An experimental study of the learnability of congestion control

Anirudh Sivaraman, Keith Winstein, Pratiksha Thaker, Hari Balakrishnan

MIT CSAIL

May 26, 2014

Designing congestion-control protocols today

Designing congestion-control protocols today

- Formulate a mental model of the target network

Designing congestion-control protocols today

- Formulate a mental model of the target network
- Decide on the protocol's goal

Designing congestion-control protocols today

- Formulate a mental model of the target network
- Decide on the protocol's goal
- Design a protocol to achieve this goal

Designing congestion-control protocols today

- Formulate a mental model of the target network
- Decide on the protocol's goal
- Design a protocol to achieve this goal
- Can either be implicit or explicit

But, the model is always wrong!

- Lost throughput due to stochastic loss
- Bufferbloat when queues are incorrectly sized
- Incast in datacenters

Our work

- Can we formalize this design process ...

Our work

- Can we formalize this design process ...
- and quantify the consequences of model mismatch?

Contributions

- Formalize learnability in the context of congestion control
- Use it to answer:
  - Do we need to know the link speed exactly?
  - What is the cost of backwards compatibility?
  - Do we need to know the topology exactly?

Approach

- Specify a *training scenario*.
  - Topology
  - Locations of senders and receiver
  - Application workload
  - Buffer size and queuing discipline
- Specify an *objective function*.
- Synthesize protocol automatically.
- Evaluate on a *testing scenario* inside ns-2

Automated protocol synthesis

- Find best protocol, given an imperfect network model
- The problem is hard to solve in general
- Rely on Remy[1] to produce congestion-control protocols.

[1] KW and Hari Balakrishnan, **TCP ex Machina: Computer-Generated Congestion Control**, *SIGCOMM 2013*

Caveats and non-goals

- Very simple, controlled experiments
- Results could change with better protocol-design tools
- Not trying to understand Remy's internals

Is using Remy reasonable?

- Training scenario:

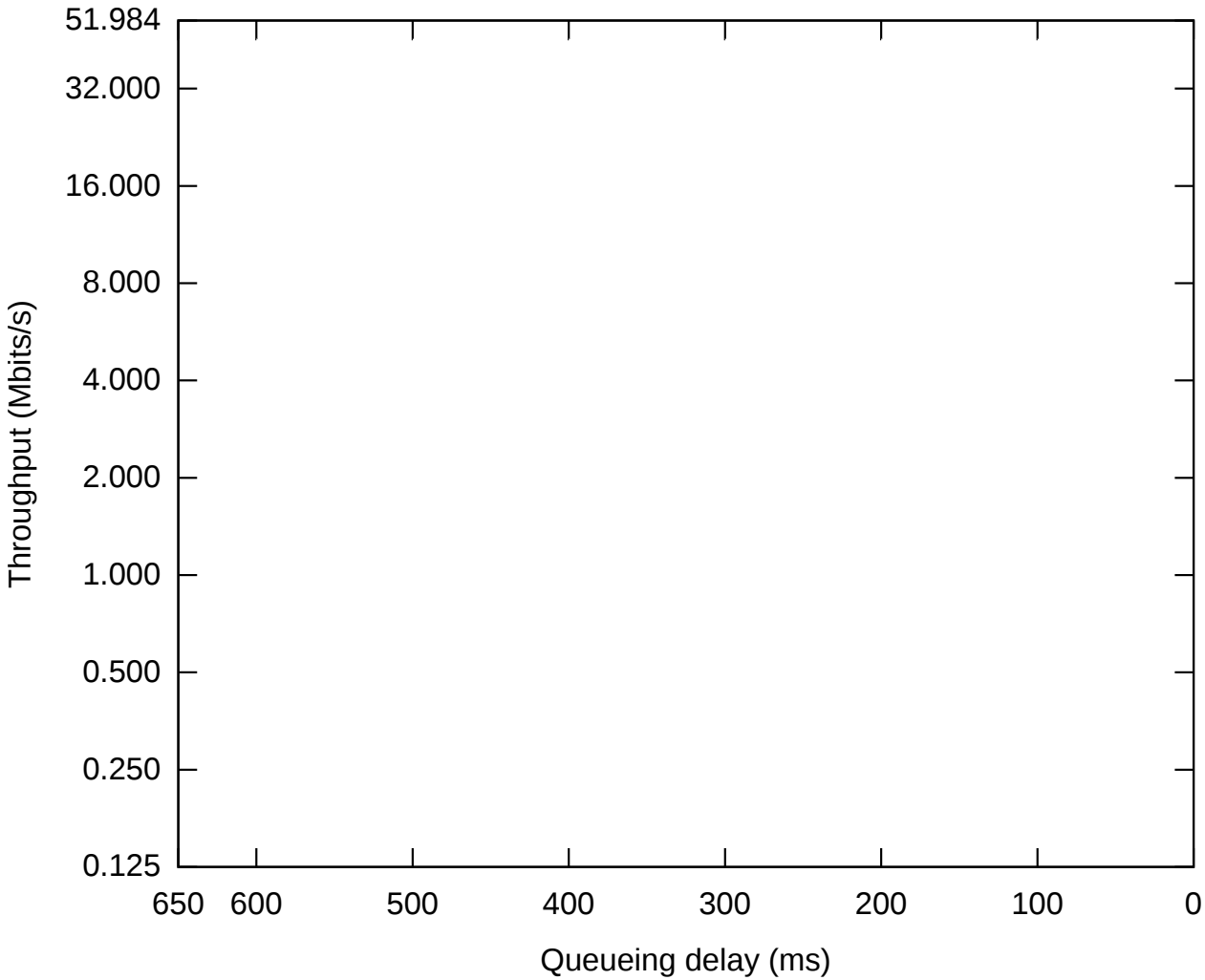
Link speed	32 Mbits/sec
Minimum RTT	150 ms
Topology	Dumbbell
Number of senders	2
Workload	1 sec ON/OFF times
Buffer size	5 BDP
Objective function	$\sum$ log(throughput) - log(delay)

- Testing scenario identical to training scenario

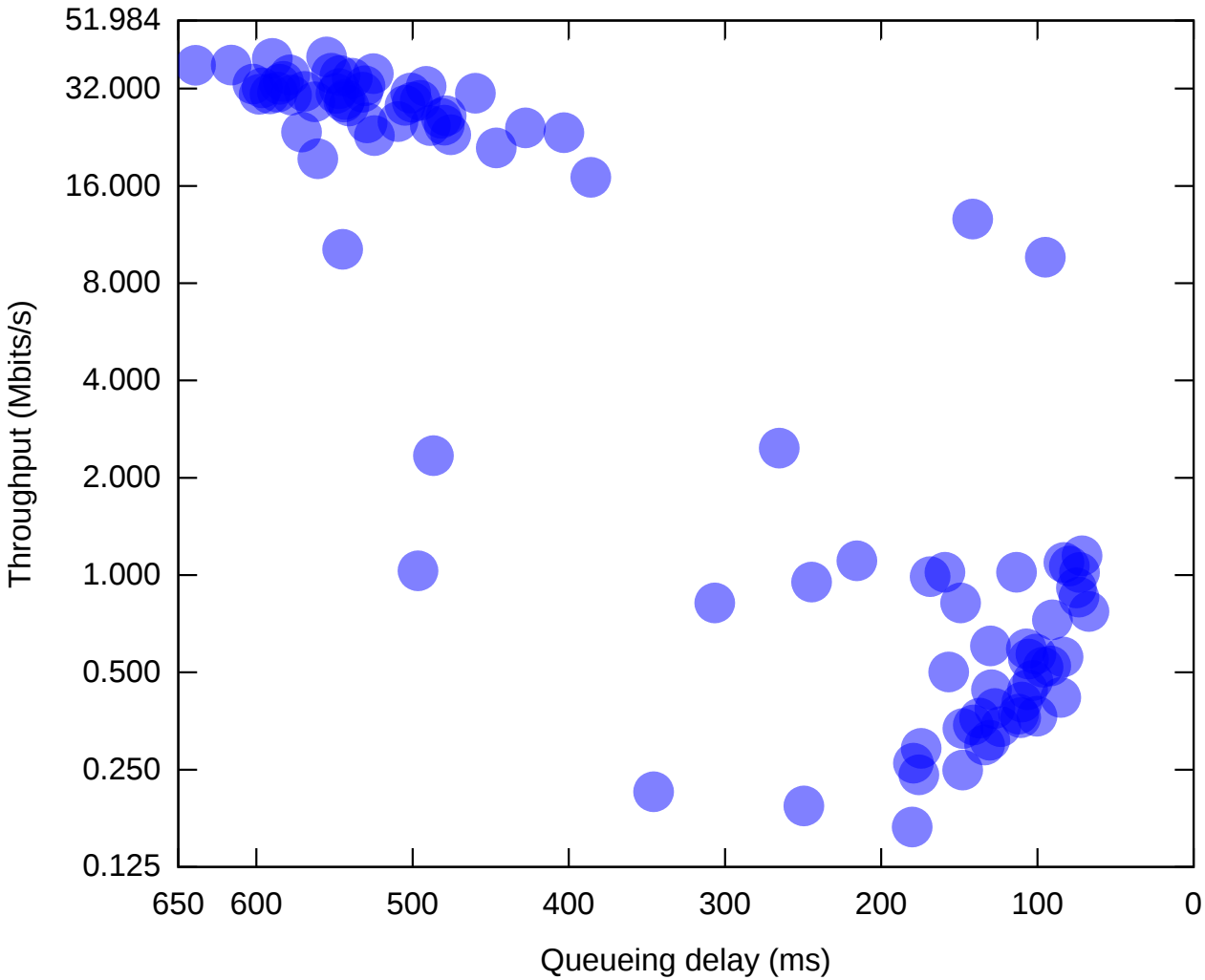
Is using Remy reasonable?

- A hypothetical centralized scheme (CEN)
  - Every time a sender turns ON/OFF, solve $\sum$ log (throughput)
  - Set each sender's rate using obtained solution
  - Zero queuing delay

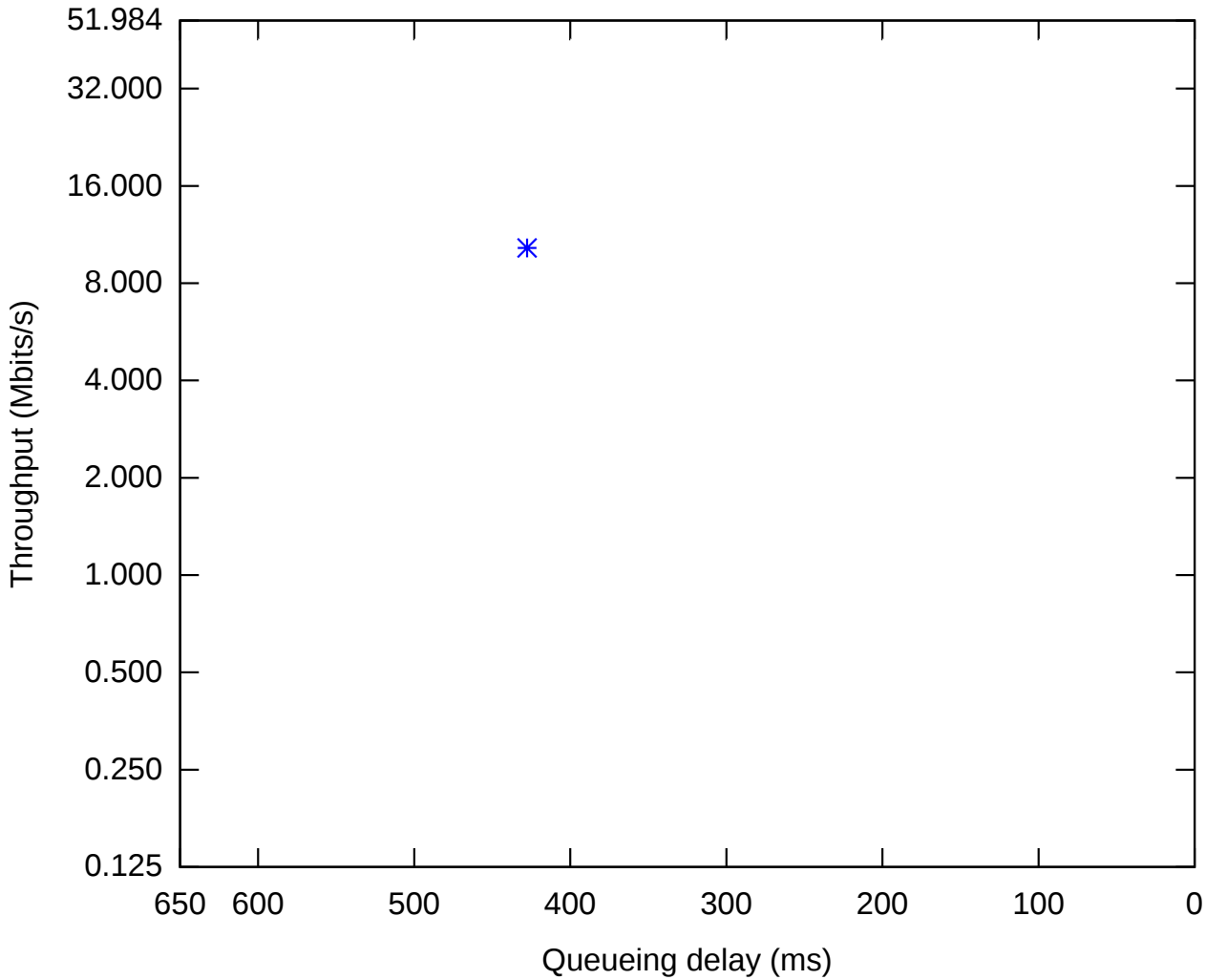
Is using Remy reasonable?



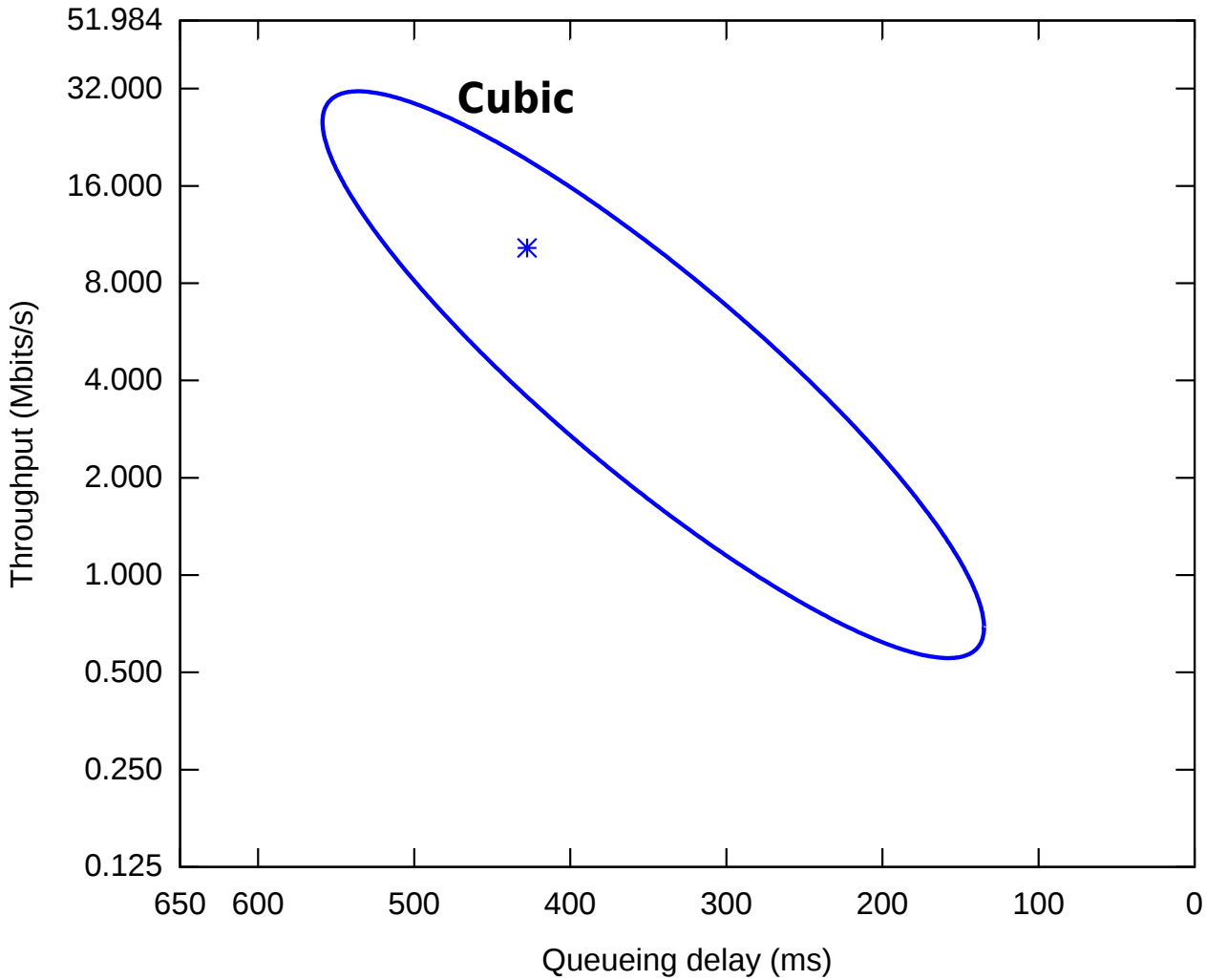
Is using Remy reasonable?



