The Pantheon of Congestion Control

Keith Winstein

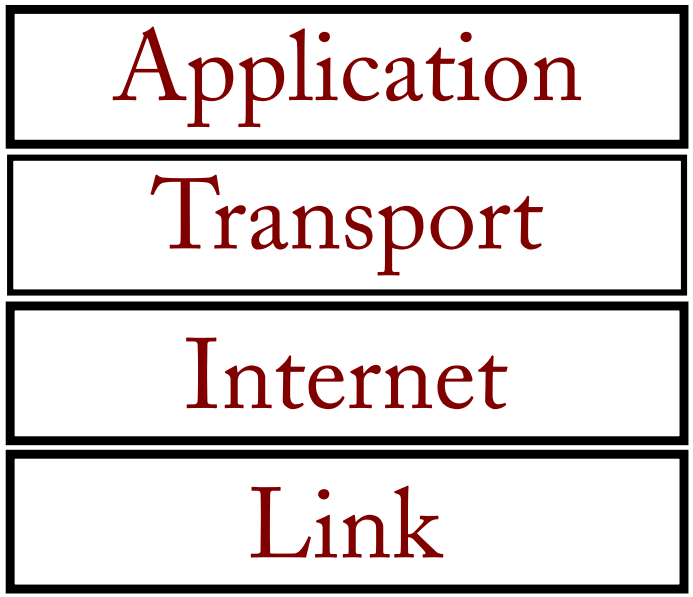
Stanford University

July 21, 2016

Includes joint work with Greg Hill, Yu Yan, Ravi Netravali,
Anirudh Sivaraman, Pratiksha Thaker, Pauline Varley, Hari Balakrishnan.

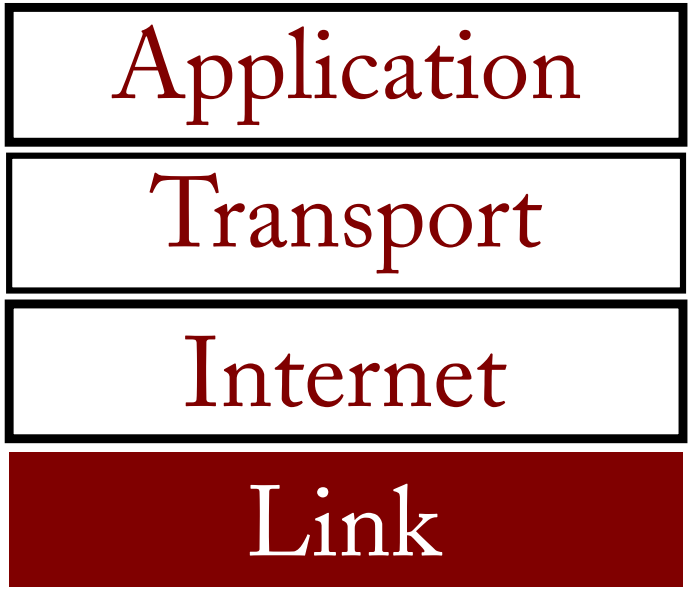
The Internet evolves

In 20 years, computer networks have seen dramatic change:



The Internet evolves

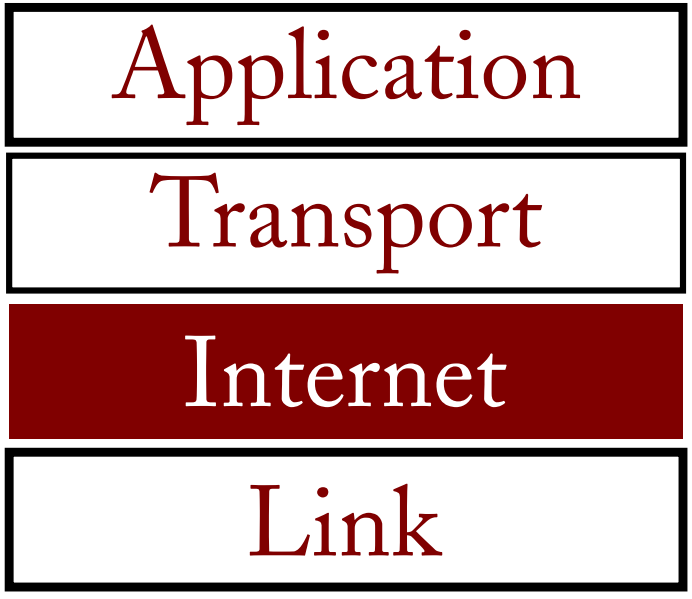
In 20 years, computer networks have seen dramatic change:



- Wi-Fi
- Cellular networks
- Terrible cellular networks
- Satellites
- Datacenters
- 10 GigE
- Transoceanic links

The Internet evolves

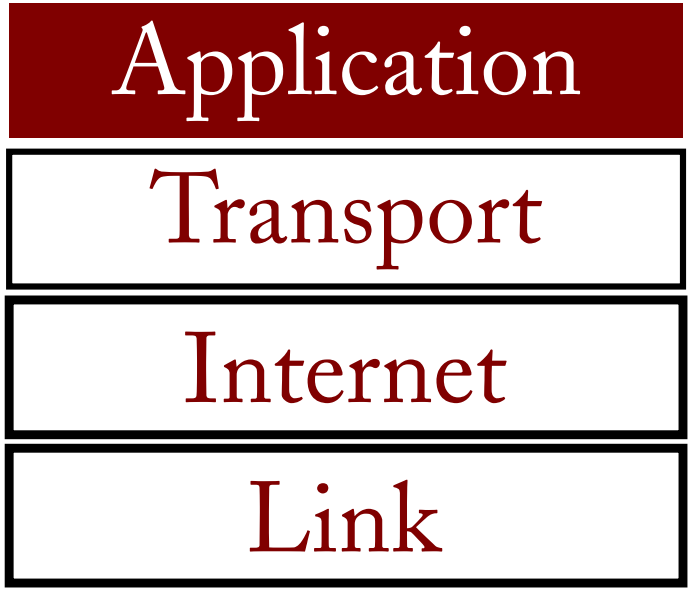
In 20 years, computer networks have seen dramatic change:



- Ubiquitous mobility

The Internet evolves

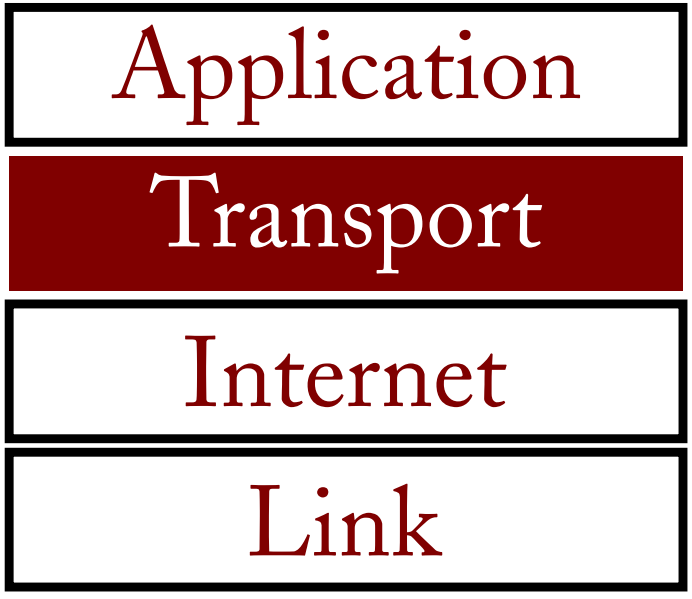
In 20 years, computer networks have seen dramatic change:



- Short flows (Web)
- Streaming video (YouTube/Netflix)
- Conferencing (Skype/Facetime)

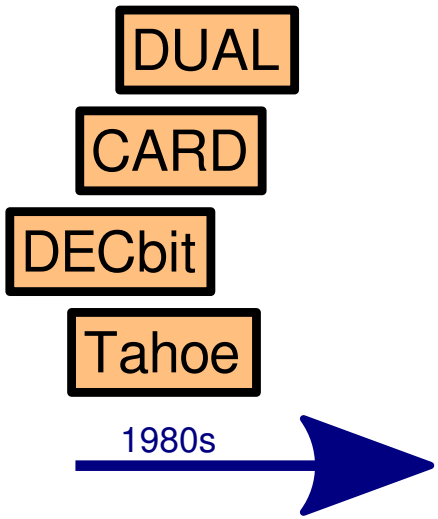
The Internet evolves

In 20 years, computer networks have seen dramatic change:

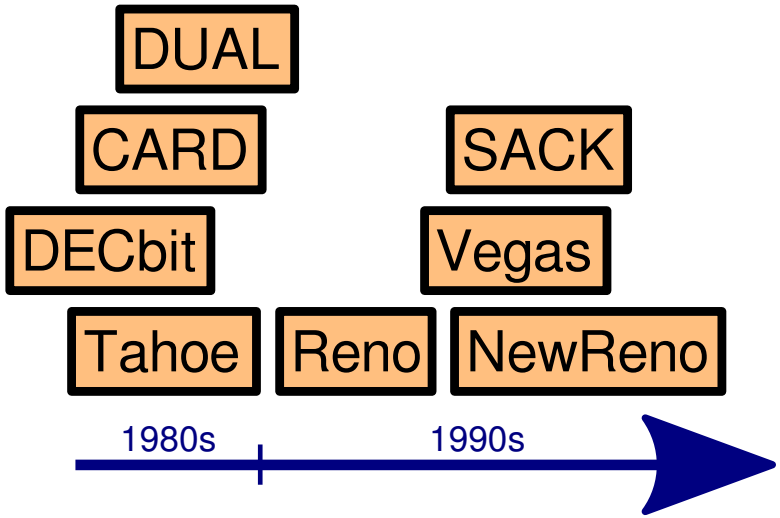


- ???

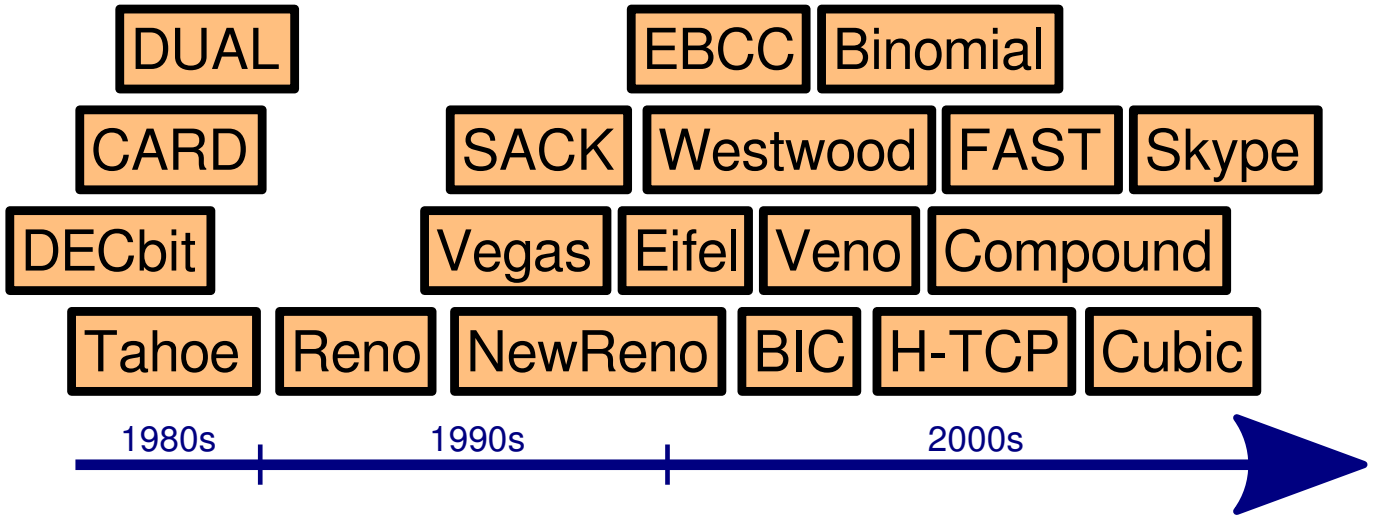
The march of congestion-control protocols



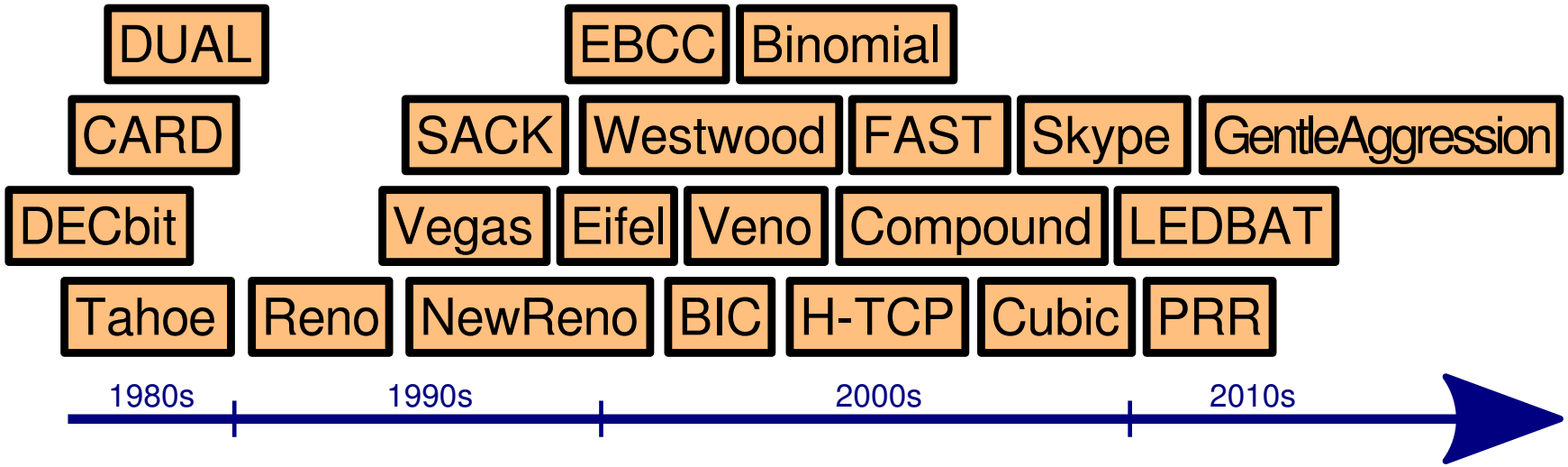
The march of congestion-control protocols



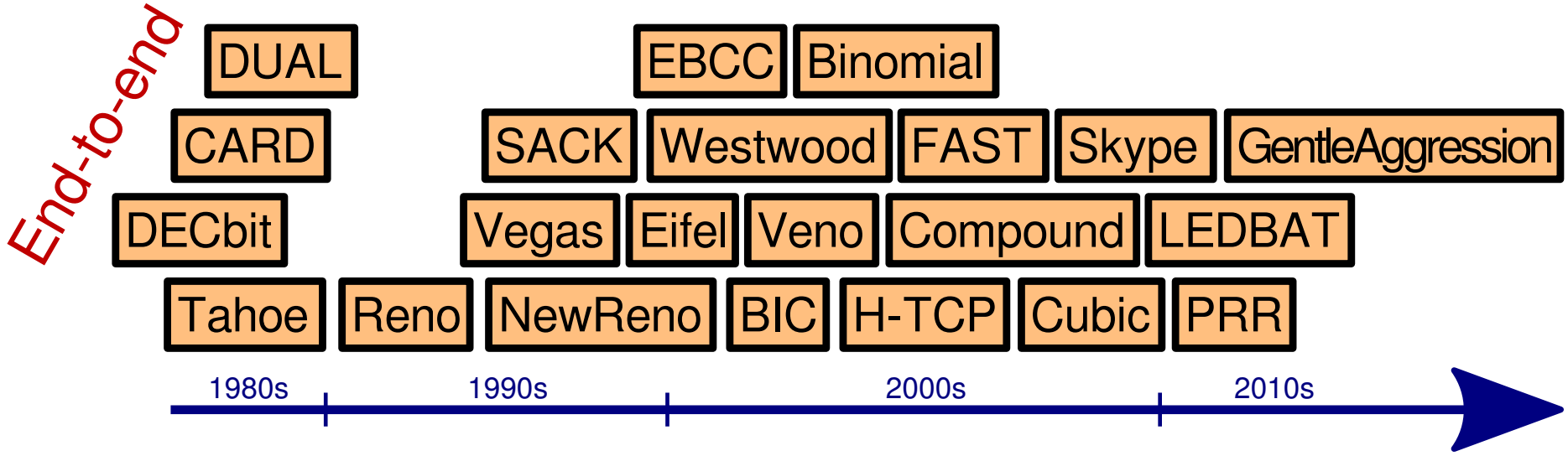
The march of congestion-control protocols



