TCP CONGESTION SIGNATURES

Srikanth Sundaresan (Princeton Univ.)
Amogh Dhamdhere (CAIDA/UCSD)
kc Claffy (CAIDA/UCSD)
Mark Allman (ICSI)

INTERNATIONAL COMPUTER SCIENCE INSTITUTE

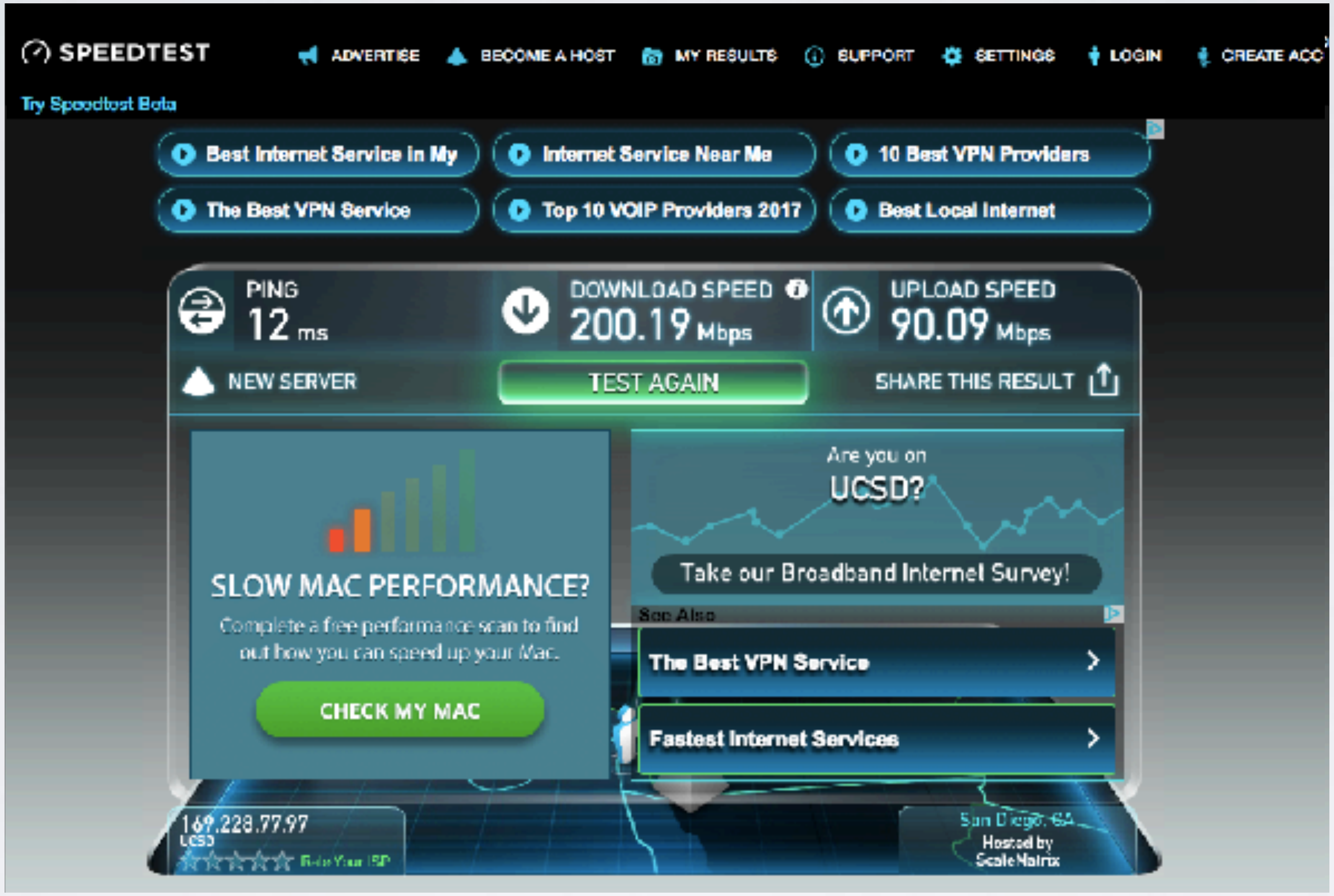
PRINCETON UNIVERSITY

caida

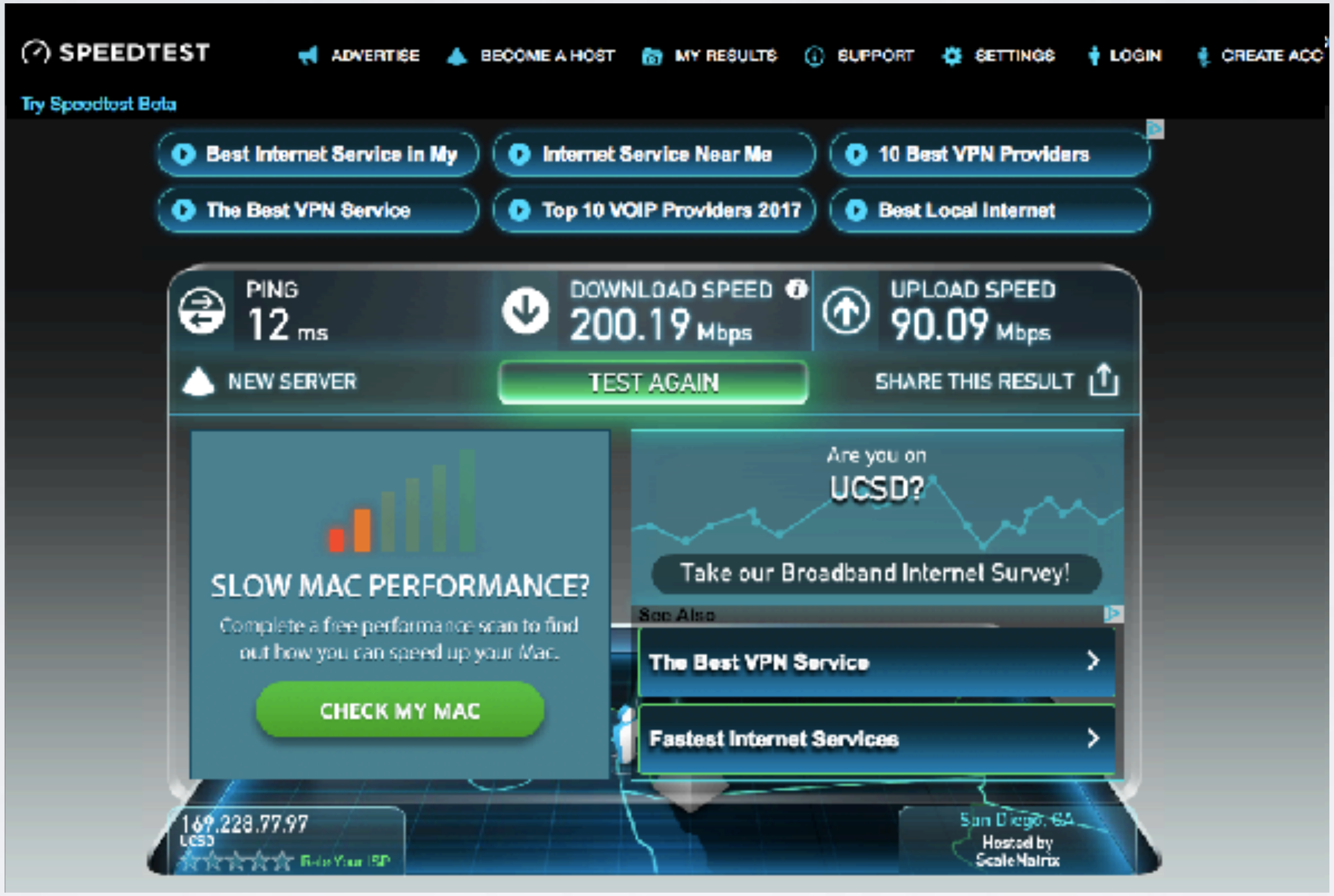
Typical Speed Tests Don't Tell Us Much

Typical Speed Tests Don't Tell Us Much

Typical Speed Tests Don't Tell Us Much

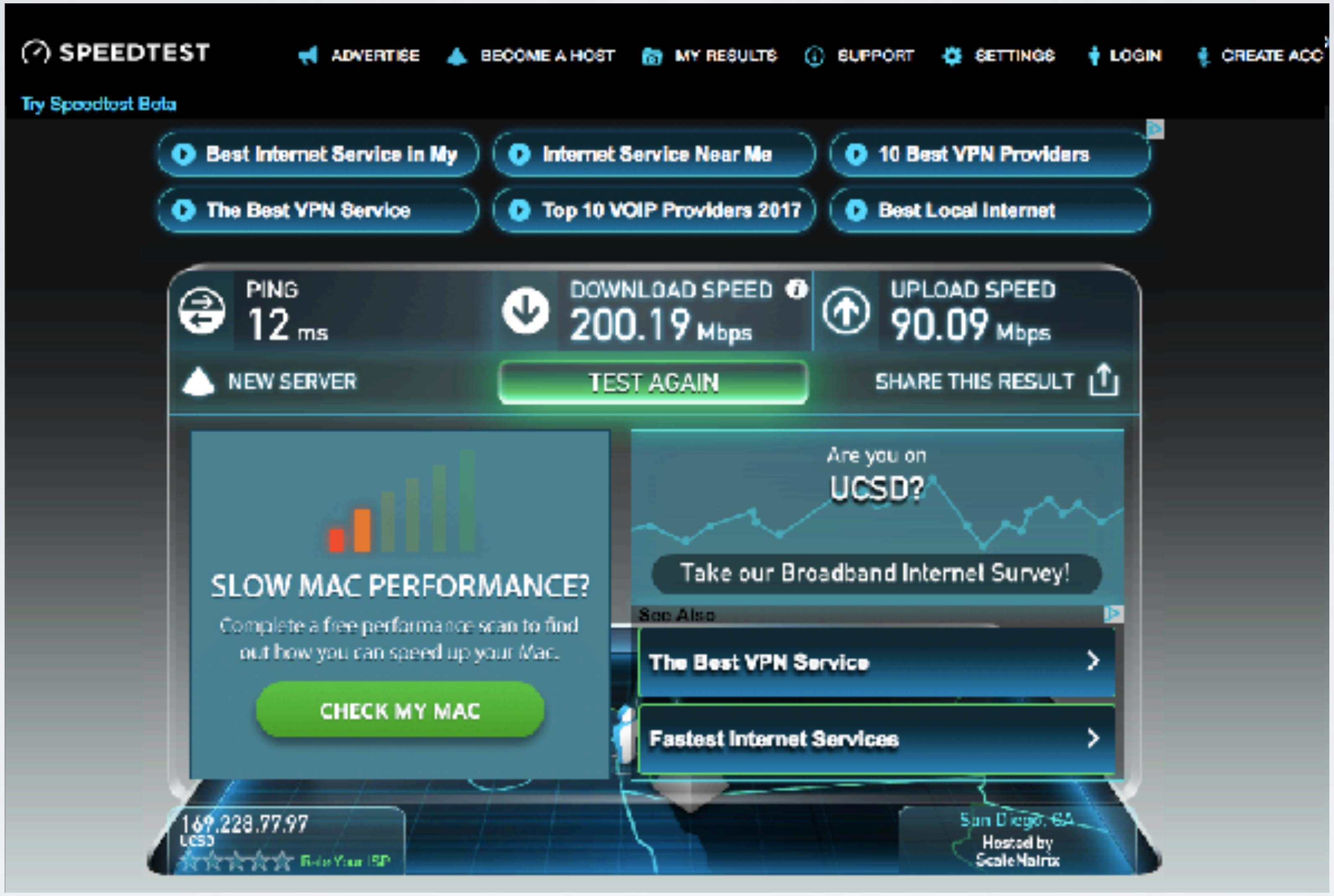


Typical Speed Tests Don't Tell Us Much



- Upload and download throughput measurements: no information beyond that

Typical Speed Tests Don't Tell Us Much



What type of congestion did the TCP flow experience?

Two Potential Sources of Congestion in the End-to-end Path

Two Potential Sources of Congestion in the End-to-end Path

- Self-induced congestion
  - Clear path, the flow itself induced congestion
  - eg: last-mile access link

Two Potential Sources of Congestion in the End-to-end Path

- Self-induced congestion
  - Clear path, the flow itself induced congestion
  - eg: last-mile access link
- External congestion
  - Flow starts on an already congested path
  - eg: congested interconnect

Two Potential Sources of Congestion in the End-to-end Path

- Self-induced congestion
  - Clear path, the flow itself induced congestion
  - eg: last-mile access link
- External congestion
  - Flow starts on an already congested path
  - eg: congested interconnect

Distinguishing the two cases has implications for users / ISPs / regulators

How can we distinguish the two?

- Cannot distinguish using just throughput numbers
  - Access plan rates vary widely, and are typically not available to content / speed test providers
  - eg: Speed test reports 5 Mbps – is that the access link rate (DSL), or a congested path?

How can we distinguish the two?

- Cannot distinguish using just throughput numbers
  - Access plan rates vary widely, and are typically not available to content / speed test providers
  - eg: Speed test reports 5 Mbps – is that the access link rate (DSL), or a congested path?

We can use the dynamics of TCP's startup phase, i.e., Congestion Signatures

TCP's RTT Congestion Signatures

TCP's RTT Congestion Signatures

- Flows experiencing self-induced congestion fill up an empty buffer during slow start
  - Hence increase the TCP flow RTT

TCP's RTT Congestion Signatures

- Flows experiencing self-induced congestion fill up an empty buffer during slow start
  - Hence increase the TCP flow RTT
- Externally congested flows encounter an already full buffer
  - Less potential for RTT increases

TCP's RTT Congestion Signatures

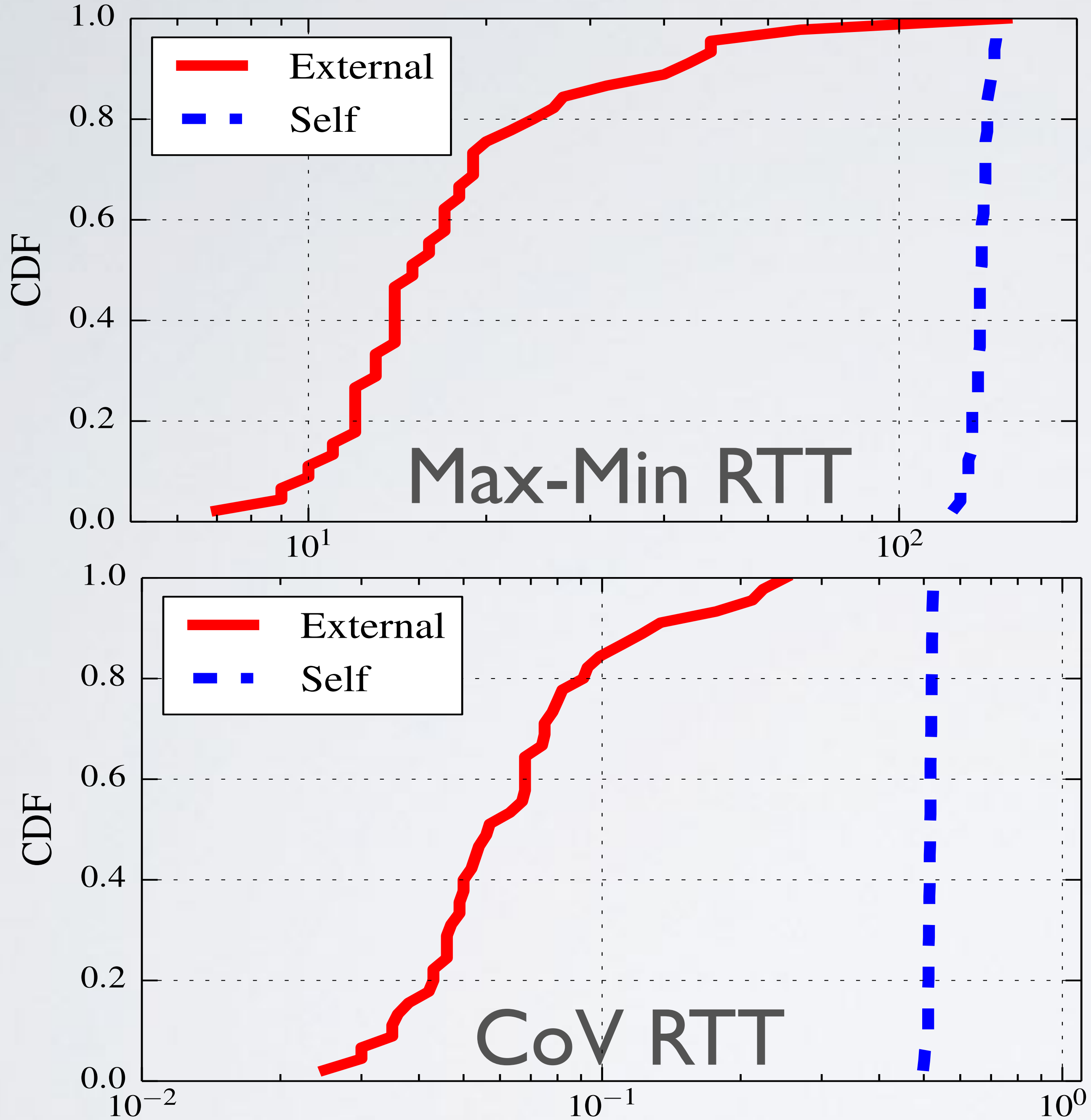
- Flows experiencing self-induced congestion fill up an empty buffer during slow start
  - Hence increase the TCP flow RTT
- Externally congested flows encounter an already full buffer
  - Less potential for RTT increases
- Self-induced congestion therefore has higher RTT variance compared to external congestion

