Characterizing Commercial Broadband Networks
or
How we fail to get simple system designs right as we try to build complex, self-(*) systems

Andreas Haeberlen, Marcel Dischinger
Krishna P. Gummadi

Networked Systems Research Group
Max Planck Institute for Software Systems
Saarbruecken, Germany

High-level problem

- The Internet is huge, complex, and rapidly changing
  – new user applications and workloads emerge over time
    - music/movie downloads, Voice-over-IP, multi-player games
  – new network infrastructures are constantly deployed
    - broadband (cable/DSL), cellular, wireless (802.11X, WiMax)

- Research agenda:
  – what are characteristics of deployed networks & apps?
  – what are the implications for future network & app designs?
  – how to implement and test new designs in practice?

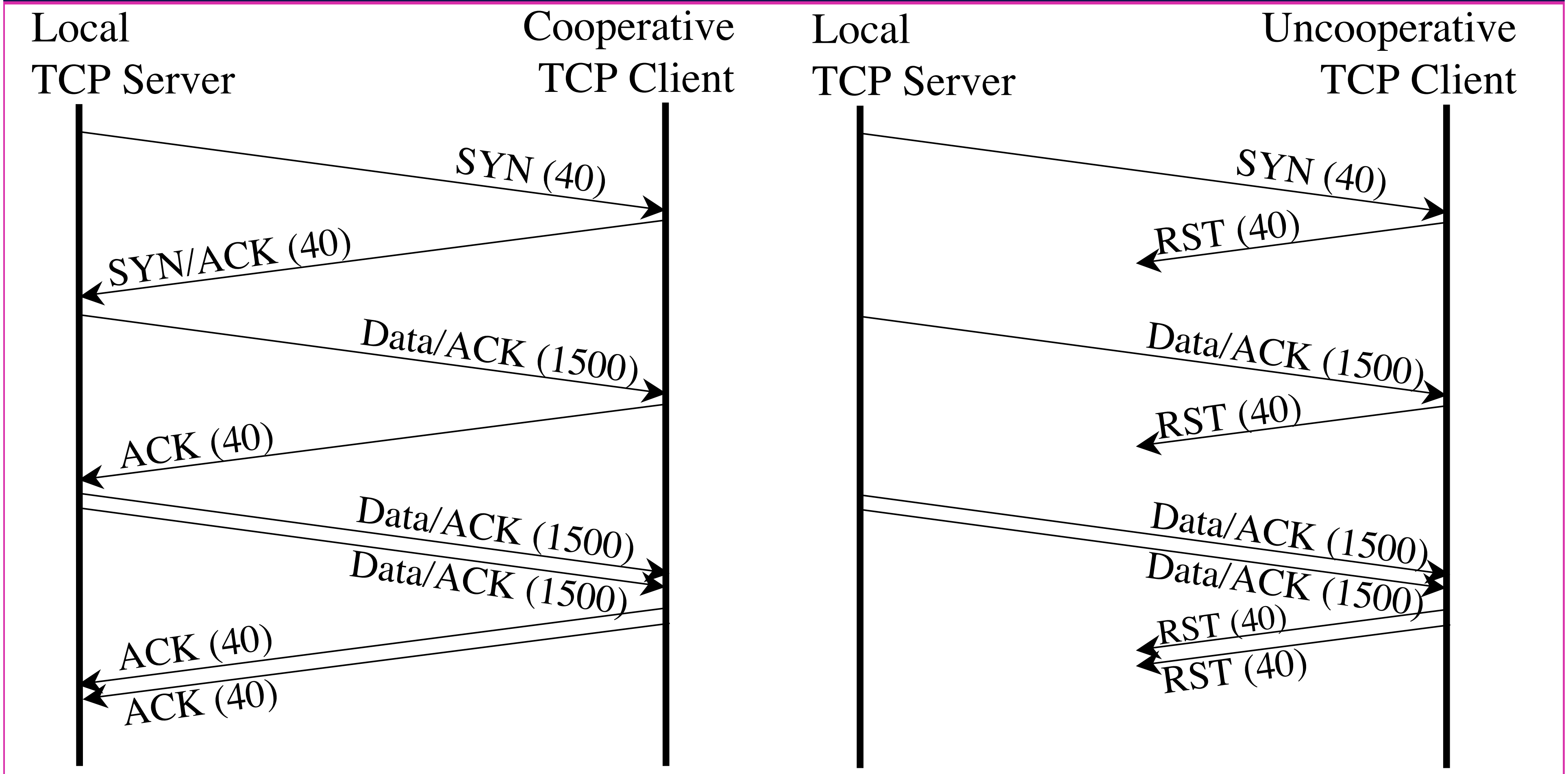
This talk:
Study of commercial broadband networks

- A large, growing fraction of Internet users connect via Cable & DSL
- A major driving force behind emerging applications
  - VoIP, online games, multimedia downloads, P2P, distance learning
- Yet, very little is known about their characteristics
- Goals:
  - measure broadband access networks characteristics
  - derive implications for the design of future networks and systems