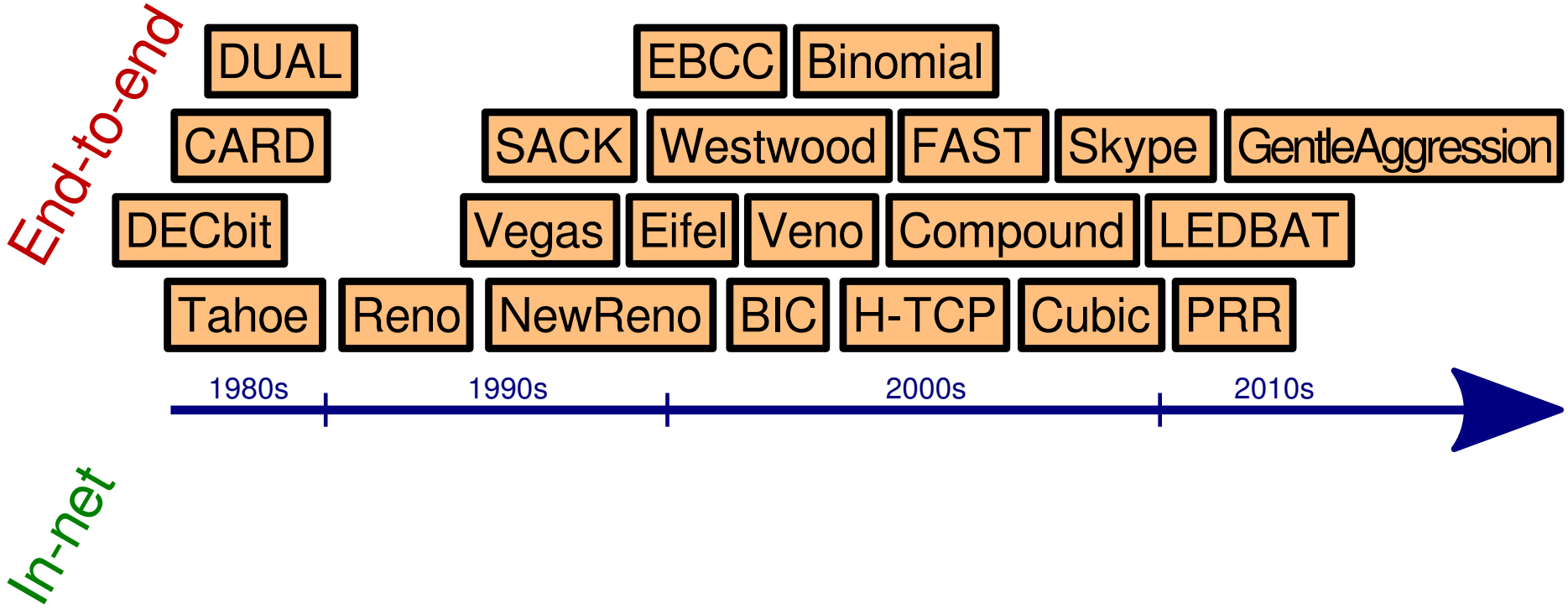
The march of congestion-control protocols



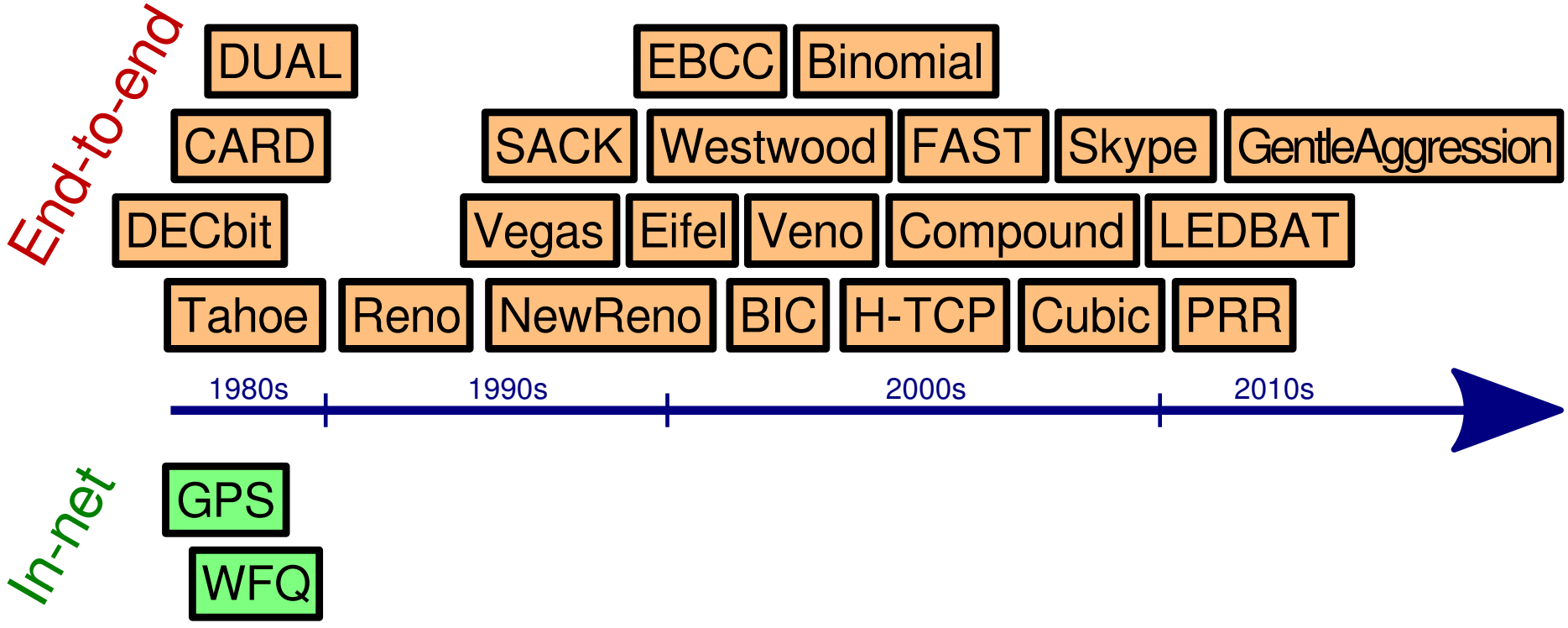
The march of congestion-control protocols



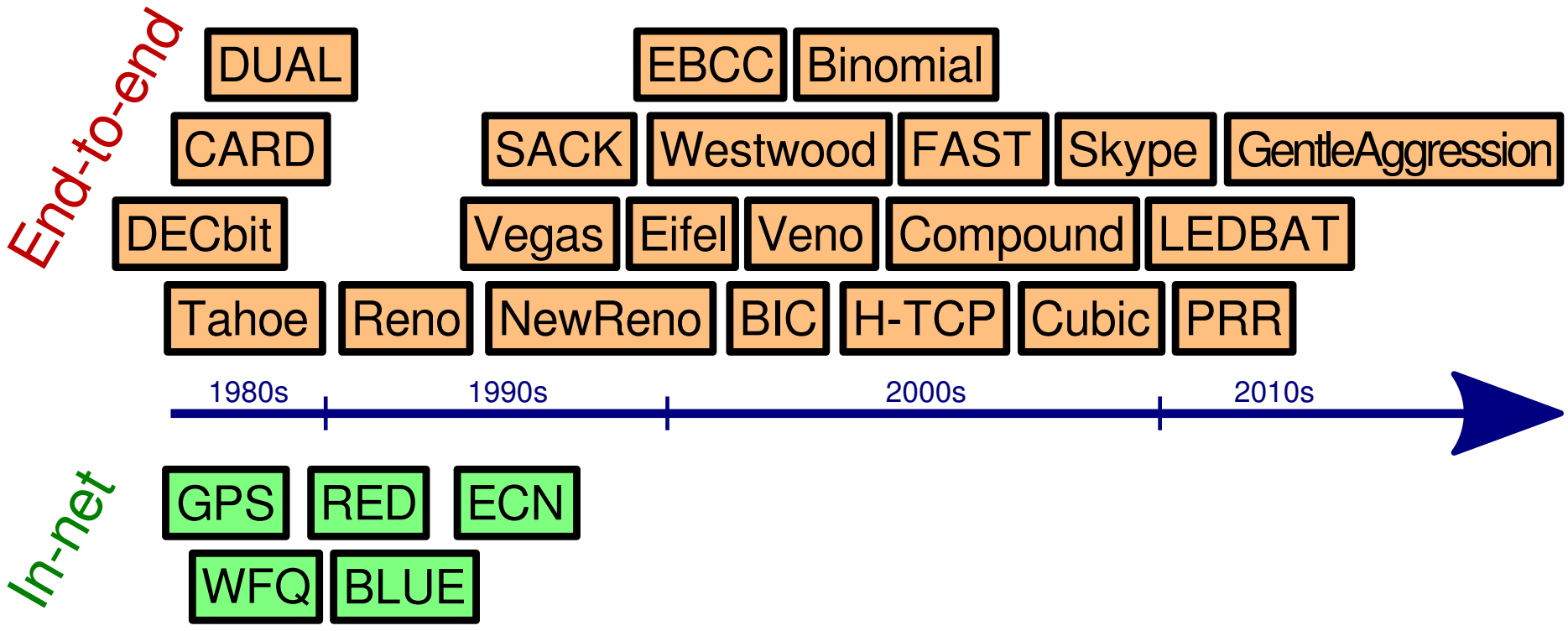
The march of congestion-control protocols



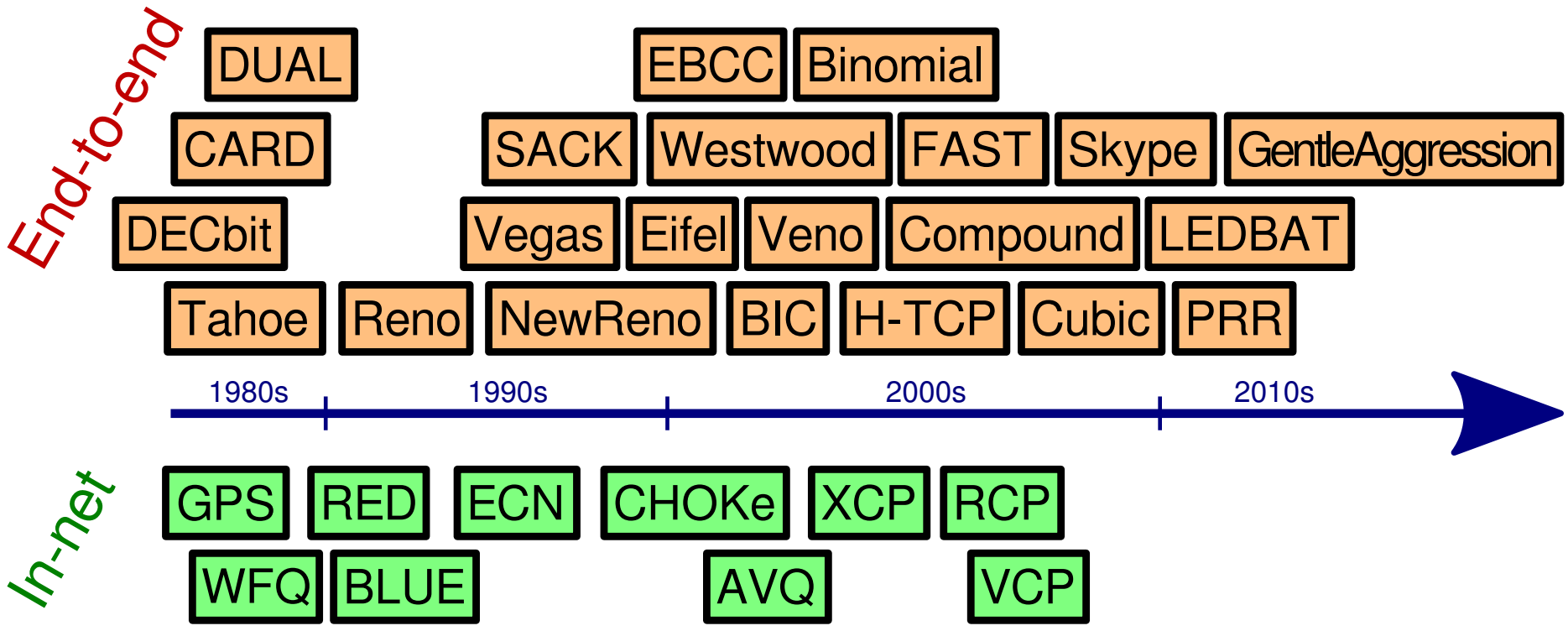
The march of congestion-control protocols



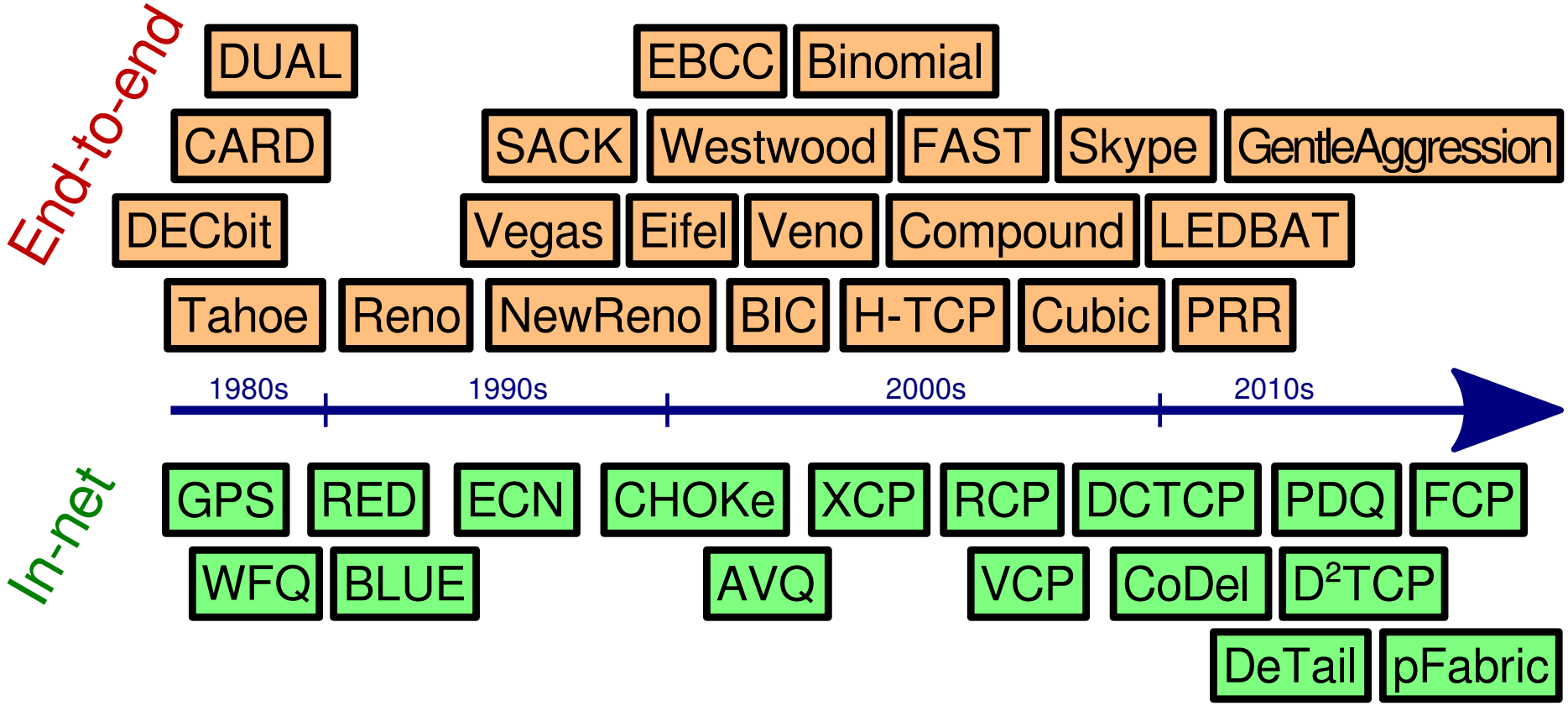
The march of congestion-control protocols