TCP's RTT Congestion Signatures

- Flows experiencing self-induced congestion fill up an empty buffer during slow start
  - Hence increase the TCP flow RTT
- Externally congested flows encounter an already full buffer
  - Less potential for RTT increases
- Self-induced congestion therefore has higher RTT variance compared to external congestion

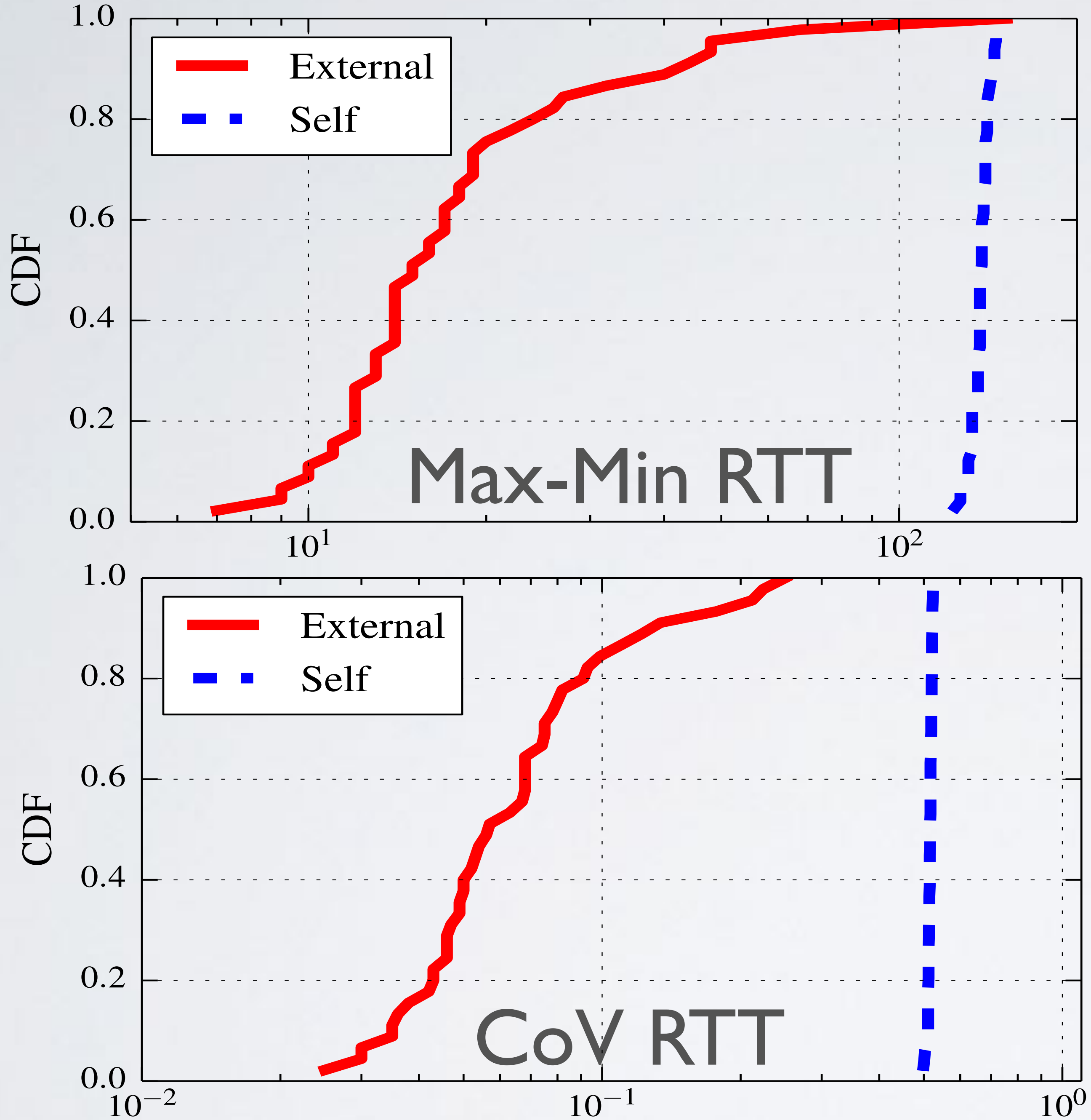
We can quantify this using Max-Min and CoV of RTT

Example Controlled Experiment



- 20 Mbps "access" link with 100 ms buffer
- 1 Gbps "interconnect" link with 50 ms buffer
- Self-induced congestion flows have higher values for both metrics and are clearly distinguishable

Example Controlled Experiment



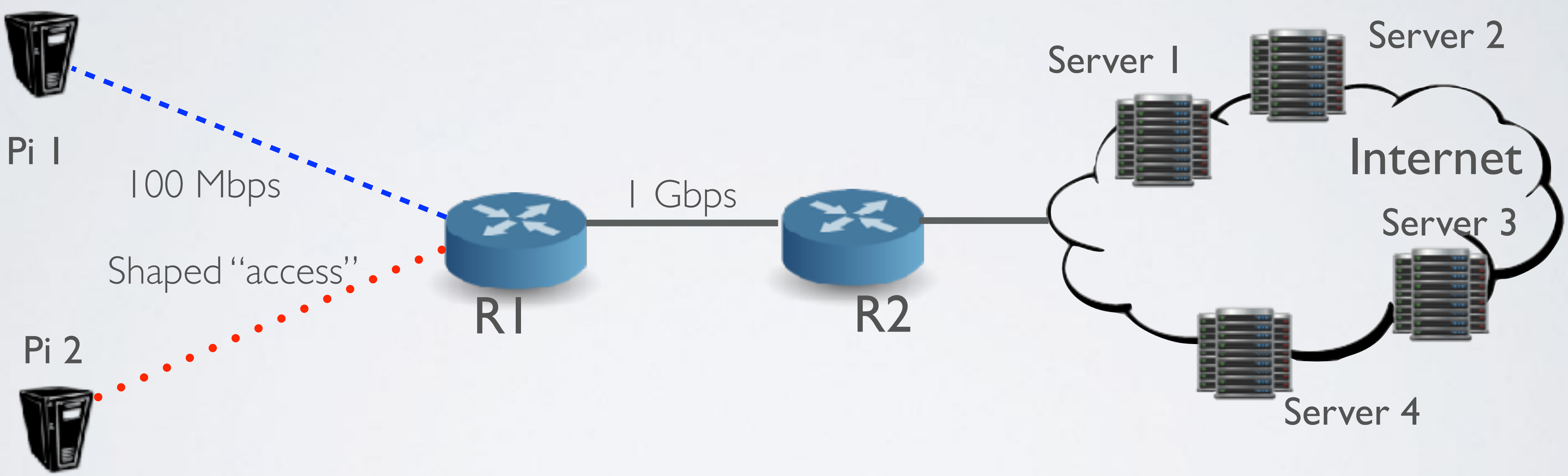
- 20 Mbps "access" link with 100 ms buffer
- 1 Gbps "interconnect" link with 50 ms buffer
- Self-induced congestion flows have higher values for both metrics and are clearly distinguishable

The two types of congestion exhibit widely contrasting behaviors

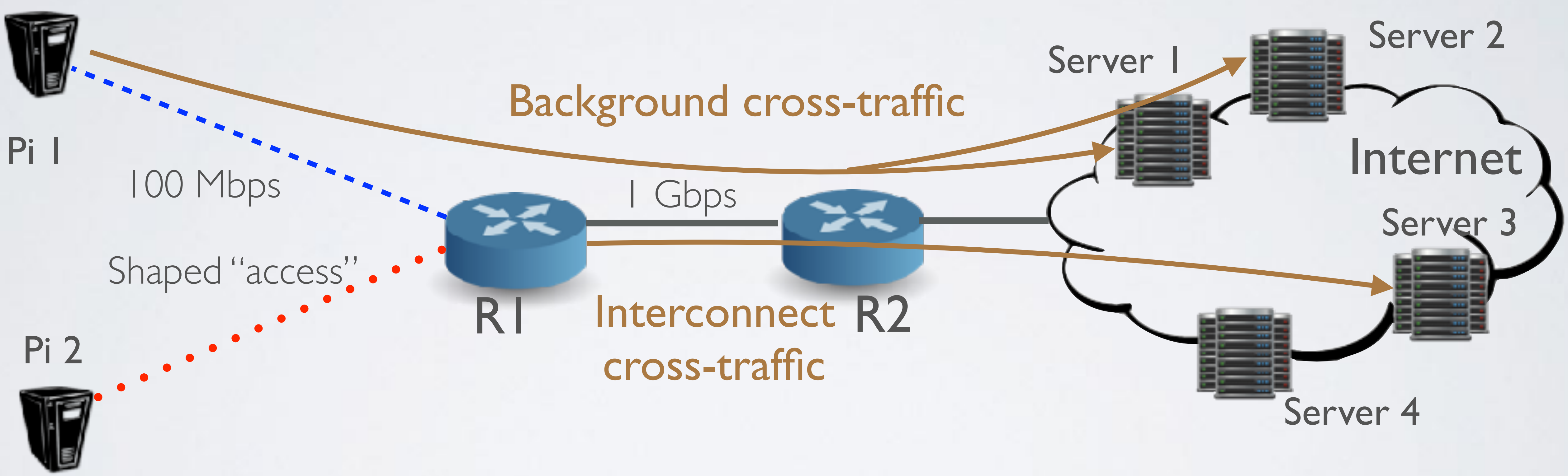
Model

- Max-min and CoV of RTT derived from RTT samples during slow start
- We feed the two metrics into a simple Decision Tree
  - We control the depth of the tree to a low value to minimize complexity
- We build the decision tree classifier using controlled experiments and apply it to real-world data

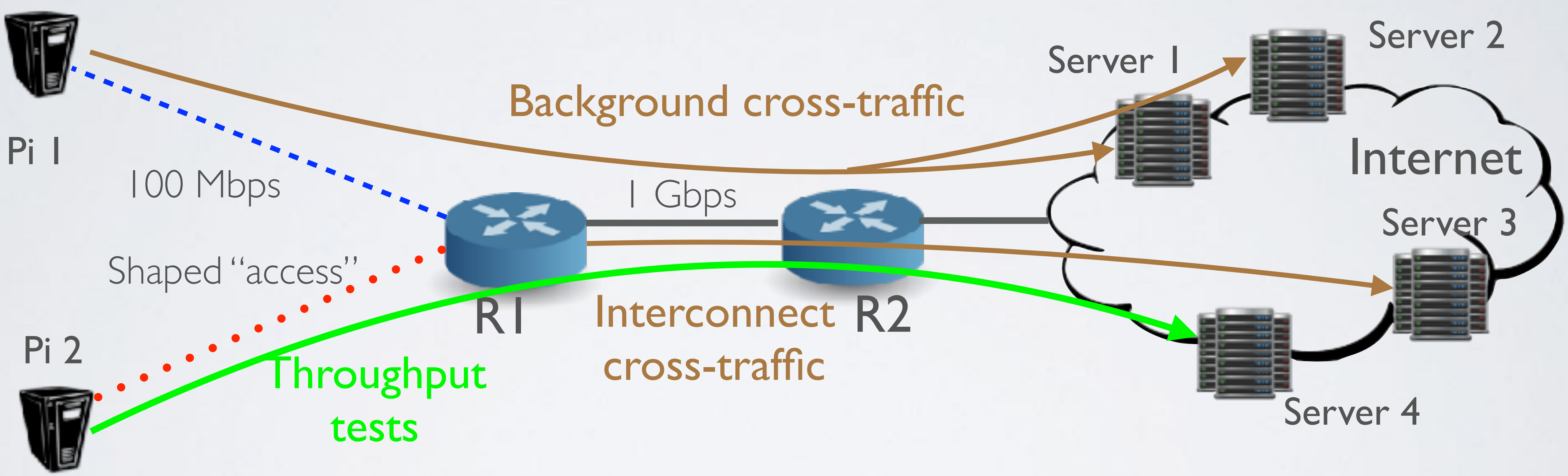
Validating the Method: Step 1 - Controlled Experiments



Validating the Method: Step 1 - Controlled Experiments



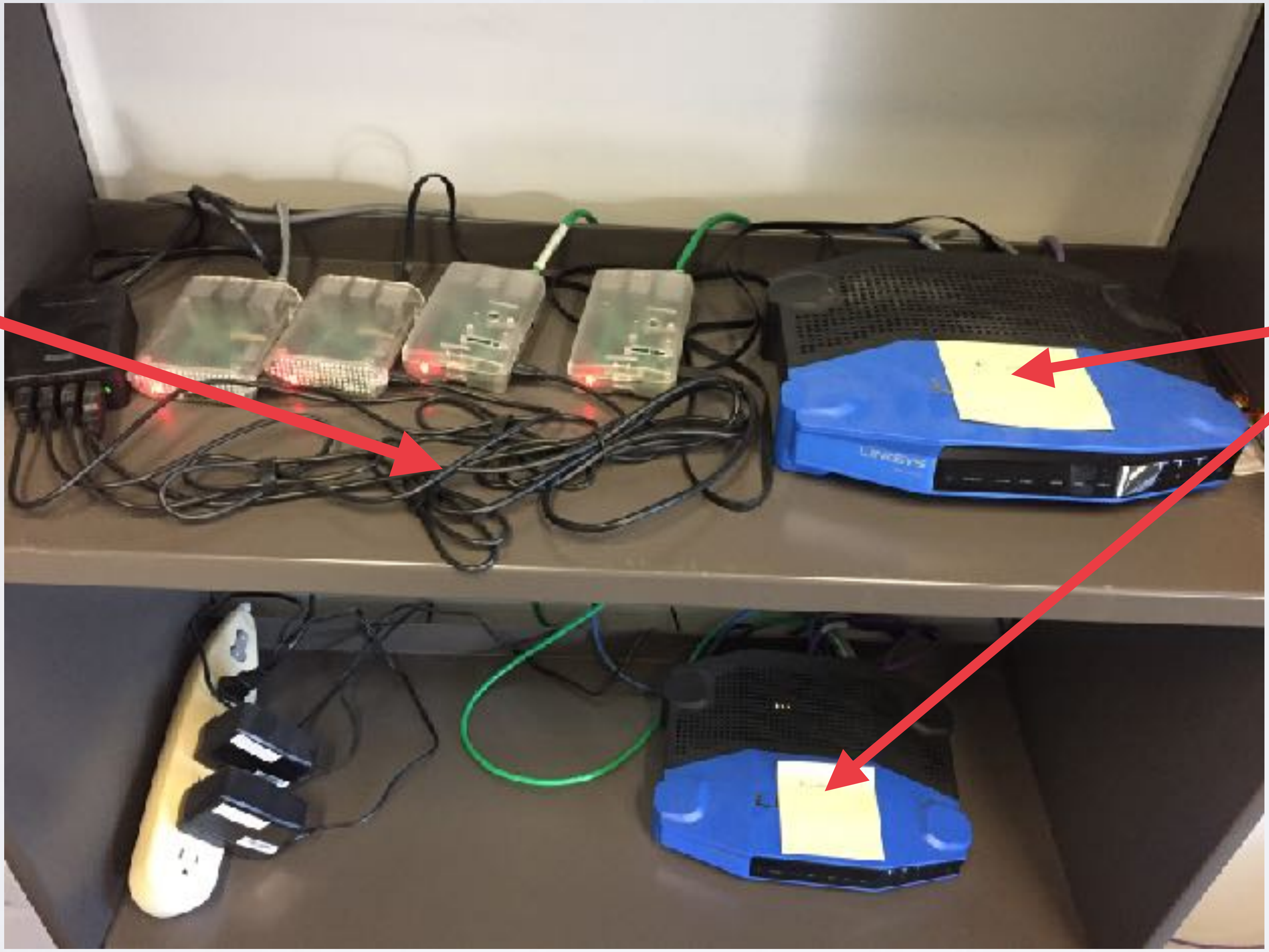
Validating the Method: Step 1 - Controlled Experiments