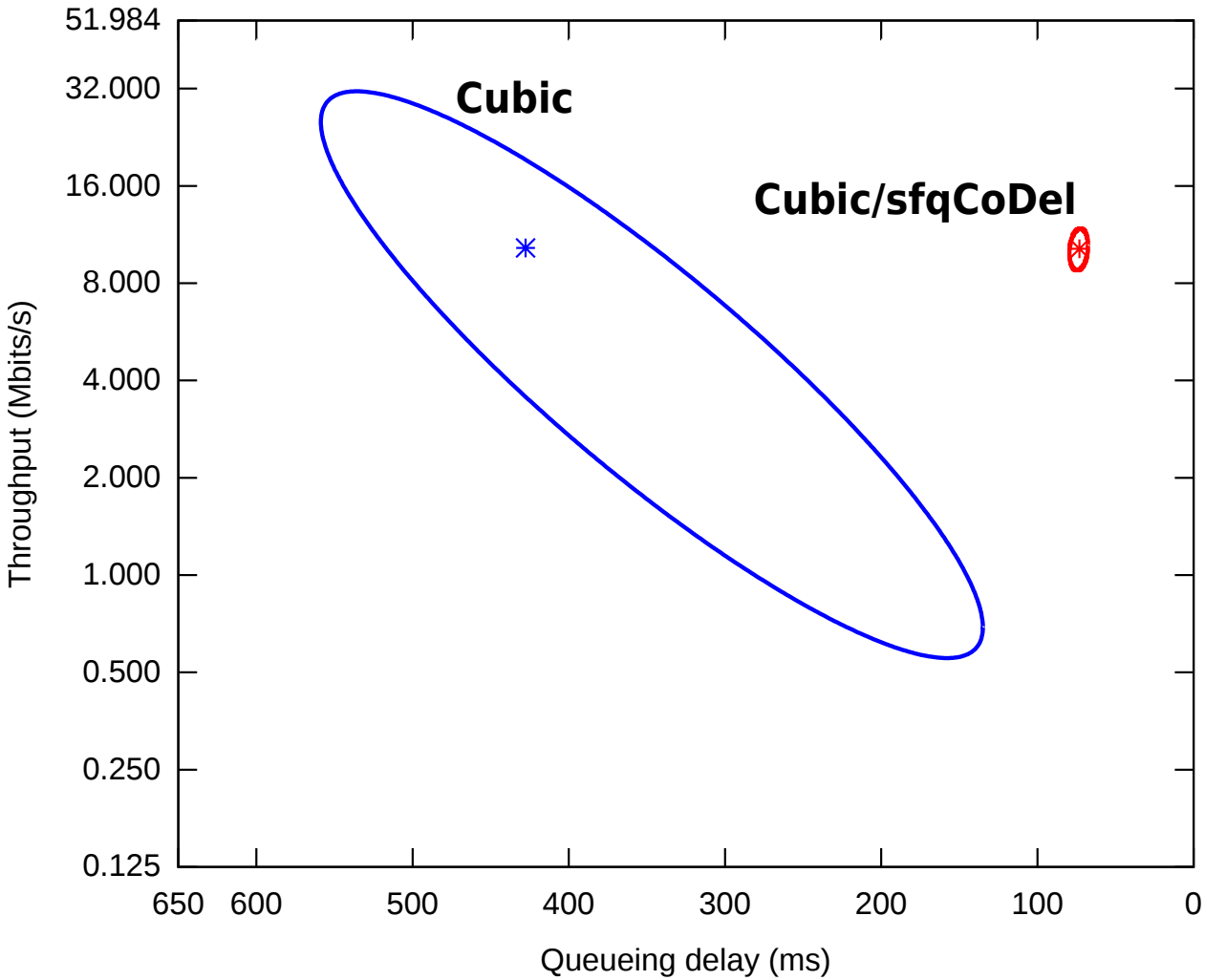
Is using Remy reasonable?



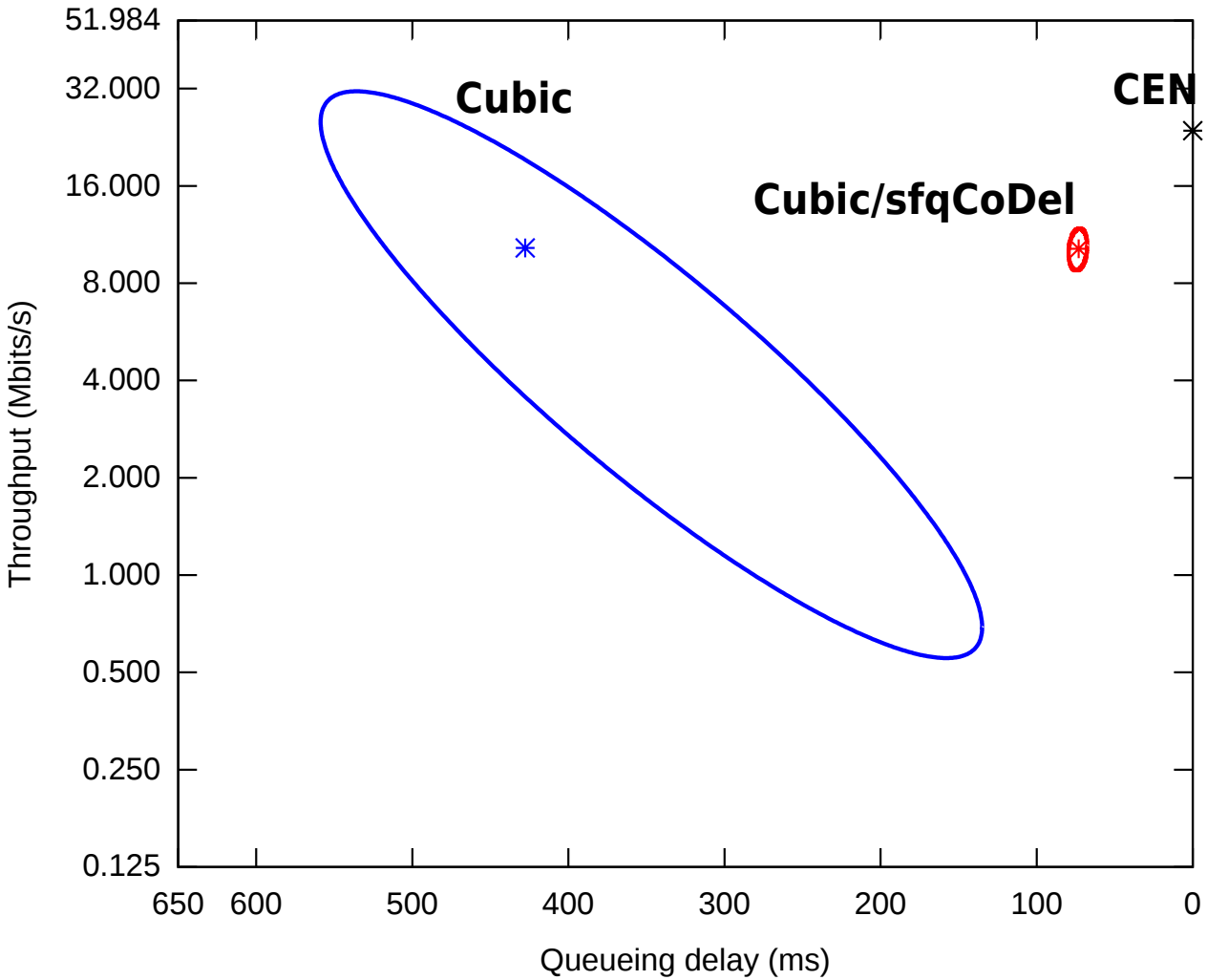
Is using Remy reasonable?



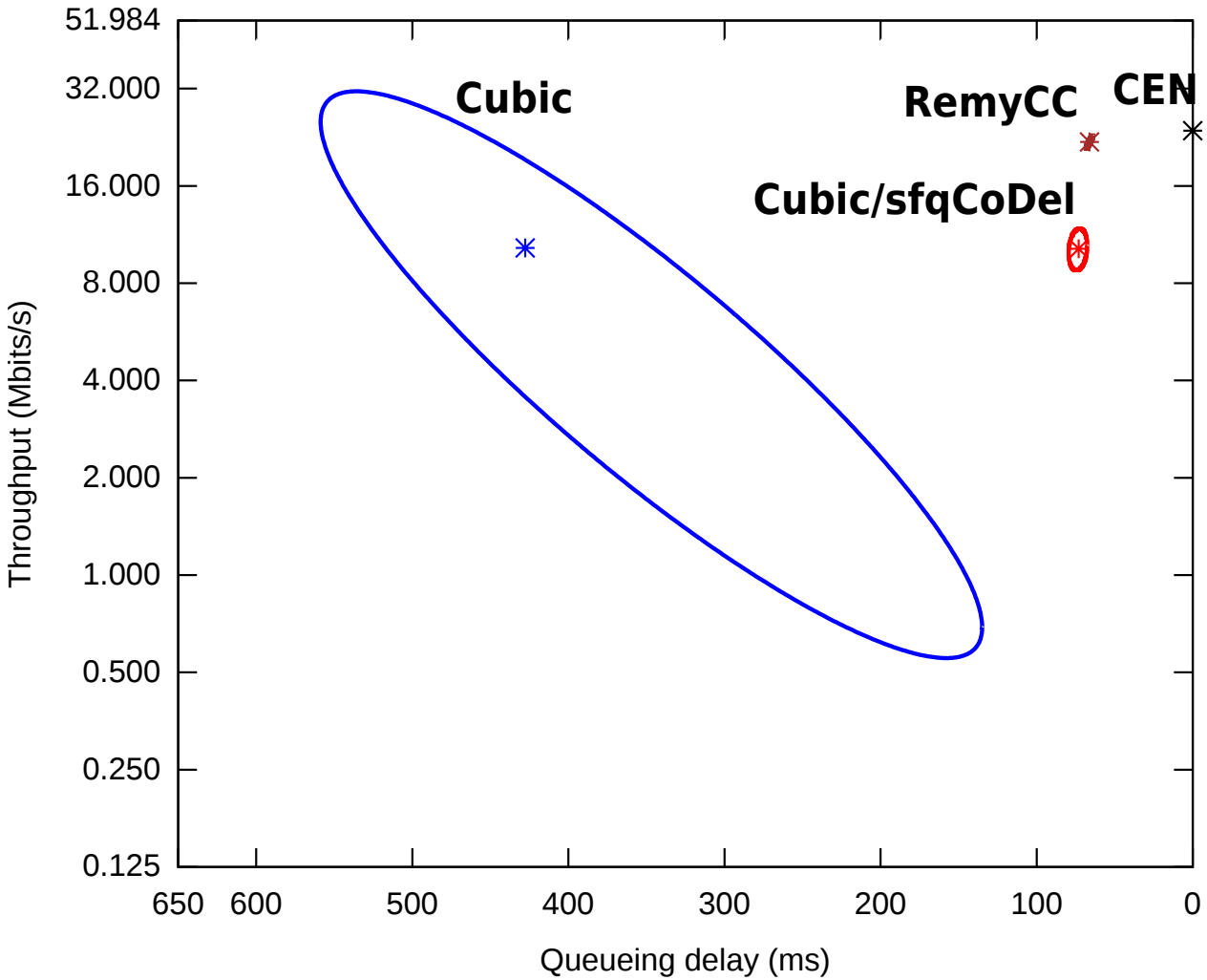
Is using Remy reasonable?



Is using Remy reasonable?



Is using Remy reasonable?



The cost of generality, or forwards-compatibility

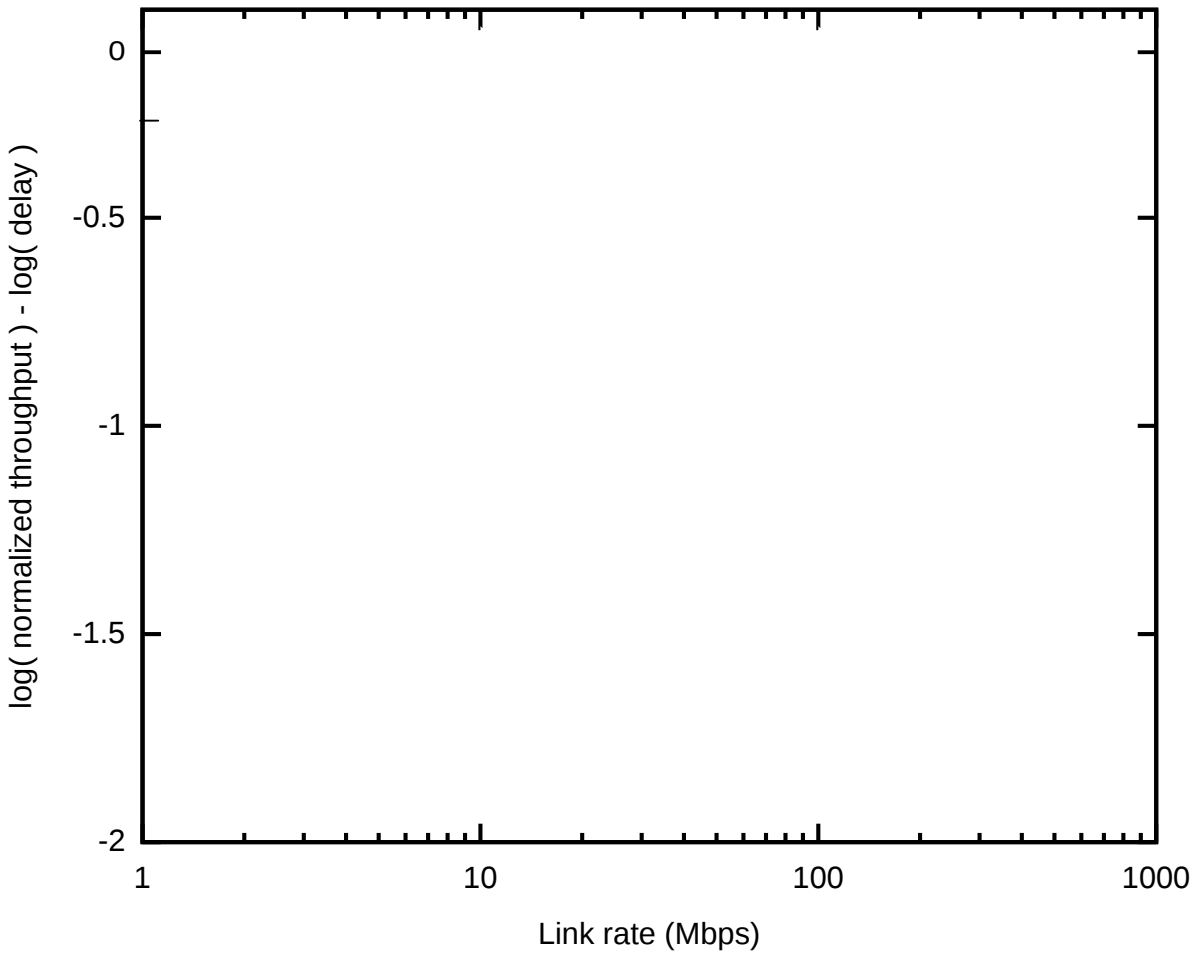
RemyCC	**Link rates**	**RTT**	**Senders**	ON/OFF time	Topology
1000x	1–1000 Mbps	150 ms	2	1 sec	Dumbbell
100x	3.2–320 Mbps	150 ms	2	1 sec	Dumbbell
10x	10–100 Mbps	150 ms	2	1 sec	Dumbbell
2x	22–44 Mbps	150 ms	2	1 sec	Dumbbell