The march of congestion-control protocols



The march of congestion-control protocols

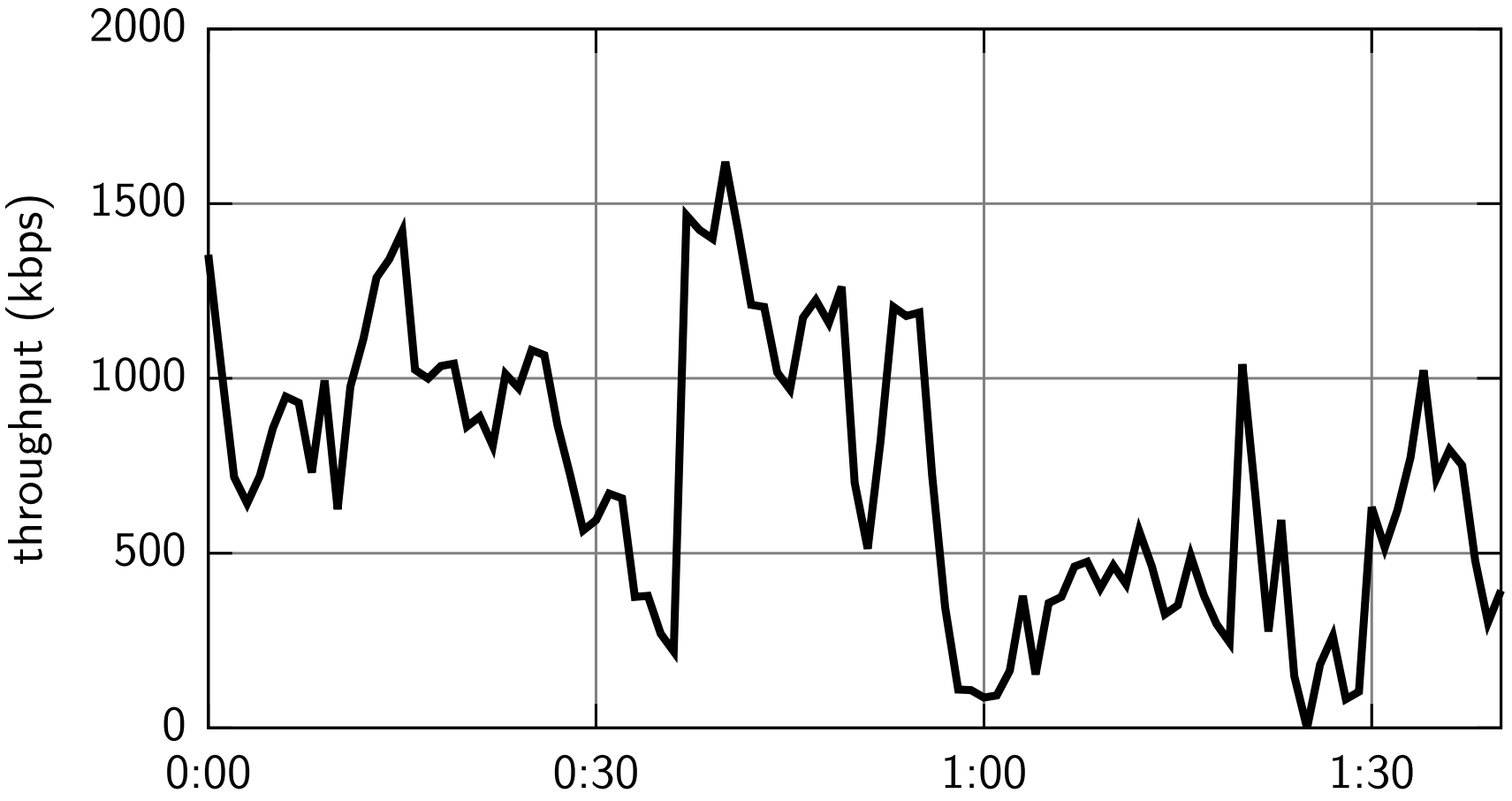


Sprout: a transport protocol designed for variability

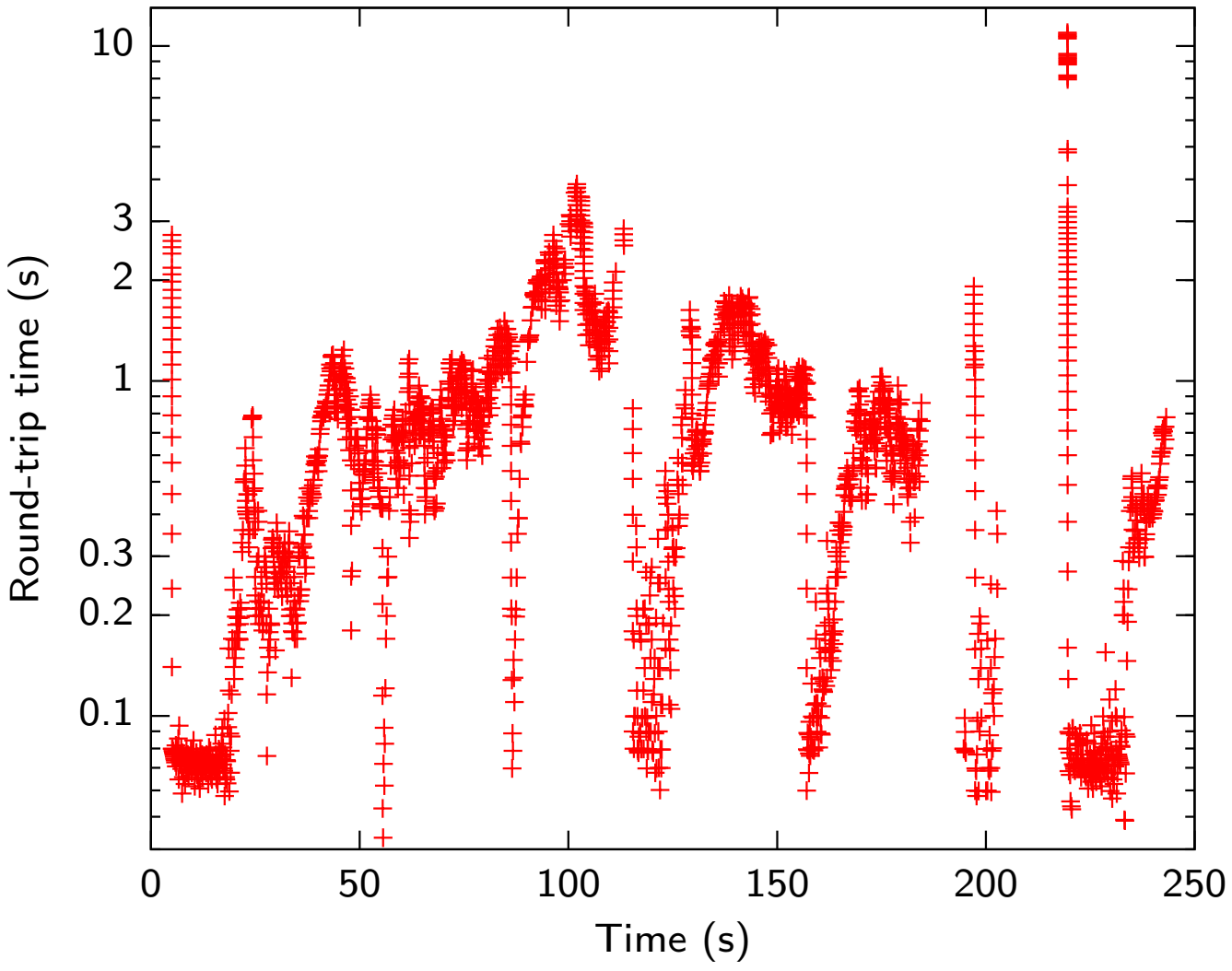
Observation:

Videoconferences perform poorly over cellular networks.

Verizon LTE throughput is highly variable



During a download, LTE delays packets for $>$ 10 seconds

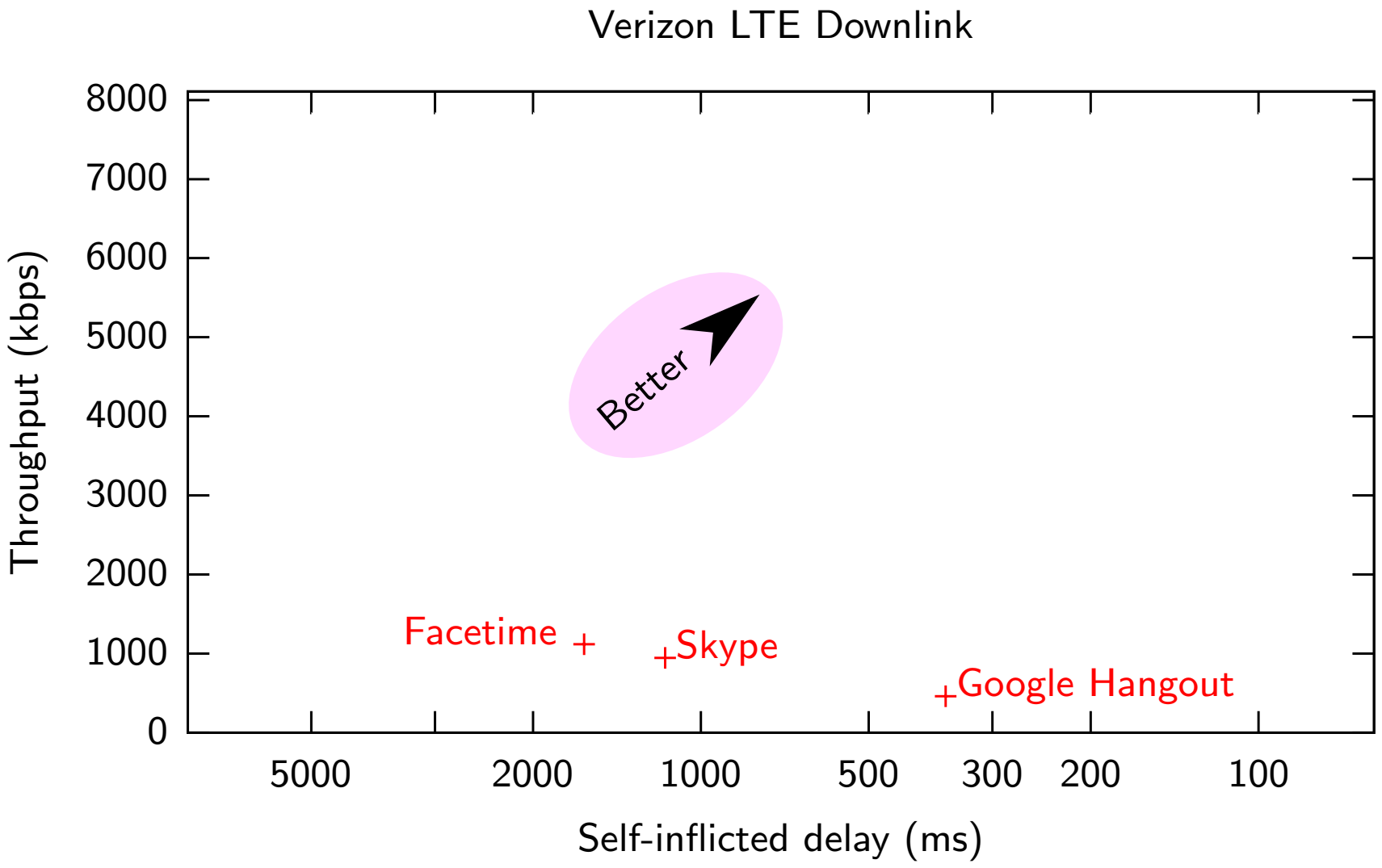


Interactive apps work poorly

- We measured cellular networks while driving:
  - **Verizon LTE**
  - Verizon 3G (1xEV-DO)
  - AT&T LTE
  - T-Mobile 3G (UMTS)
- Then ran apps across replayed network trace:
  - **Skype** (Windows 7)
  - Google Hangouts (Chrome on Windows 7)
  - Apple Facetime (OS X)

Skype's performance

Performance summary



What's wrong?

- Existing schemes **react** to congestion signals.
  - Packet loss.
  - Increase in round-trip time.
- Feedback comes too late.
- The killer: **self-inflicted queueing delay**.

Question

Can a transport protocol that's a function of a **model** and **mission** do better?

Sprout's **mission**

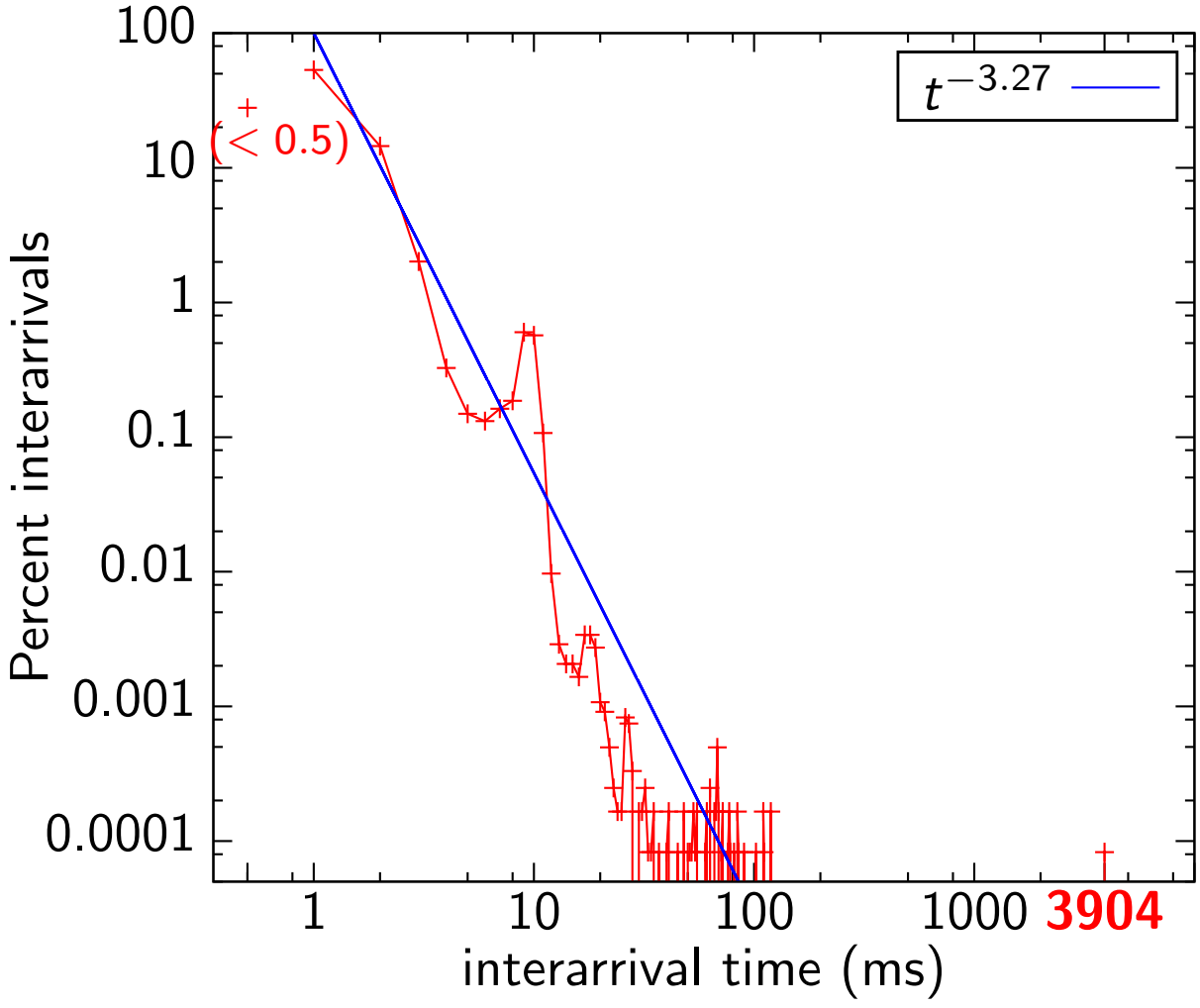
- Most throughput
- Bounded risk of delay $>$ 100 ms

KW, Anirudh Sivaraman, and Hari Balakrishnan, **Stochastic Forecasts Achieve High Throughput and Low Delay over Cellular Networks**, *USENIX NSDI 2013*