It's Real

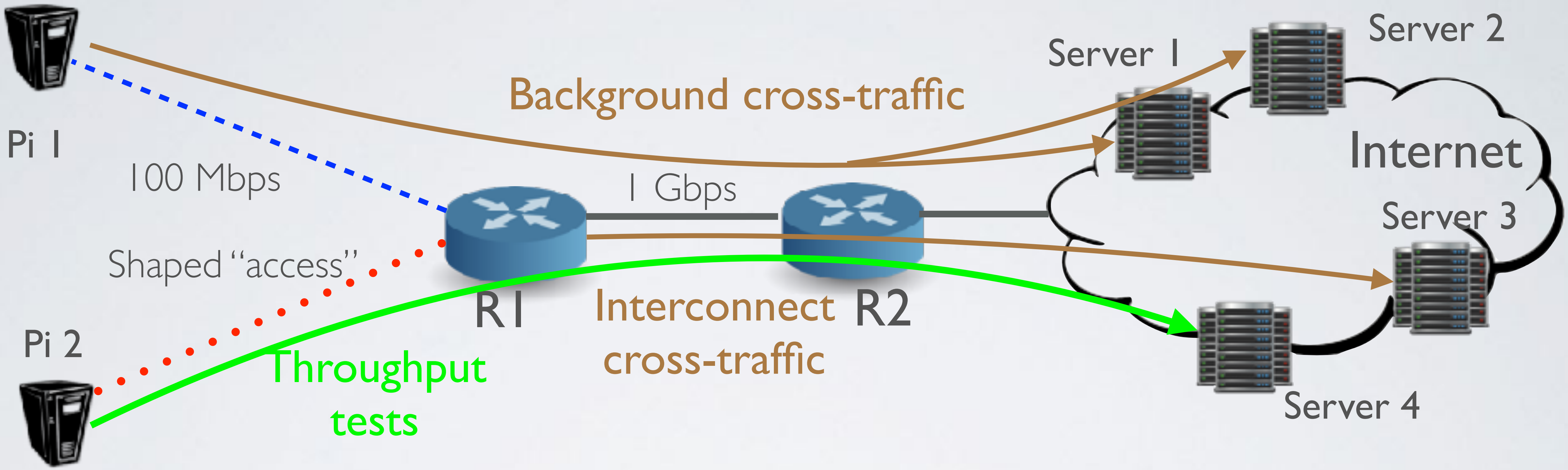
It's Real

Fantastic Cabling effort



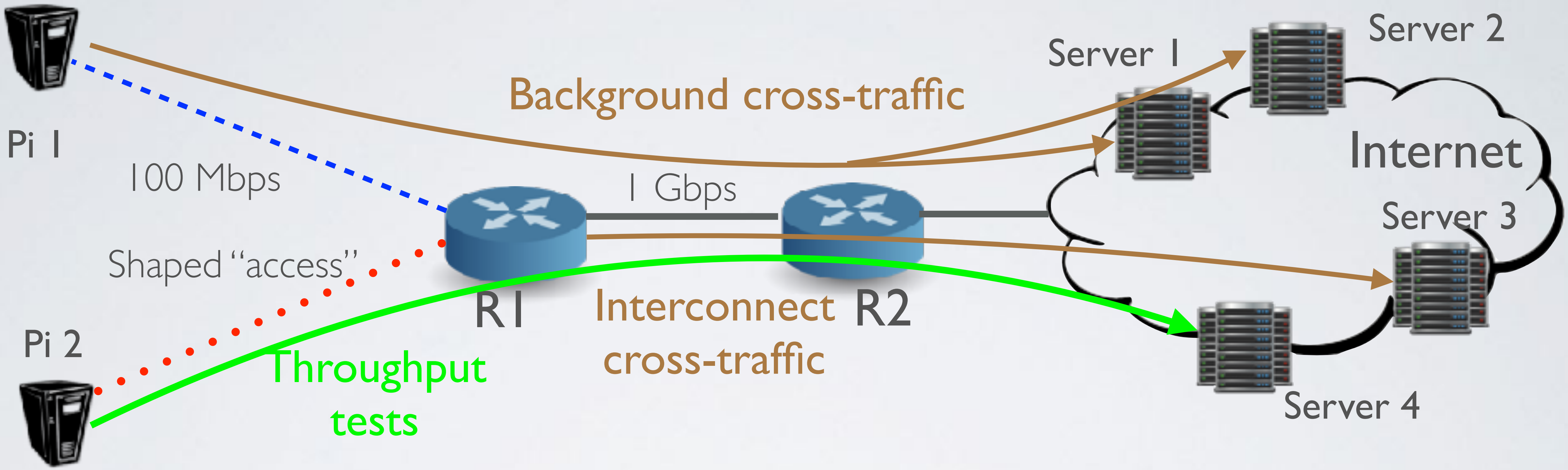
Post-it defined networking

Validating the Method: Step 1 - Controlled Experiments



- Emulated access link + "core" link
  - Wide range of access link throughputs, buffer sizes, loss rates, cross-traffic (background and congestion-inducing)
  - Can accurately label flows in training data as "self" or "externally" congested

Validating the Method: Step 1 - Controlled Experiments



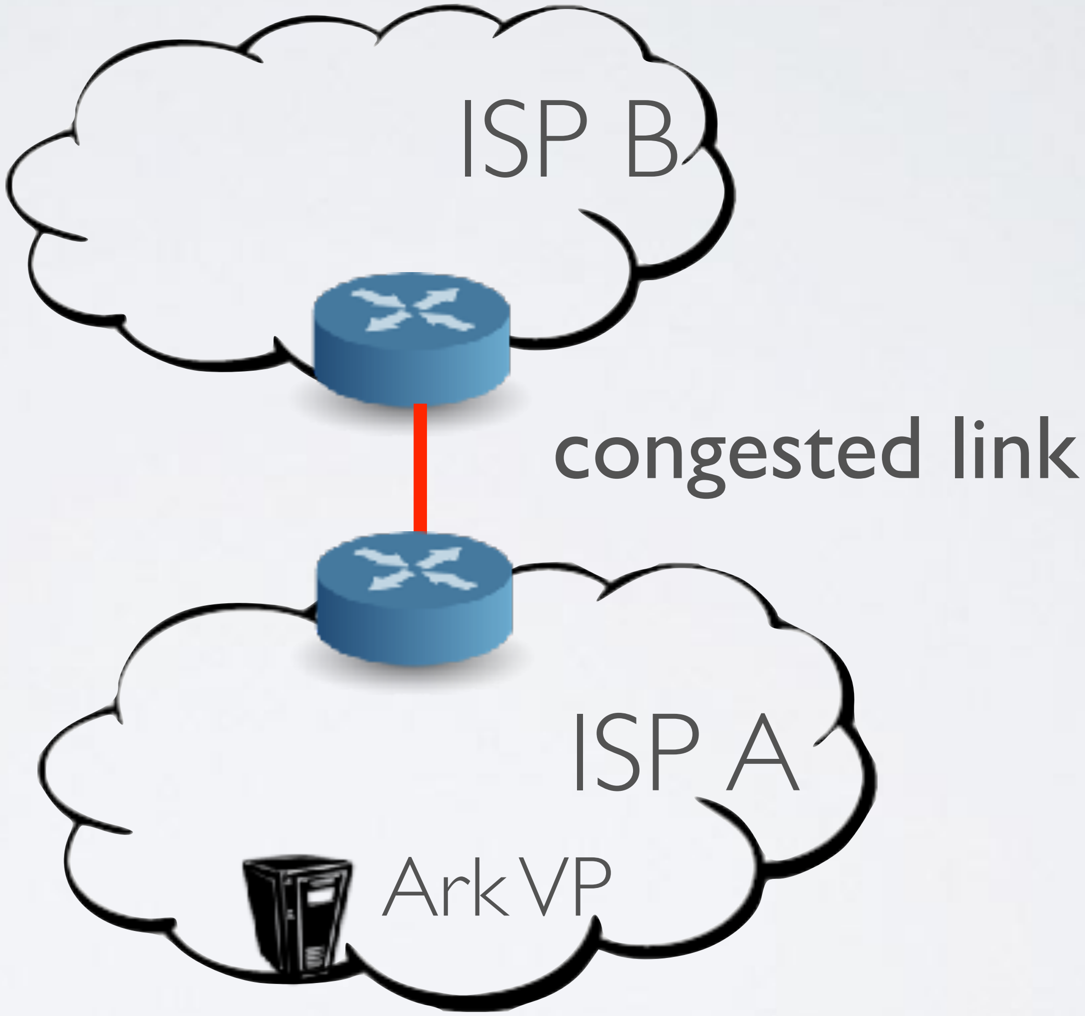
High accuracy: precision and recall > 80%
robust to model settings

Validating the Method: Step 2

- From Ark VP in ISP A identified congested link with ISP B using TSLP*

*Luckie et al. "Challenges in Inferring Internet Interdomain Congestion", IMC 2014

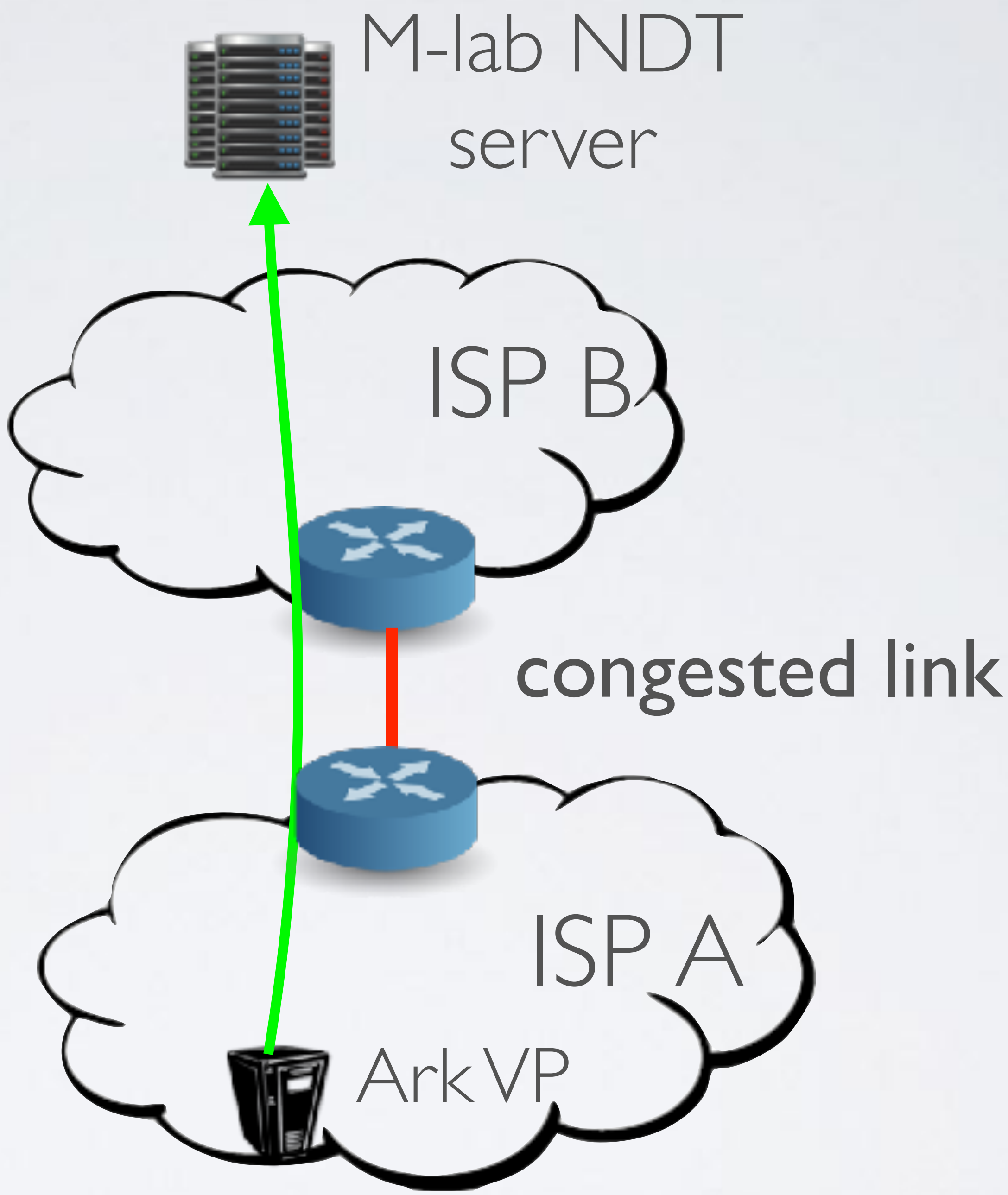
Validating the Method: Step 2



- From Ark VP in ISP A identified congested link with ISP B using TSLP*

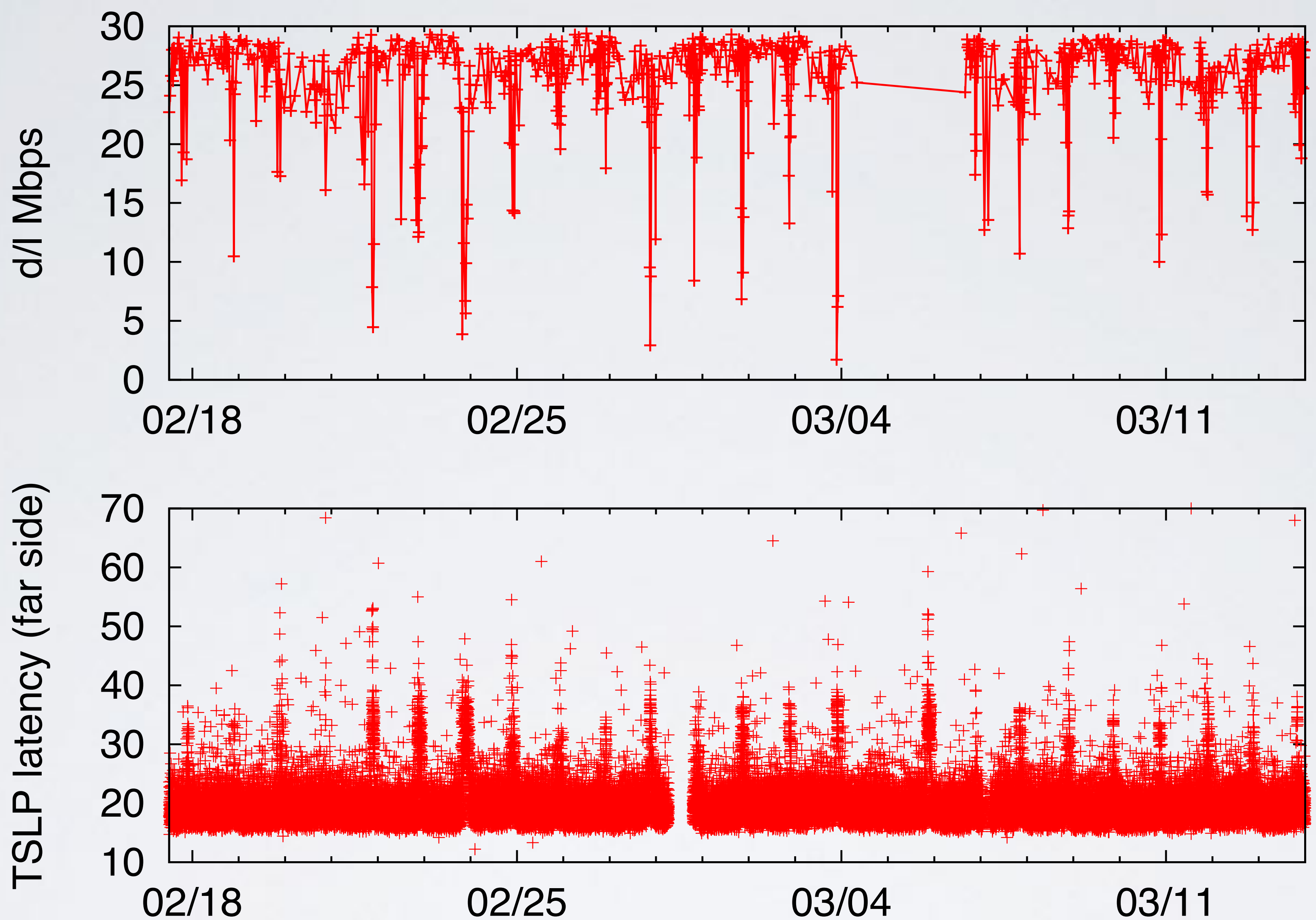
*Luckie et al. "Challenges in Inferring Internet Interdomain Congestion", IMC 2014