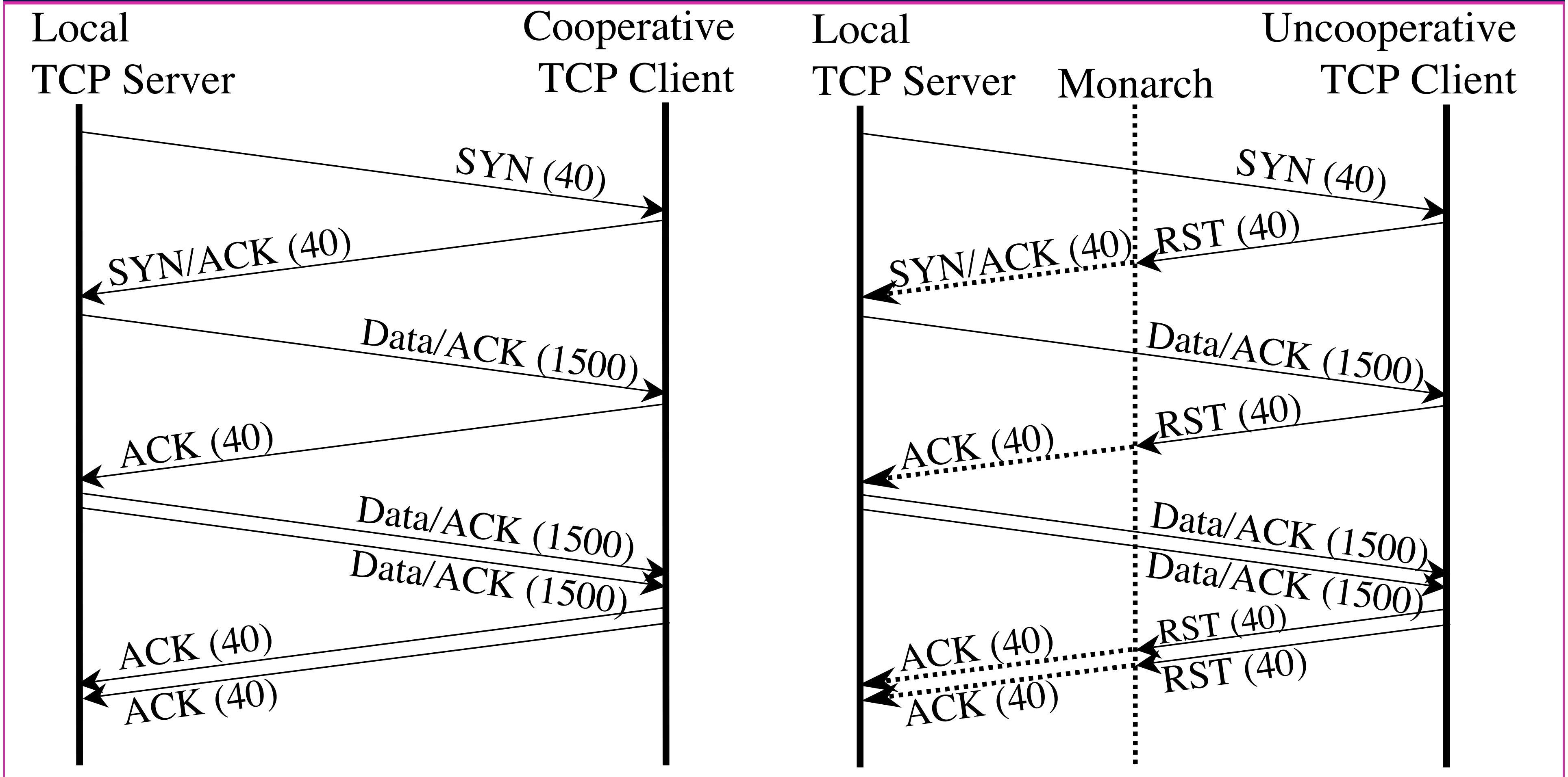
Measuring commercial broadband is challenging

- Hard to obtain lots of cooperating broadband hosts
  - prior studies limited by willingly participating nodes
    - e.g., a popular study by MSR had 25 nodes
- Challenge: How do we measure uncooperative nodes?
  - how can we generate TCP traffic to uncooperative nodes?

Monarch: Emulating TCP flows to uncooperative hosts



Monarch: Emulating TCP flows to uncooperative hosts



- Basic idea: Intercept and transform RST packets from an uncooperative remote host into ACK packets

Broadband study methodology

- TCP servers at 4 geographically dispersed locations
  - Seattle, Houston, Toronto, and Saarbruecken
- 6, 228 TCP clients from 12 major broadband providers

DSL							*Cable*				
Ameritech	***Bellsouth***	***BT***	***Pacbell***	***Qwest***	***Swbell***	***Verizon***	***Chartr***	***Chell***	***Comcast***	***Roadrunner***	***Rogers***
USA	**USA**	**UK**	**USA**	**USA**	**USA**	**USA**	**USA**	**NL**	**USA**	**USA**	**Canada**
223	**271**	**24 4**	**314**	**440**	**882**	**84**	**354**	**215**	**1445**	**1592**	**224**