Bounded risk of delay

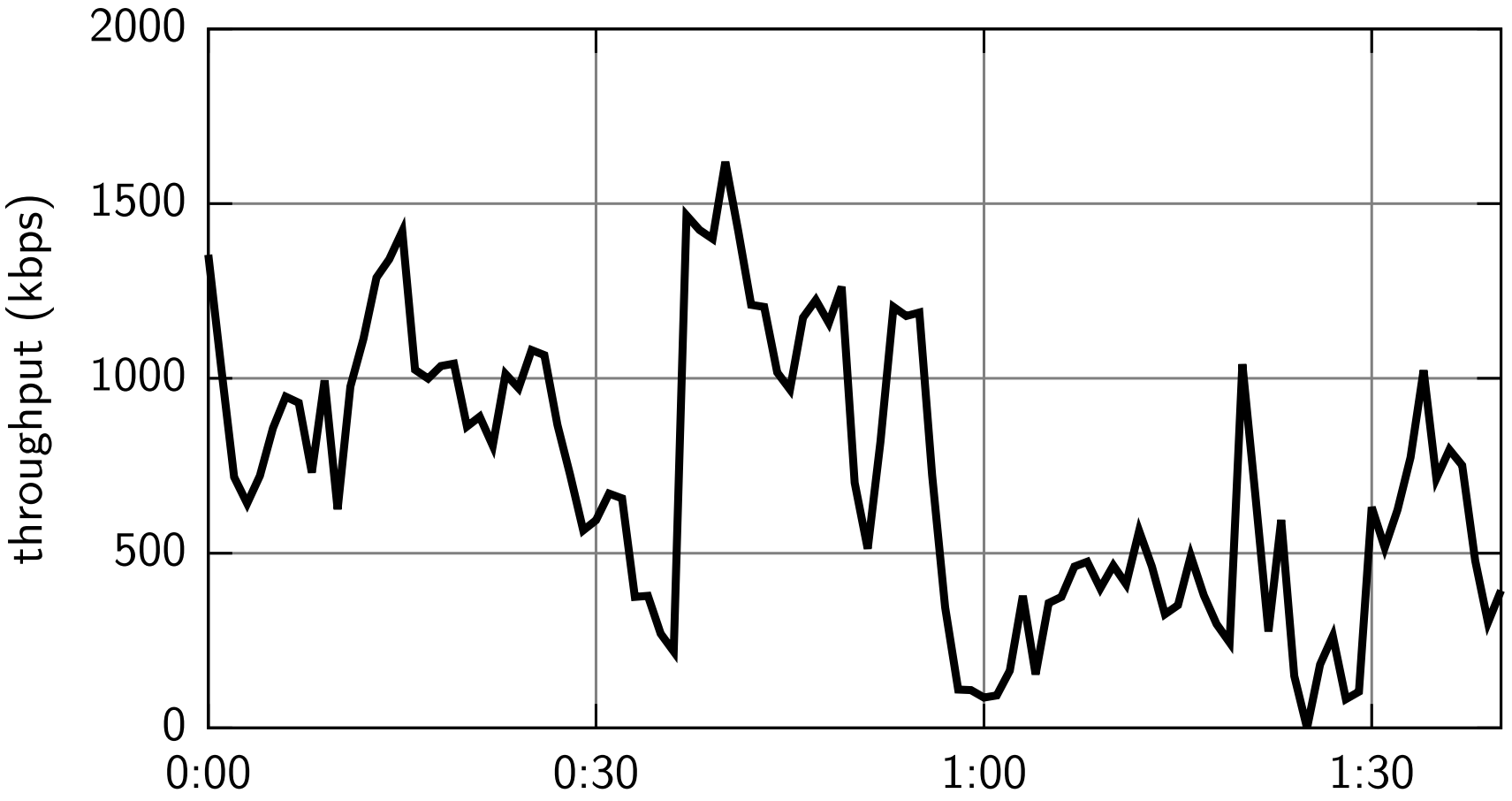
- **Model** variation in link speed
- **Infer** current link speed
- **Predict** future link speed
  - Don't wait for congestion
- **Control:** Send as much as possible, but require:
  - 95% chance all packets arrive within 100 ms

Model: packet deliveries looks like flicker noise

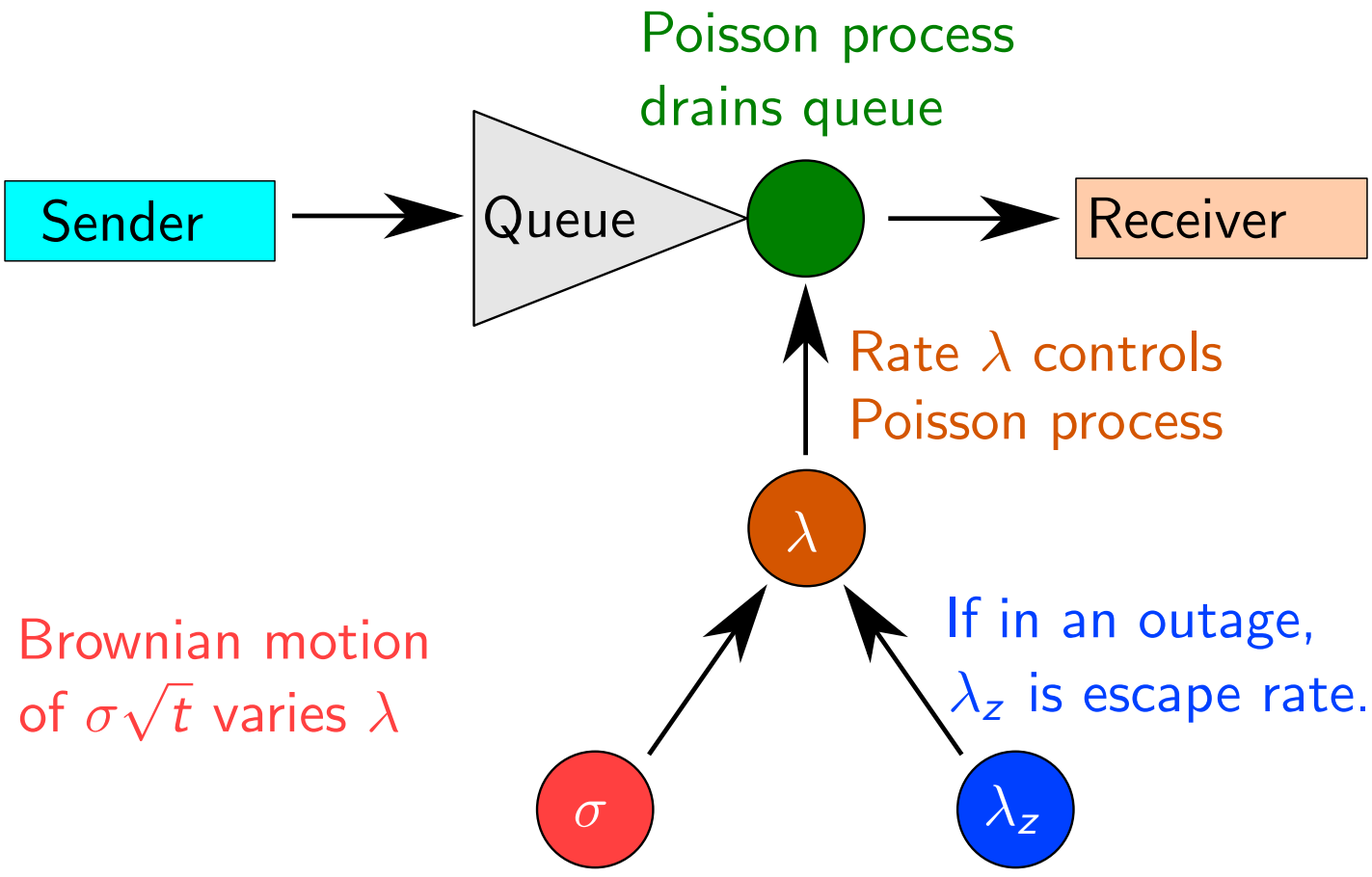


(Verizon LTE, phone stationary.)

