Identifying and Measuring Points of Congestion

Georgios Smaragdakis

A Team Effort

David Clark, Steve Bauer,
William Lehr

Kimberly Claffy, Matthew Luckie,
Amogh Dhamdhere, Bradley Huffaker

Arthur Berger, KC Ng,
Bala Chandrasekaran (intern from Duke)

Why Study Congestion?

- The Internet is not any more a "nice to have" service; network delays affect productivity

- Revenue can be sensitive to milliseconds delay (see references in [1])
  - Amazon found that 100 msec of latency cost 1% of sales.
  - Google found that delays in displaying webpages lead to revenue reduction (up to 20% for a 500 msec delay).

- Engagement of users is also sensitive to delay [2]
  - Users start to abandon a video streaming is start-up time >2 seconds.
  - Users experience re-buffering for 1% of video duration play 5% less video (and watch fewer ads).

[1] "Practical Guide to Controlled Experiments on the Web: Listen to Your Customers not to the HiPPO" KDD'07
[2] "Video Stream Quality Impacts Viewer Behavior: Inferring Causality Using Quasi-Experimental Designs" IMC'12

Why Study Congestion?

- Unintended consequences, e.g., congestion can cause errors to NTP accuracy

- Location where congestion occurs matters:
  - Congestion in access links affects users in a region
  - Congestion in transit/interconnections affects thousands of users!

- Shed light on the root causes of congestion (economic, technical, architectural) towards building a better Internet/inform policymakers.

Congestion: Anecdote or Evidence?

ars technica

Why YouTube buffers: The secret deals that make—and break—online video

THE WALL STREET JOURNAL.
WSJ

Netflix to Pay Comcast for Smoother Streaming

TIME

Netflix's Disputes With Verizon, Comcast Under Investigation

ars technica

Netflix war is over, but money disputes
still harm Internet users.

REUTERS

Europe's competition watchdog is investigating some of the region's biggest telecoms companies over whether they abused their market position

Building a Congestion Measurement Platform

Objectives:

- Collect and analyze data to provide *unbiased* evidence of congestion.
- Develop tools to construct a detailed *"heat map"* of congestion per city, peering location, the interconnection between two networks.

Focus on:

- Persistent congestion; clear daily patterns that span multiple days

Requirements:

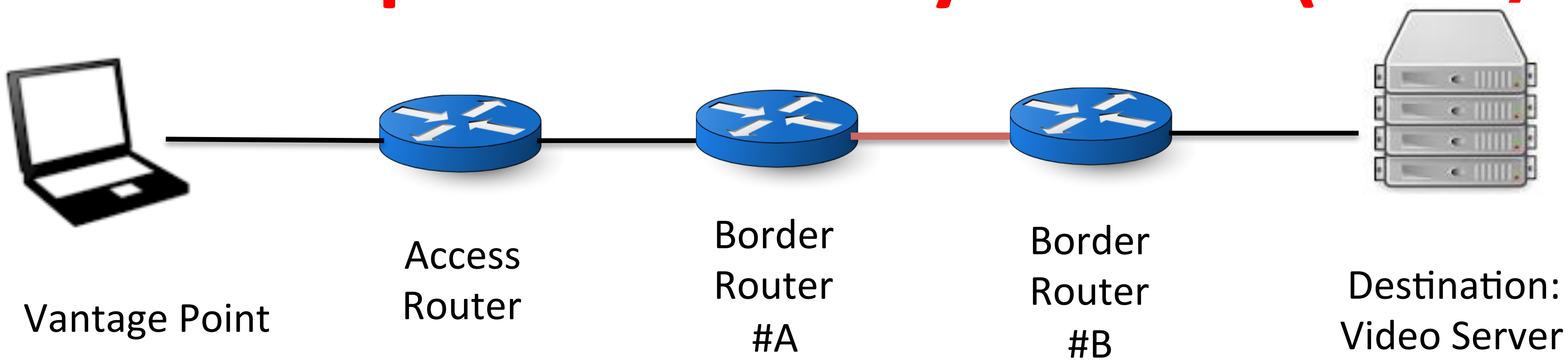
- Large-scale but lightweight measurements