Table : Training scenarios for forwards-compatibility experiment

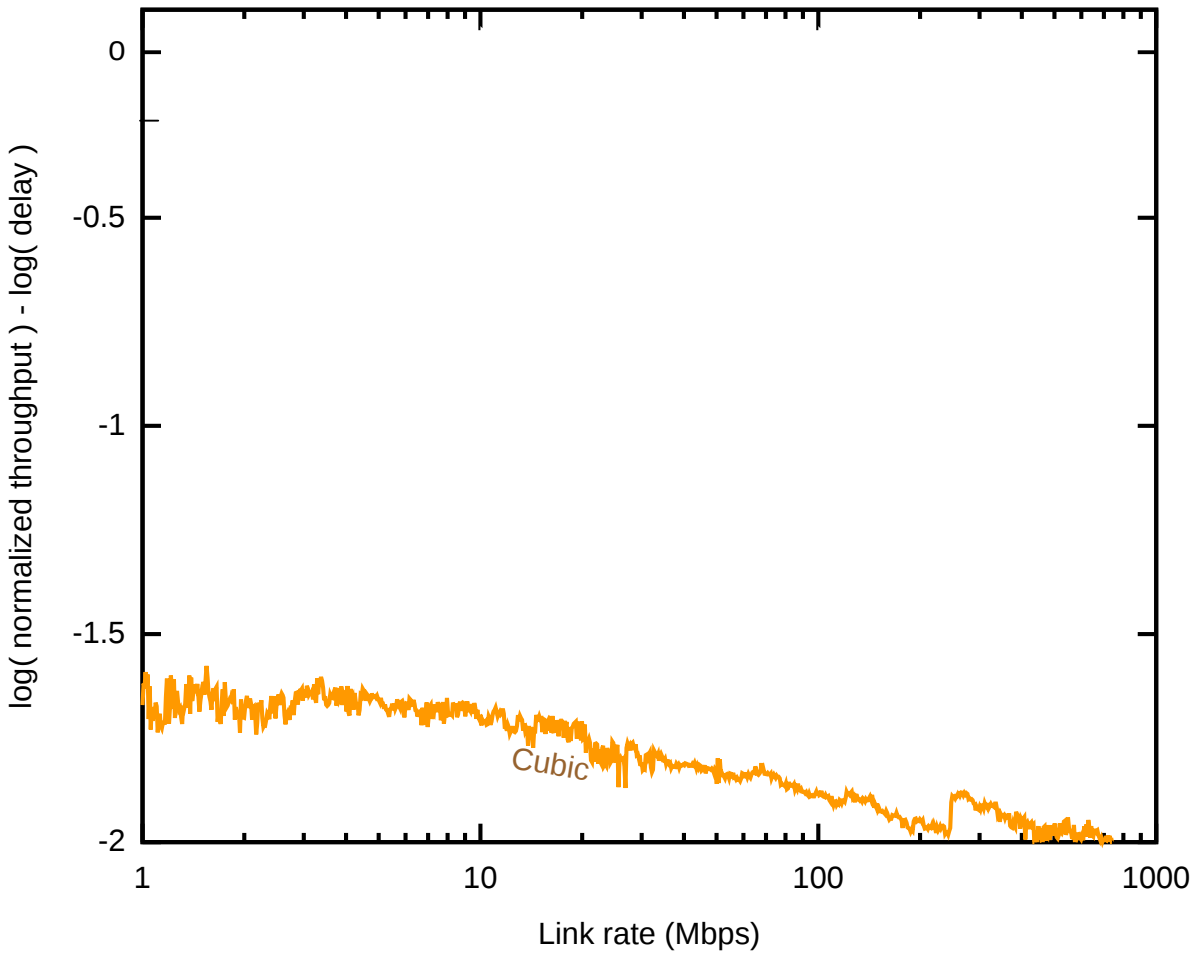
Link rates	**RTT**	**Senders**	ON/OFF time	Topology
1–1000 Mbps	150 ms	2	1 sec	Dumbbell

Table : Testing scenarios for forwards-compatibility experiment

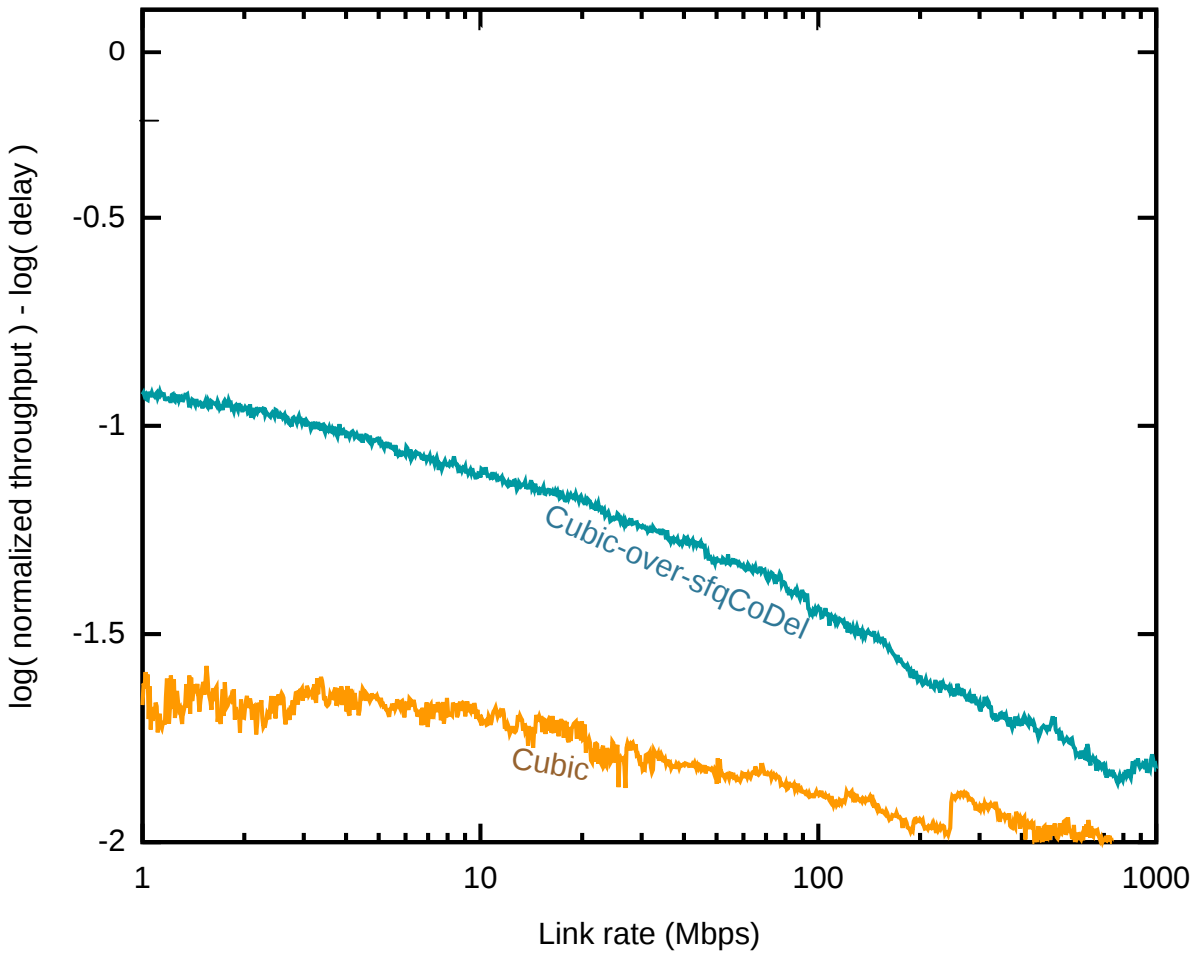
The cost of generality, or forwards-compatibility



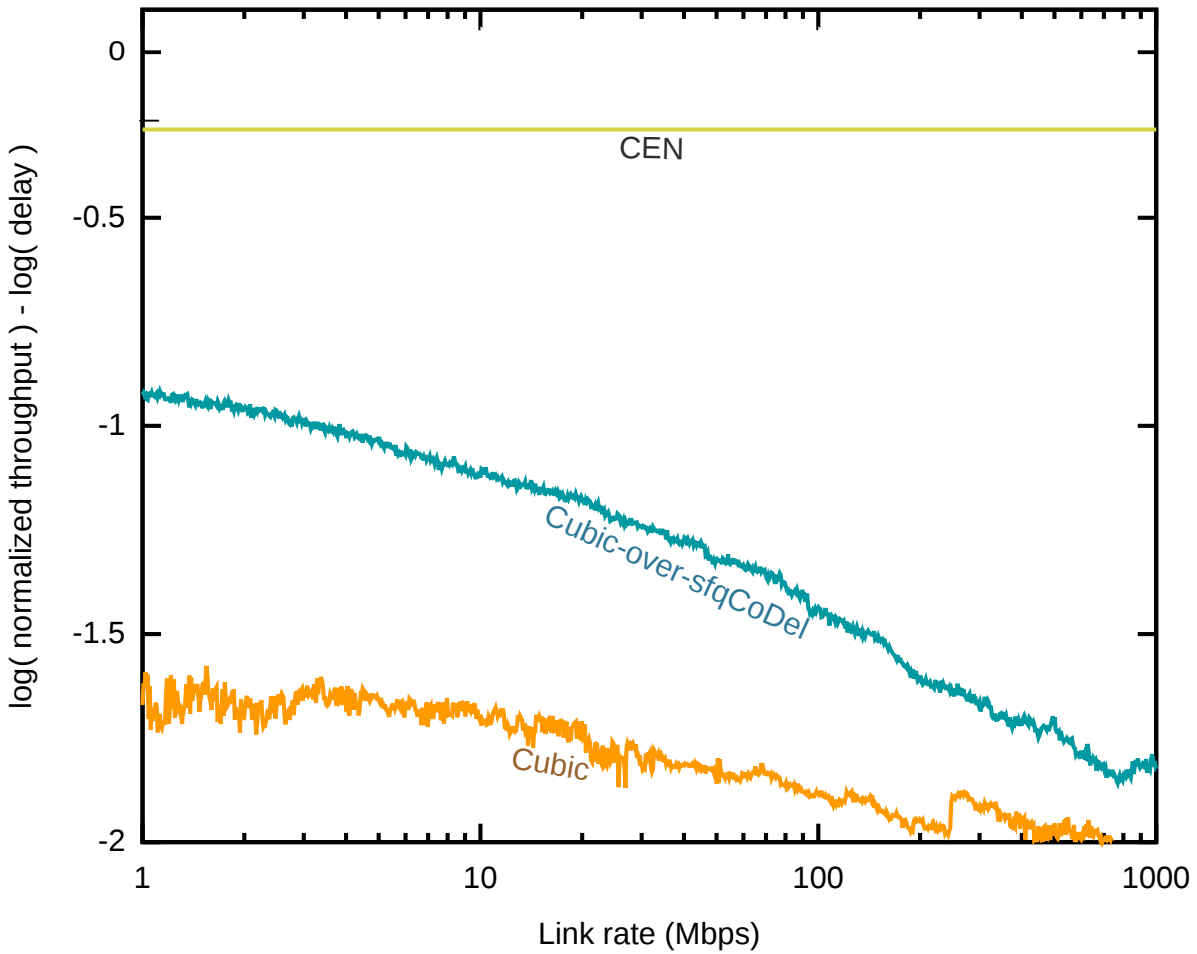
The cost of generality, or forwards-compatibility



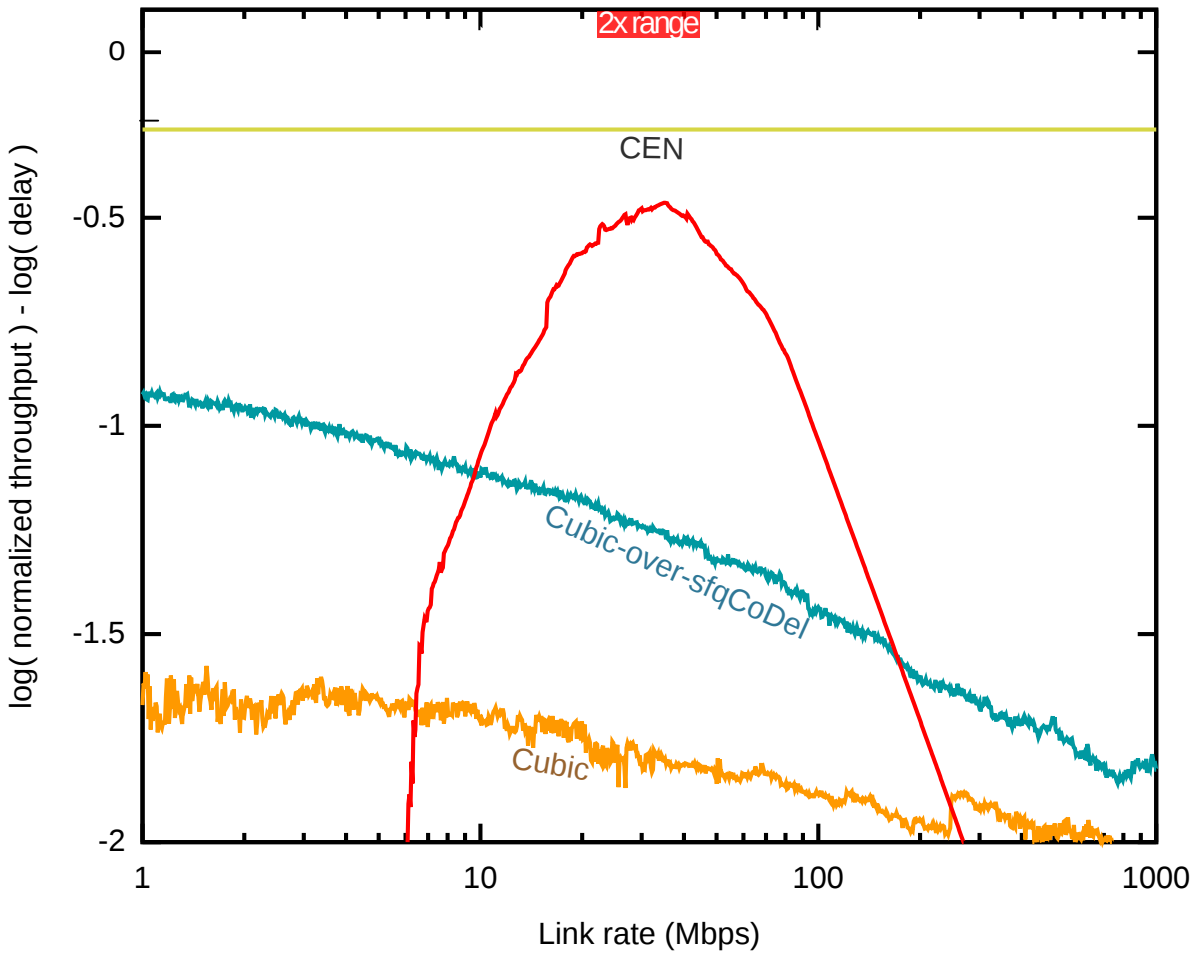
The cost of generality, or forwards-compatibility



