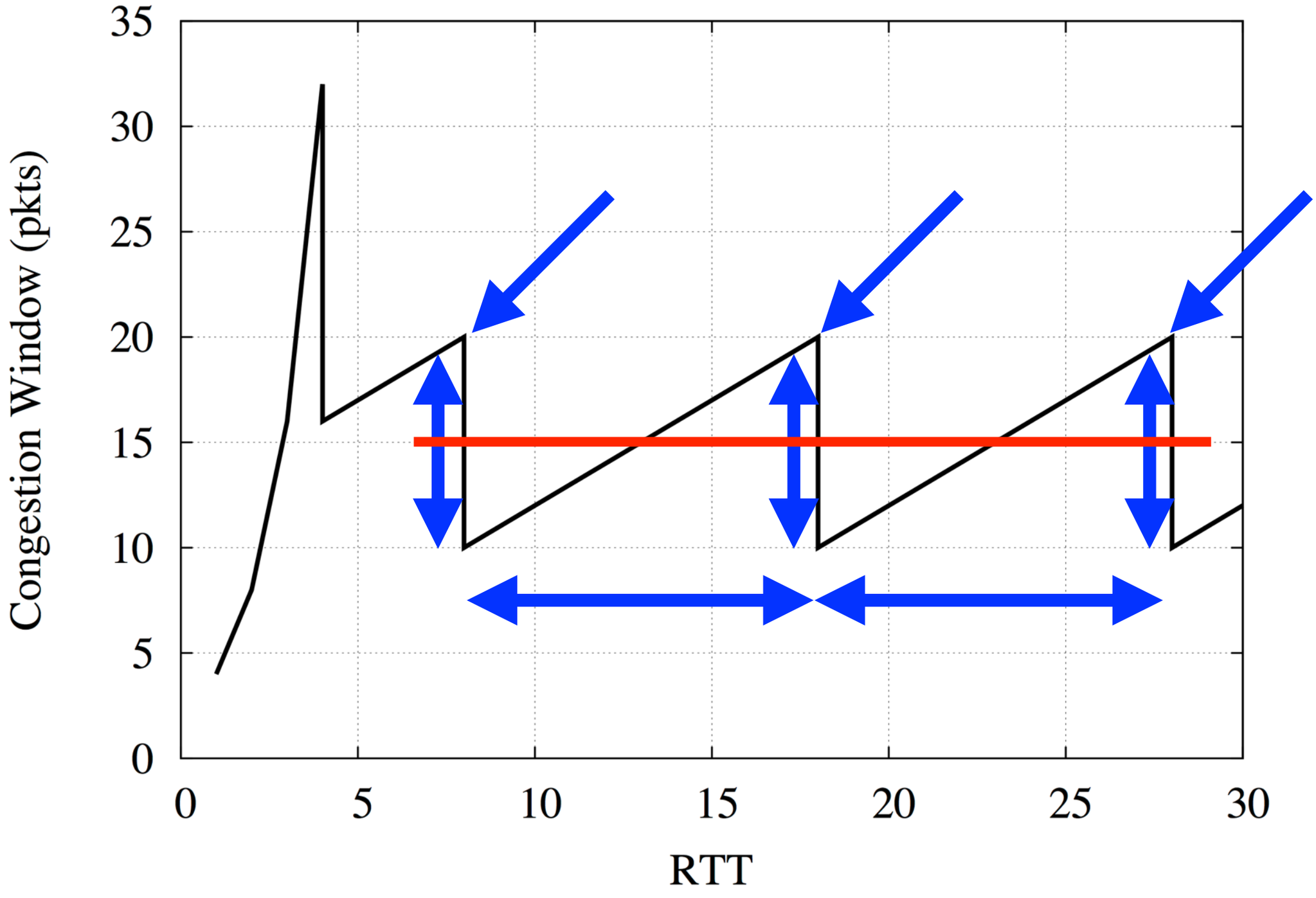
CCZ Connections Without Loss



CCZ Connections Without Loss

- 12% of connections constrained by advertised window
- 45% of connections constrained by a sender-side buffer