Validating the Method: Step 2



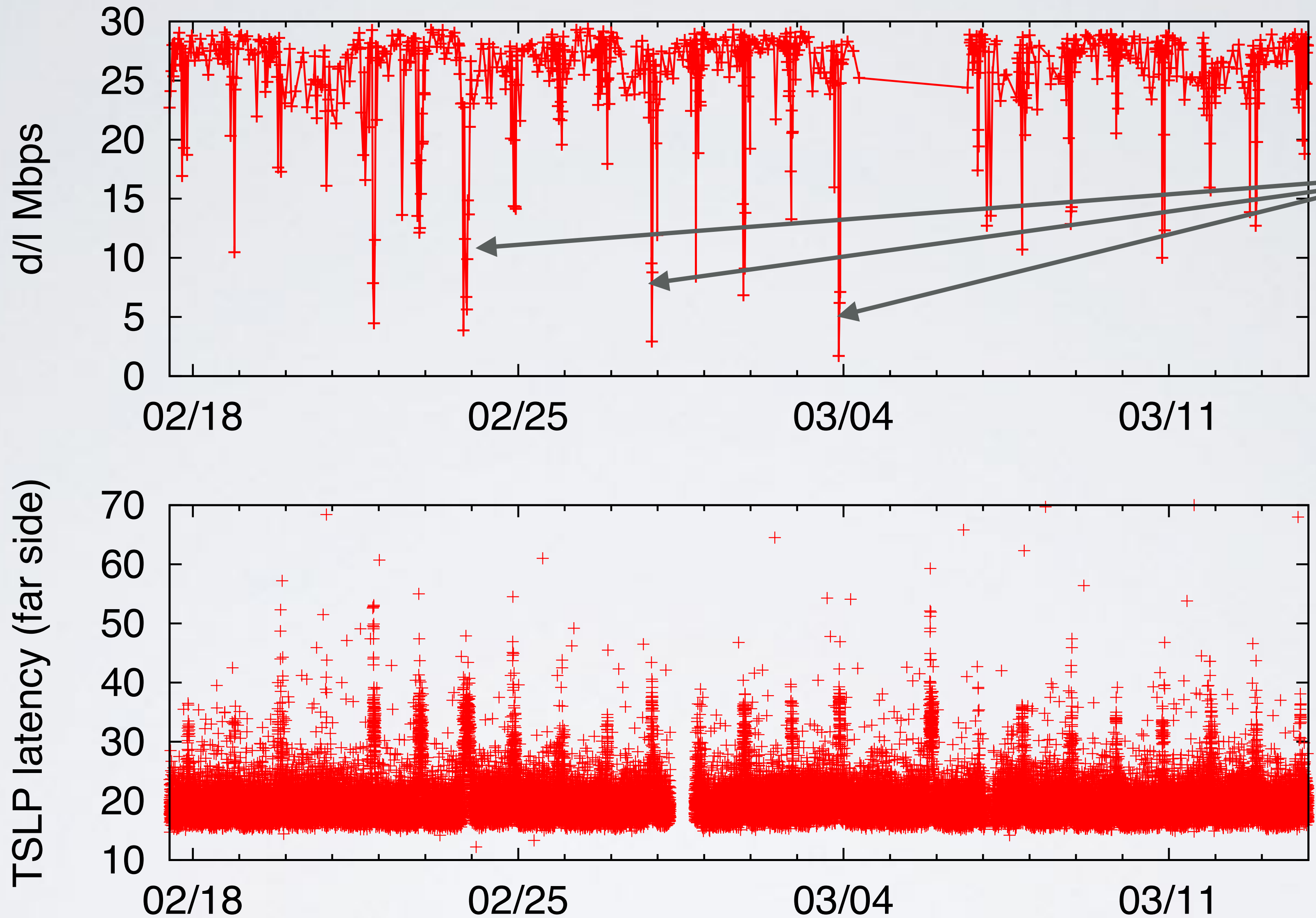
- Periodic NDT tests from Ark VP to M-Lab NDT server "behind" the congested interdomain link

Validation of the Method: Step 2



Strong correlation between throughput and TSLP latency: flows during elevated TSLP latency labeled as "externally" congested

Validation of the Method: Step 2

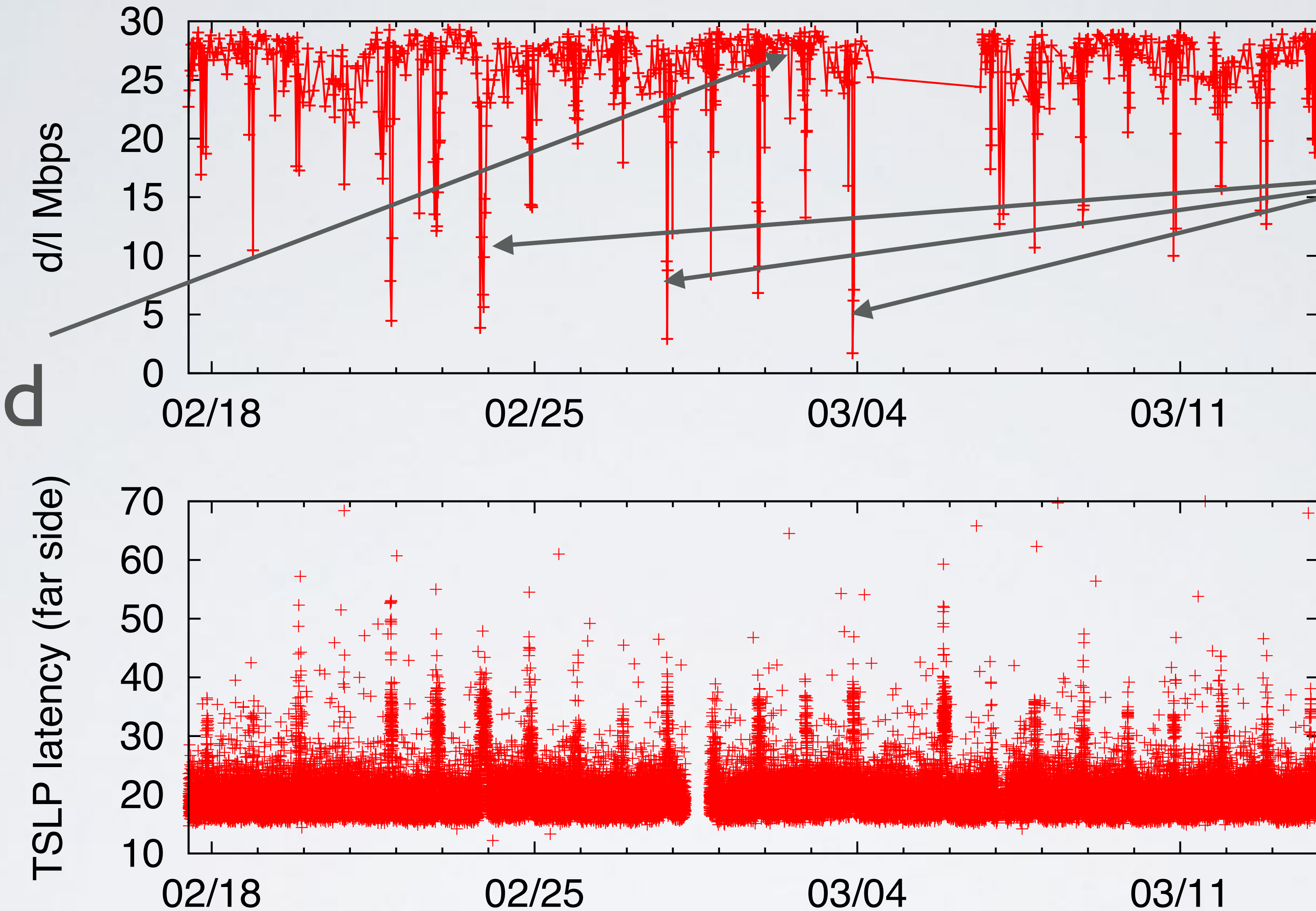


"Externally" congested

Strong correlation between throughput and TSLP latency: flows during elevated TSLP latency labeled as "externally" congested

Validation of the Method: Step 2

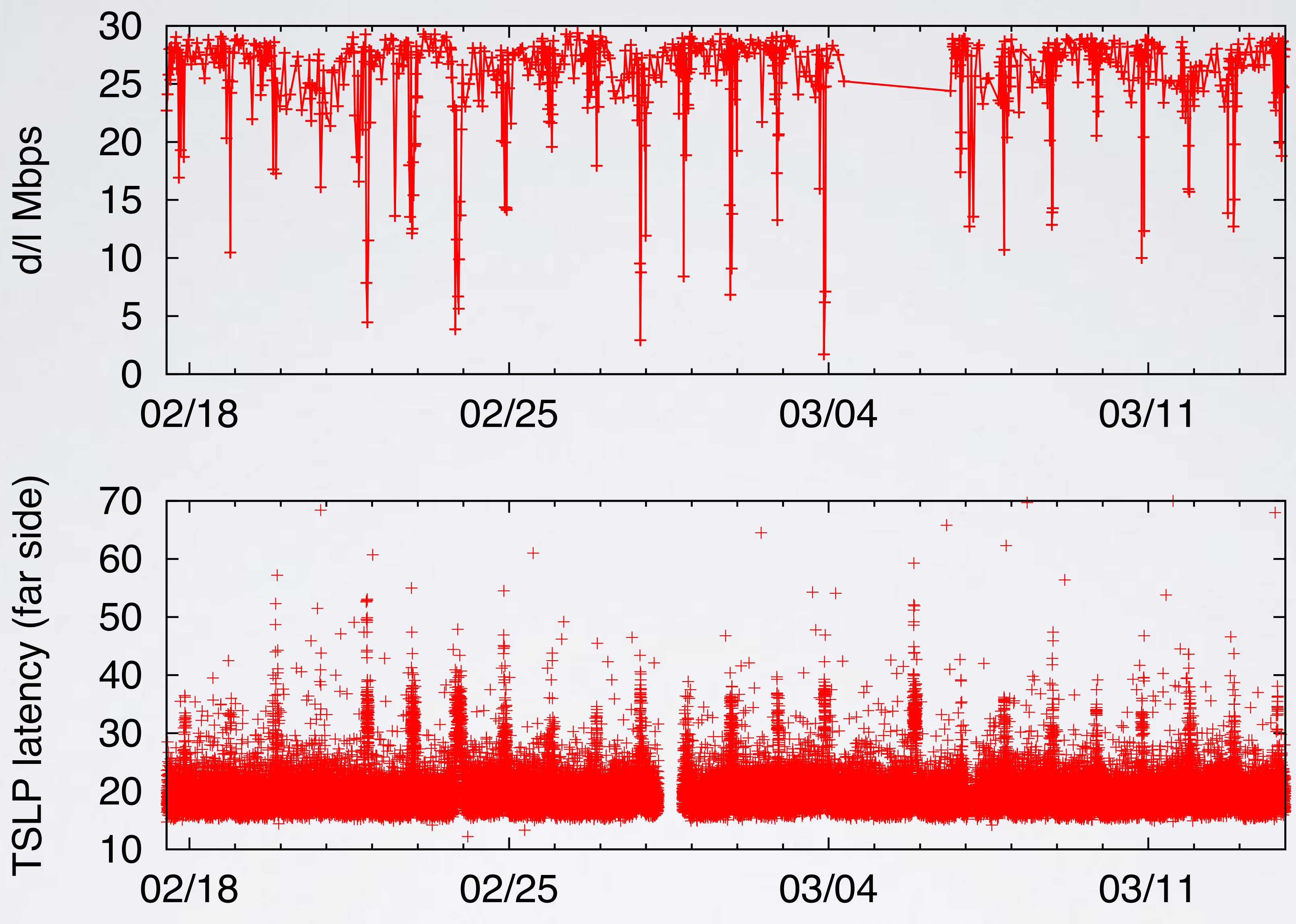
"self" congested



"Externally" congested

Strong correlation between throughput and TSLP latency: flows during elevated TSLP latency labeled as "externally" congested

Validation of the Method: Step 2

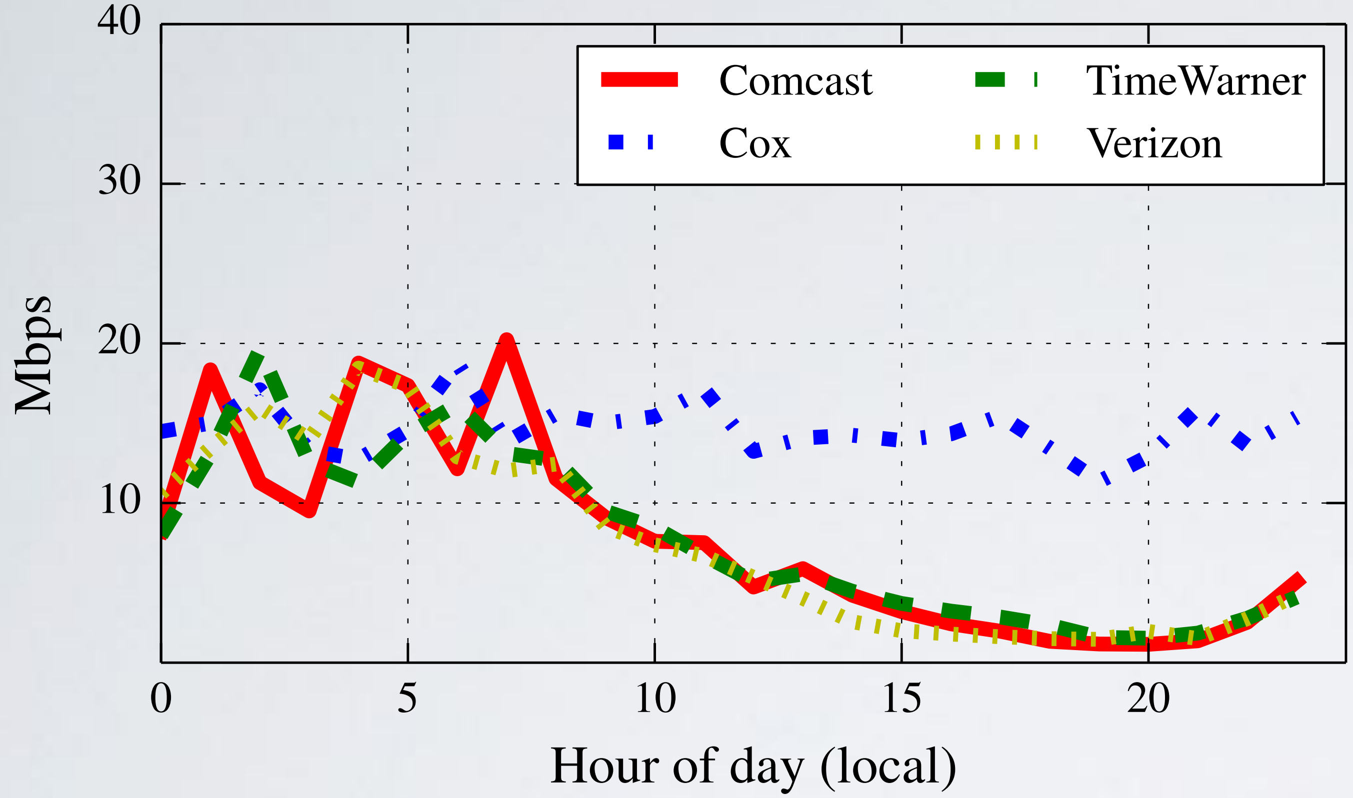


75%+ accuracy in detecting external congestion,
100% accuracy for self-induced congestion