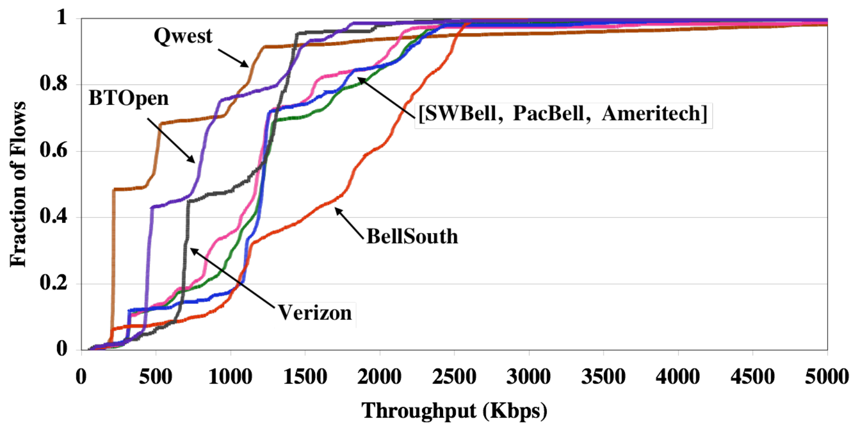
- Analyzed Monarch TCP traces from servers to clients to infer: (a) throughput, (b) latency & queueing delay, (c) packet loss

Analysis driven by three questions

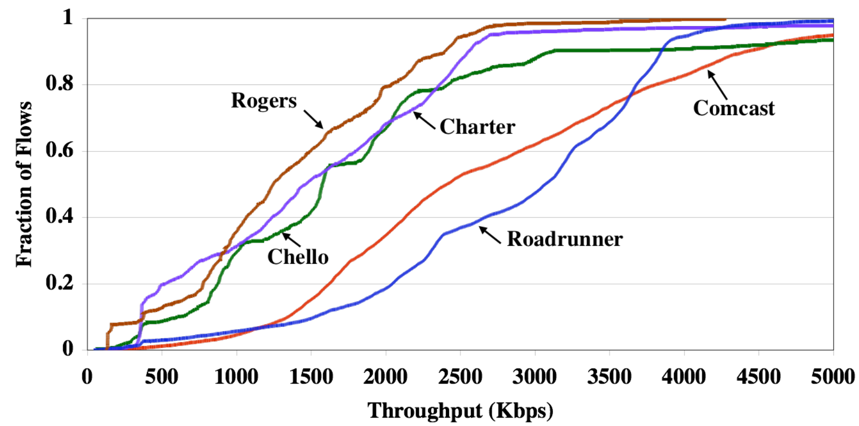
Flow characteristics of interest: throughput, latency, loss

- How do flow characteristics vary across different ISPs?
- How does server location affect flow characteristics?
- What factors drive the differences or similarities?

How do throughputs differ across providers?