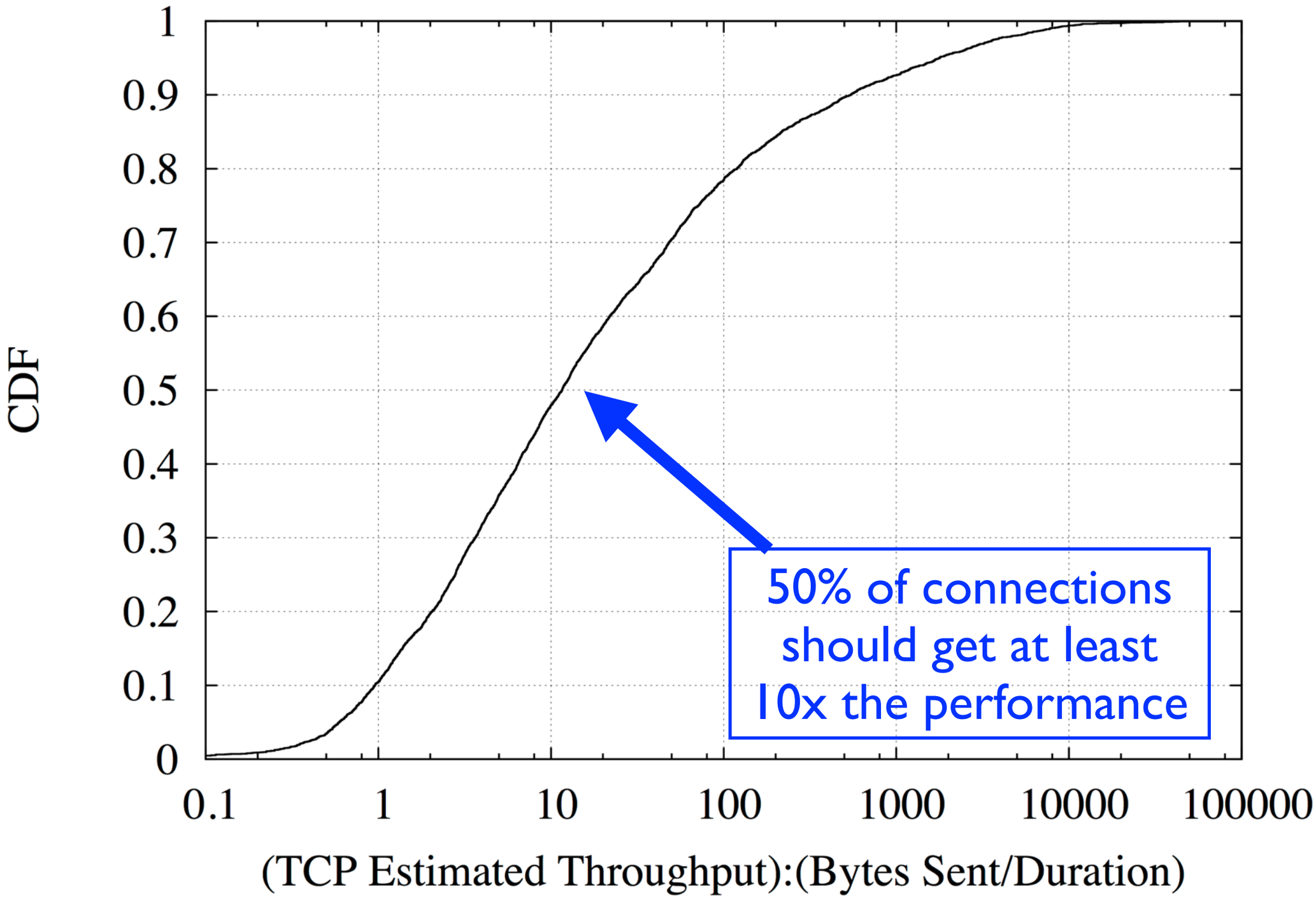
Congestion Control



General TCP Performance Model