Identifying and Measuring Points of Congestion

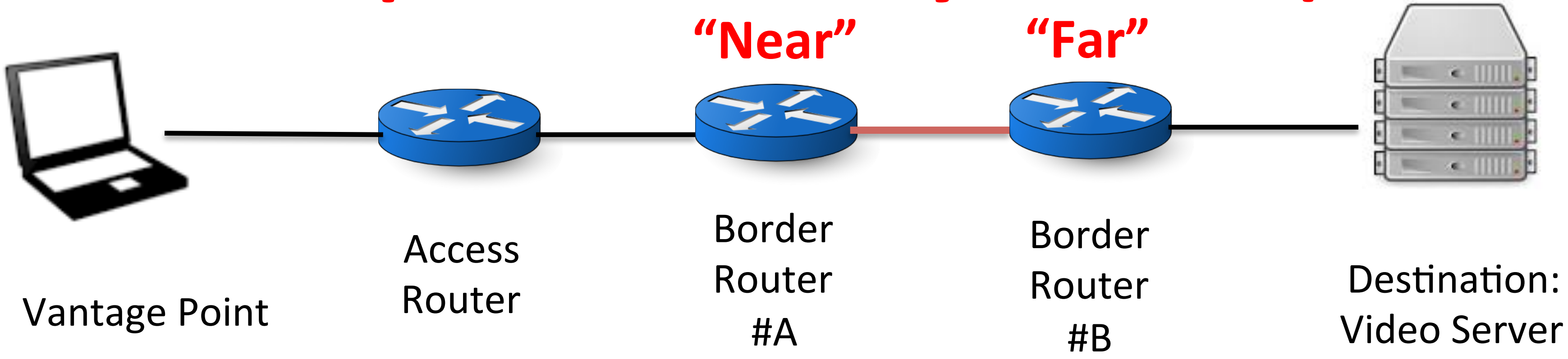
Part I: Targeted Interconnections or Internal Links

- “Challenges in Inferring Internet Interdomain Congestion”, Luckie et al., IMC 2014
- “Measurement and Analysis of Internet Interconnection and Congestion” Clark at al., TPRC 2014

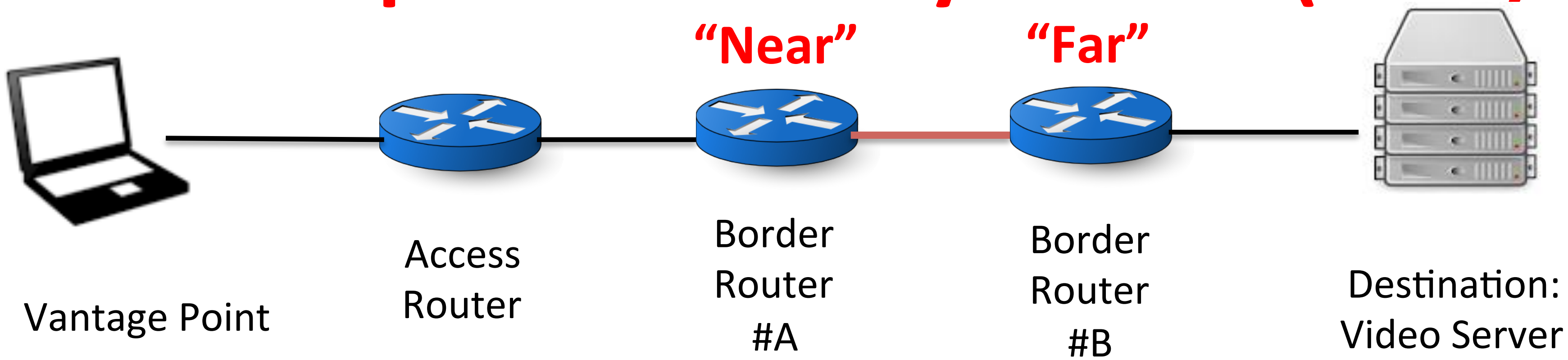
Methodology:
Time Sequence Latency Probes (TSLP)



Methodology: Time Sequence Latency Probes (TSLP)