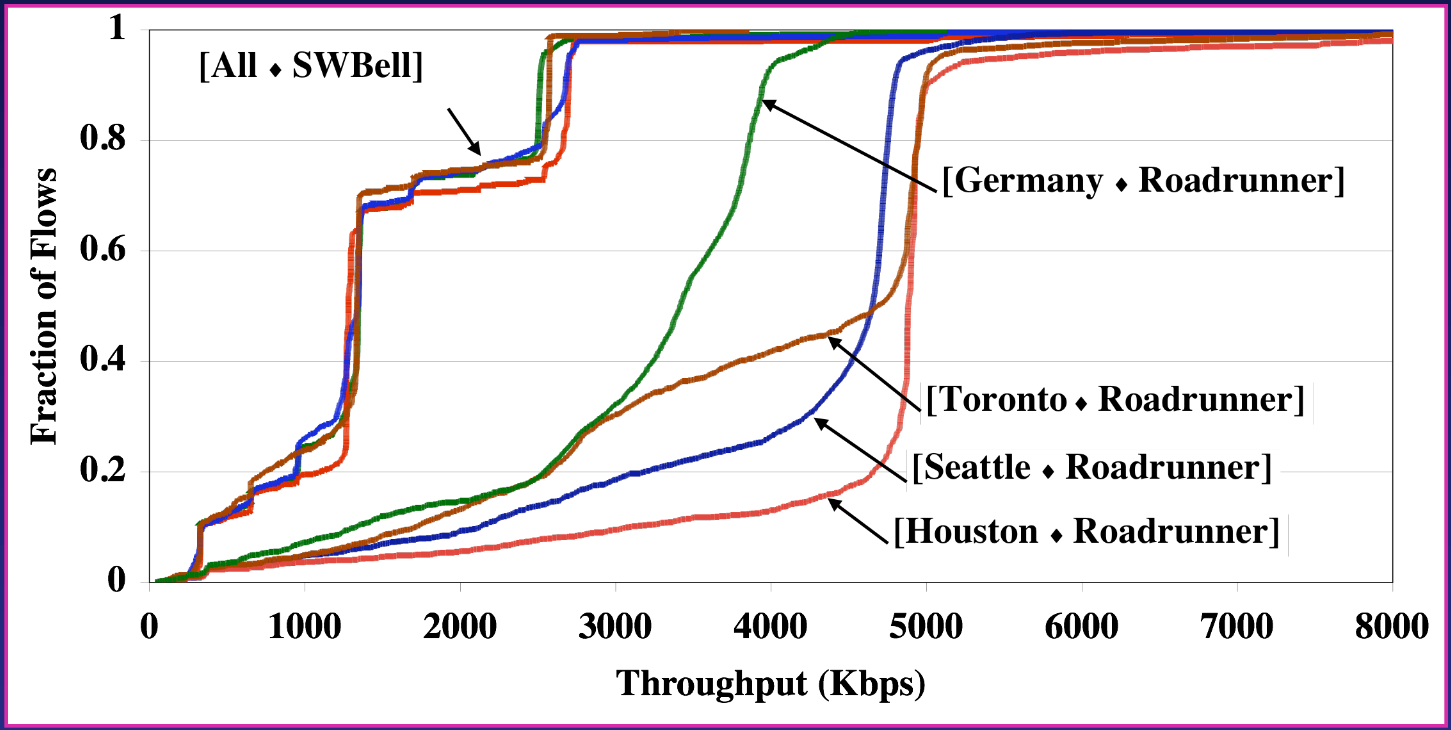
DSL



Cable

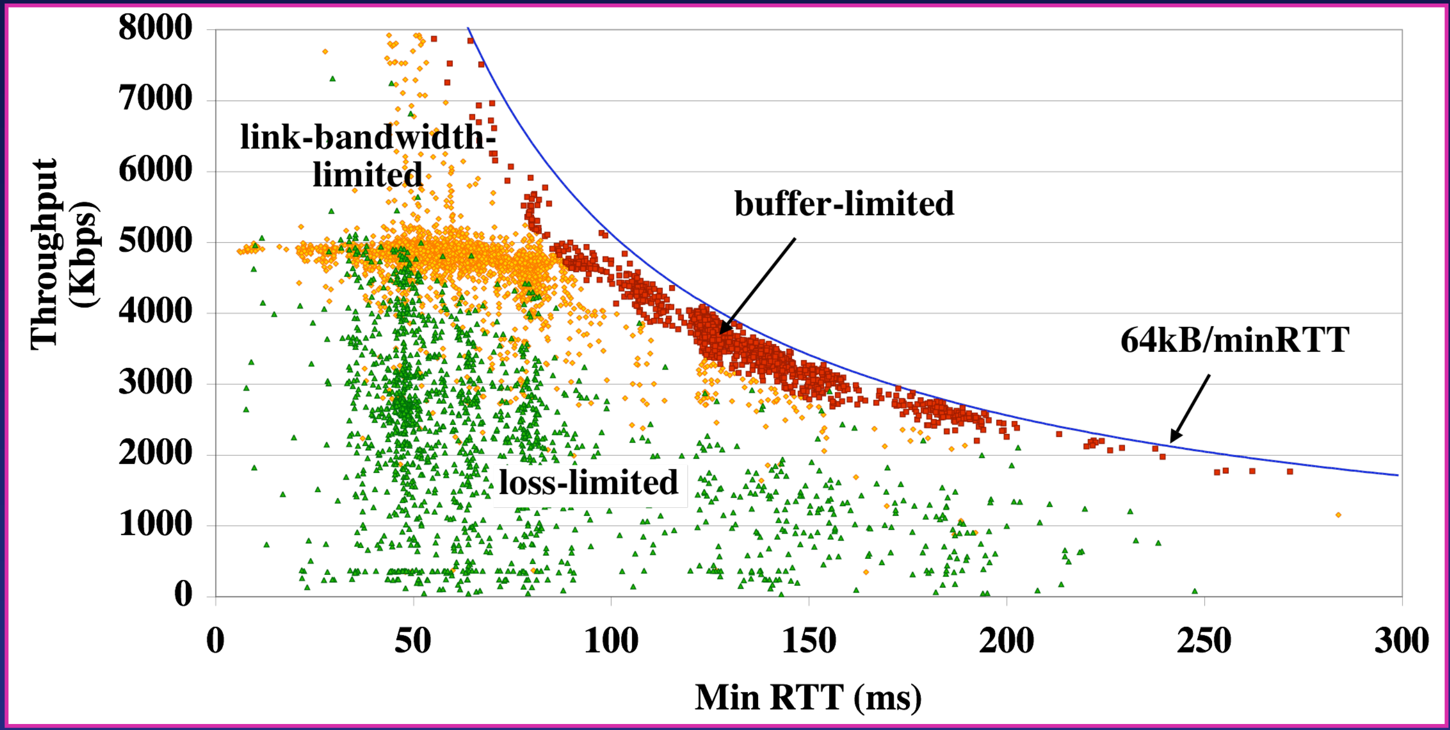
- DSL exhibits discrete throughput levels
- Cable throughputs are higher, but continuous

How does server location affect throughput?