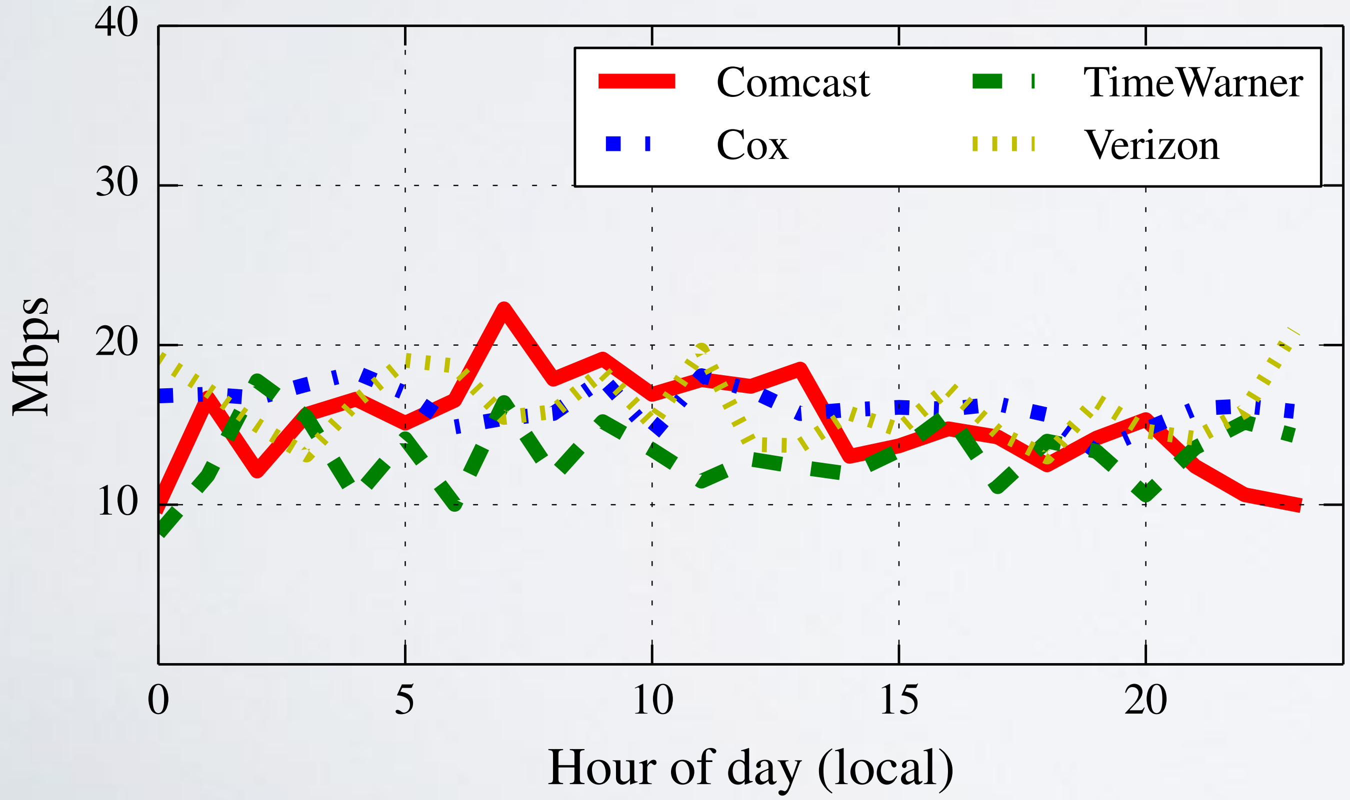
Validation of the Method: Step 3

- We use Measurement Lab's NDT test data for real-world validation
- Cogent interconnect issue in late 2013/early 2014
  - NDT tests to Cogent servers saw significant drops in throughput during peak hours
  - Several major U.S. ISPs were affected, except Cox
  - The problem was identified as congested interconnects

Using the M-lab Data



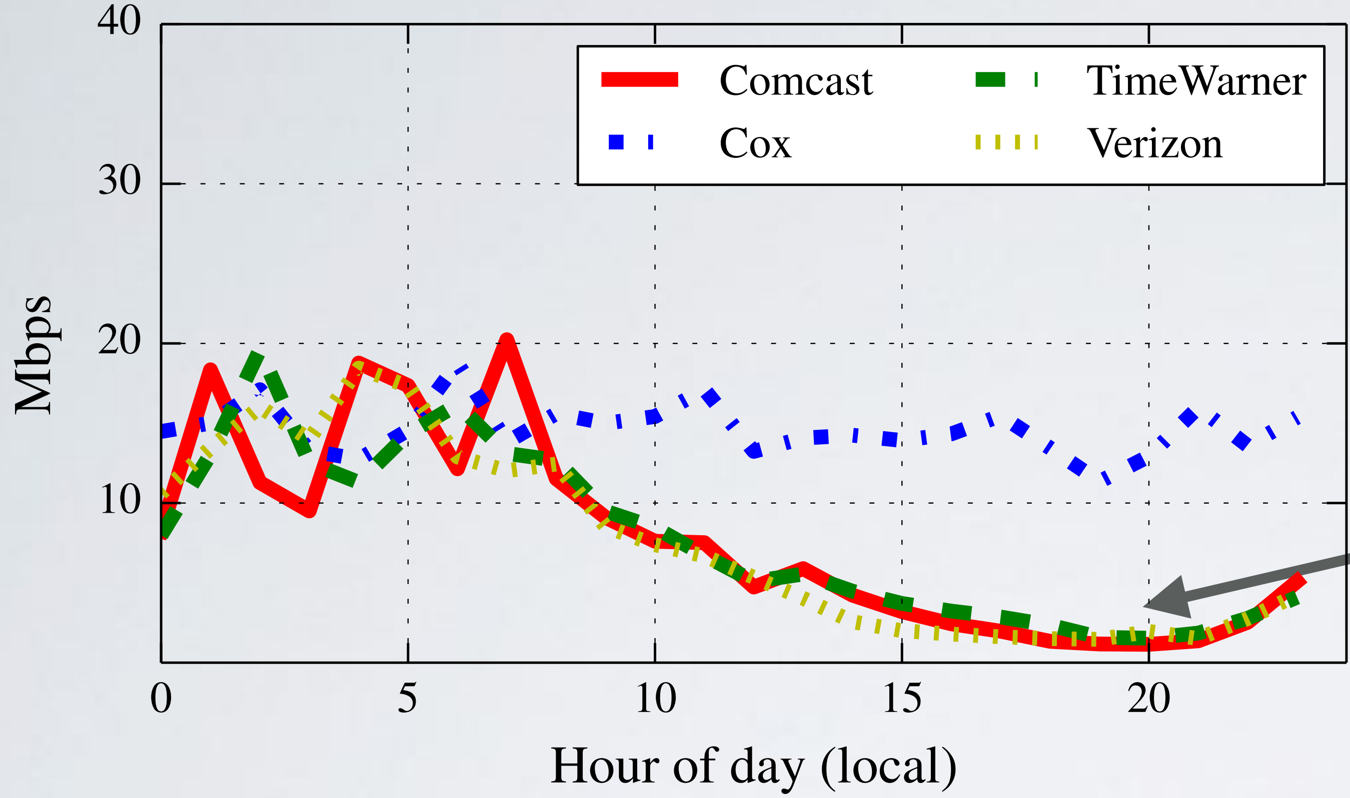
January 2014



April 2014

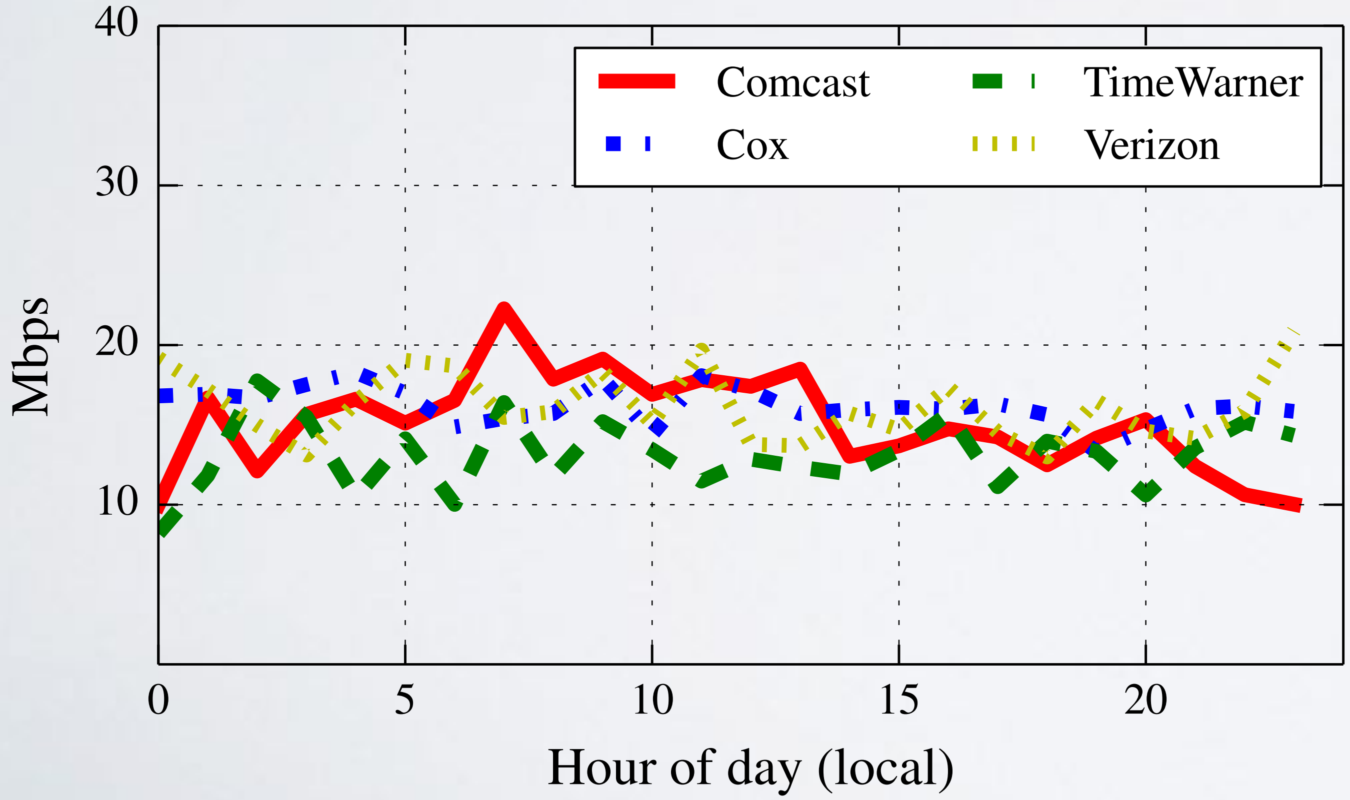
caida

Using the M-lab Data



January 2014

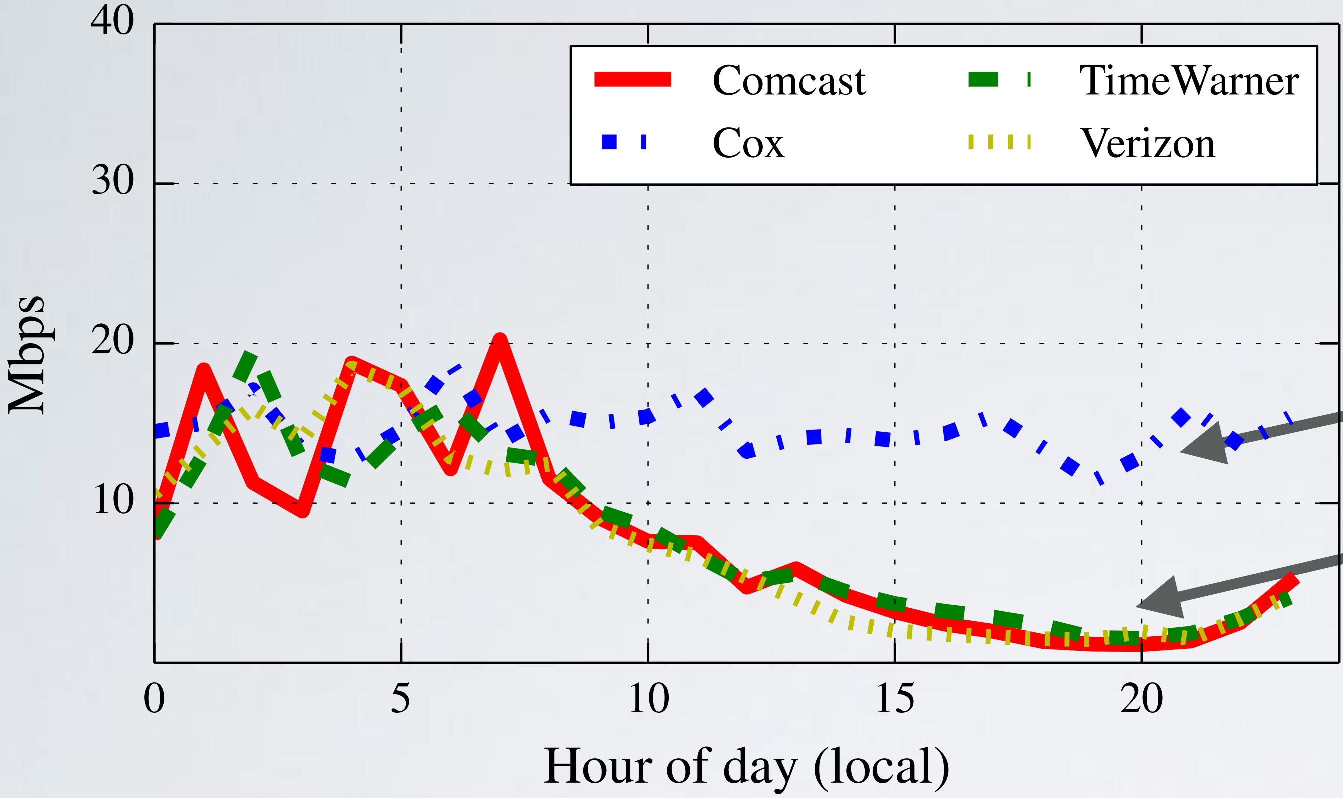
Drop in peak-hour throughput for for Comcast, TWC, Verizon