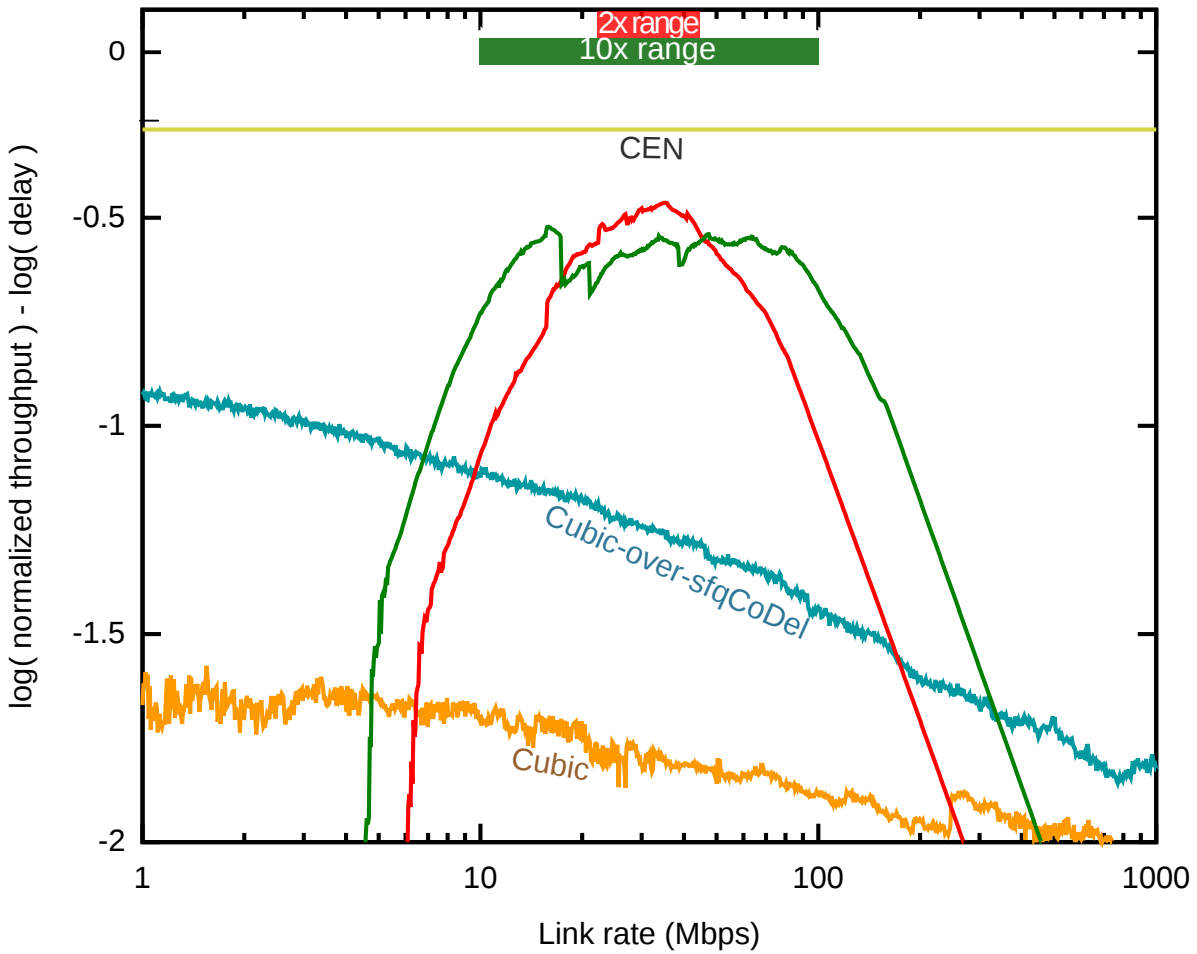
The cost of generality, or forwards-compatibility



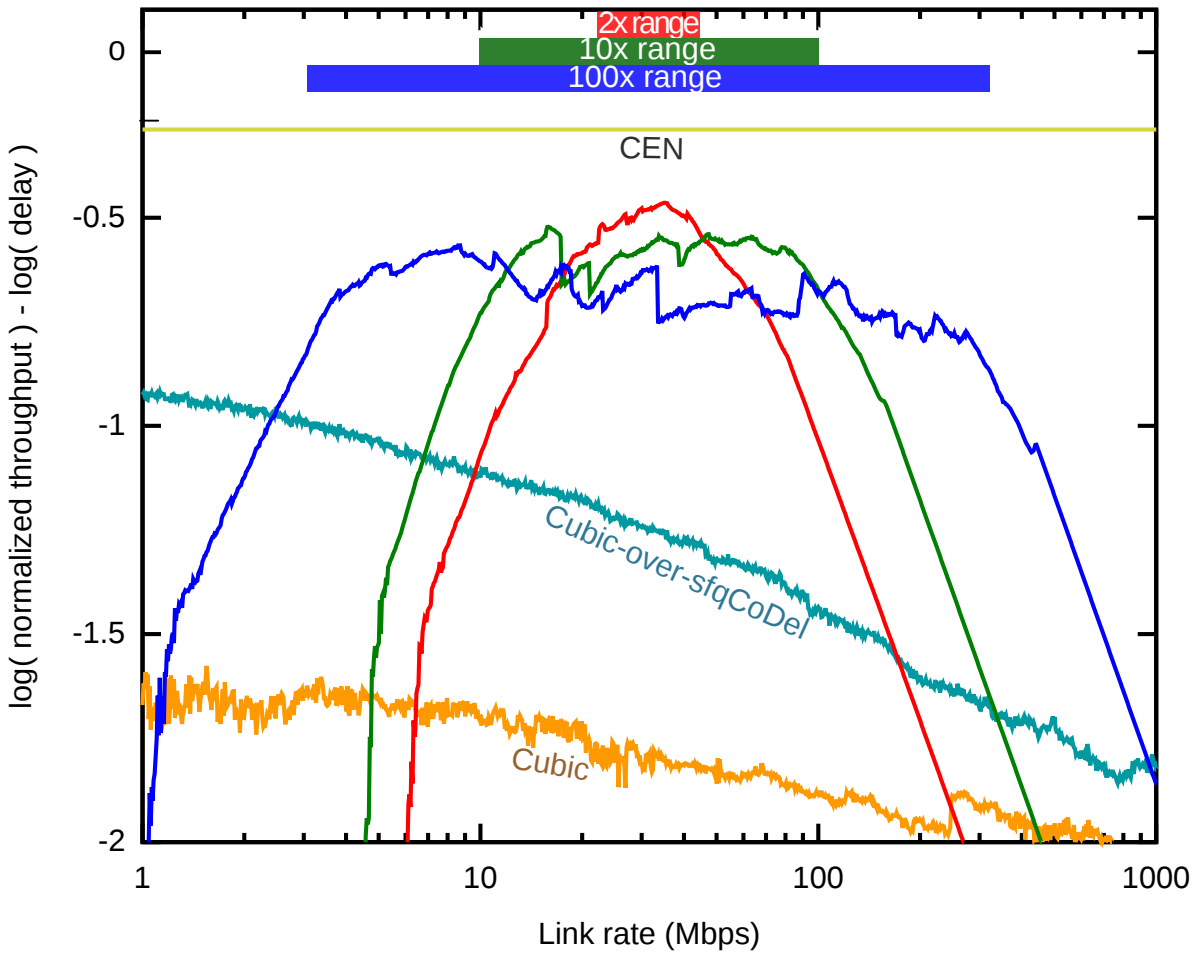
The cost of generality, or forwards-compatibility



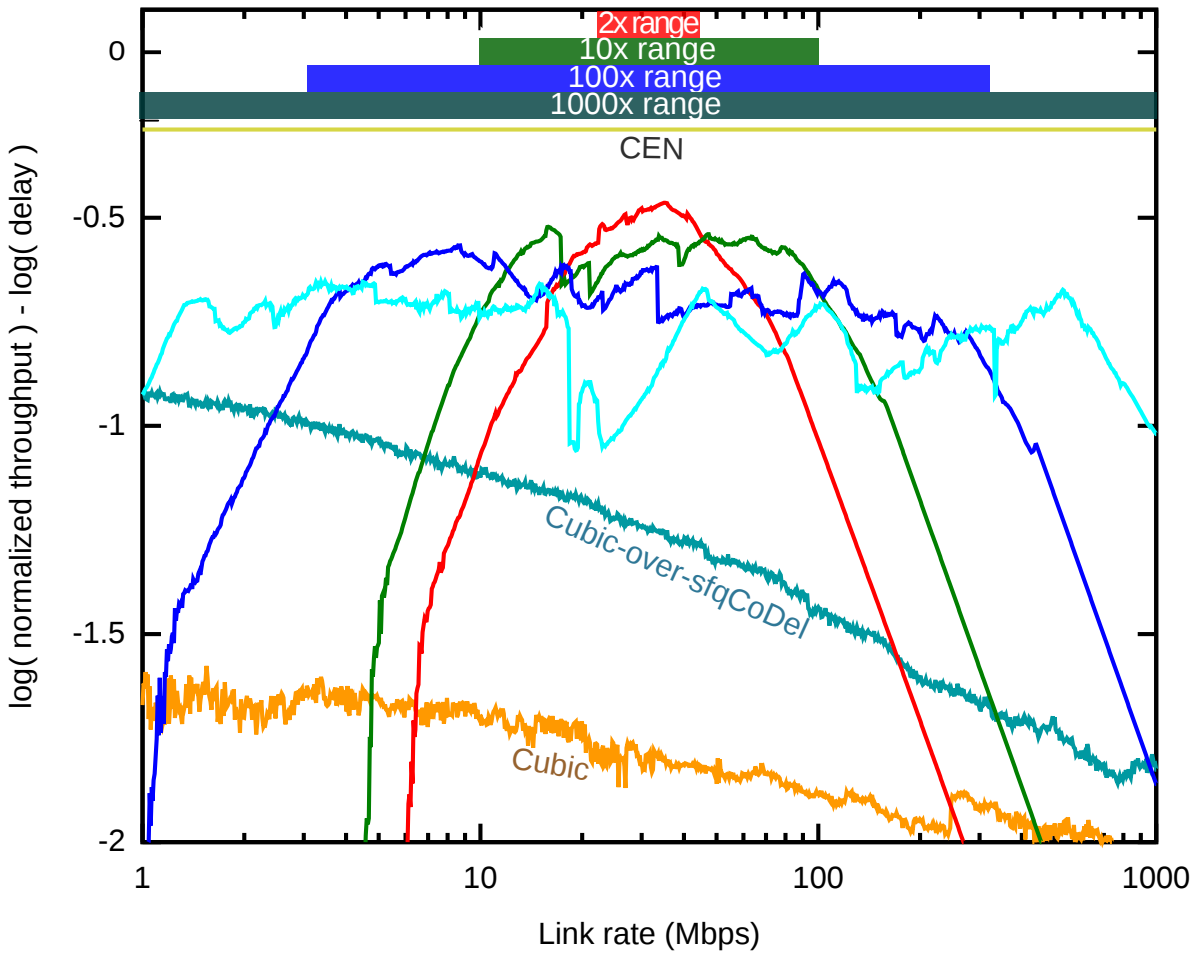
The cost of generality, or forwards-compatibility



The cost of generality, or forwards-compatibility



The cost of generality, or forwards-compatibility



Can we design a RemyCC that is TCP aware?

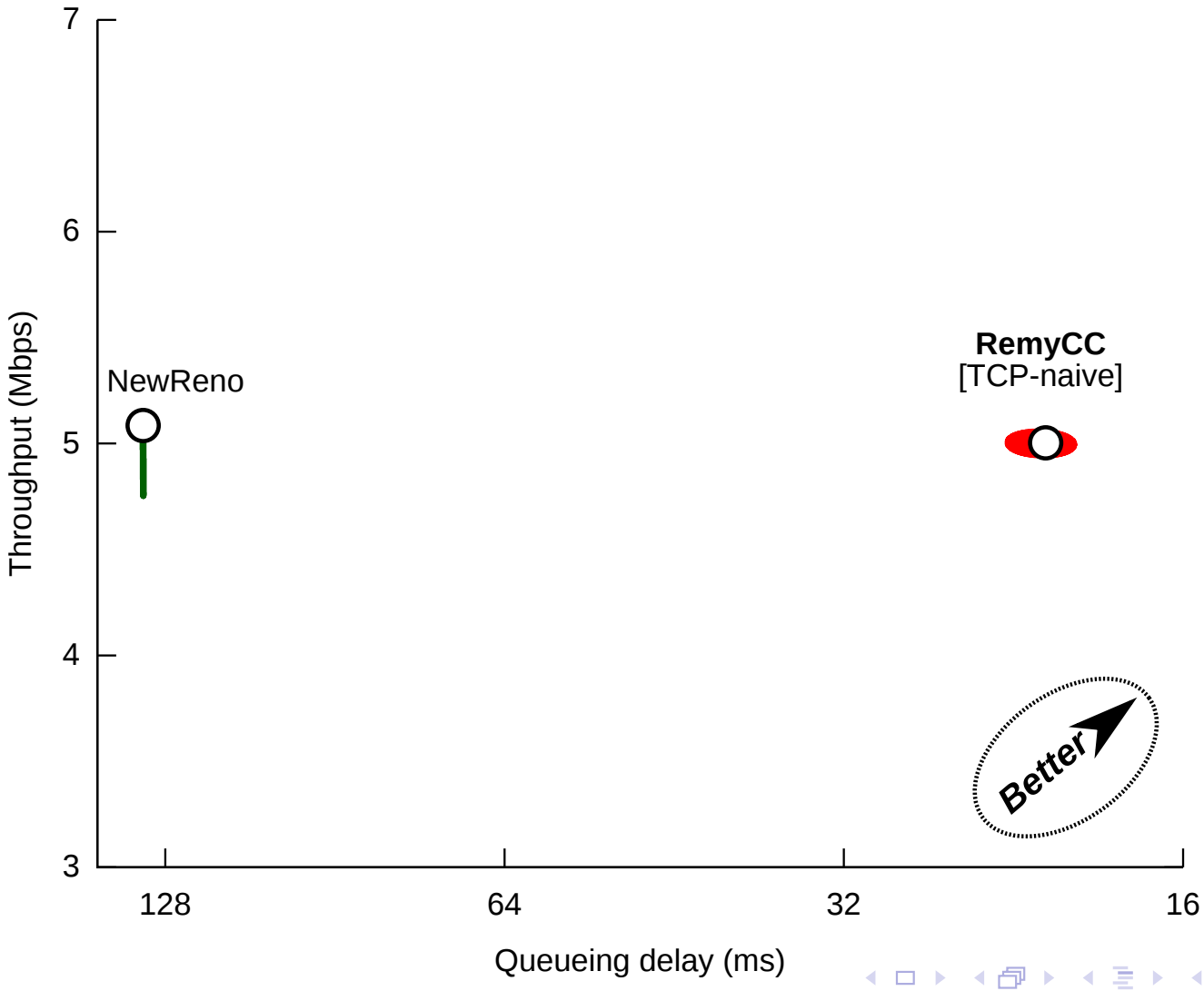
RemyCC	**Link rates**	**RTT**	**Senders**	ON/OFF time	Topology
TCP-aware	9–11 Mbps	100 ms	2 Remy 1 Remy, 1 AIMD	5 sec ON/OFF 5 sec ON, 10 ms OFF	Dumbbell
TCP-naive	9–11 Mbps	100 ms	2 Remy	5 sec ON/OFF 5 sec ON, 10 ms OFF	Dumbbell