Model: average rate looks like random walk



Sprout's **model**: G/M/1 queue



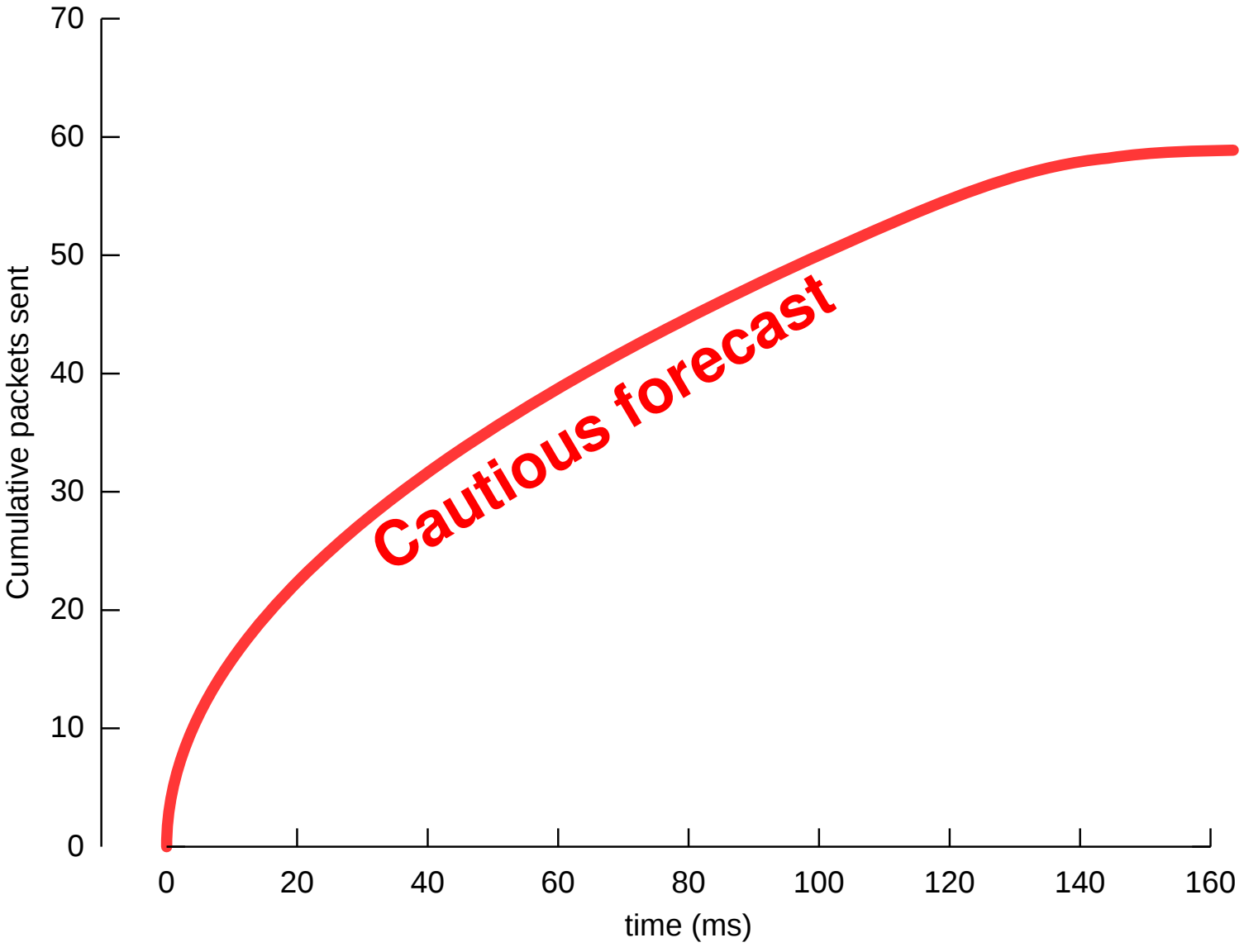
Infer: current link speed

- **Observe** packets received every τ
- **Update** $\mathrm{P}(\lambda)$

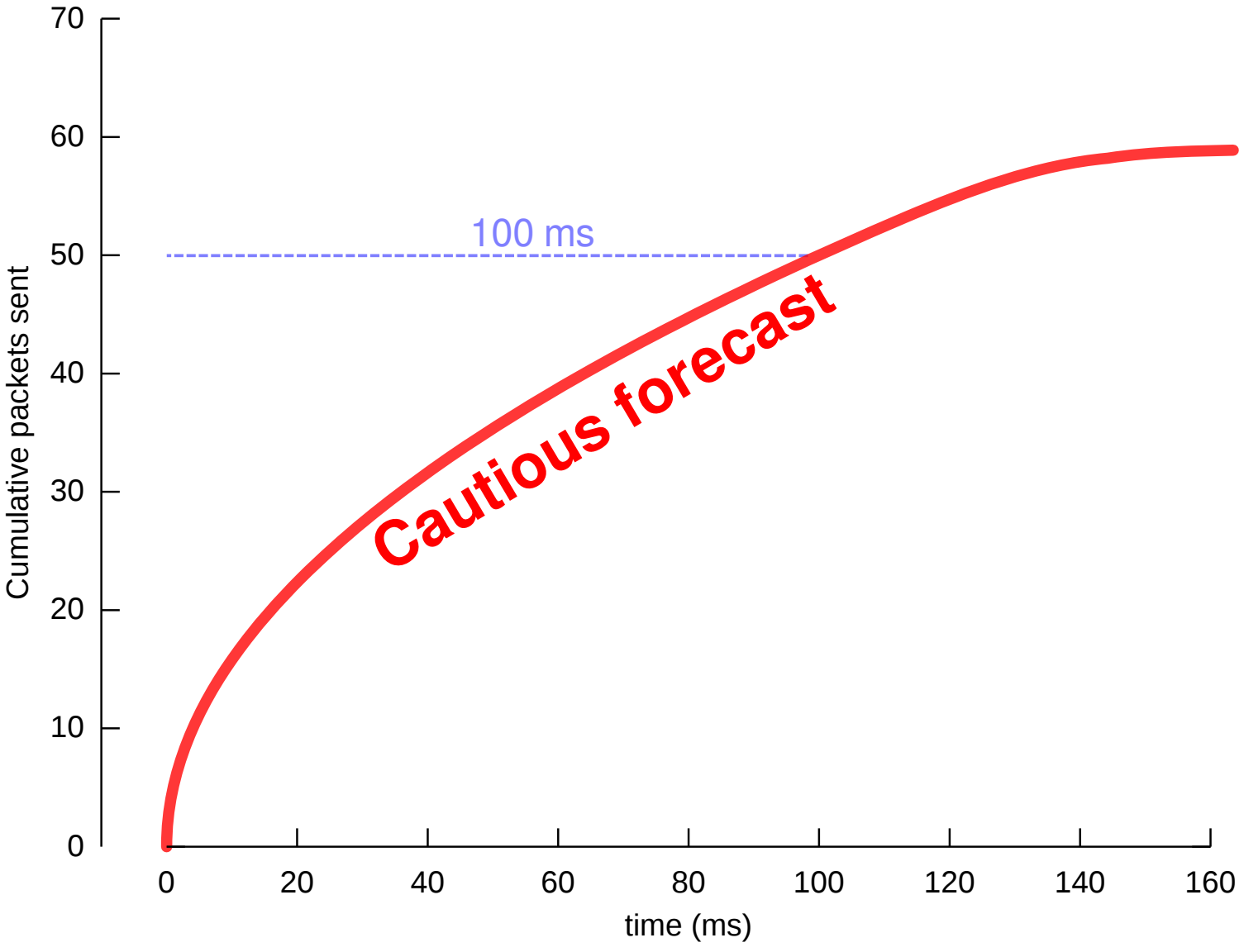
Predict: future link speed

- **Evolve** model forward
- **Forecast** 5th percentile cumulative packets

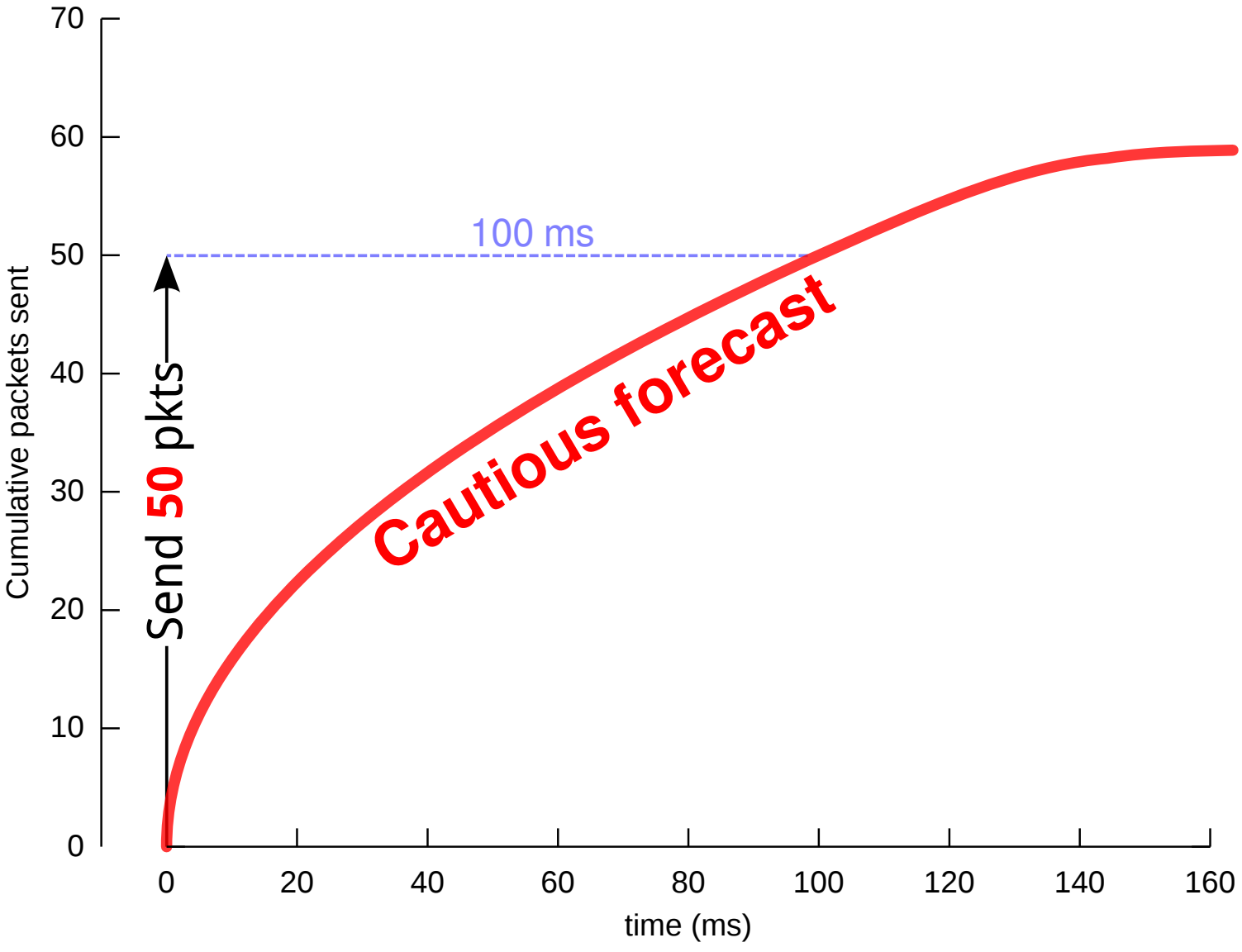
Control: fill up 100 ms forecast window



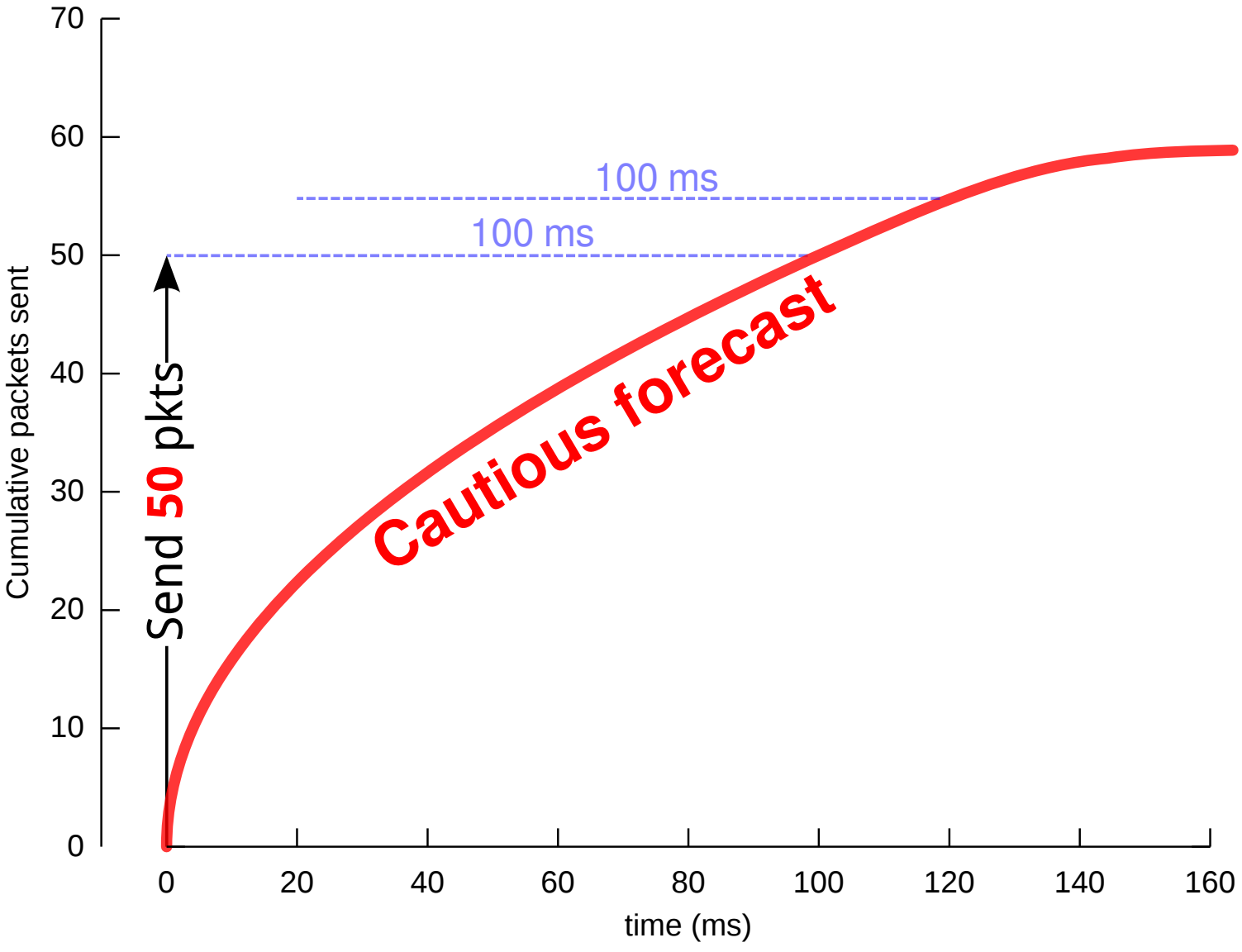
Control: fill up 100 ms forecast window



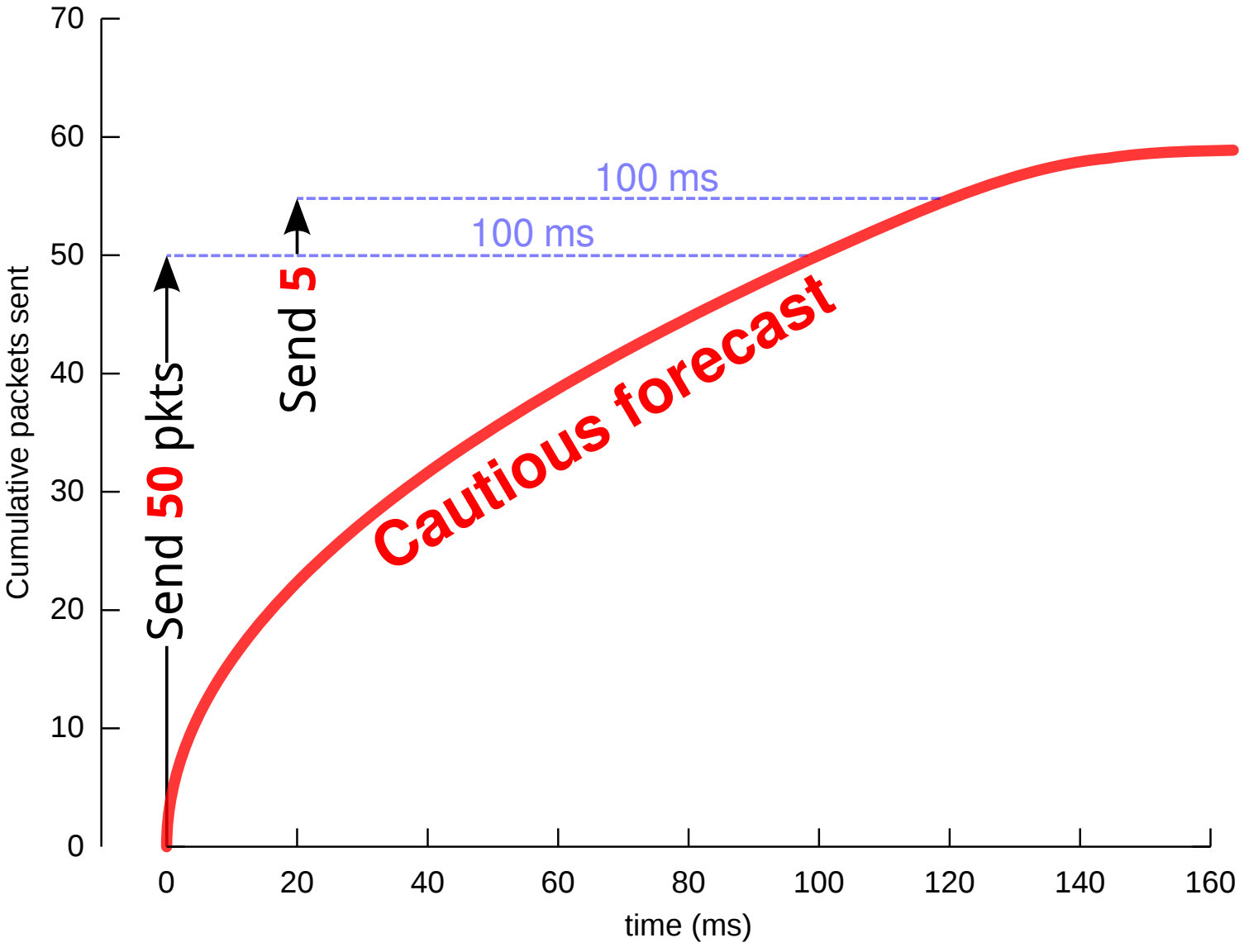
Control: fill up 100 ms forecast window



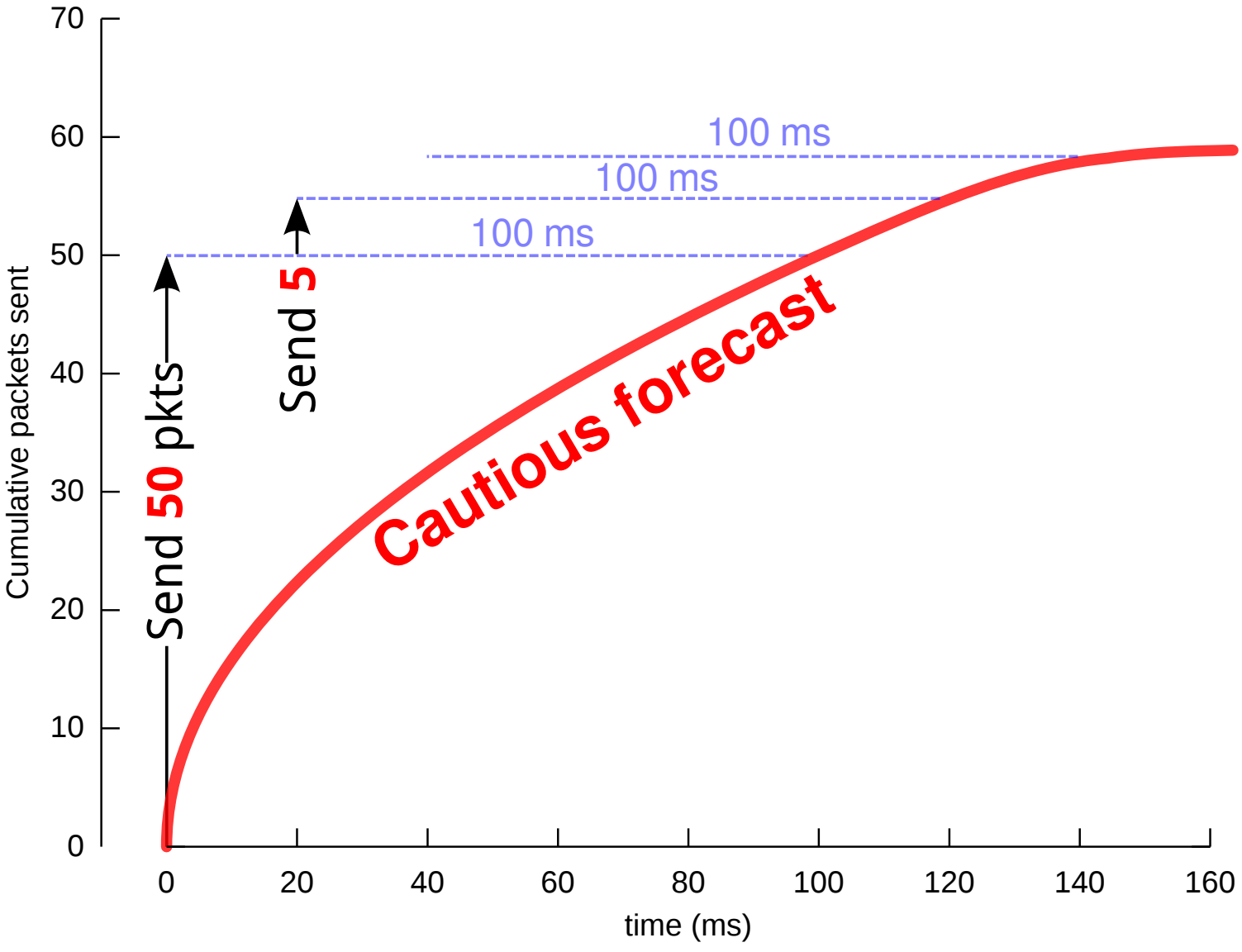
Control: fill up 100 ms forecast window



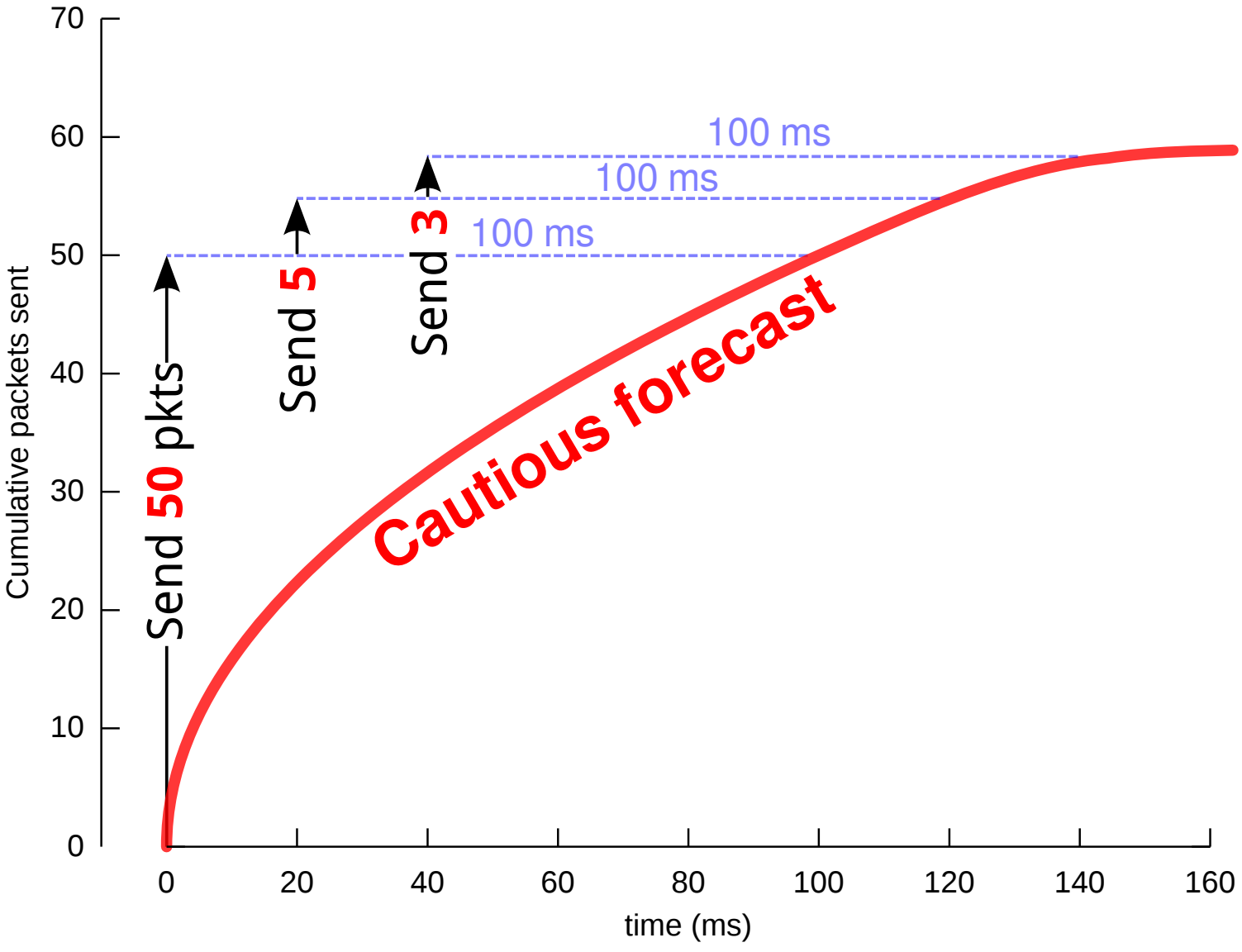
Control: fill up 100 ms forecast window



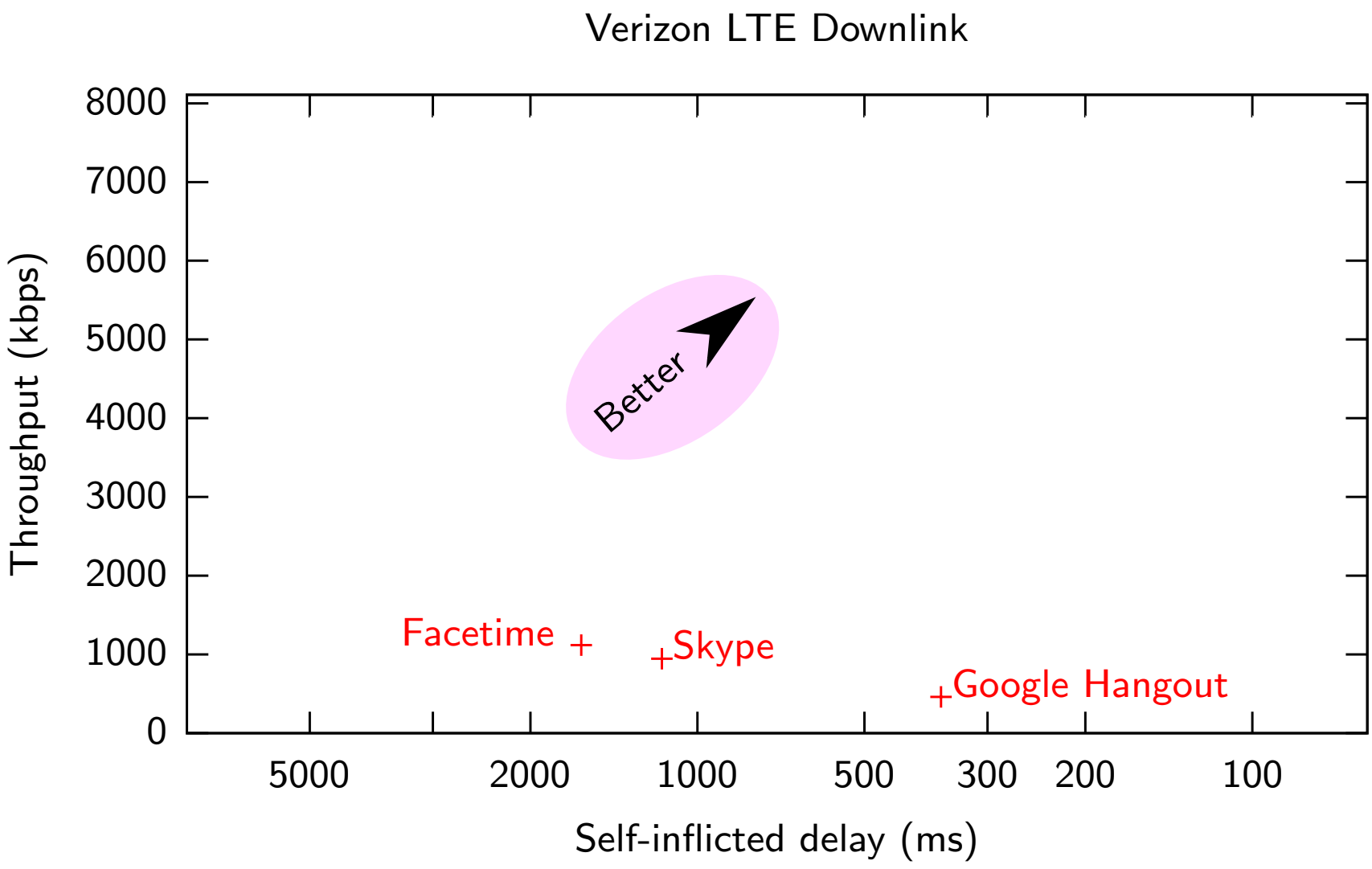
Control: fill up 100 ms forecast window