April 2014

caida

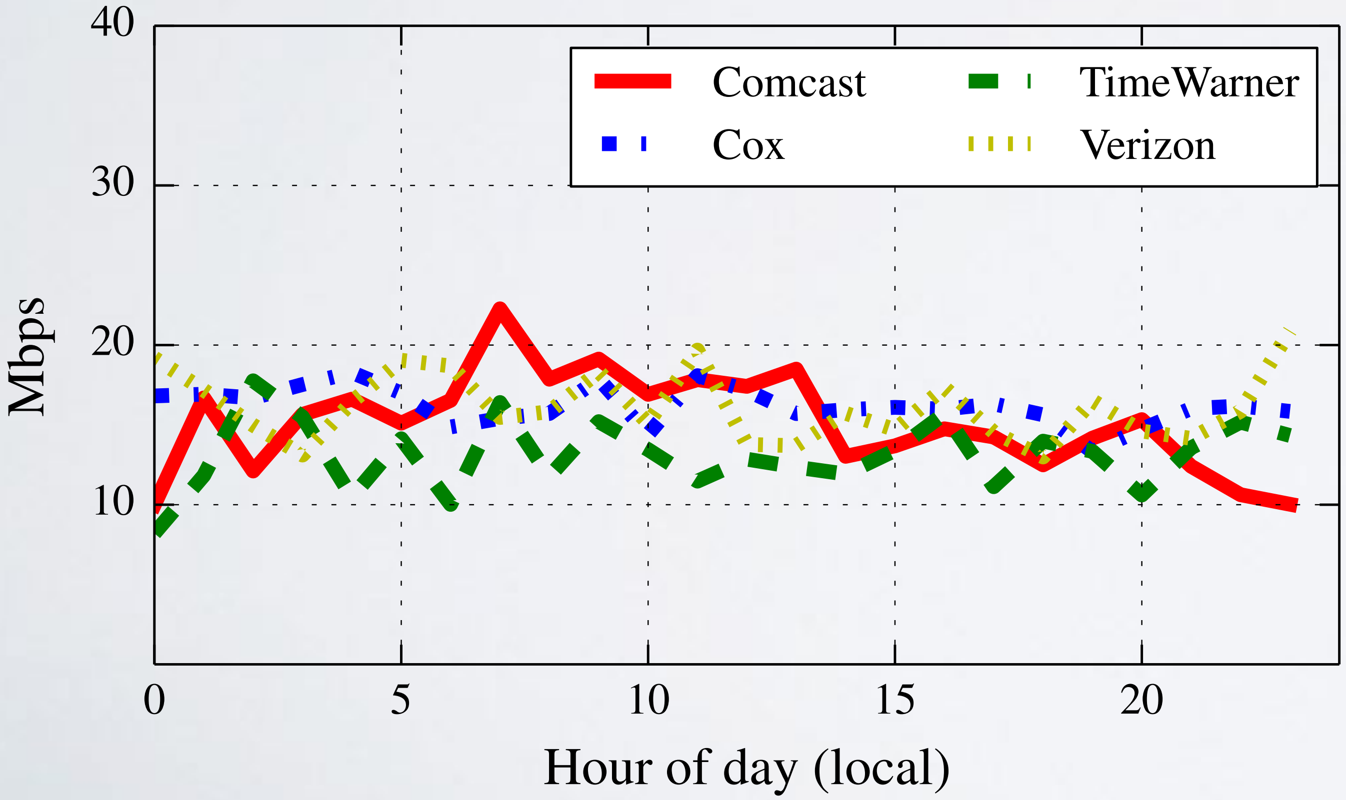
Using the M-lab Data



January 2014

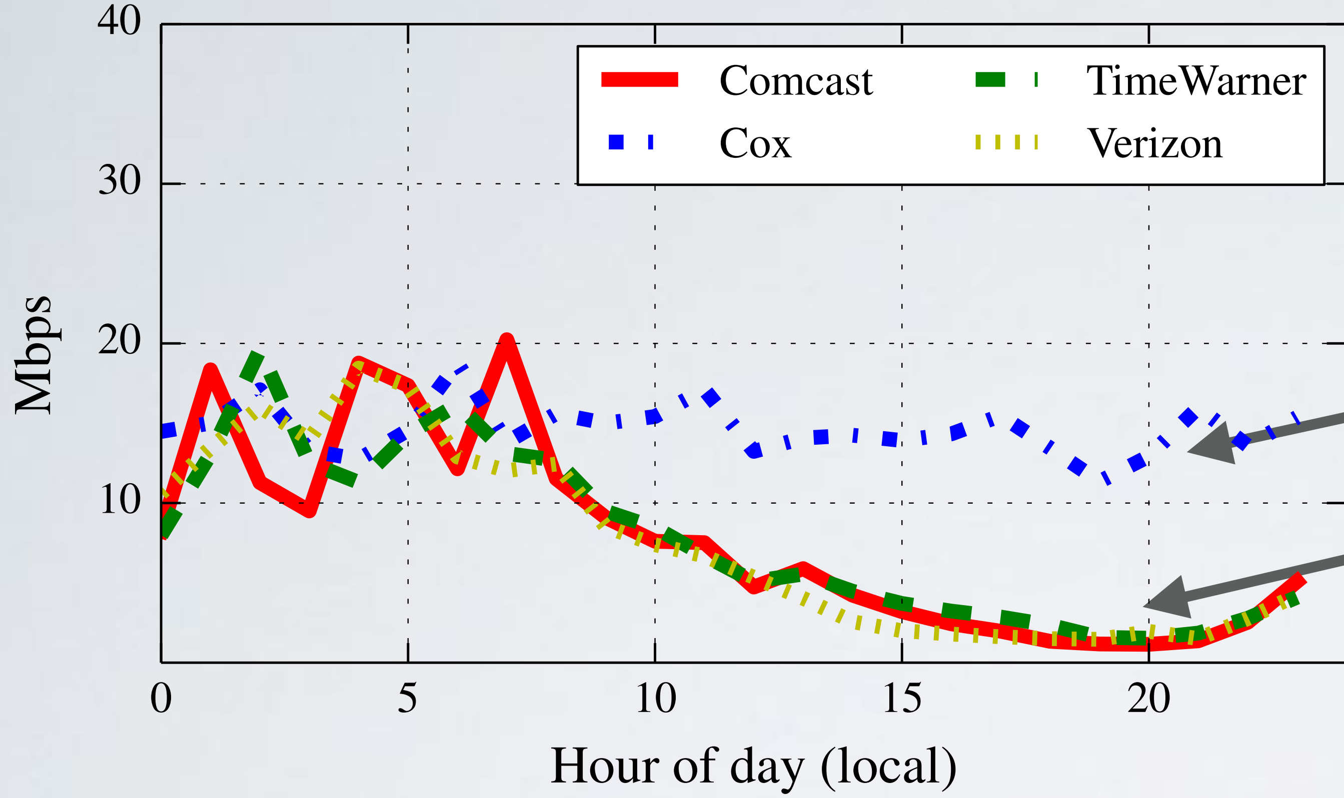
Cox not affected

Drop in peak-hour throughput for for Comcast, TWC, Verizon



April 2014

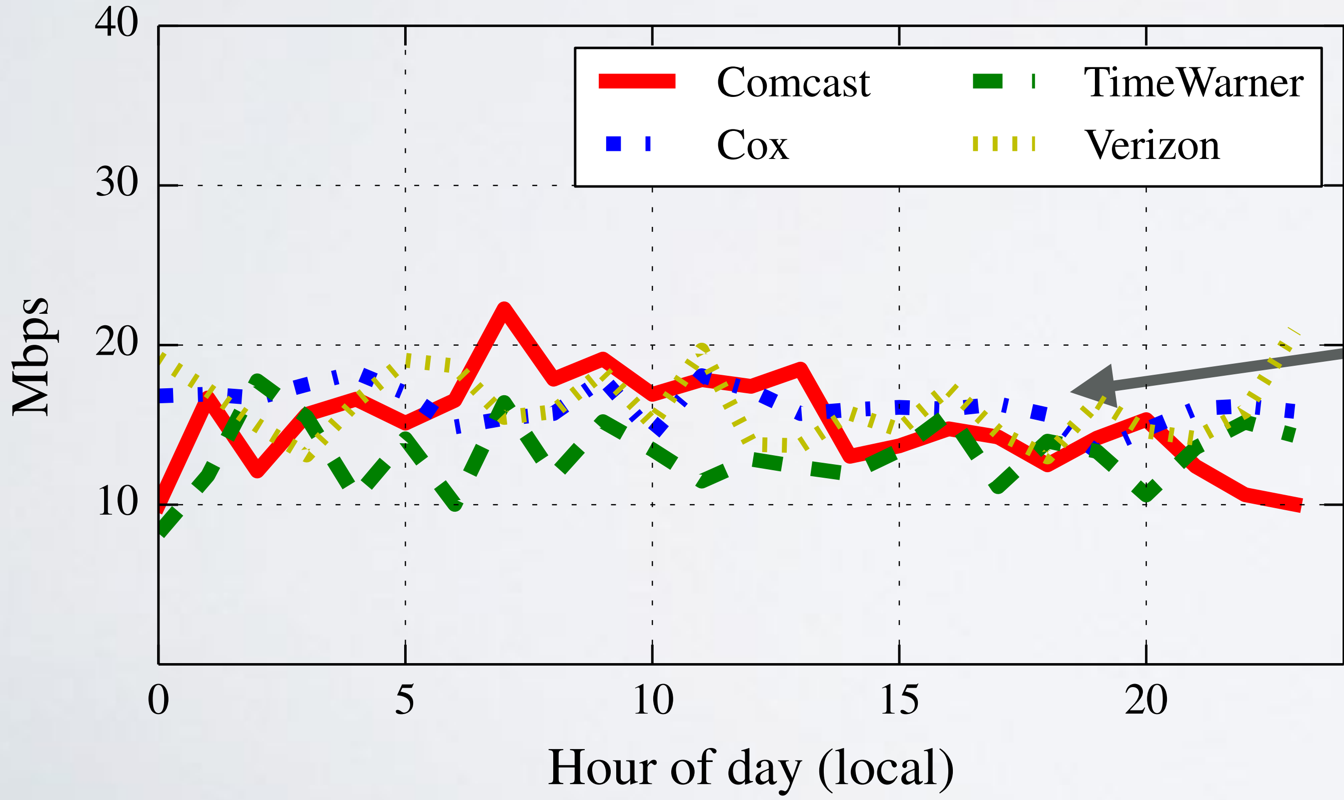
Using the M-lab Data



January 2014

Cox not affected

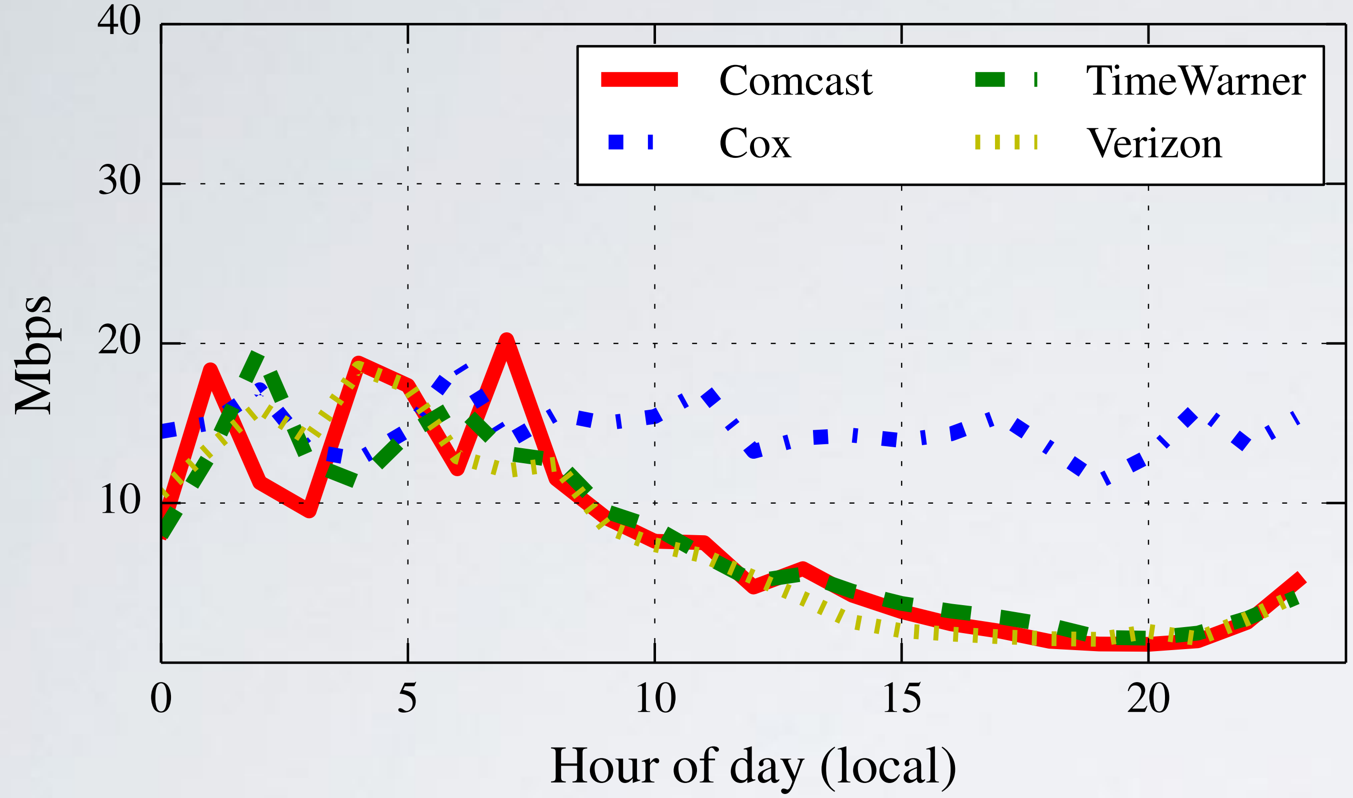
Drop in peak-hour throughput for for Comcast, TWC, Verizon



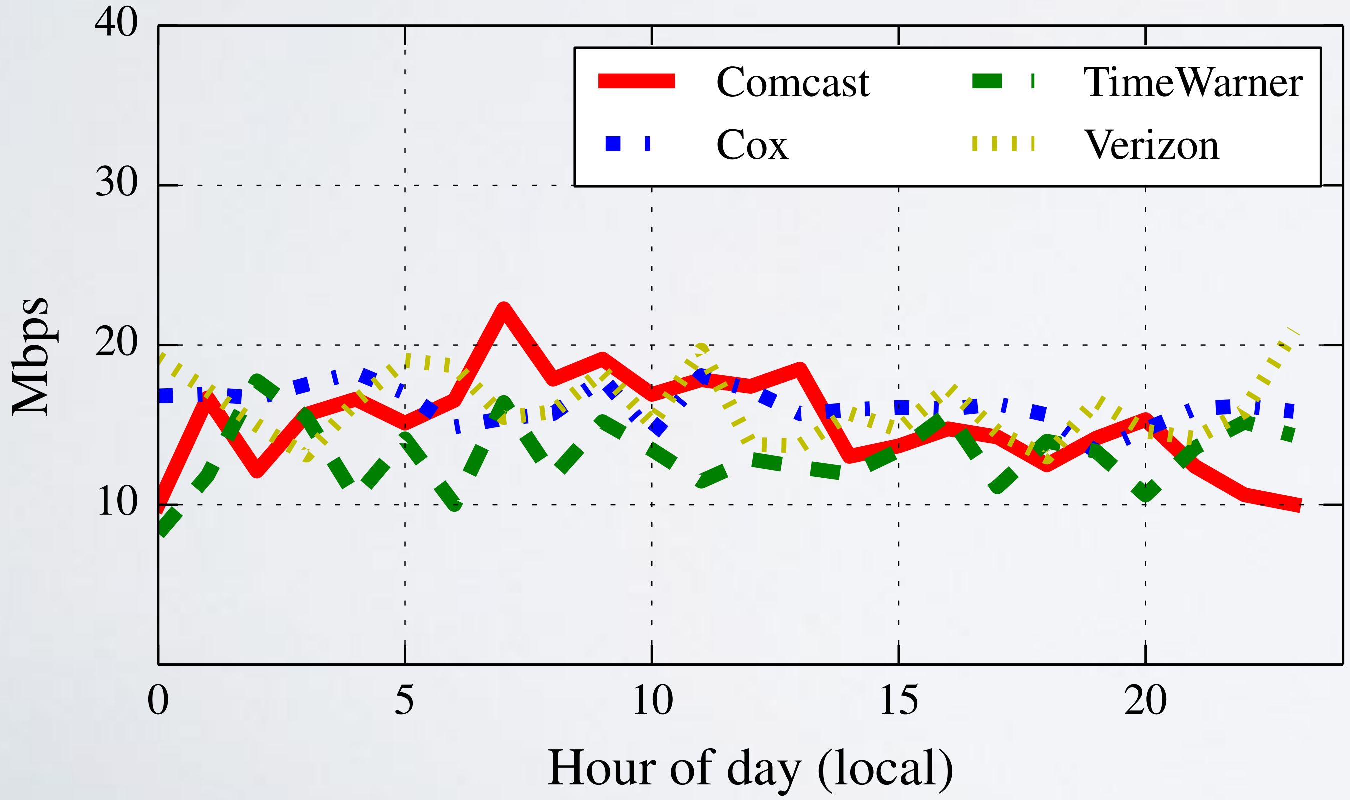
April 2014

Interconnection dispute resolved; no diurnal effect

Using the M-lab Data



Peak hour tests in Jan/Feb 2014 are likely "externally" congested



Off-peak tests in Mar/Apr 2014 are likely "self" congested

But didn't you just say it's hard to infer congestion using throughput tests??

But didn't you just say it's hard to infer congestion using throughput tests??

- Yes :)

But didn't you just say it's hard to infer congestion using throughput tests??

- Yes :)

- For that reason, our labeling is broad and coarse. All tests labeled "external" may not be traversing congested interconnects

But didn't you just say it's hard to infer congestion using throughput tests??