Table : Training scenarios

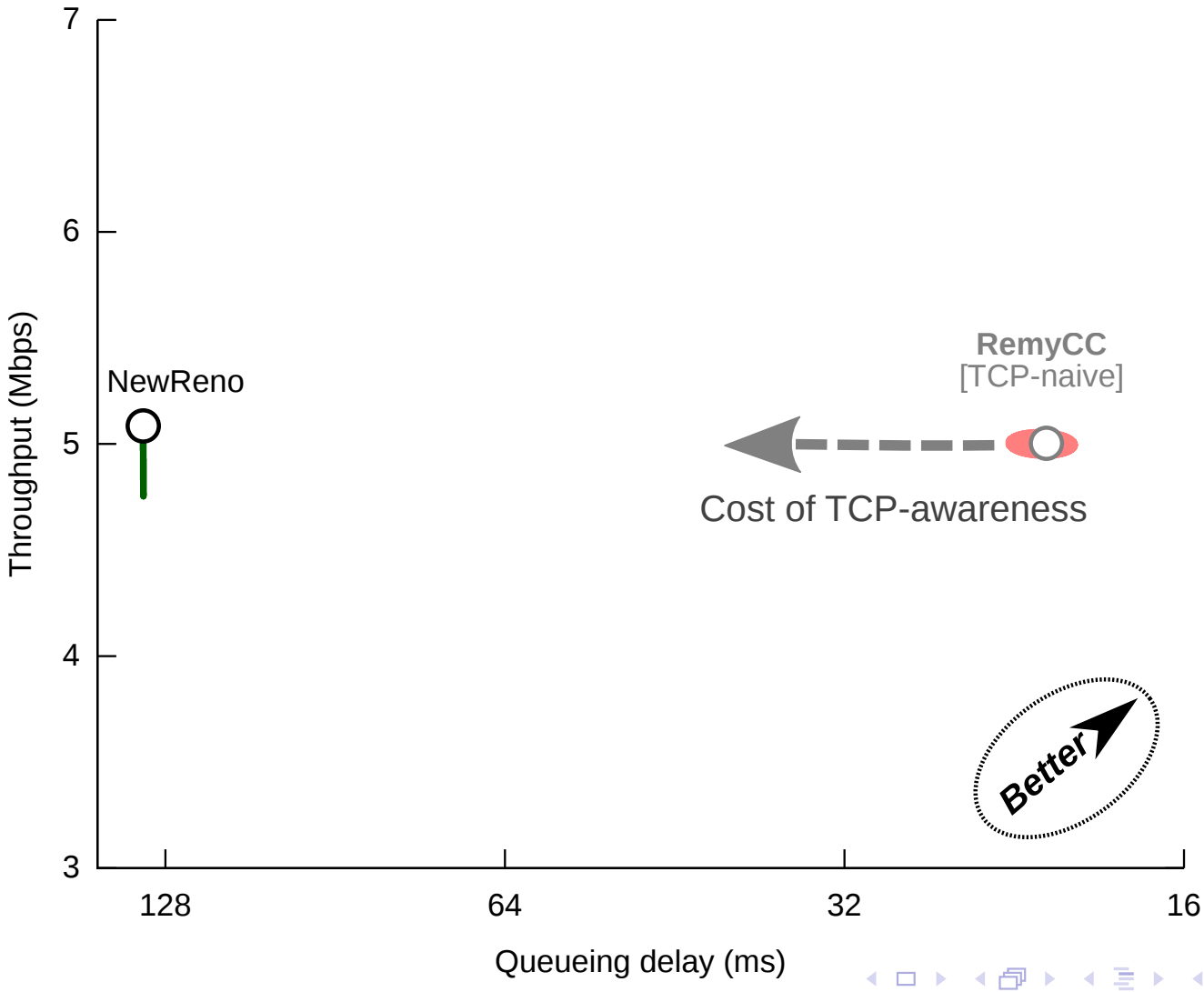
Link rates	**RTT**	**Senders**	ON/OFF time	Topology
10 Mbps	100 ms	2 TCP-aware	5 sec ON, 10 ms OFF	Dumbbell
10 Mbps	100 ms	2 TCP-naive	5 sec ON, 10 ms OFF	Dumbbell
10 Mbps	100 ms	TCP-aware, AIMD	5 sec ON, 10 ms OFF	Dumbbell
10 Mbps	100 ms	TCP-naive, AIMD	5 sec ON, 10 ms OFF	Dumbbell
10 Mbps	100 ms	2 AIMD	5 sec ON, 10 ms OFF	Dumbbell