- Server location affects cable more than DSL
  - closer server yields more bandwidth for cable

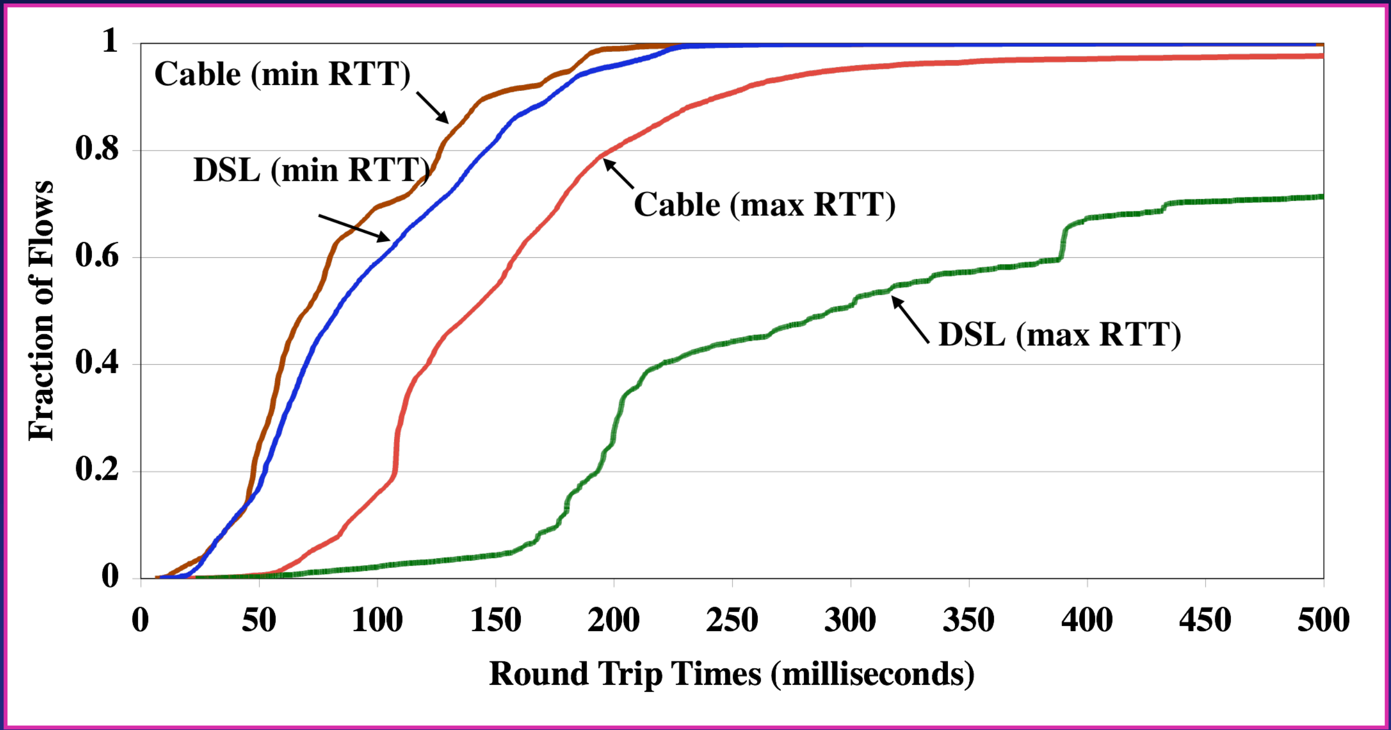
Why do cable throughputs differ from DSL?