- Yes :)

- For that reason, our labeling is broad and coarse. All tests labeled "external" may not be traversing congested interconnects

- We do not expect the technique to identify all peak hour tests as externally congested, and vice versa

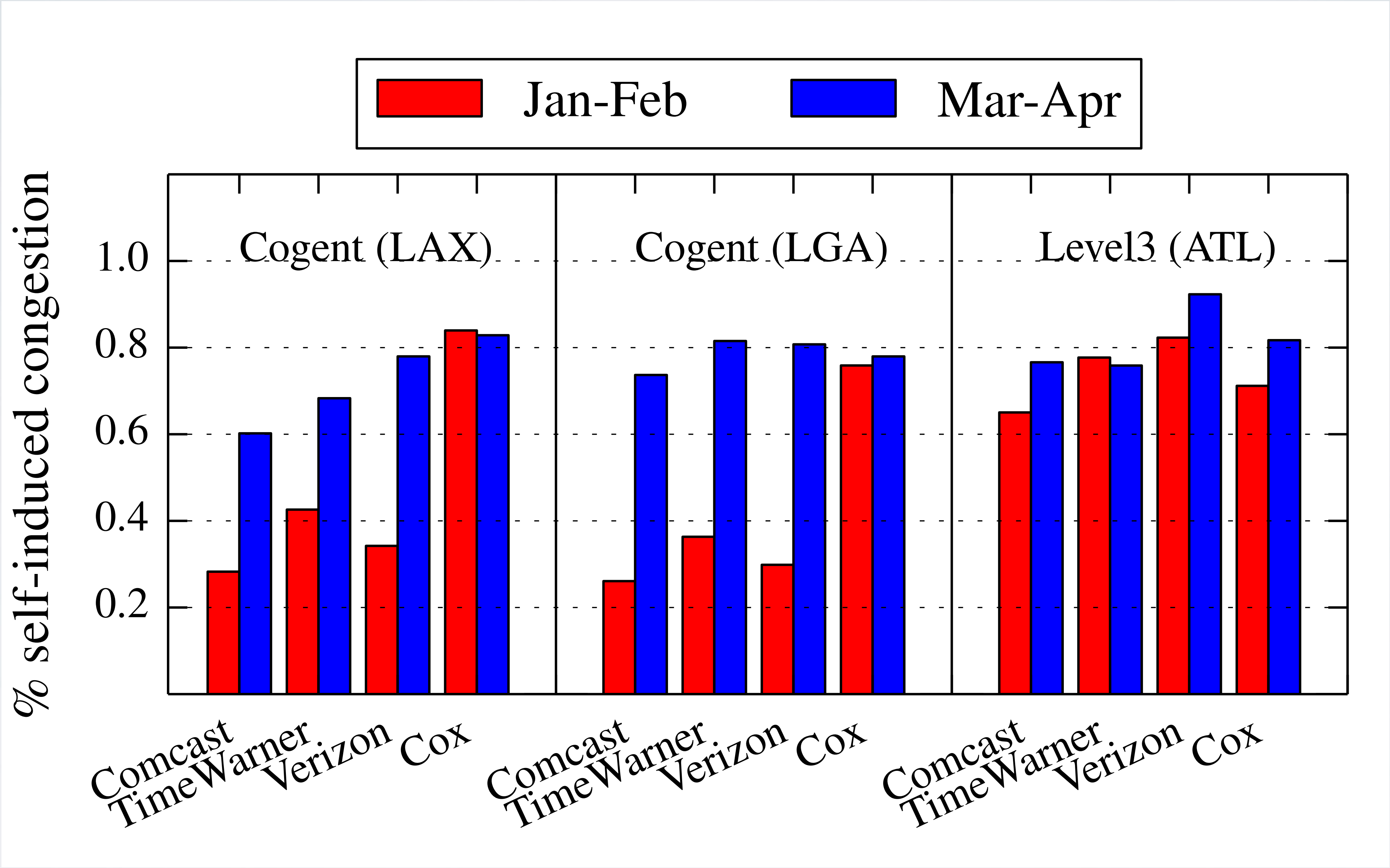
  - Looking for qualitative differences

But didn't you just say it's hard to infer congestion using throughput tests??

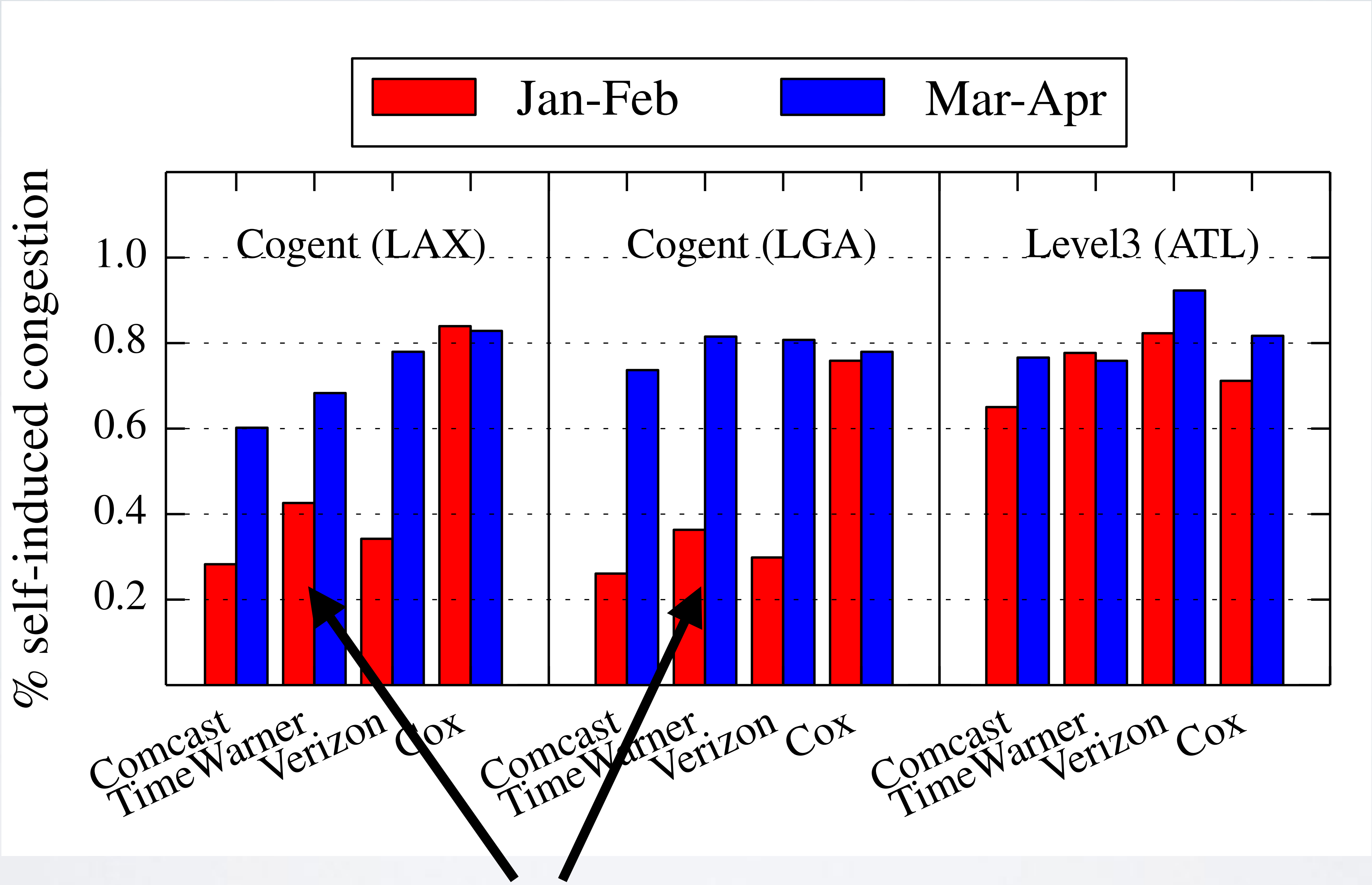
- Yes :)
- For that reason, our labeling is broad and coarse. All tests labeled "external" may not be traversing congested interconnects
- We do not expect the technique to identify all peak hour tests as externally congested, and vice versa
  - Looking for qualitative differences
- The general observations about congestion were verified by other sources, e.g., CAIDA's TSLP measurements

caida

Applying the Model to M-lab data

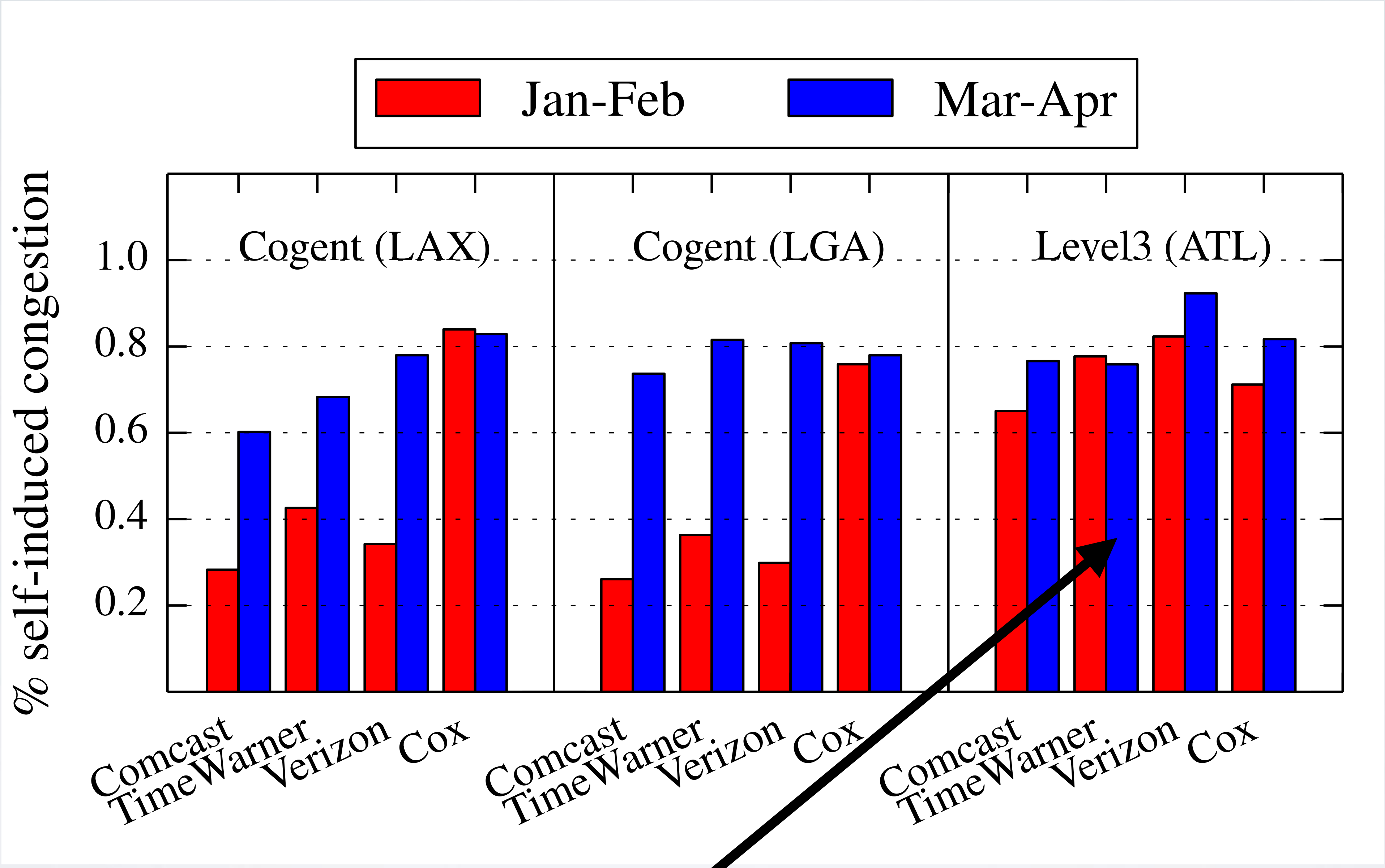


Applying the Model to M-lab data



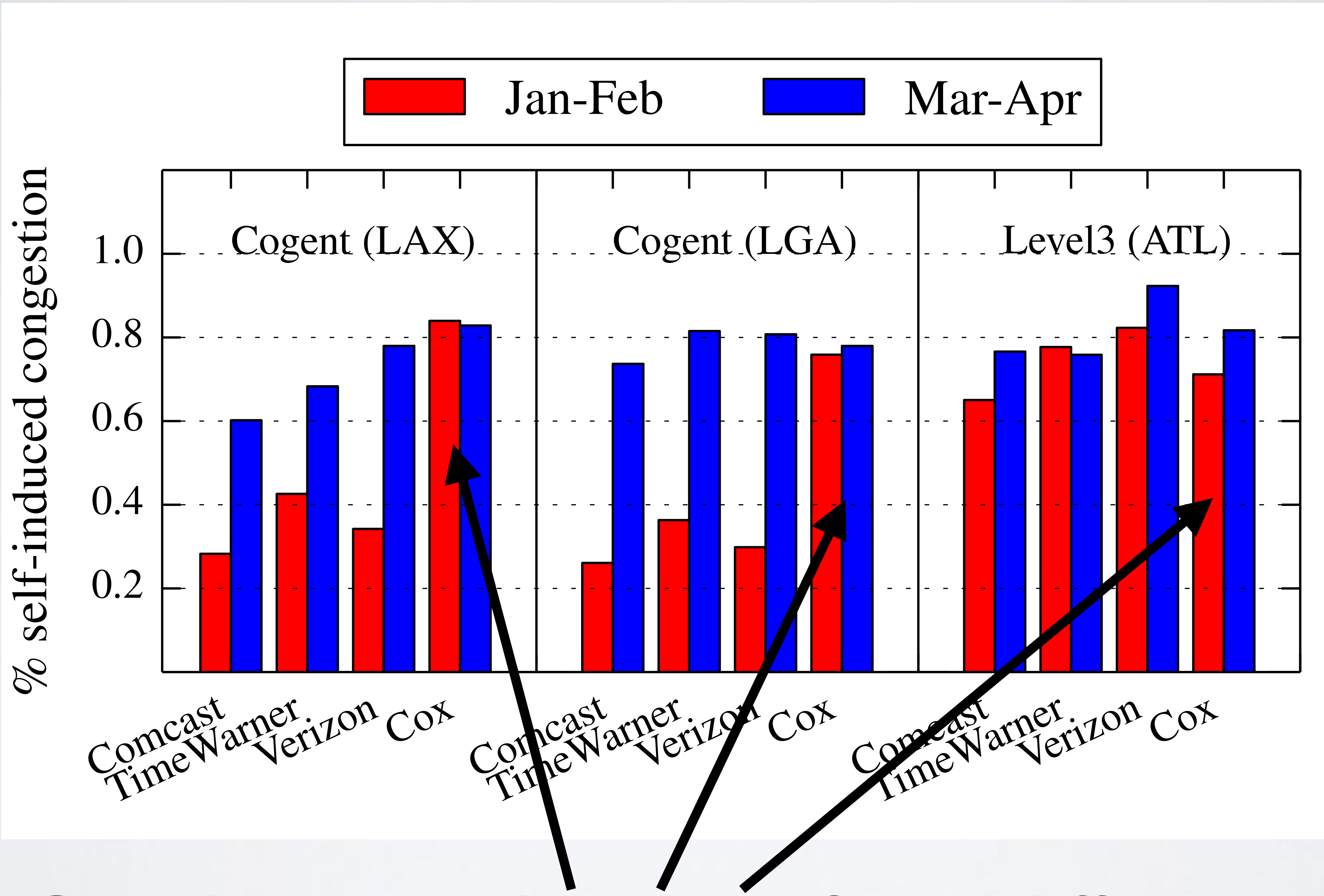
Much lower incidences of self-induced congestion for Cogent in Jan/Feb 2014 as compared to Mar/Apr

Applying the Model to M-lab data



Level3 does not show significant differences, was not affected by interconnection disputes