Table : Testing scenarios

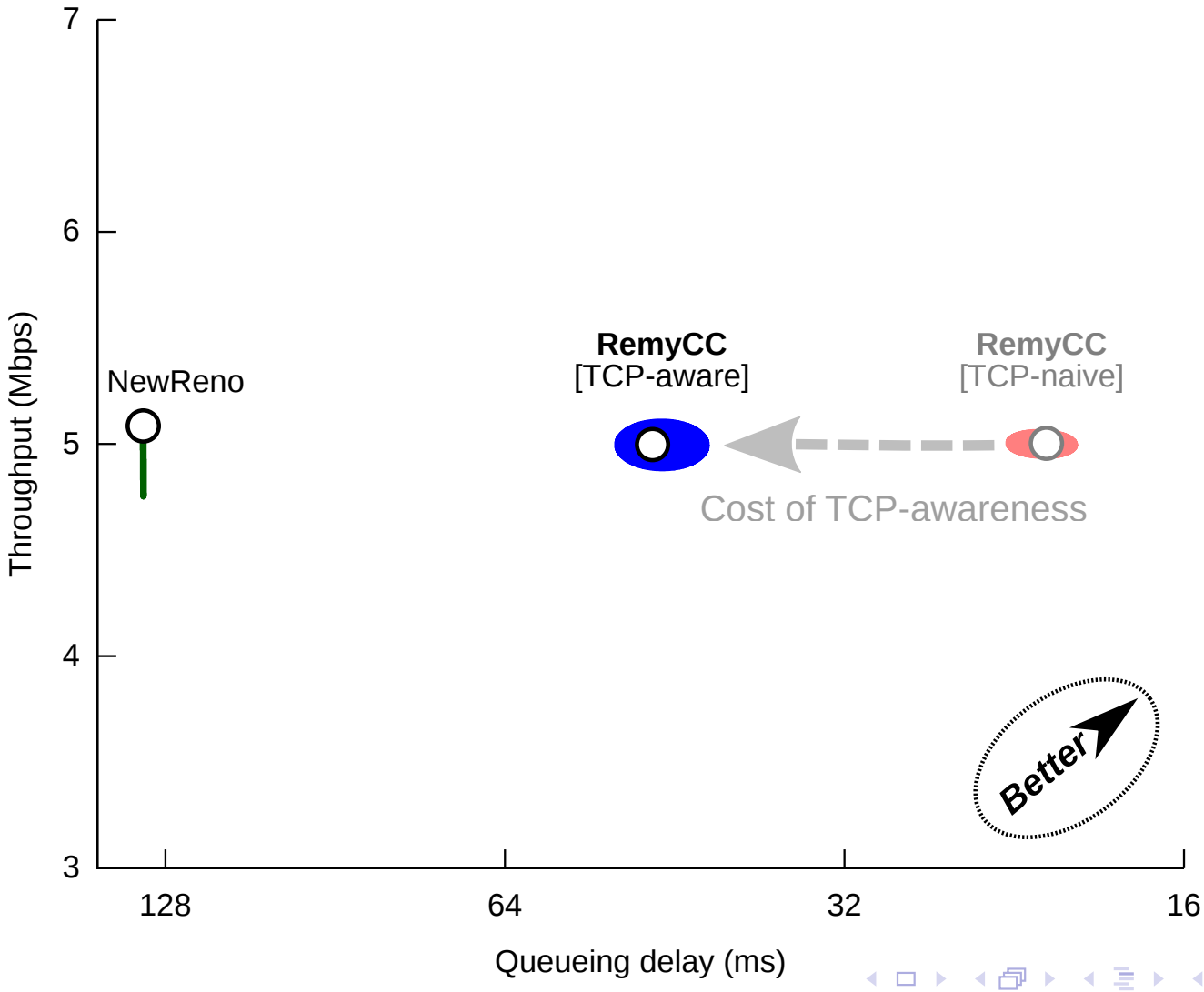
RemyCC competing against itself



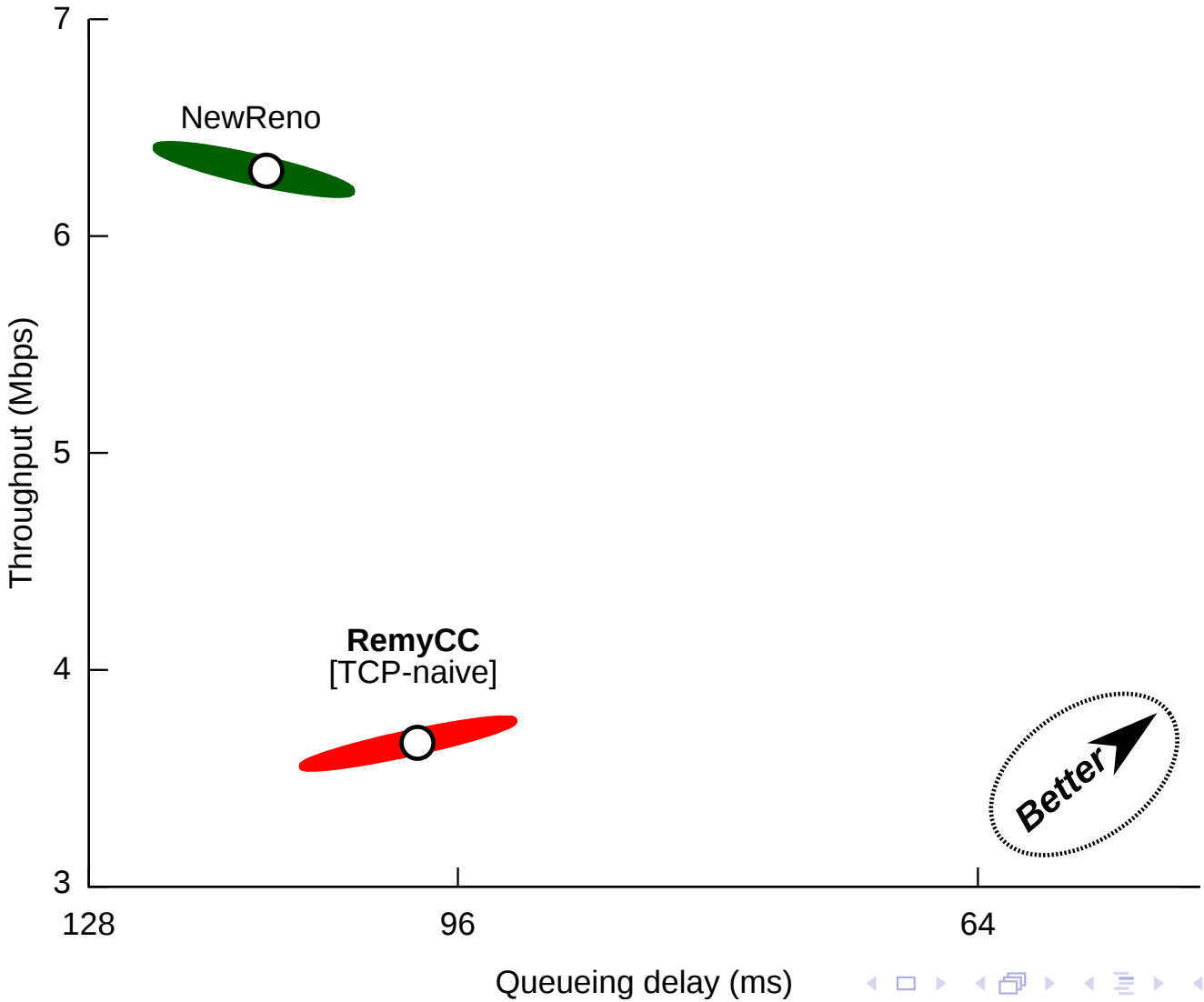
RemyCC competing against itself



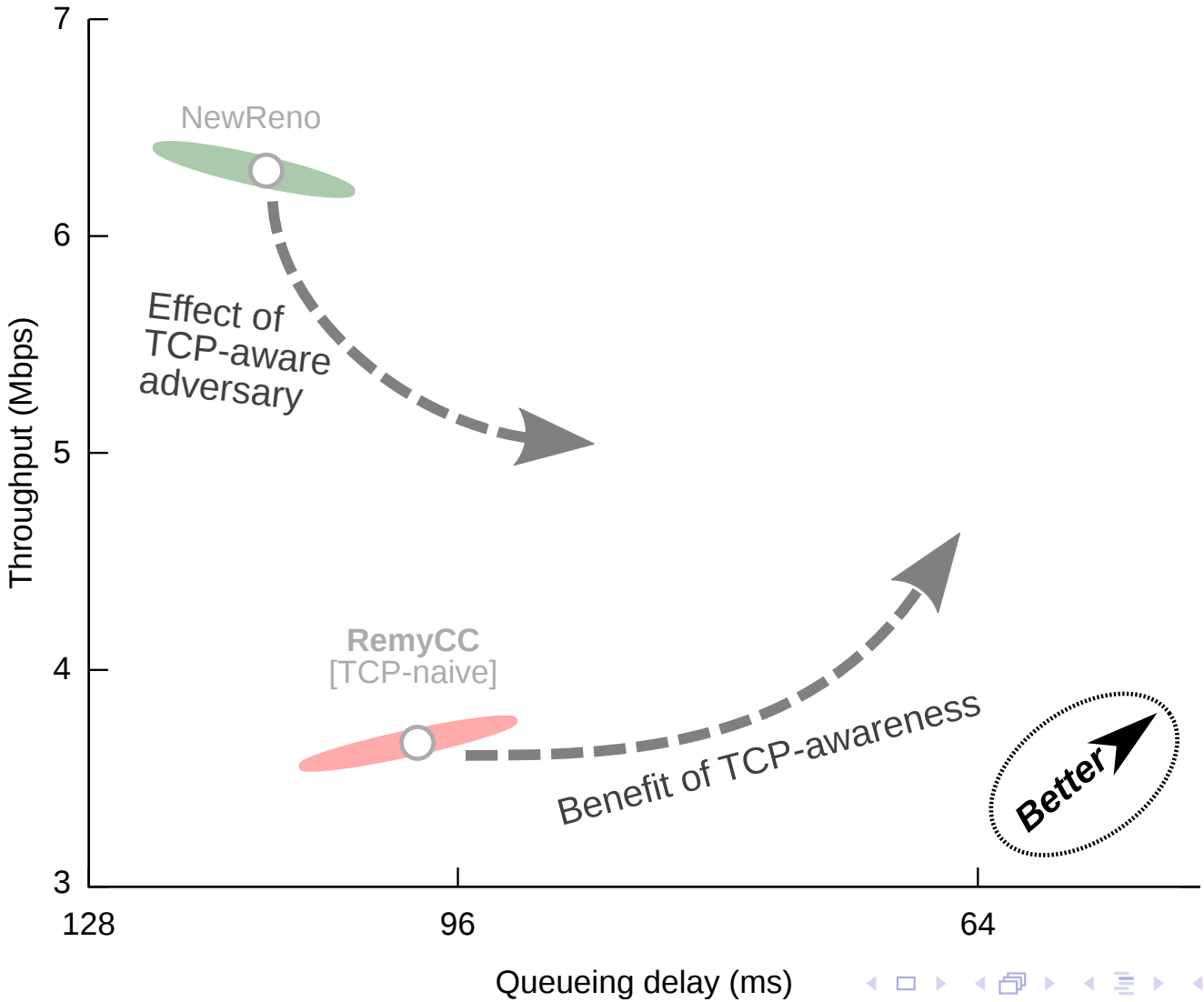
RemyCC competing against itself



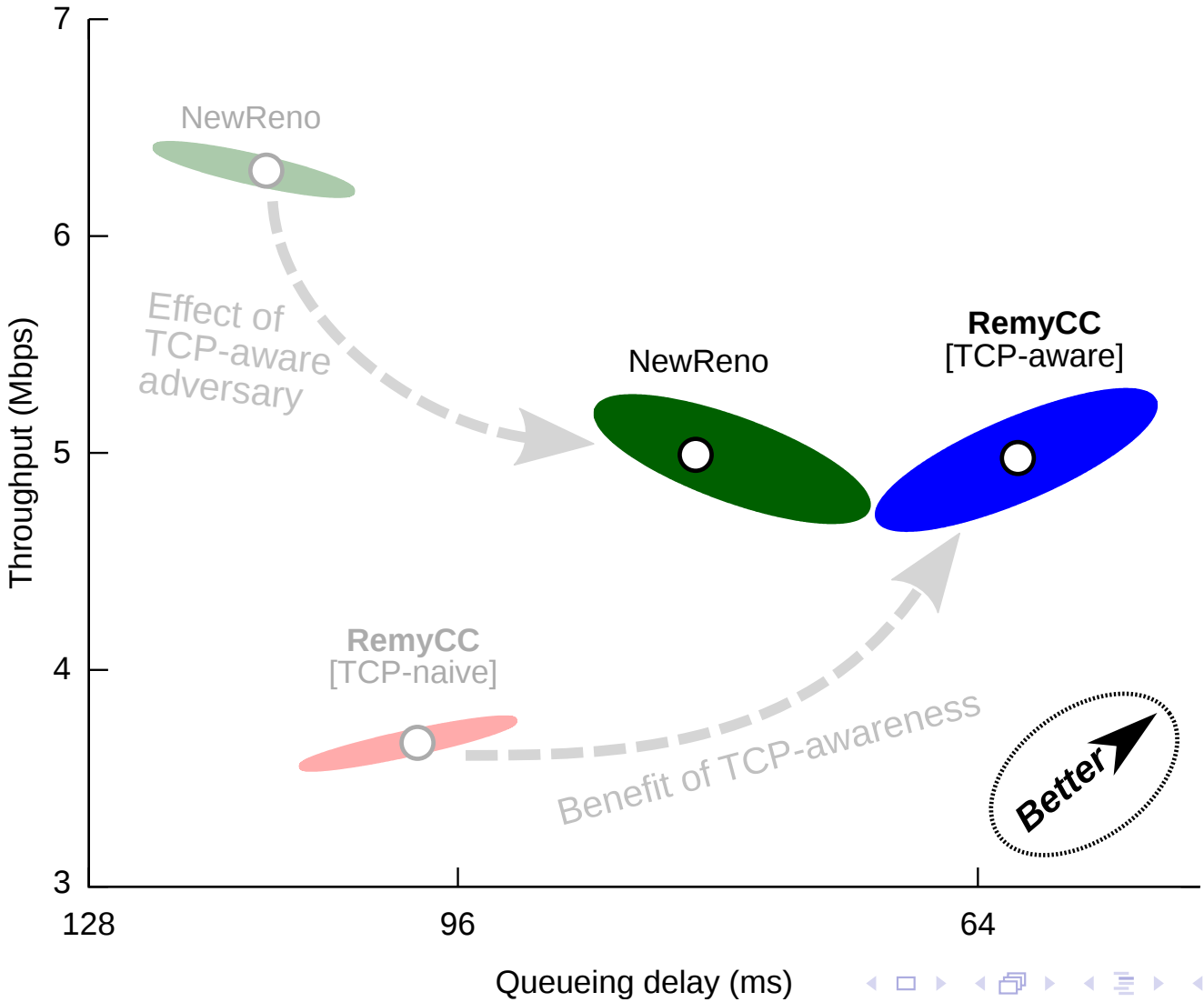
RemyCC competing against TCP NewReno



RemyCC competing against TCP NewReno

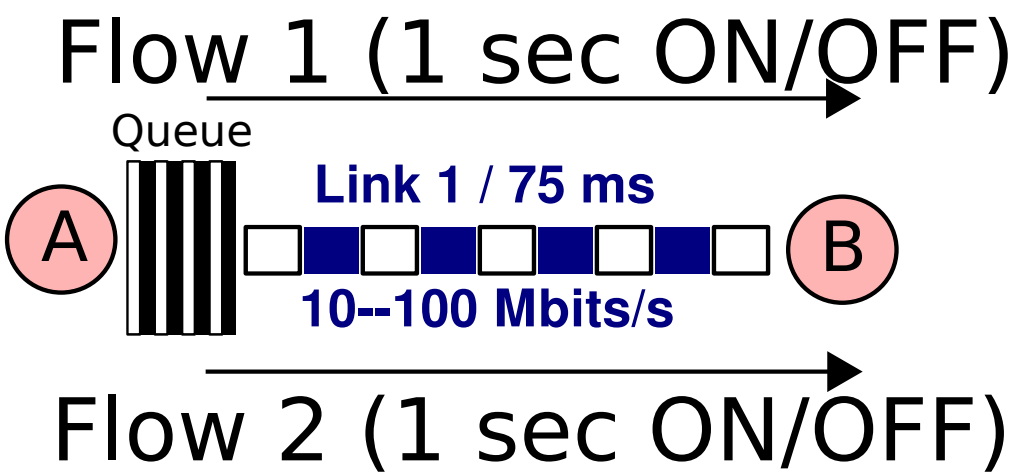


RemyCC competing against TCP NewReno



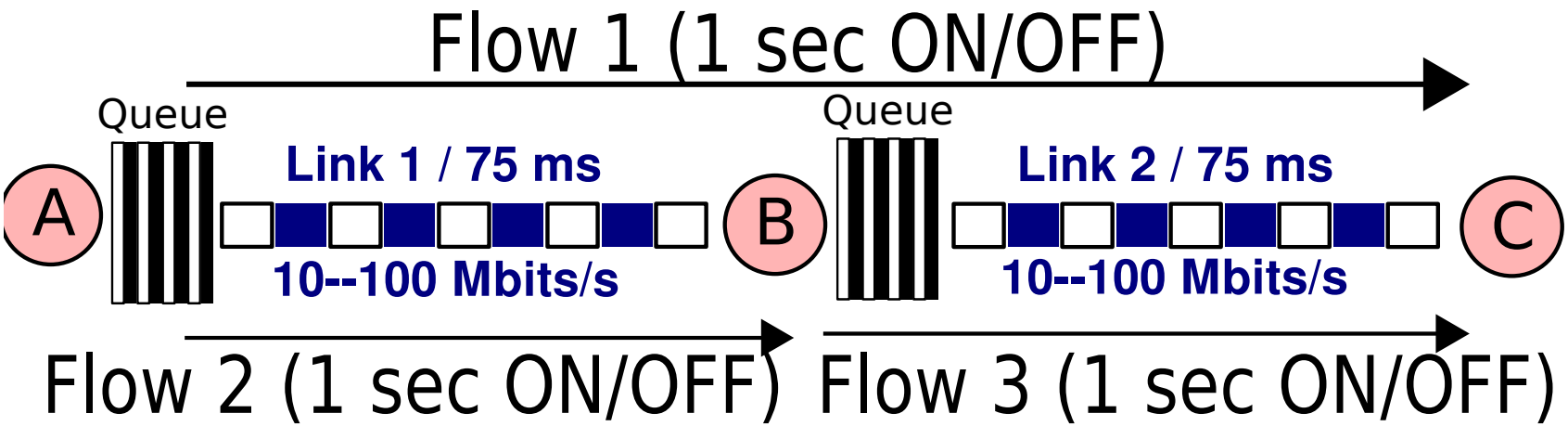
When the model is wrong about the topology

One bottleneck

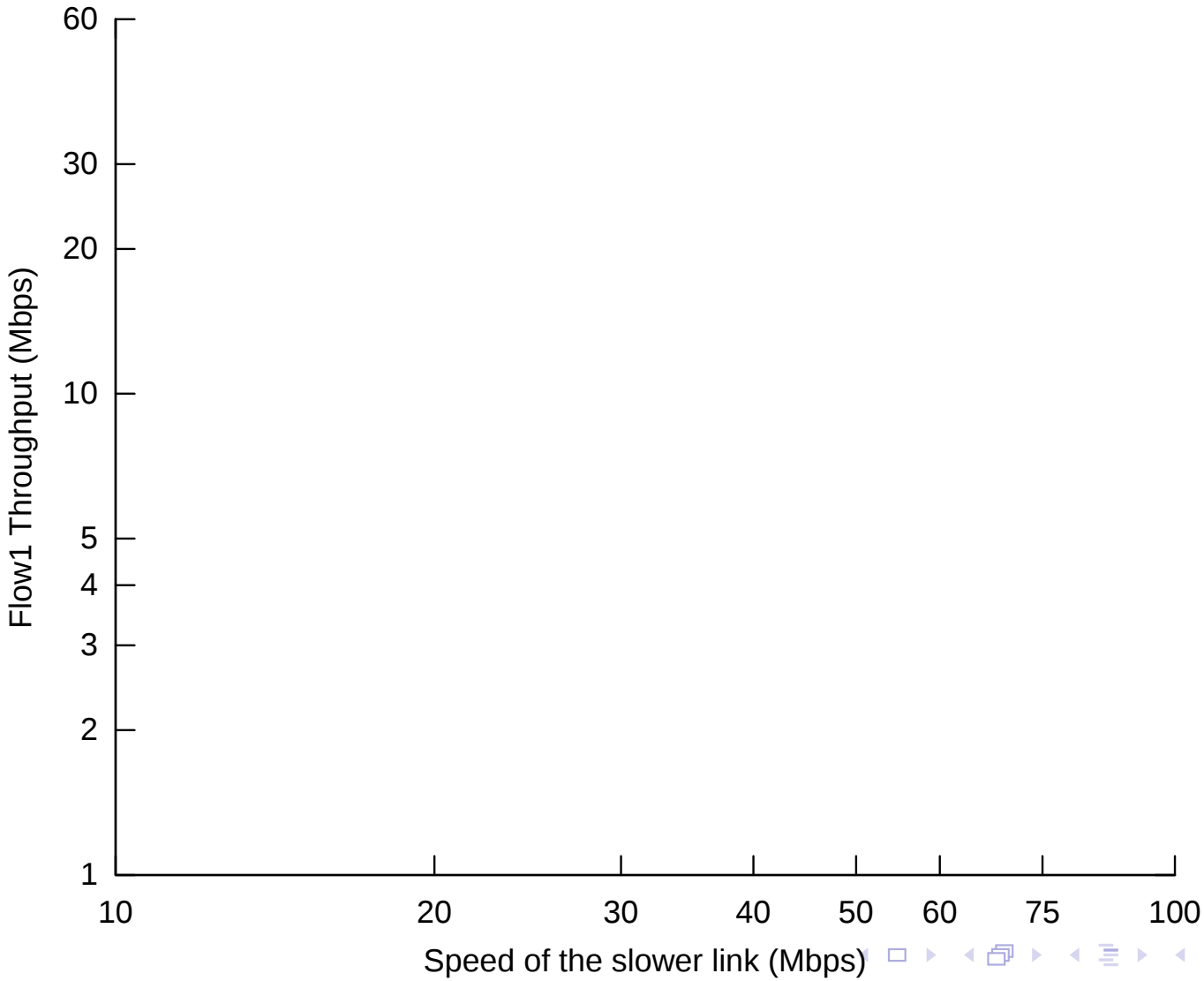


When the model is wrong about the topology

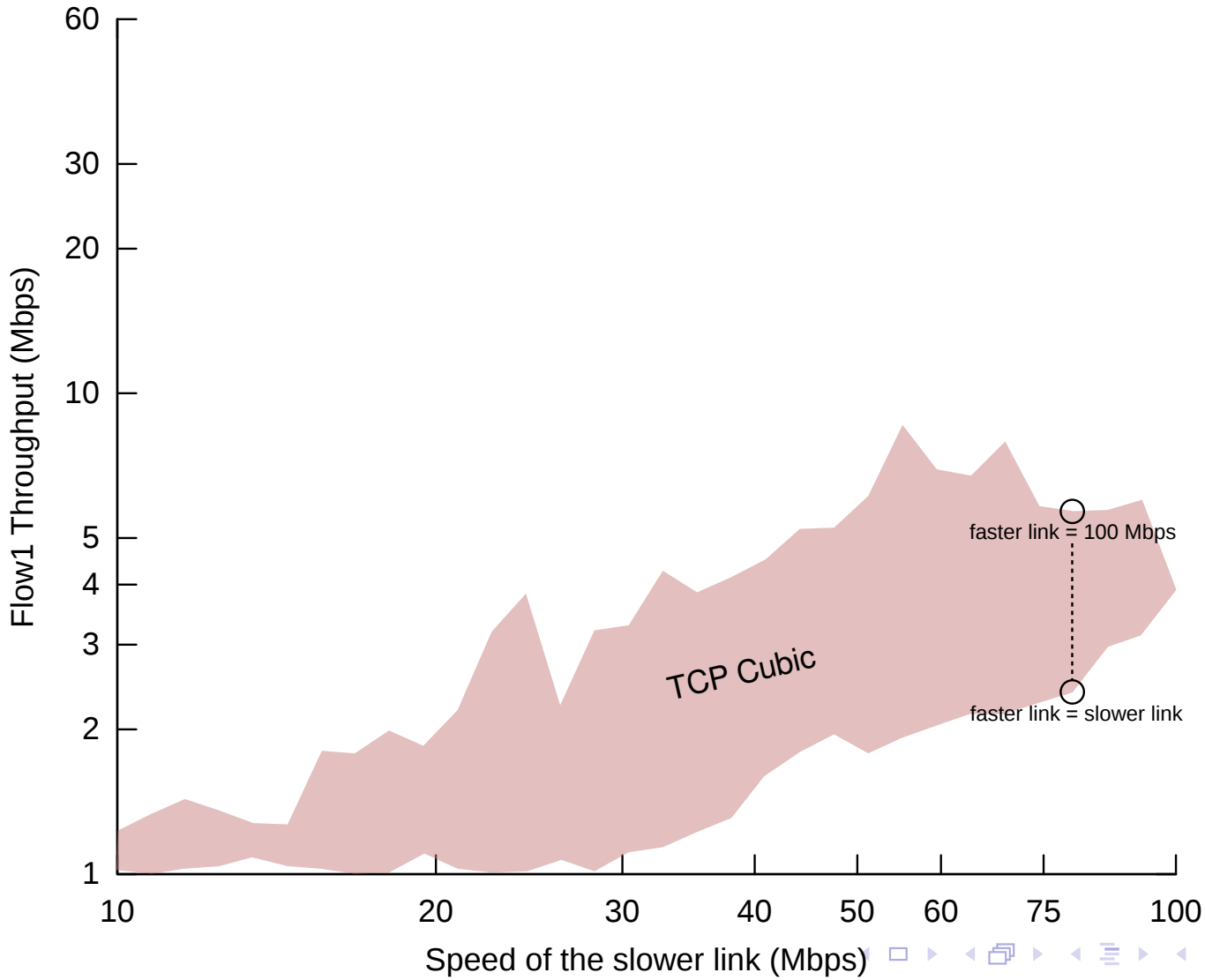
Two bottlenecks



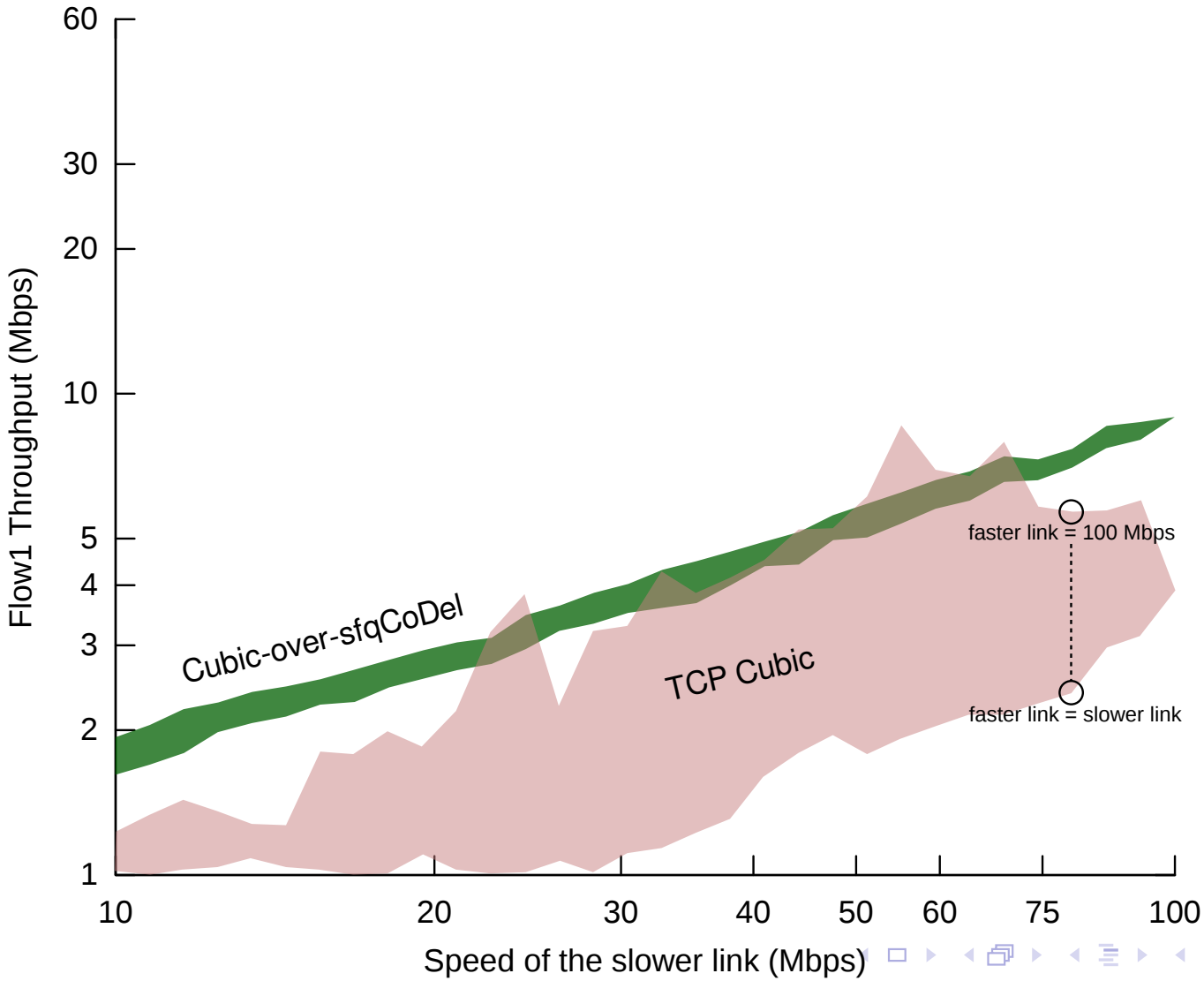
When the model is wrong about the topology