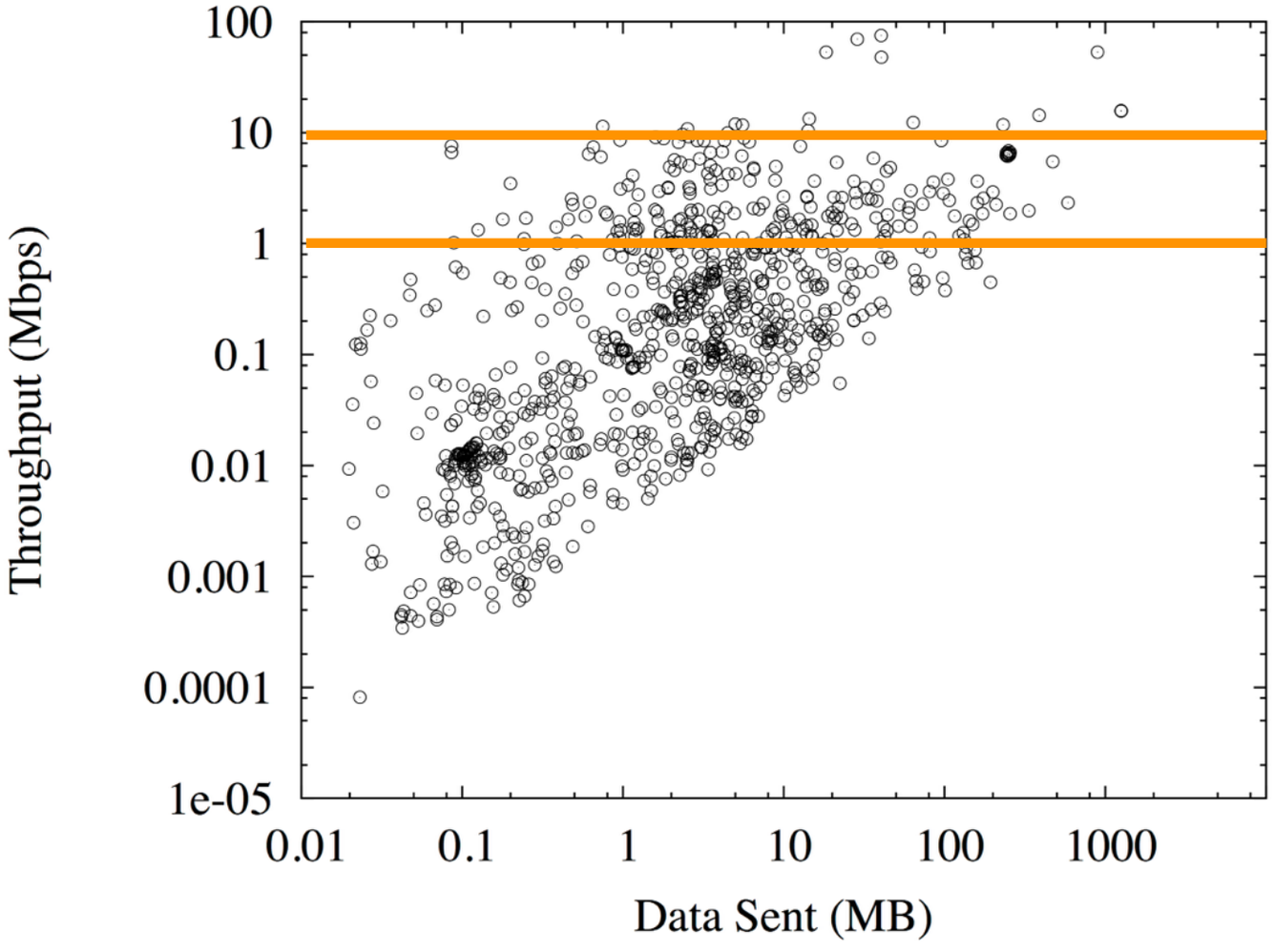
$$BW \propto \frac{MSS}{RTT \times \sqrt{p}}$$