3 factors limit throughput: access link capacity, TCP buffer size, loss

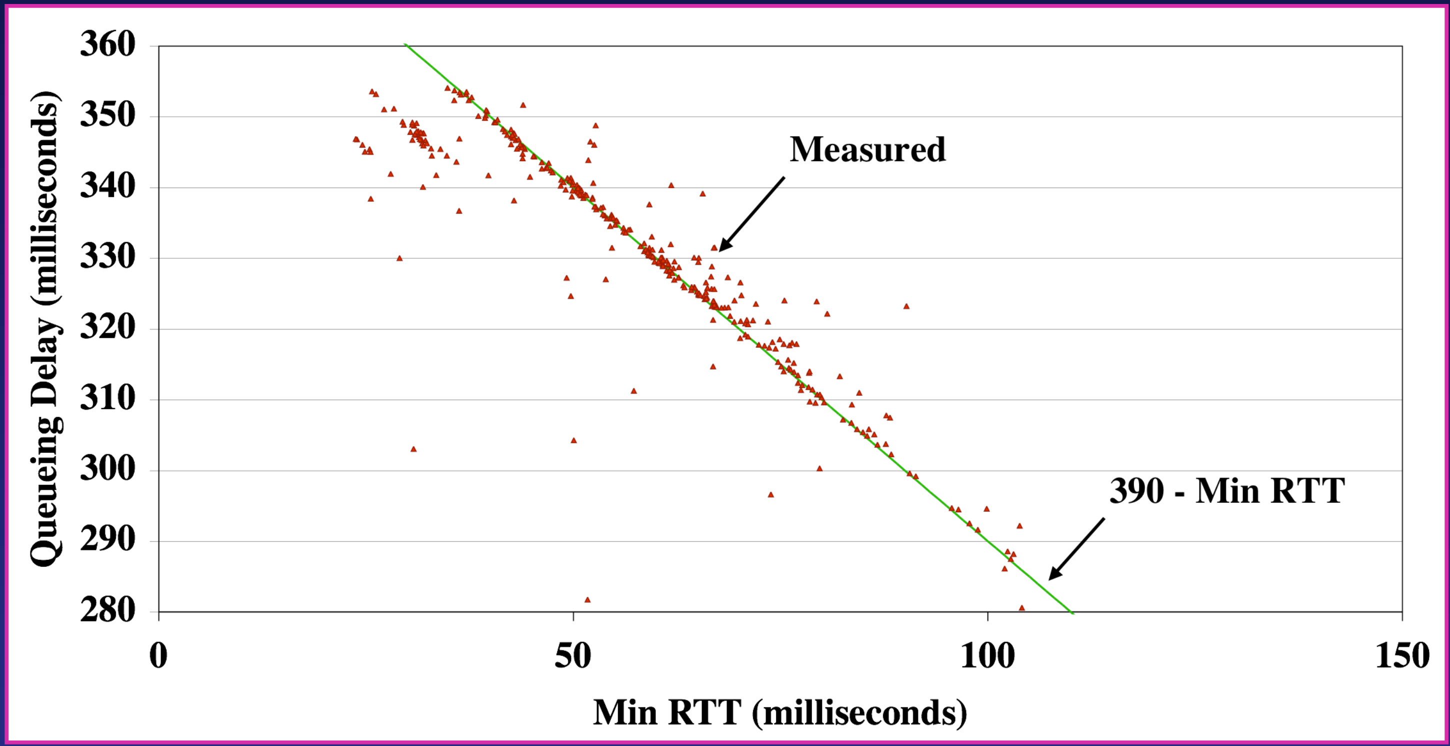
- most DSL flows are limited by access link capacities
- a majority of cable flows are limited by link capacities
- a significant fraction of cable hosts are affected by TCP buffer size & loss

How do queueing delays vary across providers?



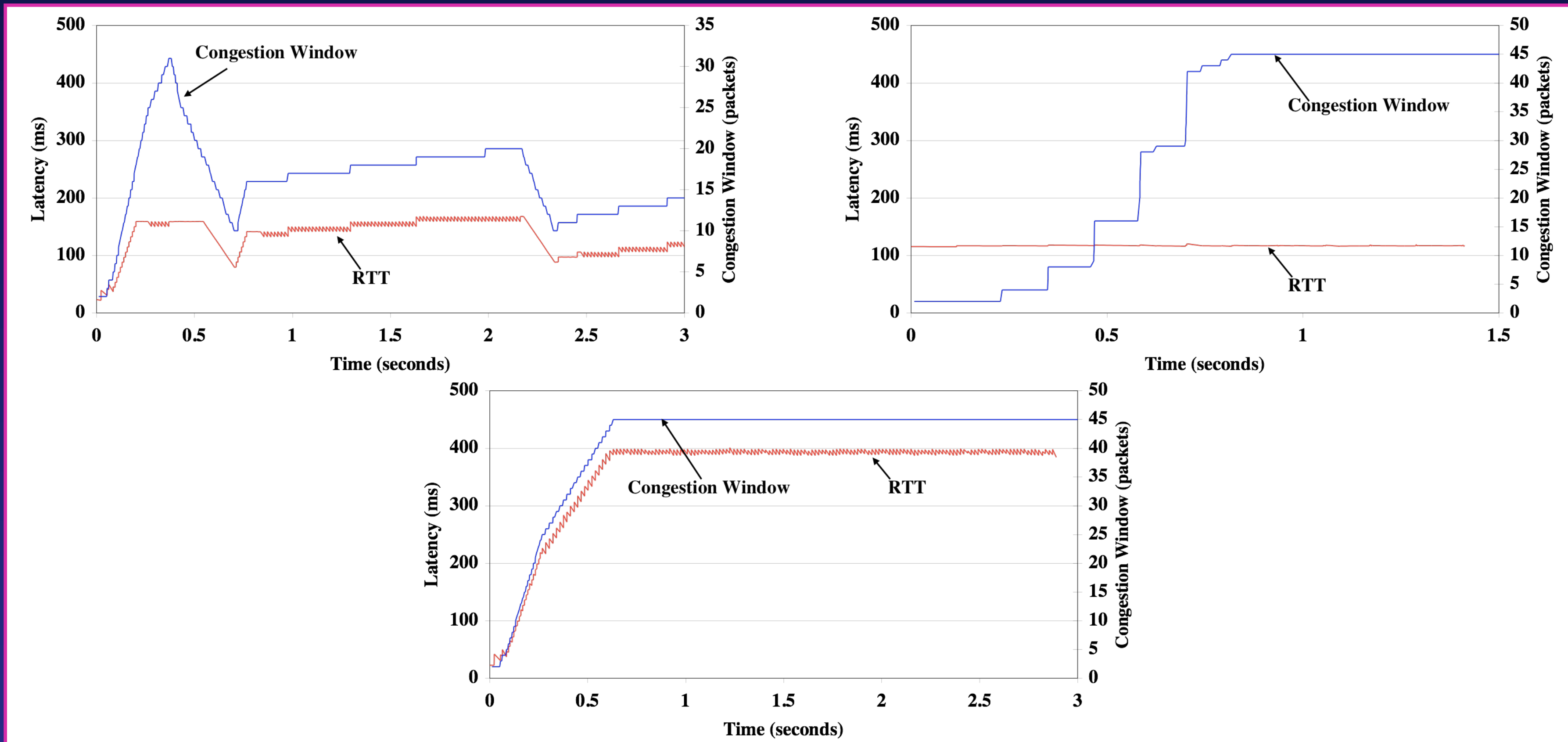
- Queueing delays are several times larger than end-to-end RTTs!
- DSL sees significantly higher queueing delays than cable