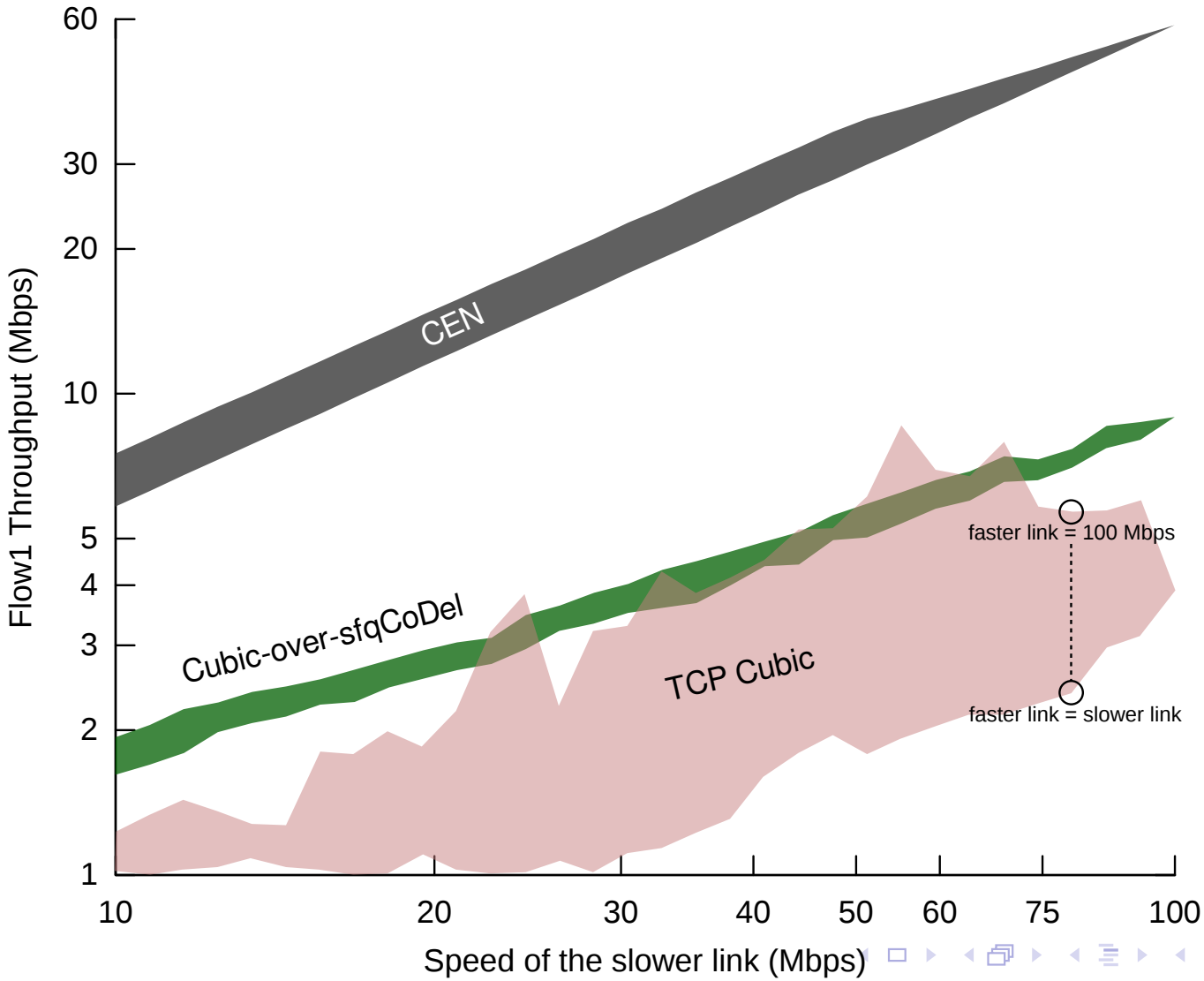
When the model is wrong about the topology



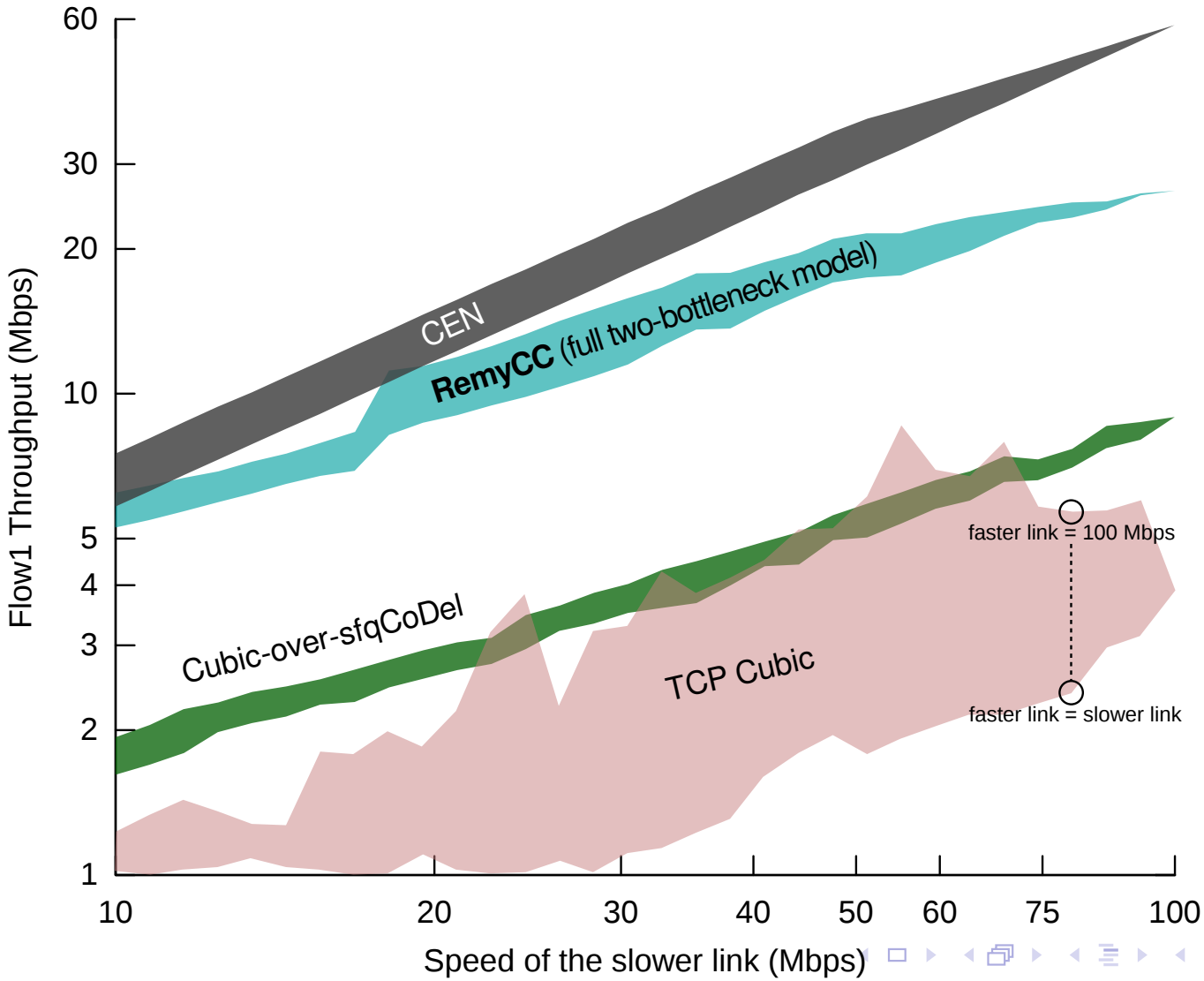
When the model is wrong about the topology



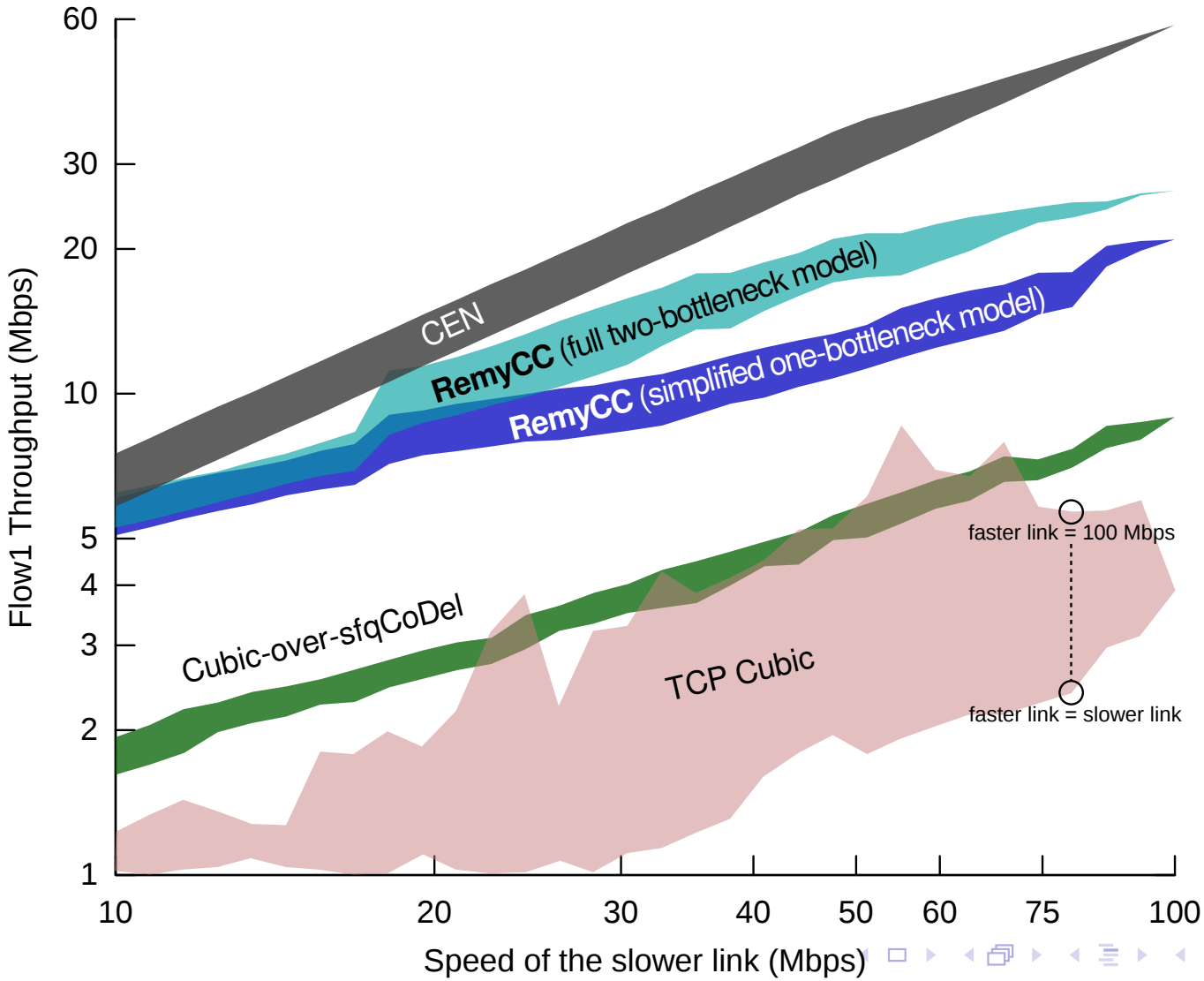
When the model is wrong about the topology



When the model is wrong about the topology



When the model is wrong about the topology



Related Work

- Probably approximately correct learning
- Transfer learning
- Machine-generated congestion control

Limitations and future work

- Generalizability to more complex topologies?
- Better characterization of gap from optimal
- Do results change if we learn in-network behavior as well?
- Model mismatches between simulation and the real world

Backup slides

Can applications with different objectives coexist?

- Tpt. Sender: A throughput-intensive sender

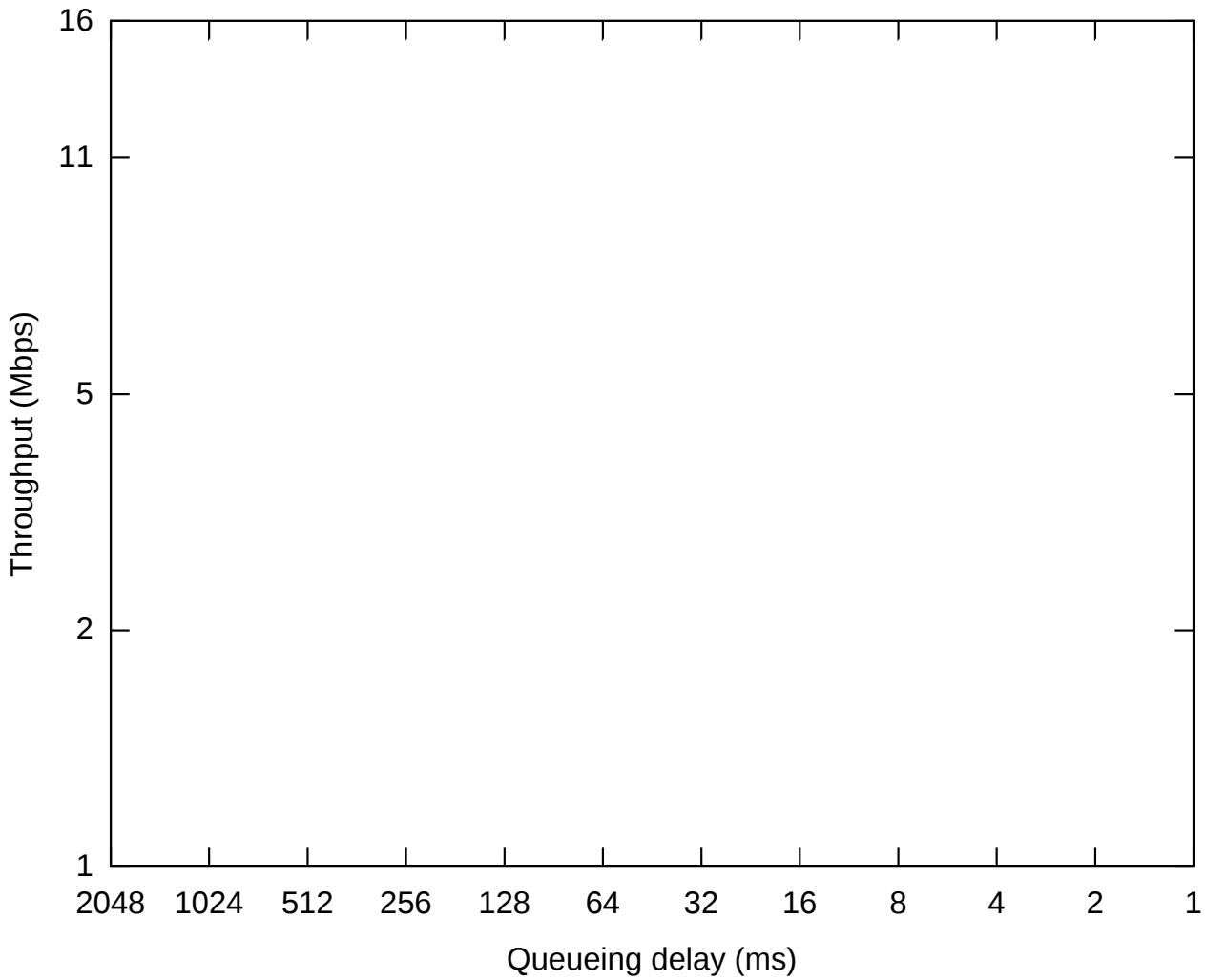
$$log(throughput) - 0.1 * log(delay) \quad (1)$$