CCZ Connections With Loss

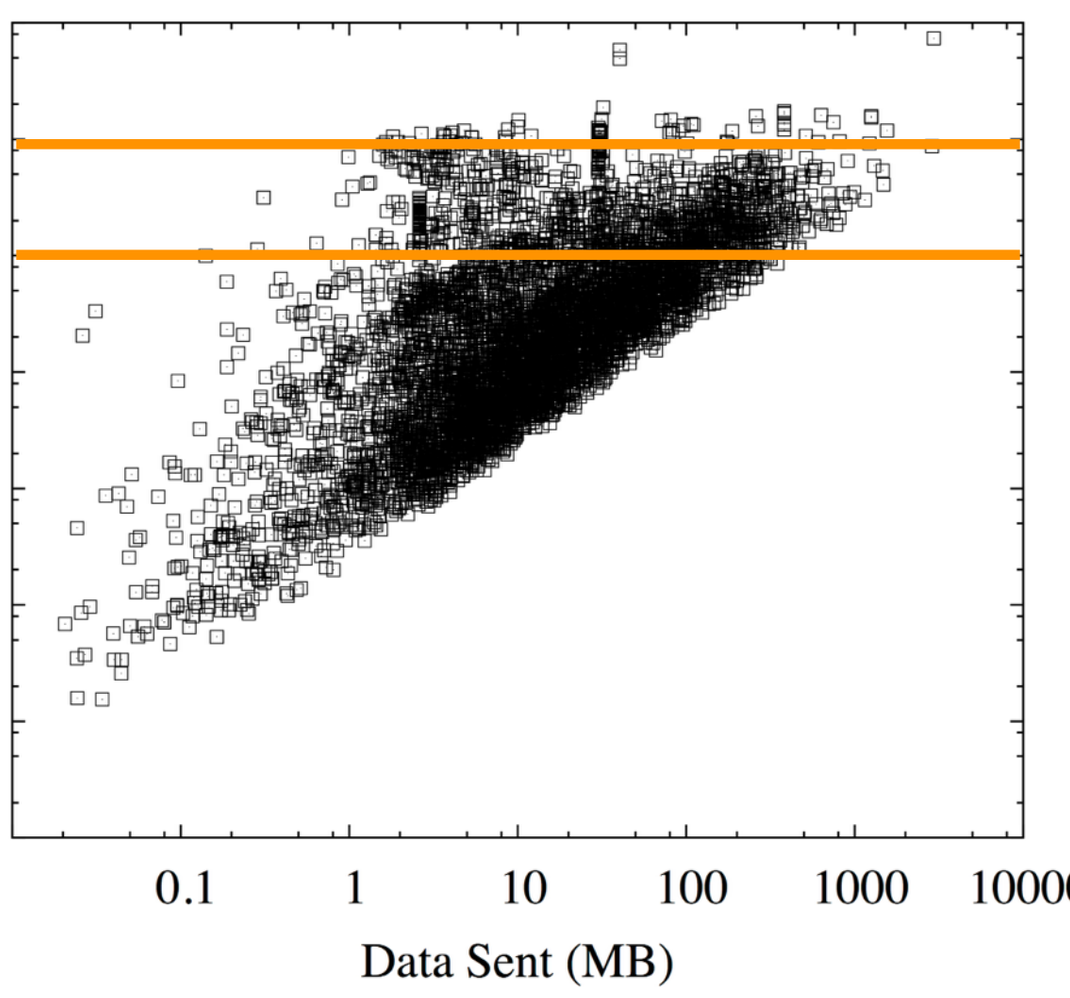


CCZ Performance

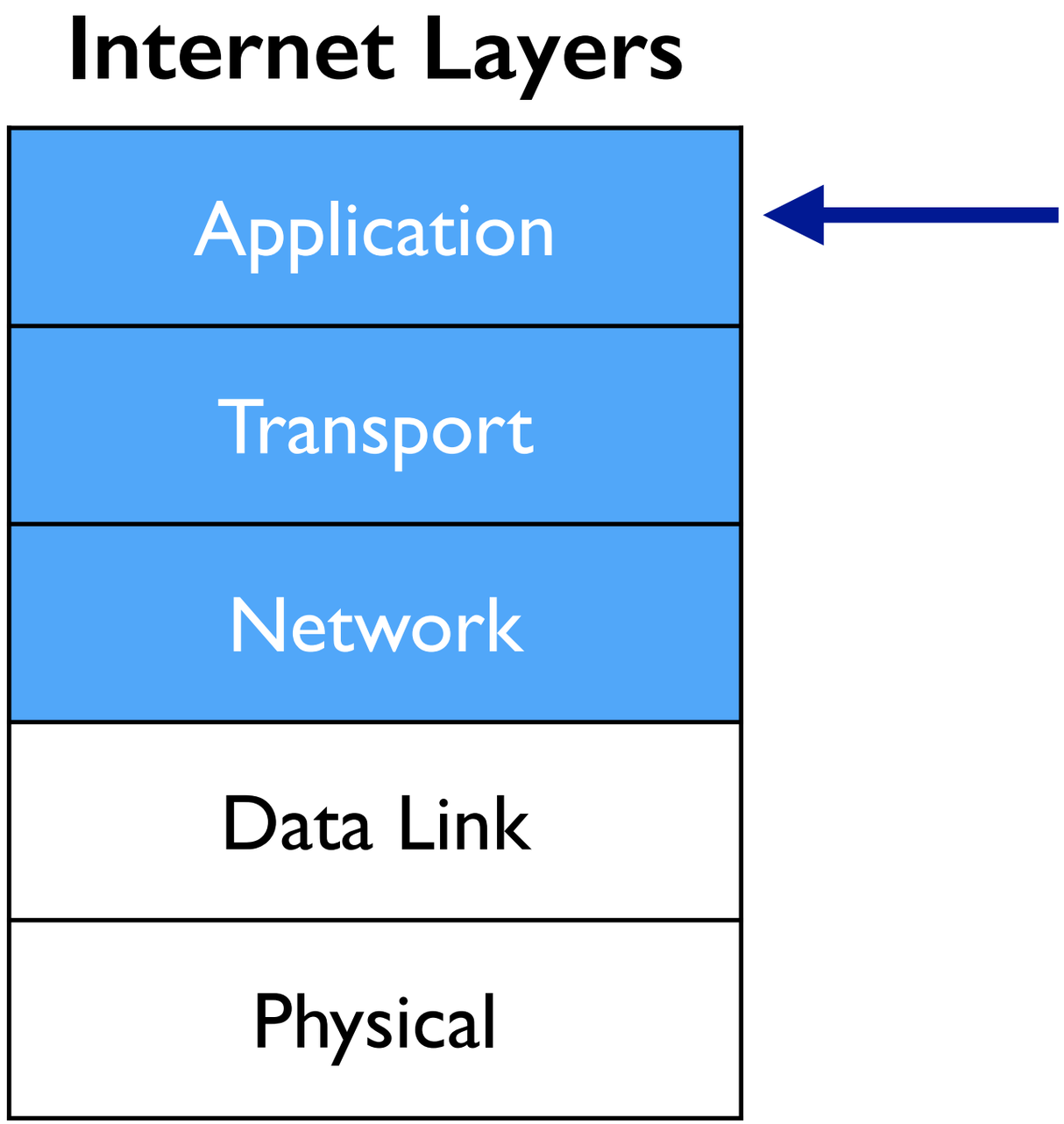
Connections Without Loss



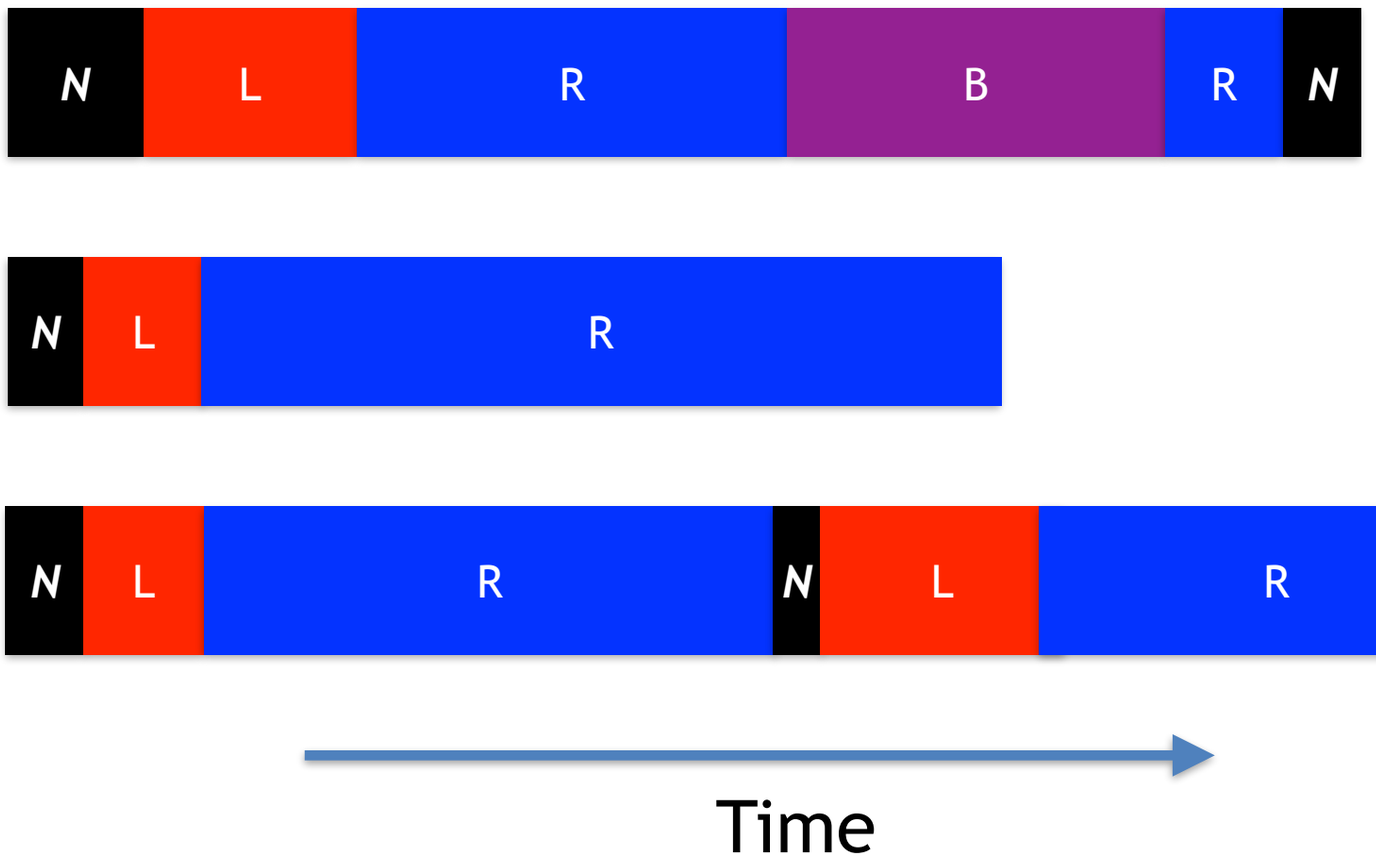
Connections With Loss



It's A Long Way To The Top …



Application Patterns



Application Use

- We find internal silence in …
  - … 28% of connections at ICSI border
  - … 37% of connections at CCZ border

Thinking About Capacity

Internet Layers

Application
Transport
Network
Data Link
Physical

Acknowledgments

- TCP performance work from Matt Sargent (Case EECS)

- Work funded in part by NSF

- Data sources: Case CCZ, LBNL, ICSI

Questions?
Comments?

References:

http://www.icir.org/mallman/perf.html

Mark Allman, *mallman@icir.org*
http://www.icir.org/mallman/