How does server location affect queueing delay?



- For some ISPs, queueing delays decrease with path RTT!
  - distant server is better than a closer server

What factors affect queueing delay?

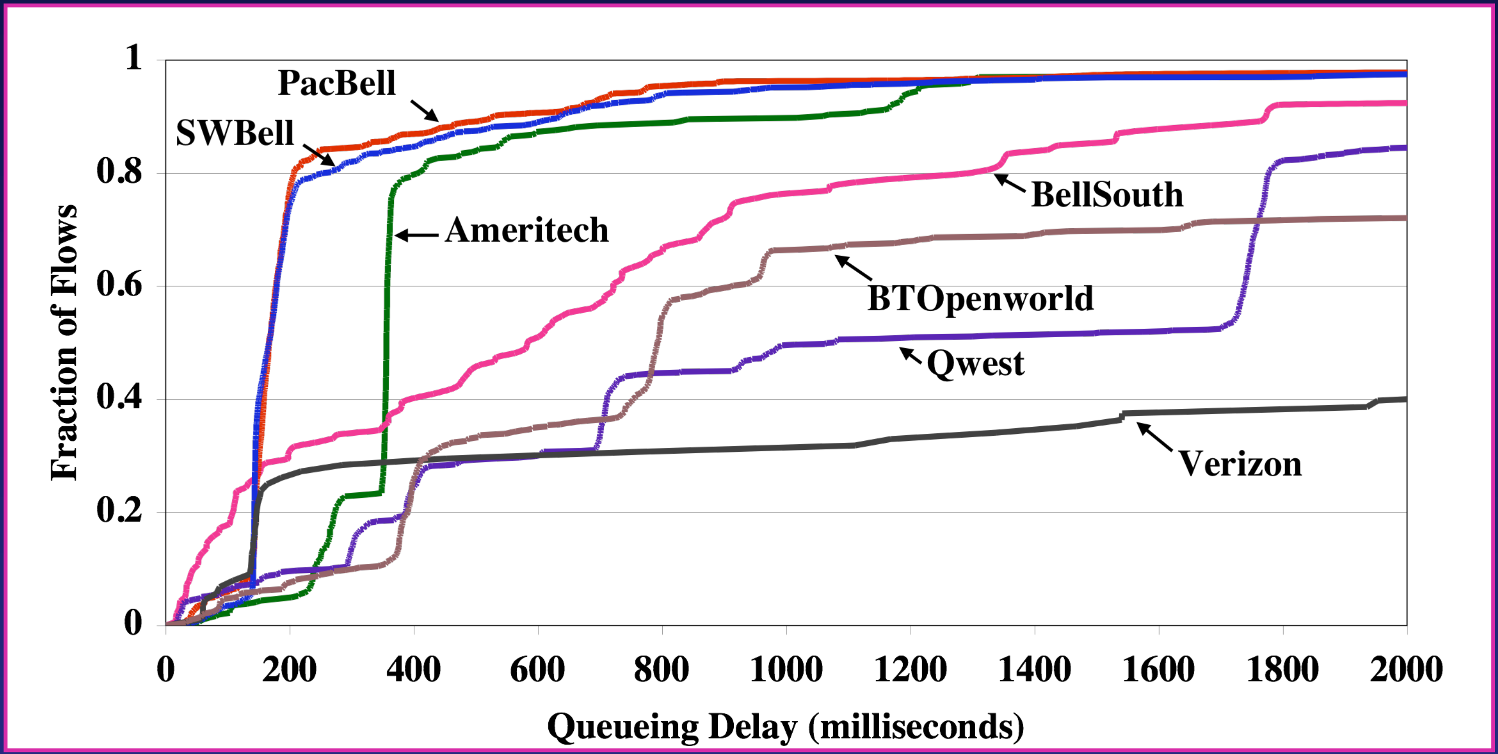


- Queueing delay depends on the factor limiting throughput:
  - loss limited -- small, variable queueing delay
  - send buffer -- no queueing delay
  - link capacity -- large, constant queueing delay (decreases with RTT)

Summary

- A majority of broadband flows are limited by access link capacity
  - unlike academic networks
- Consequently, broadband TCP flows suffer from large, constant queueing delays
- Next, re-examine widely prevalent design practices:
  - ISPs/router designers deploying larger router queues
  - TCP designers using larger send/recv buffers
  - CDNs deploying servers close to end hosts

Should ISPs deploy larger router buffers?