- Lat. Sender: A latency-sensitive sender

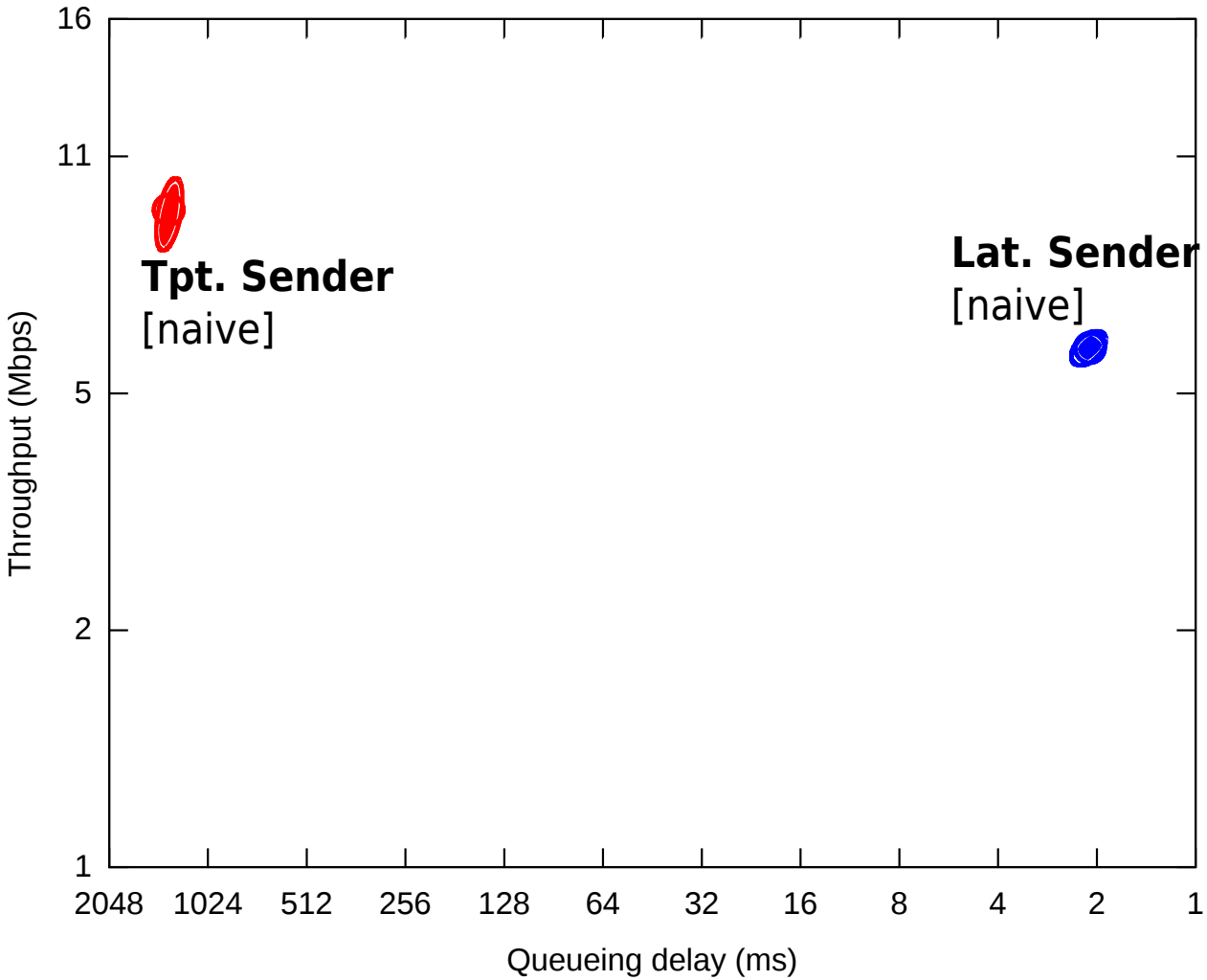
$$log(throughput) - 10.0 * log(delay) \quad (2)$$

- Running over a FIFO queue

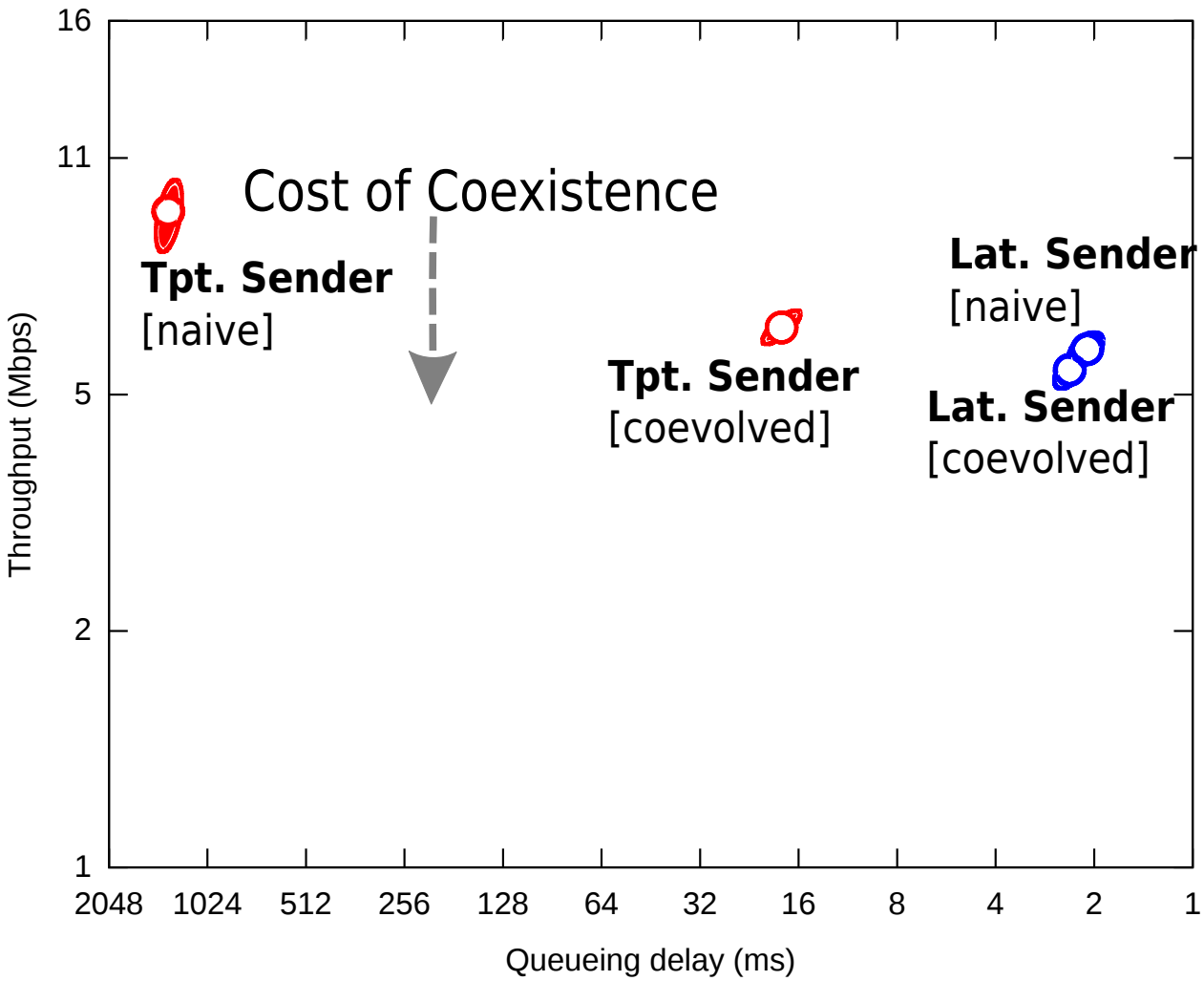
Training for diversity has a cost ...



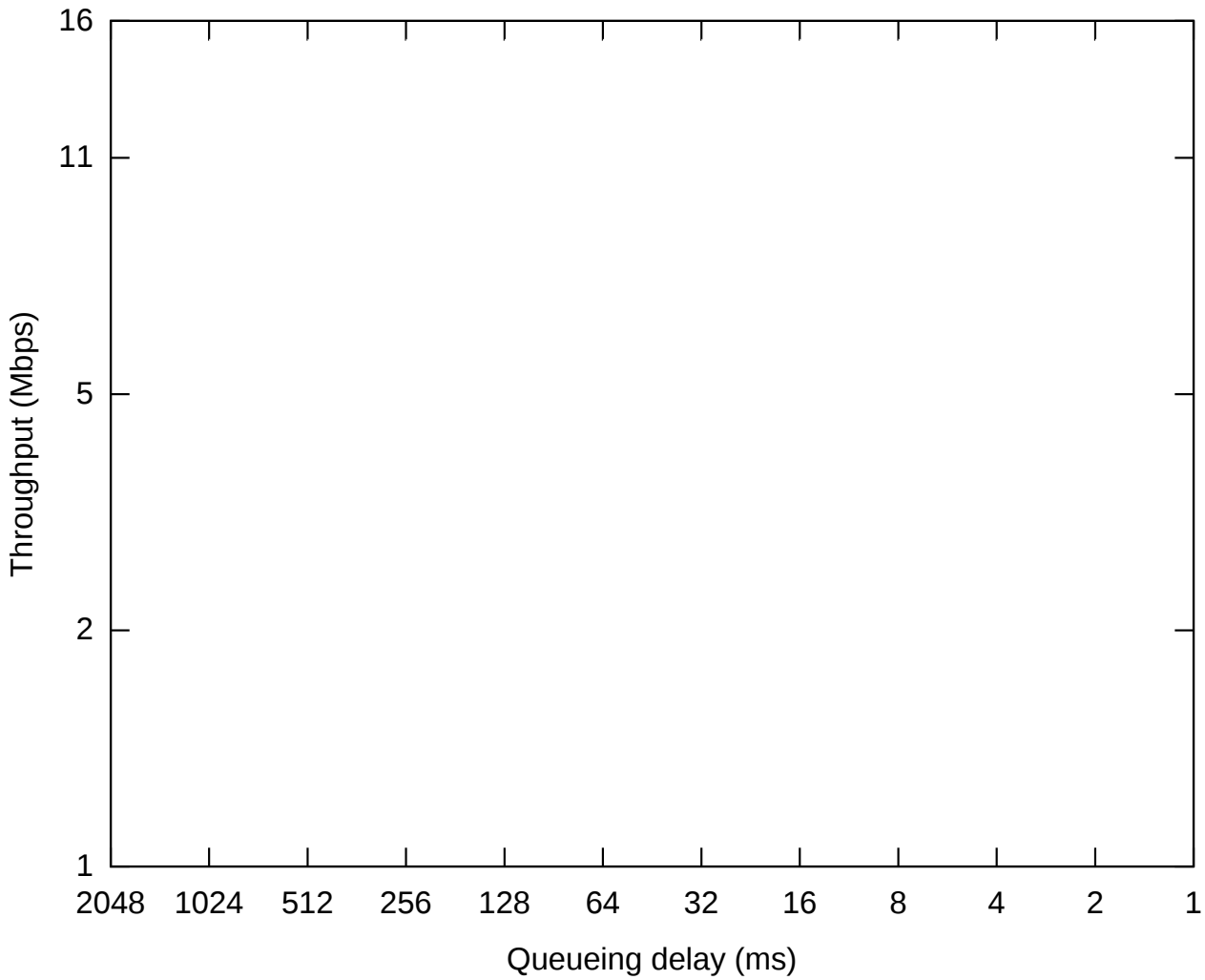
Training for diversity has a cost ...



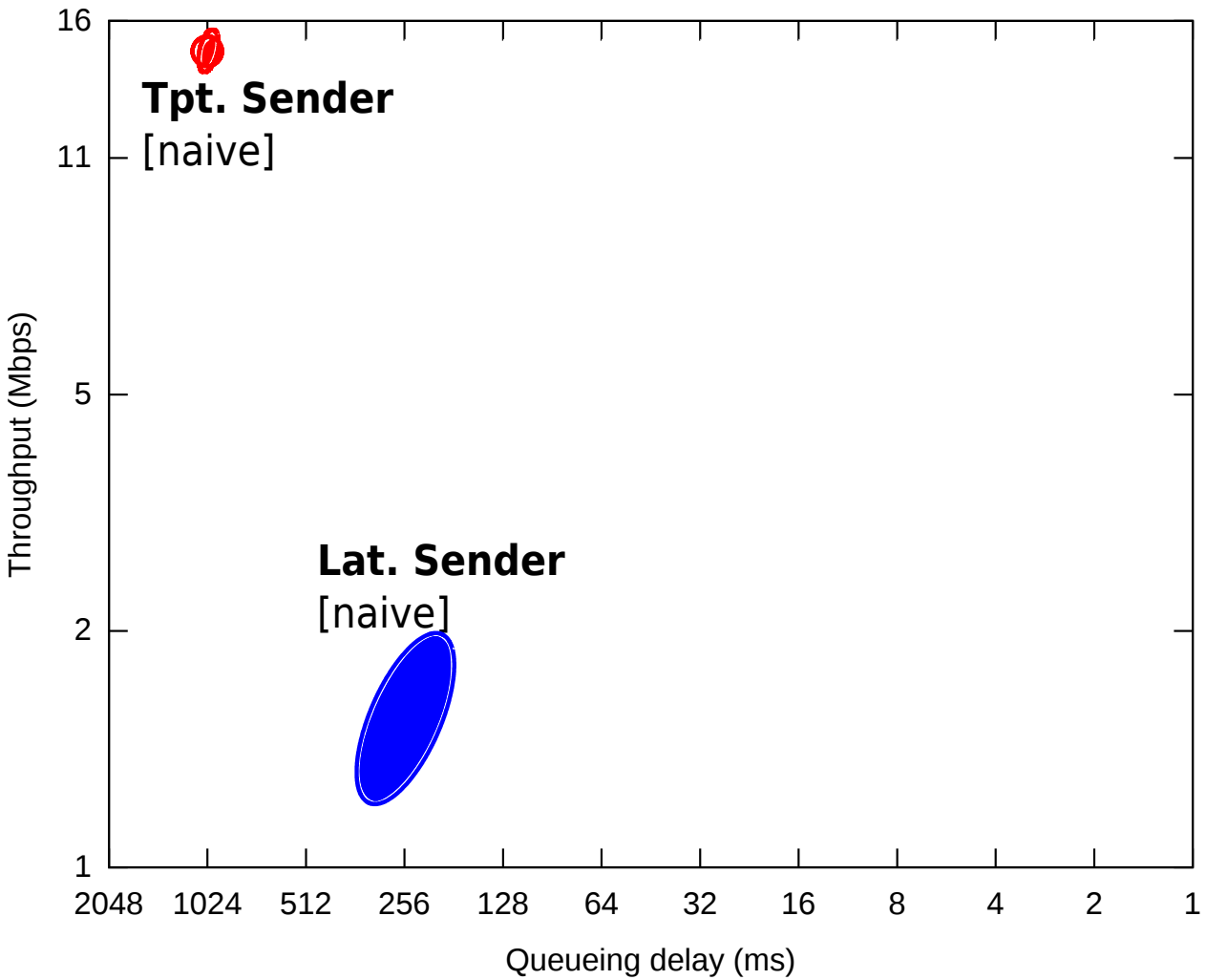
Training for diversity has a cost ...



but, benefits the docile sender



but, benefits the docile sender



but, benefits the docile sender