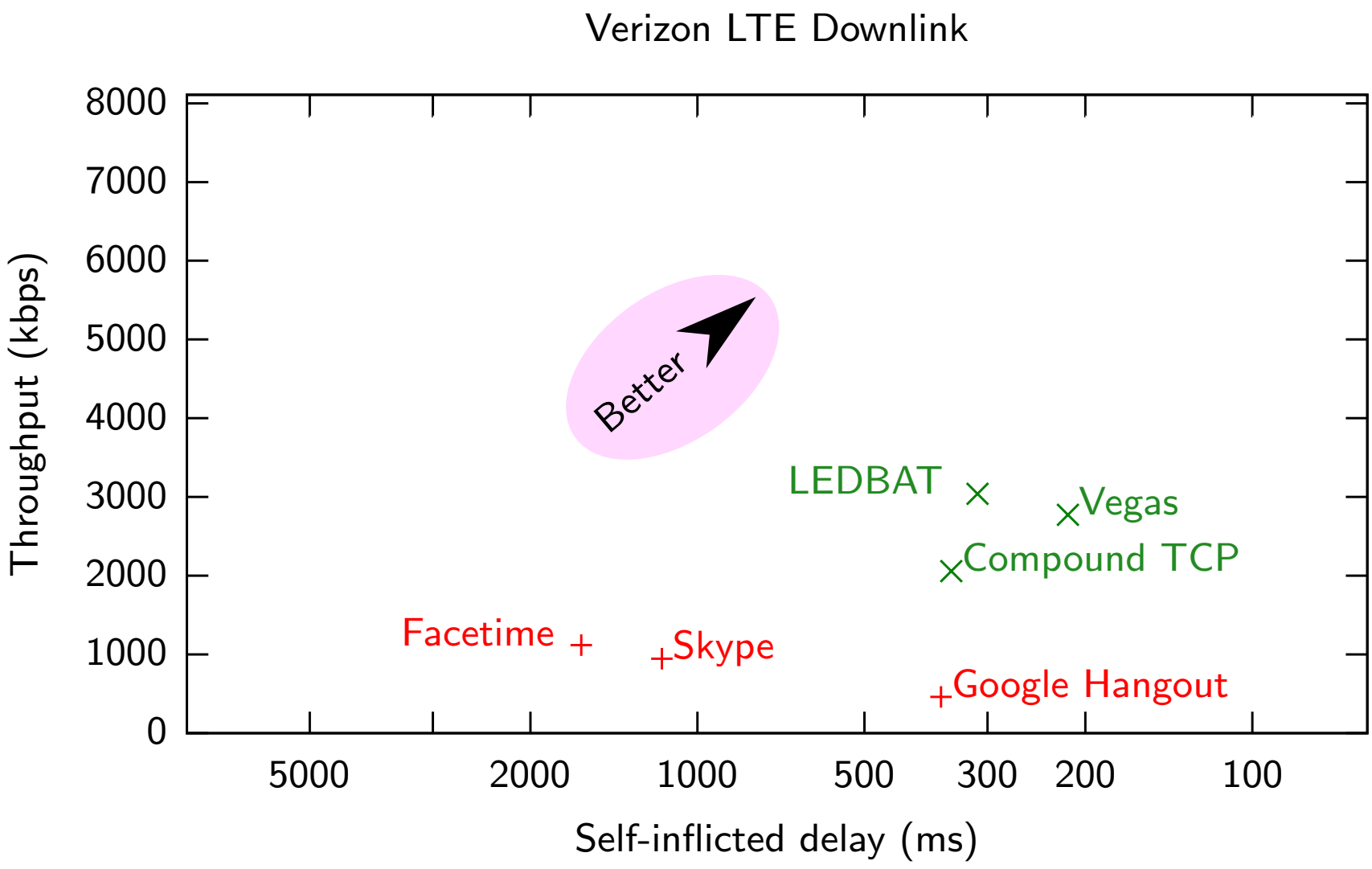


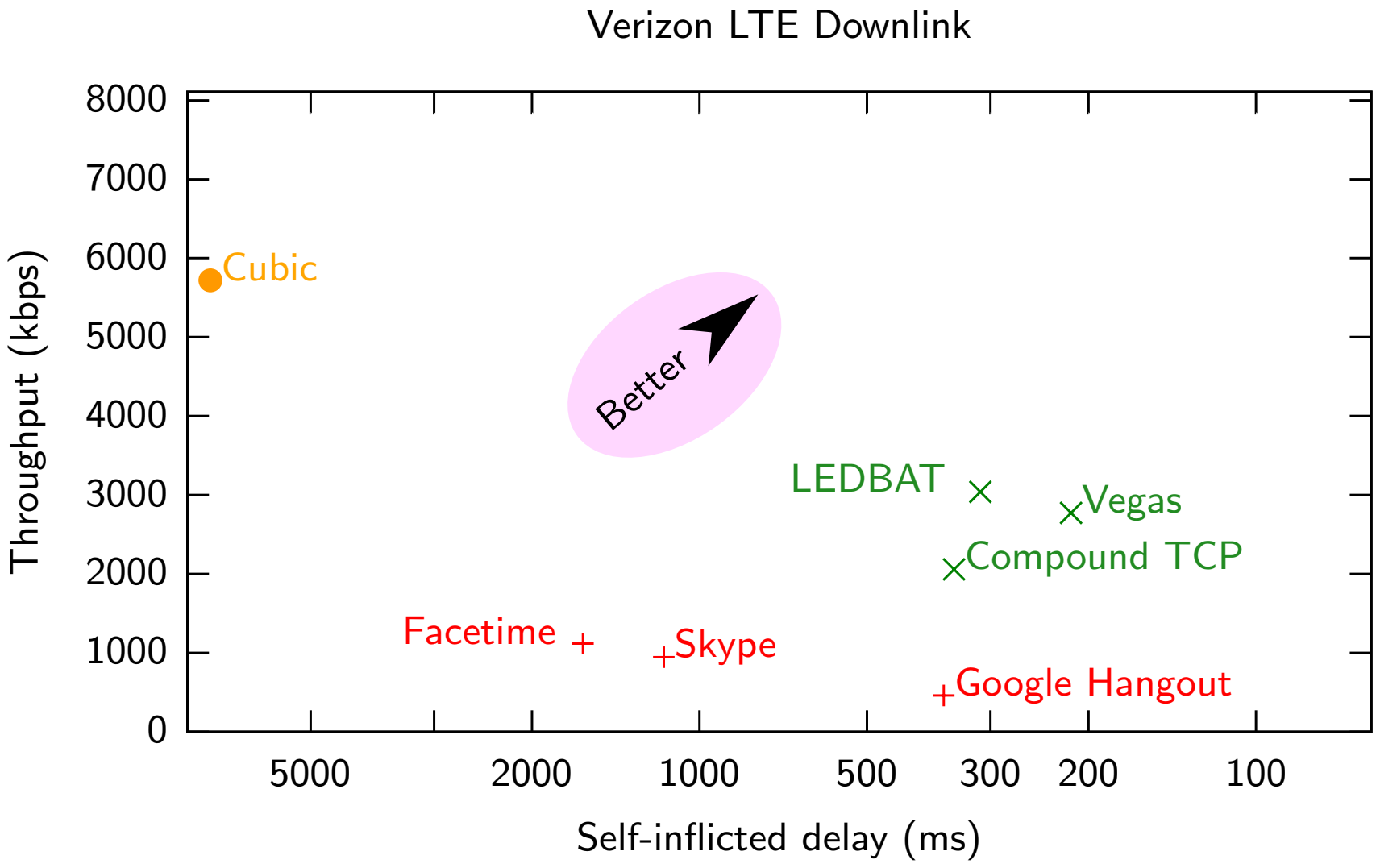
Control: fill up 100 ms forecast window

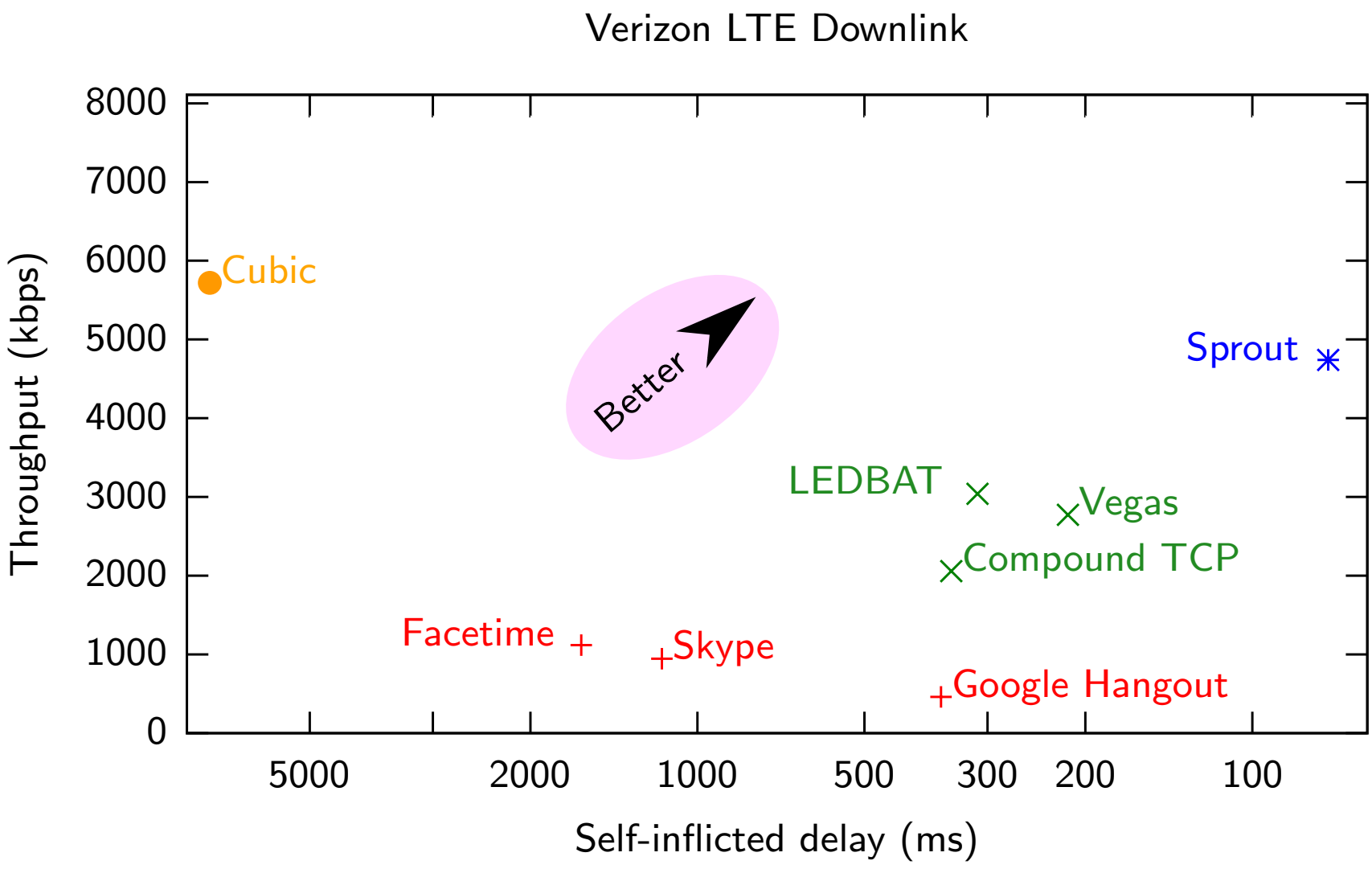


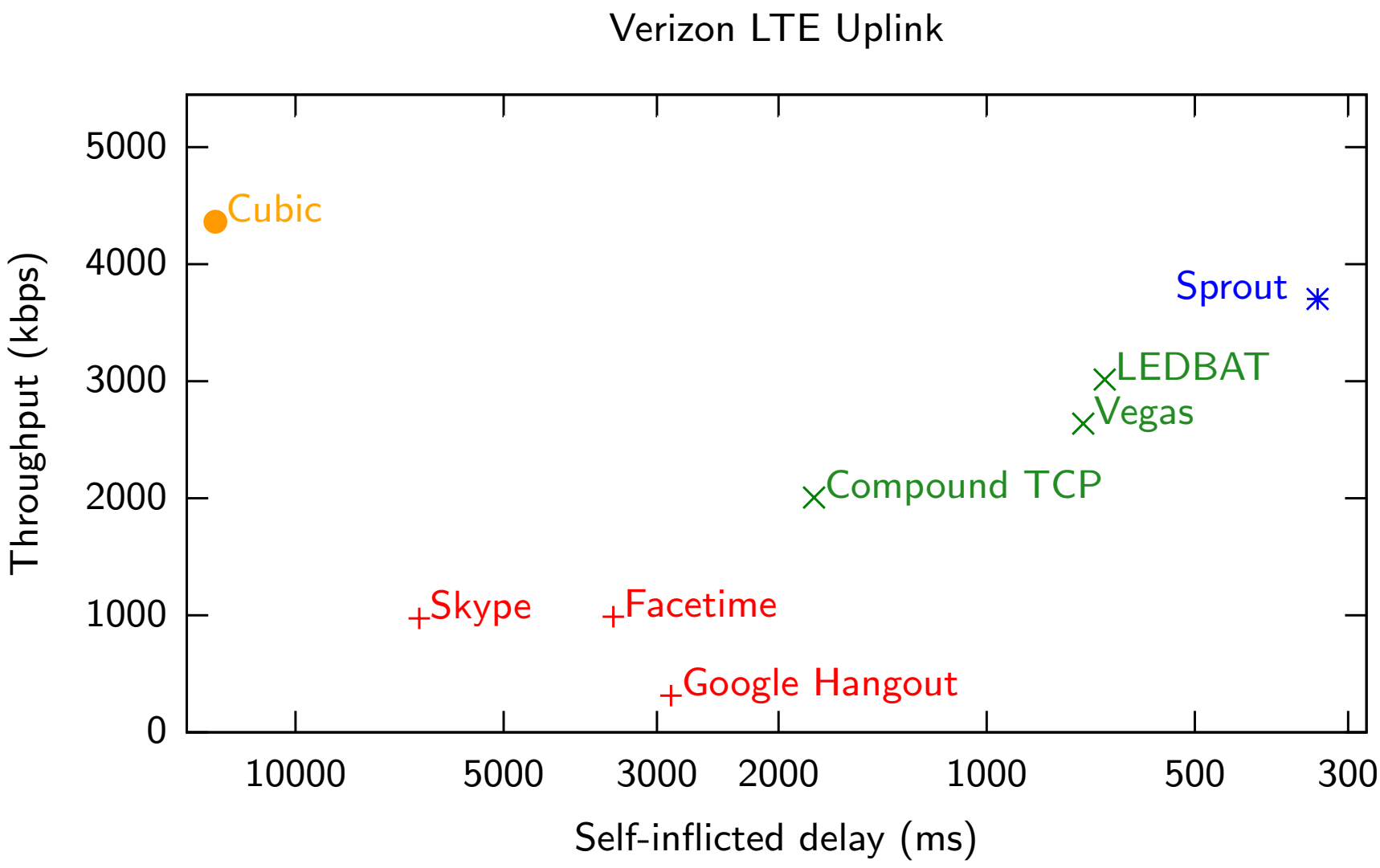
Sprout's results


Verizon LTE Downlink
Throughput (kbps)
8000
7000
6000
5000
4000
3000
2000
1000
0
Better
Facetime
Skype
Google Hangout
5000
2000
1000
500
300
200
100
Self-inflicted delay (ms)


Verizon LTE Downlink
Throughput (kbps)
8000
7000
6000
5000
4000
3000
2000
1000
0
5000
2000
1000
500
300
200
100
Self-inflicted delay (ms)
Better
LEDBAT
Vegas
Compound TCP
Facetime
Skype
Google Hangout


Verizon LTE Downlink
Throughput (kbps)
8000
7000
6000
5000
4000
3000
2000
1000
0
Cubic
Better
LEDBAT
Vegas
Compound TCP
Facetime
Skype
Google Hangout
5000
2000
1000
500
300
200
100
Self-inflicted delay (ms)


Verizon LTE Downlink
Throughput (kbps)
8000
7000