Applying the Model to M-lab data

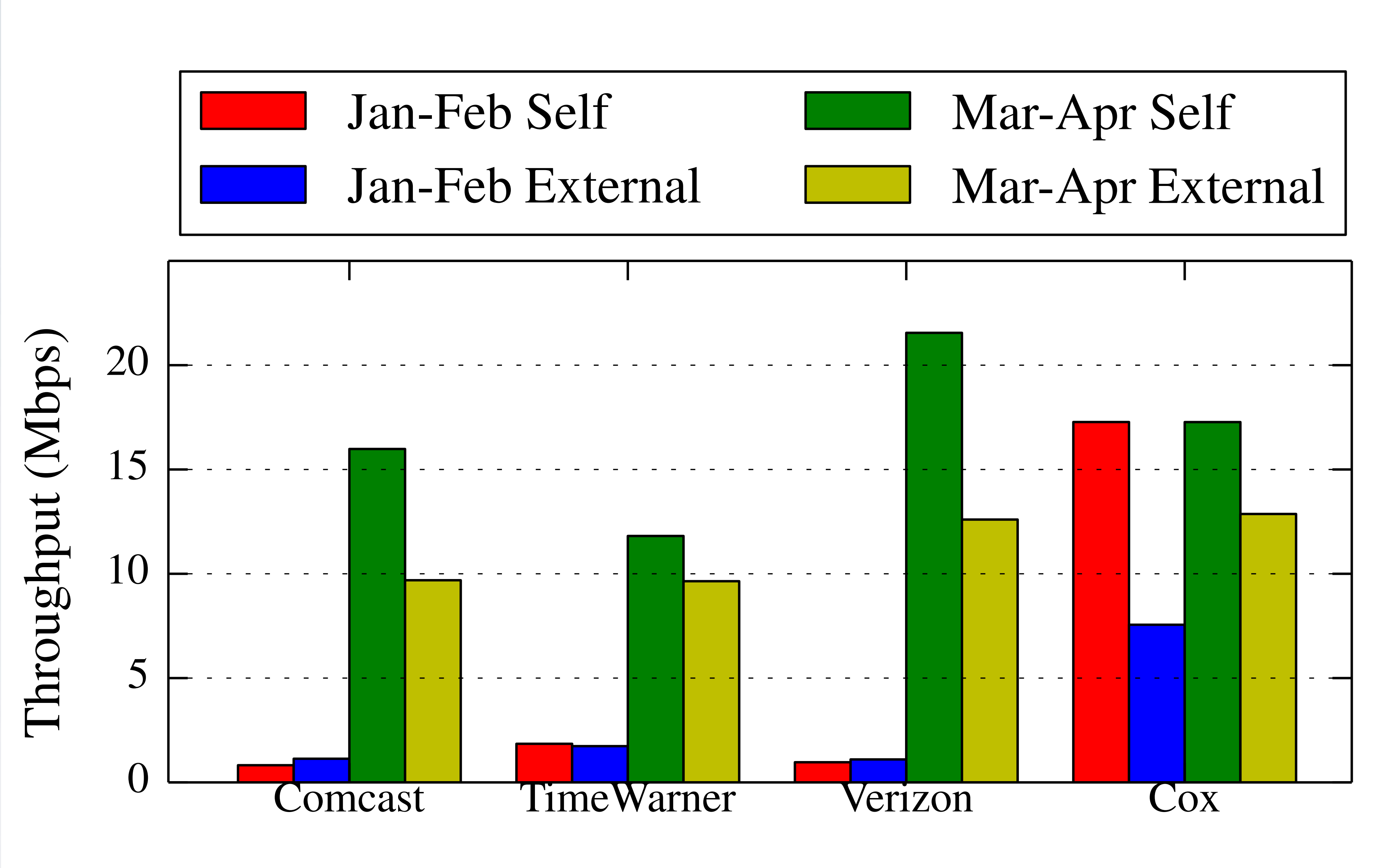


Cox does not show significant differences, was not affected by interconnection disputes

Looking at Throughput

- What throughput should we observe for "self" and "external" congested flows?
- With congested interconnects affecting many flows, both "self" and "external" should see similar throughput
- Without congested interconnects affecting many flows, "self" congested throughput should follow access link speeds, generally higher than "externally" congested

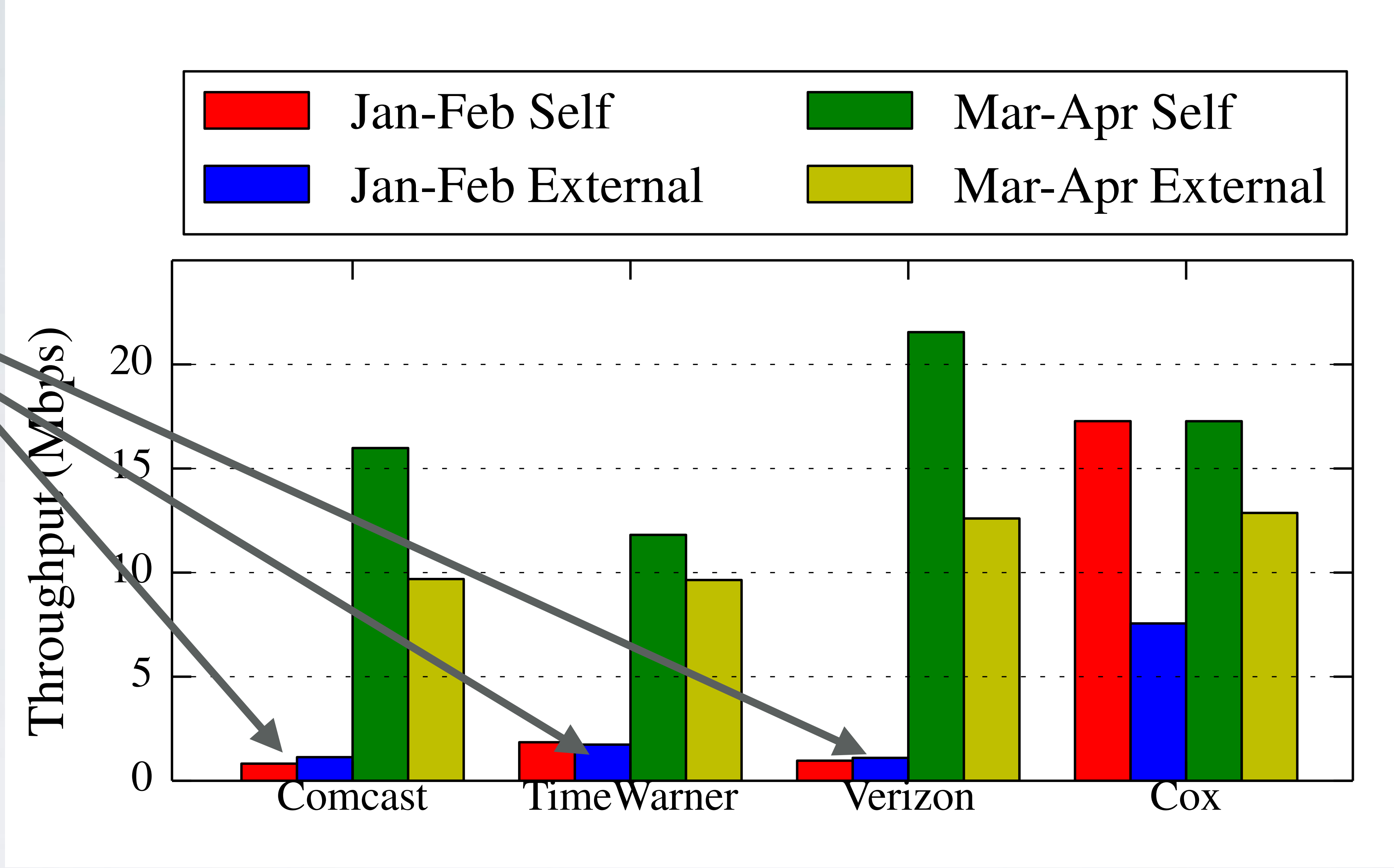
Looking at Throughput



- Avg. throughput of self-induced congestion flows significantly higher than externally congested in Mar-Apr (no interconnection disputes)

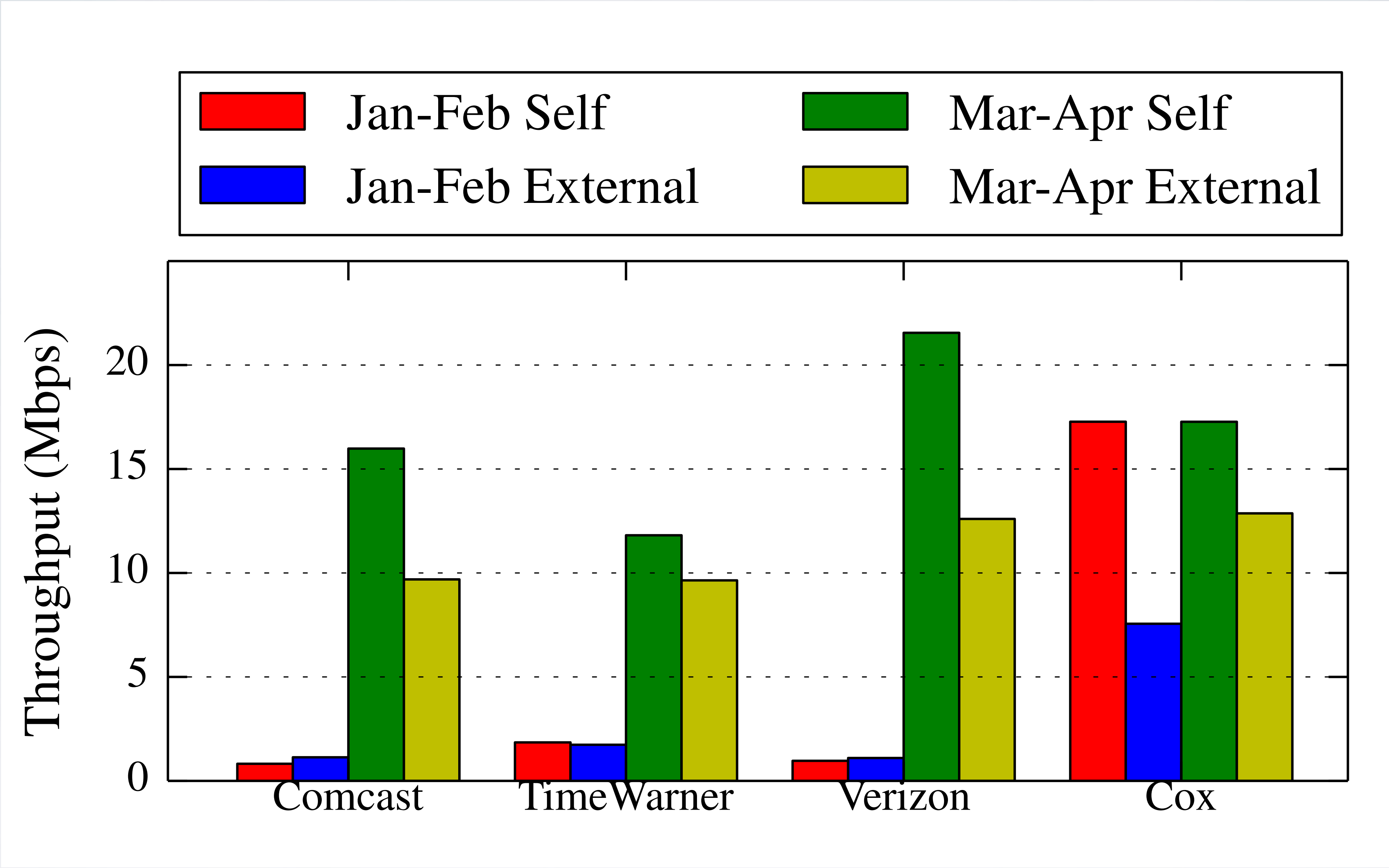
Looking at Throughput

Both "self" and "external" get similar throughput



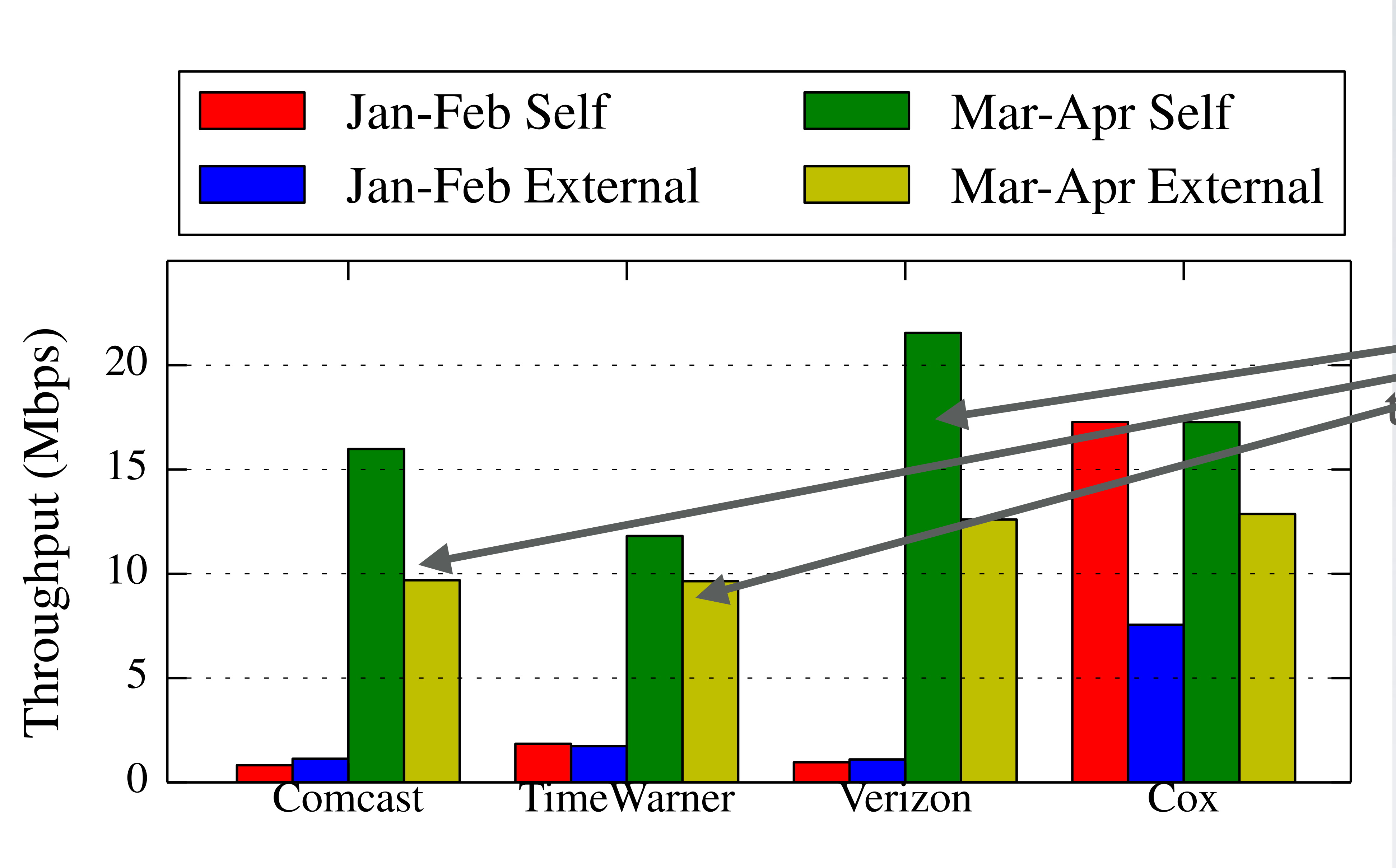
- Avg. throughput of self-induced congestion flows significantly higher than externally congested in Mar-Apr (no interconnection disputes)

Looking at Throughput



- Avg. throughput of self-induced congestion flows significantly higher than externally congested in Mar-Apr (no interconnection disputes)

Looking at Throughput



"Self" get higher throughput than "external"

- Avg. throughput of self-induced congestion flows significantly higher than externally congested in Mar-Apr (no interconnection disputes)

caida

Takeaways

- It is possible to distinguish two kinds of congestion: self-induced vs. externally congested
- The difference is important to identify the solution
  - Upgrade service plan? Or talk to ISP?
  - Also for regulatory purposes
- Simple, accurate technique using RTT during TCP slow start dynamics
  - Can be easily computed using packet captures or other tools such as Web100 (future work)

Limitations

- Relies on buffering effect
  - May not work on TCP variants that minimize buffer occupancy, e.g., BBR
- Only uses slow start dynamics
  - Might be confounded by flows that perform one way during slow start but differently afterward
- Real-world validation relies on coarsely labeled data
  - It would be great to validate on more real-world data!

Thanks!
Questions?