- ISPs already deploy very large queues
  - some deploy “bounded delay” queues > 150 msec
  - some deploy “bounded buffer” queues; at times > 2 secs

- Much larger than RTT of an average flow, as required by rule of thumb

Why does loss affect throughput severely?

- Networked systems well engineered to react to loss
  - ISPs deploy long queues
  - TCP is well engineered to handle loss
- Why is loss still a problem?

	Overall Loss Rate	**Queue < 50ms Loss Rate**
DSL	*1.63 %*	*0.13%*
Cable	*0.62 %*	*0.30%*

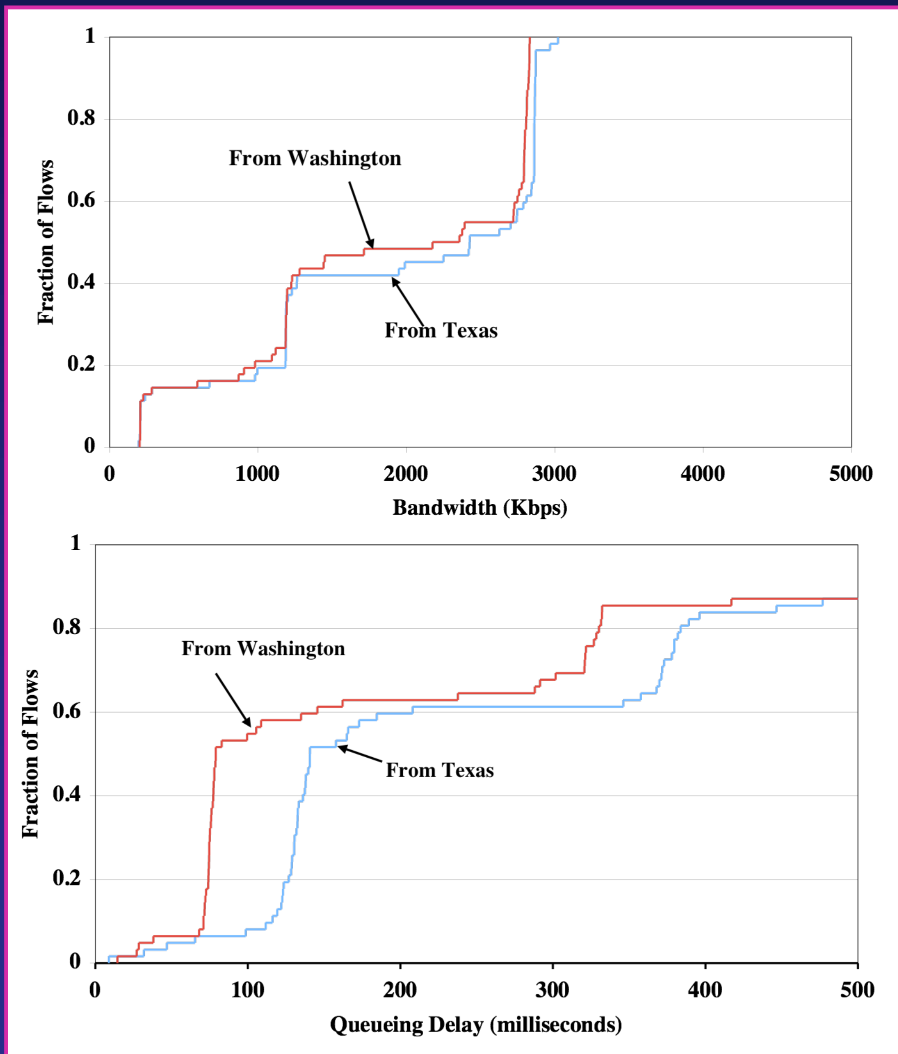
- Heretical question: Are packet losses in the Internet primarily due to queue overflows?

Should we use larger TCP buffers?

Max. vs. Min. RTT Ratio	***64 KB Send Buffer***		***512 KB Send Buffer***	
	50th percentile	*90th percentile*	*50th percentile*	*90th percentile*
DSL	*3.838*	*15.481*	*5.209*	*28.563*
Cable	*1.753*	*4.040*	*2.760*	*7.077*

- RTTs can increase by a factor of 28 with 512KB buffers!
  - without appreciable gain in throughput

How massively should CDNs replicate servers?



- A closer server can lead to larger queueing delays, without any gain in throughput

- Closest server is not always a good choice!

Conclusions

- Internet is always rapidly changing
  - shift in workloads towards bulk data & real-time traffic
  - shift in bottlenecks from network edges to access links
- What are the characteristics of deployed networks/apps?
  - TCP congestion control is ill-suited for broadband networks
- What are the implications for future networks/systems designs?
  - smaller router queues
  - latency-sensitive & receiver-driven congestion control
- How to test new designs in practice?
  - Monarch allows for testing new TCP implementations in practice