6000
5000
4000
3000
2000
1000
0
Cubic
Better
Sprout
LEDBAT
Vegas
Compound TCP
Facetime
Skype
Google Hangout
5000
2000
1000
500
300
200
100
Self-inflicted delay (ms)

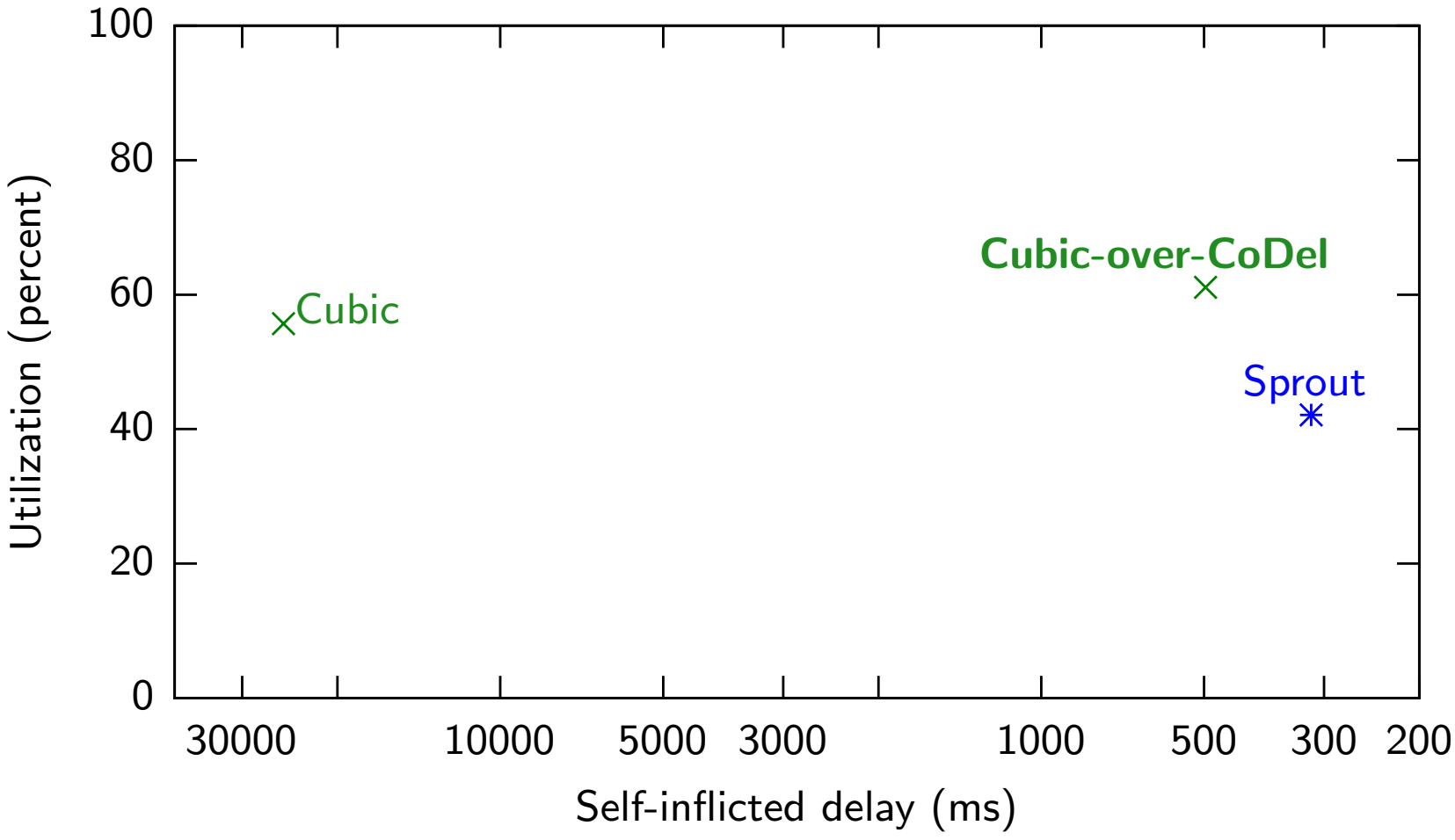

Verizon LTE Uplink
Throughput (kbps)
5000
4000
3000
2000
1000
0
Cubic
Sprout
LEDBAT
Vegas
Compound TCP
Skype
Facetime
Google Hangout
10000
5000
3000
2000
1000
500
300
Self-inflicted delay (ms)

Overall results on 8 links

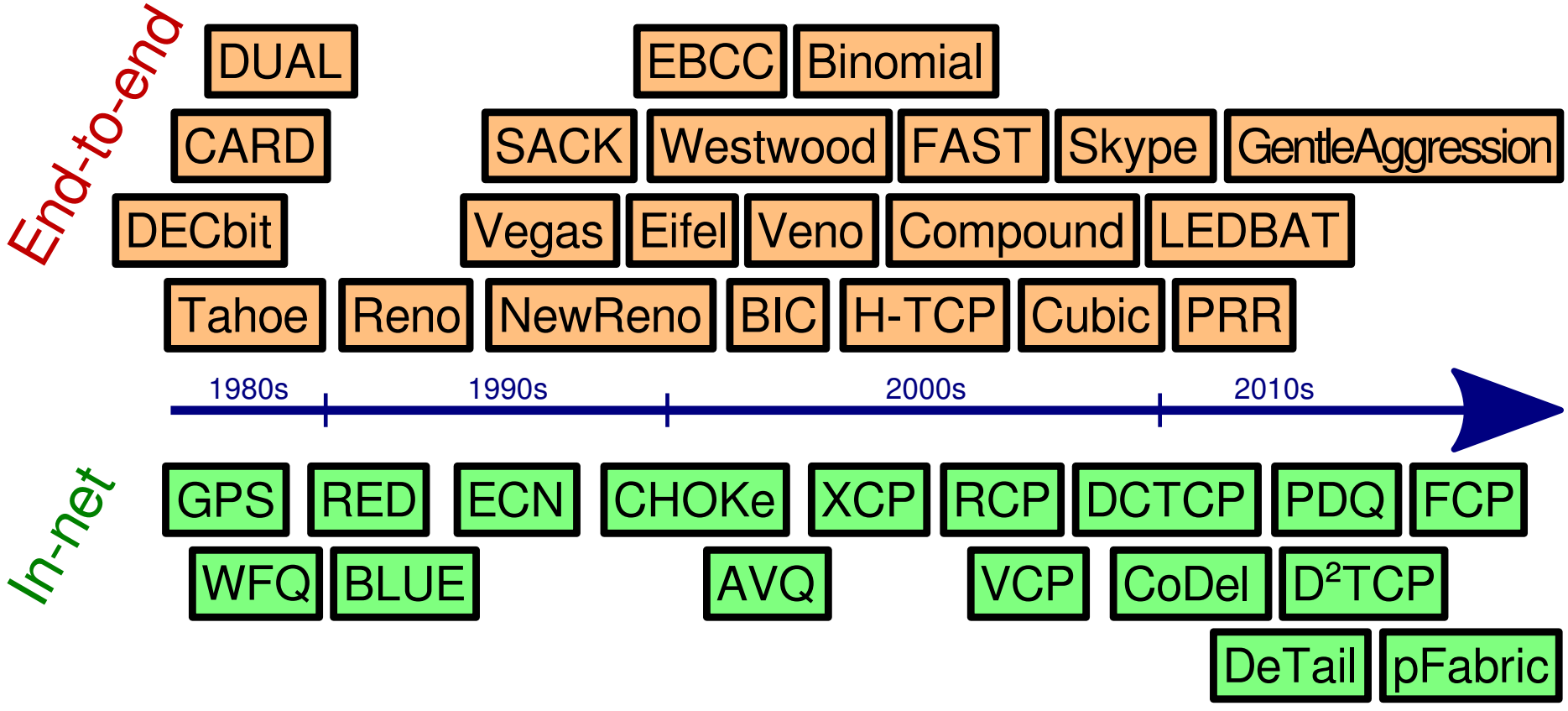
Verizon 3G/LTE, AT&T LTE, T-Mobile 3G uplink and downlink:

Sprout vs.	Avg. speedup	Delay reduction
Skype	2.2×	7.9×
Hangout	4.4×	7.2×
Facetime	1.9×	8.7×
Compound	1.3×	4.8×
TCP Vegas	1.1×	2.1×
LEDBAT	Same	2.8×
Cubic	0.91×	79×

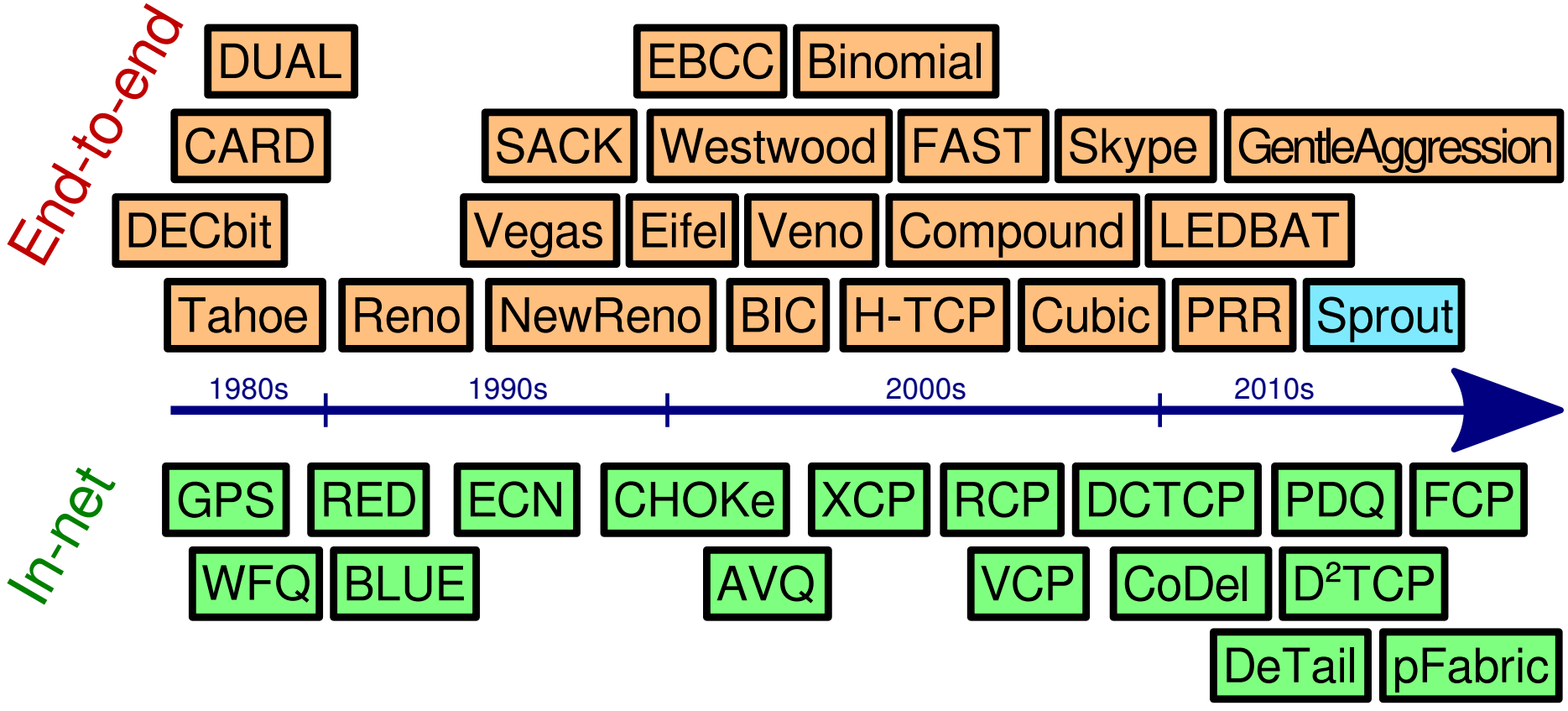
Sprout is end-to-end, but comparable to in-net control



Sprout's mark



Sprout's mark



Question

Can a protocol do even better?

Congestion-control contest (2013–2016)

- Turnkey network emulator, evaluation
- Sender, receiver run in Linux containers
- **Mission**: maximize throughput/delay
- 4th prize: $20
- 3rd prize: $30
- 2nd prize: $40
- (If beat Sprout) 1st prize:

Congestion-control contest (2013–2016)

- Turnkey network emulator, evaluation
- Sender, receiver run in Linux containers
- **Mission**: maximize throughput/delay
- 4th prize: $20
- 3rd prize: $30
- 2nd prize: $40
- (If beat Sprout) 1st prize: **Co-authorship on future paper**

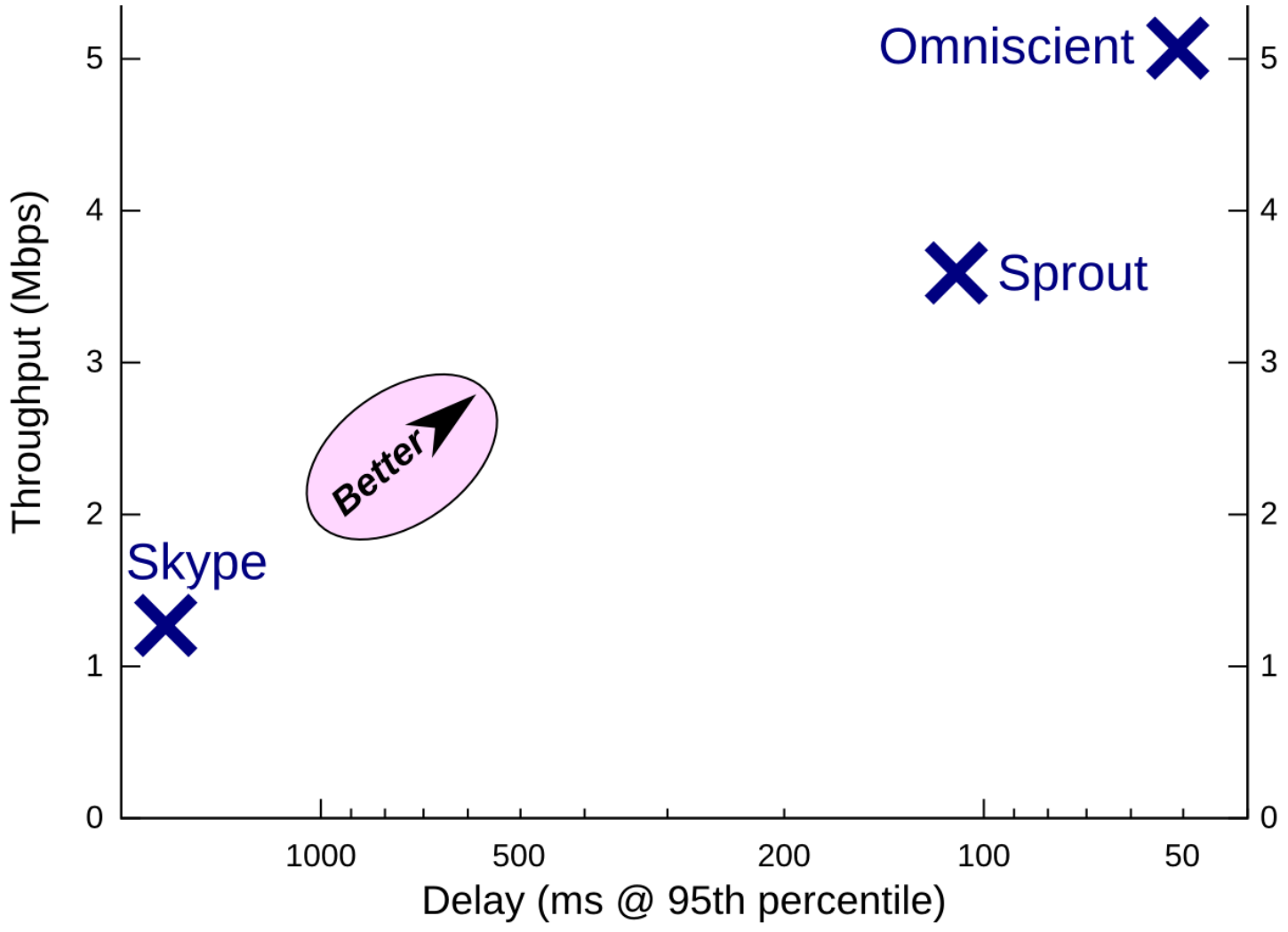
Congestion-control contest (2013–2016)

- Turnkey network emulator, evaluation
- Sender, receiver run in Linux containers
- **Mission**: maximize throughput/delay
- 4th prize: $20
- 3rd prize: $30
- 2nd prize: $40
- (If beat Sprout) 1st prize: **Co-authorship on future paper**

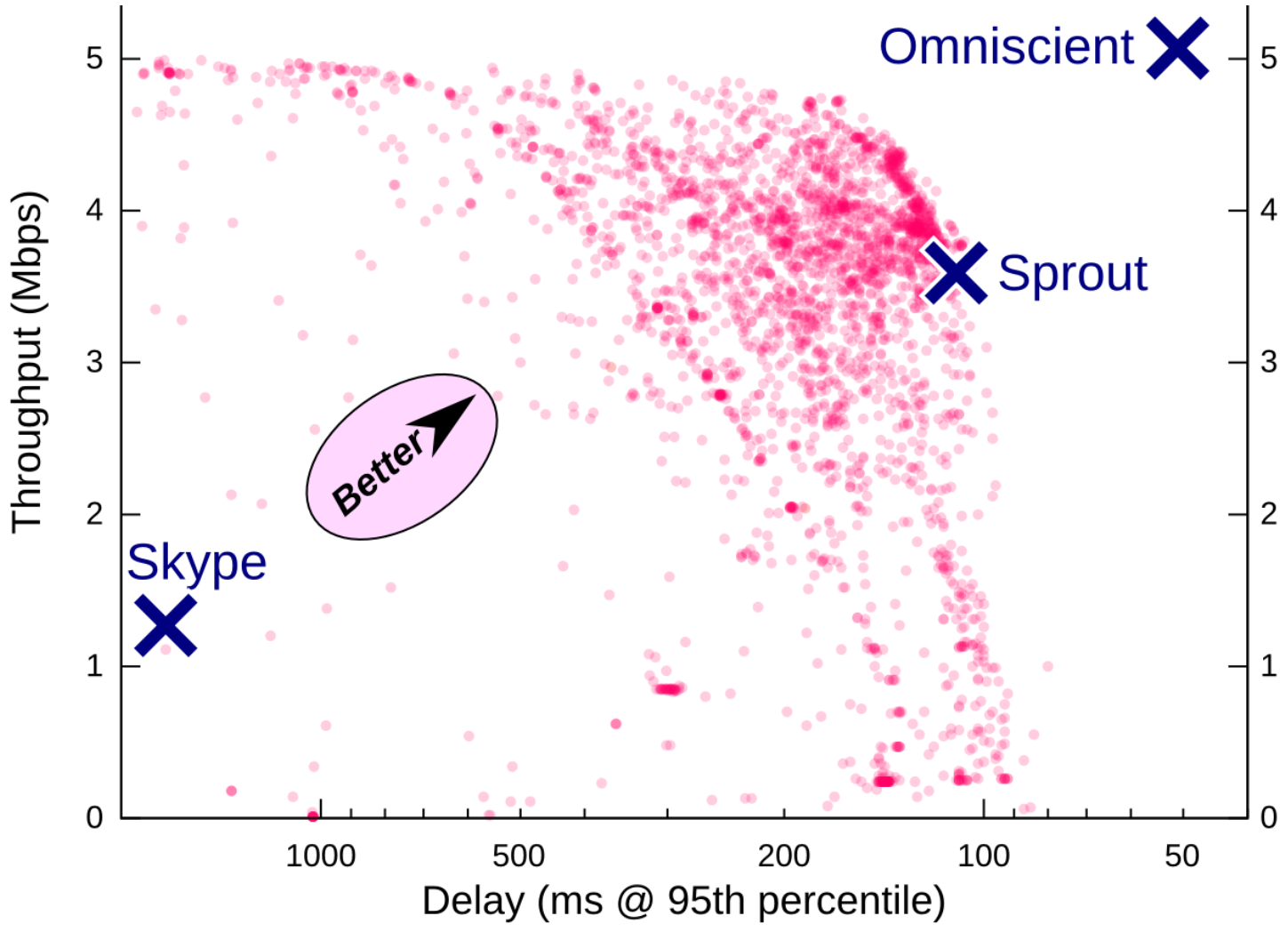
Anirudh Sivaraman, KW, Pauline Varley, Somak Das, Joshua Ma, Ameesh Goyal, João Batalha, and Hari Balakrishnan, **Protocol Design Contests**, *ACM SIGCOMM Computer Communications Review*, July 2014

[mahimahi demo]

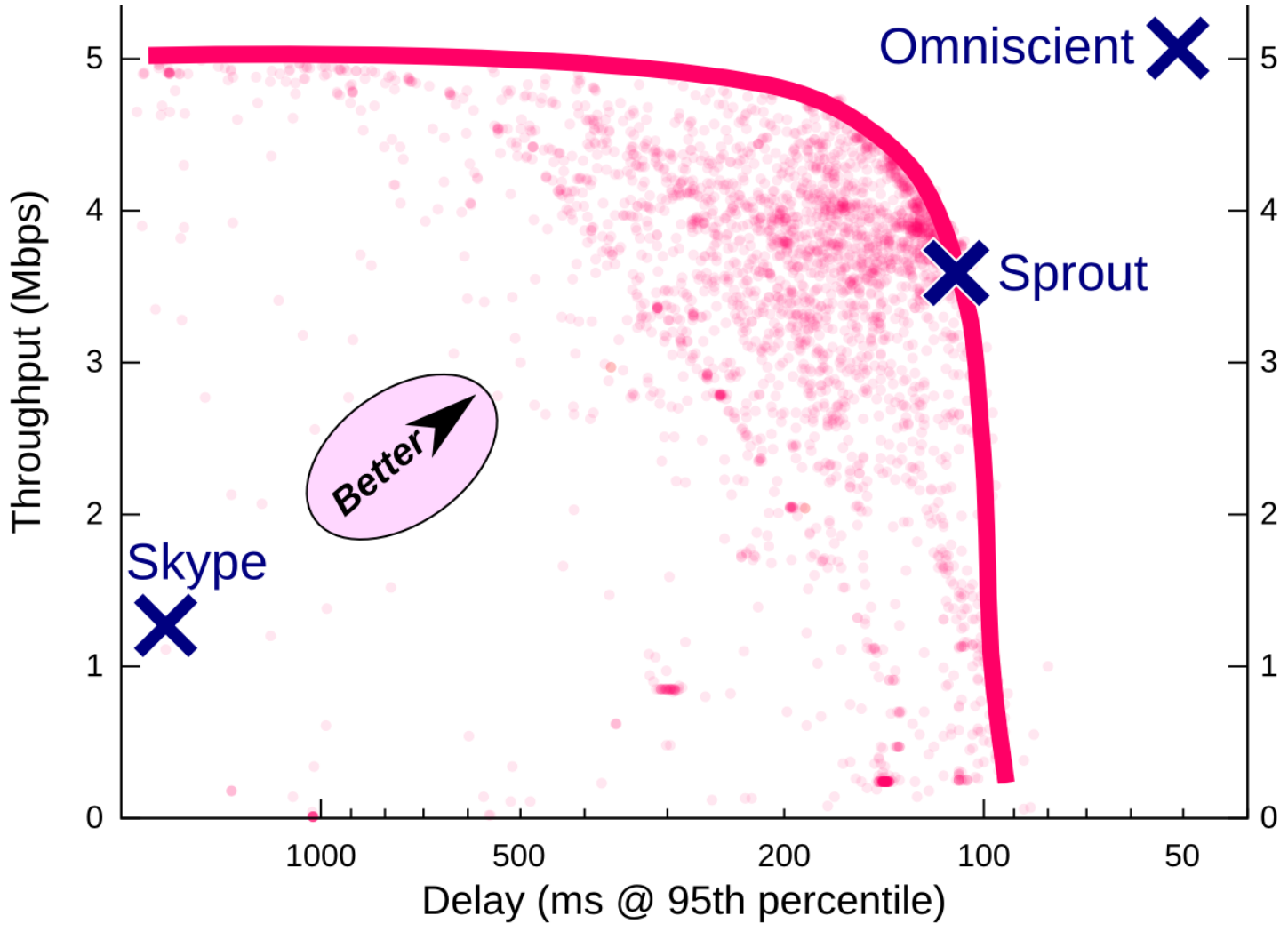
Baseline


Omniscient
Sprout
Skype
Better
Throughput (Mbps)
Delay (ms @ 95th percentile)
0
1
2
3
4
5
1000
500
200
100
50

Land of 3,000 student protocols



Sprout was on the frontier



[contest website]

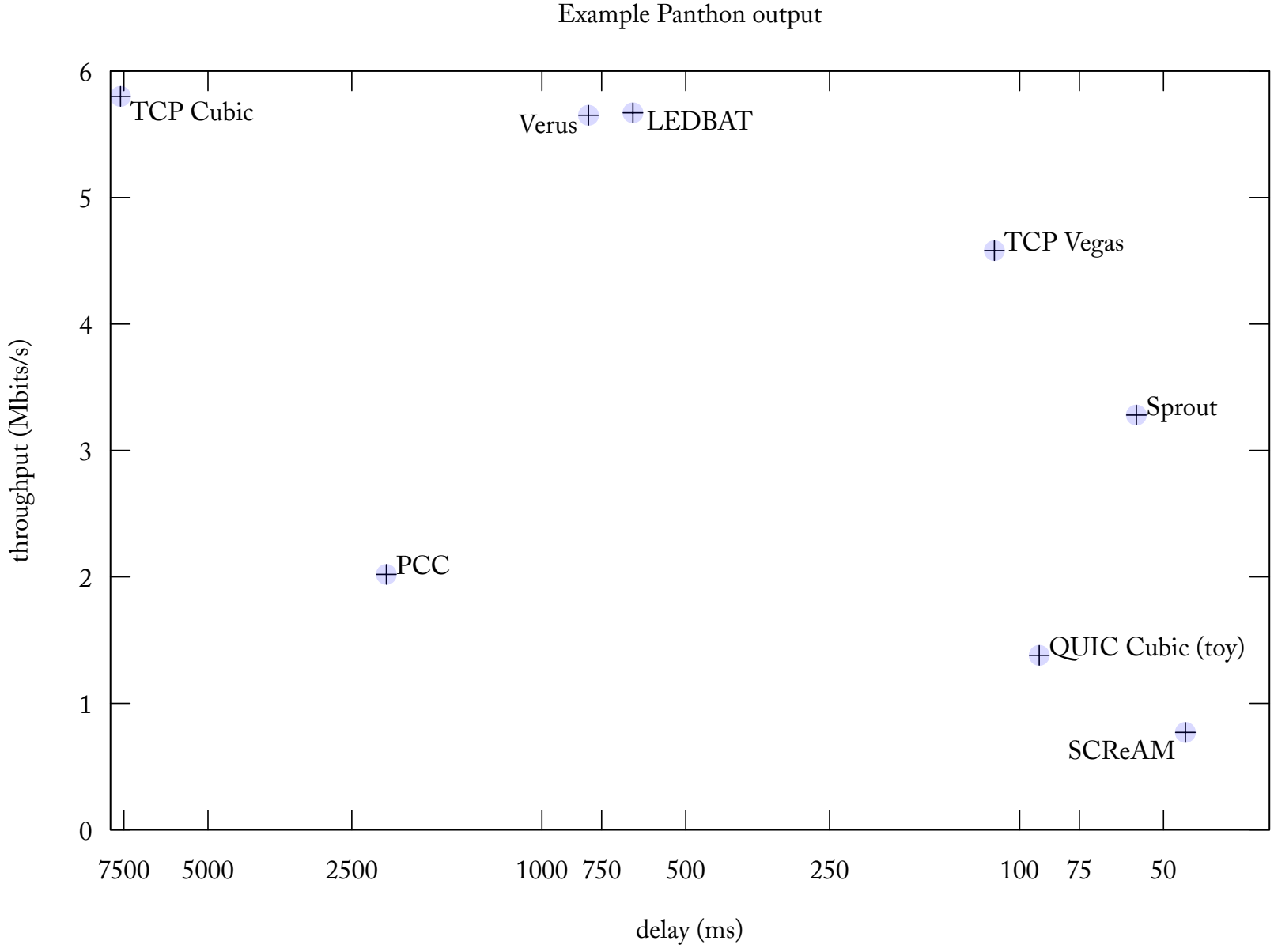
What does program synthesis tell us?

- The best algorithms use *multiple* signals.
  - Delay *vs.* ECN *vs.* loss is a false choice.
- The best algorithms use *multiple* behaviors.
  - Used both pacing and ack-clocked transmission within an RTT.

Congestion control in the real world

- Rural African medical clinics
- E.g., flow cytometer, glucose monitor, TB assay
- Cellular network often too flaky to upload
- For outages: **clinic employee on motorcycle to hilltop**

[pantheon demo]


Example Panthon output
throughput (Mbits/s)
delay (ms)
0
1
2
3
4
5
6
7500
5000
2500
1000
750
500
250
100
75
50
TCP Cubic
Verus
LEDBAT
TCP Vegas
Sprout
PCC
QUIC Cubic (toy)
SCReAM

[netobservatory.io]

Redundant Array of Incentivized Links

- How to solve outages?
- Proposal: encrypt a bundle and pay bounty to **anybody** who first uploads it
- Incentivize villagers to use their own creativity

The Pantheon of Congestion Control

- Let's all work together to test each other's algorithms
- Let's build emulators that approximate real-world results
- Weekly real-world game for anybody who wants to play
- Users don't even expect the system to work. Let's fix it!

Keith Winstein
keithw@cs.stanford.edu