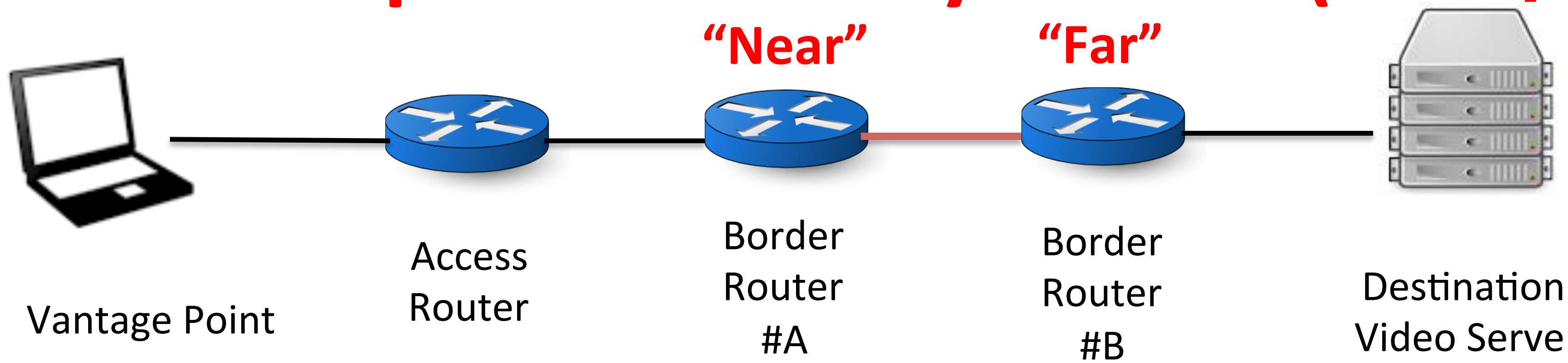


Methodology:
Time Sequence Latency Probes (TSLP)

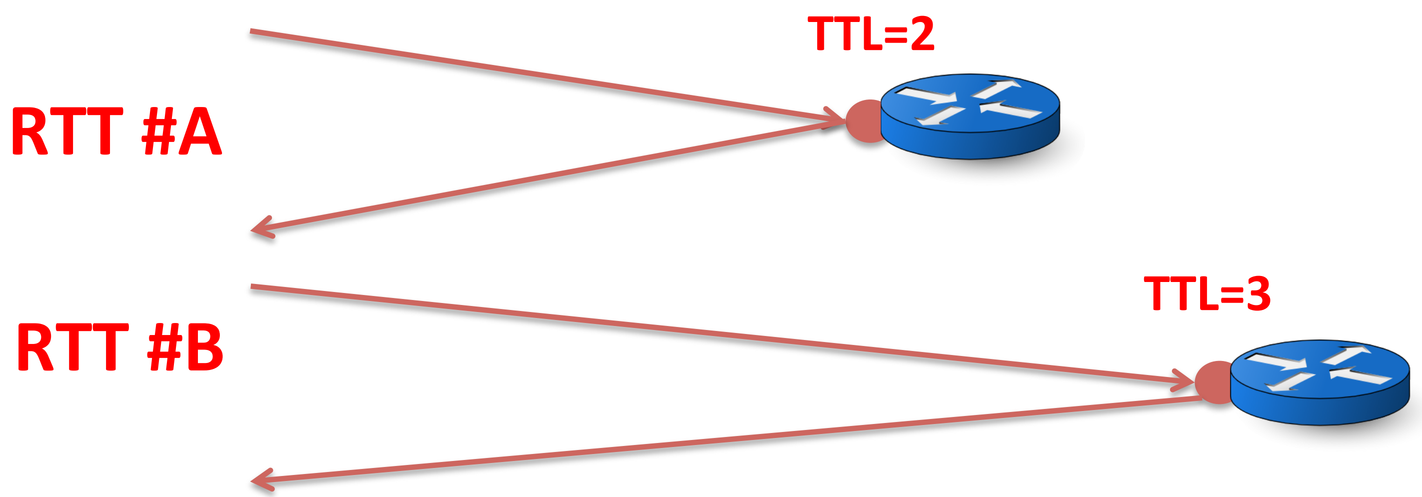


Send TTL-limited packets that expire in the “Near” and “Far” router

Methodology:
Time Sequence Latency Probes (TSLP)

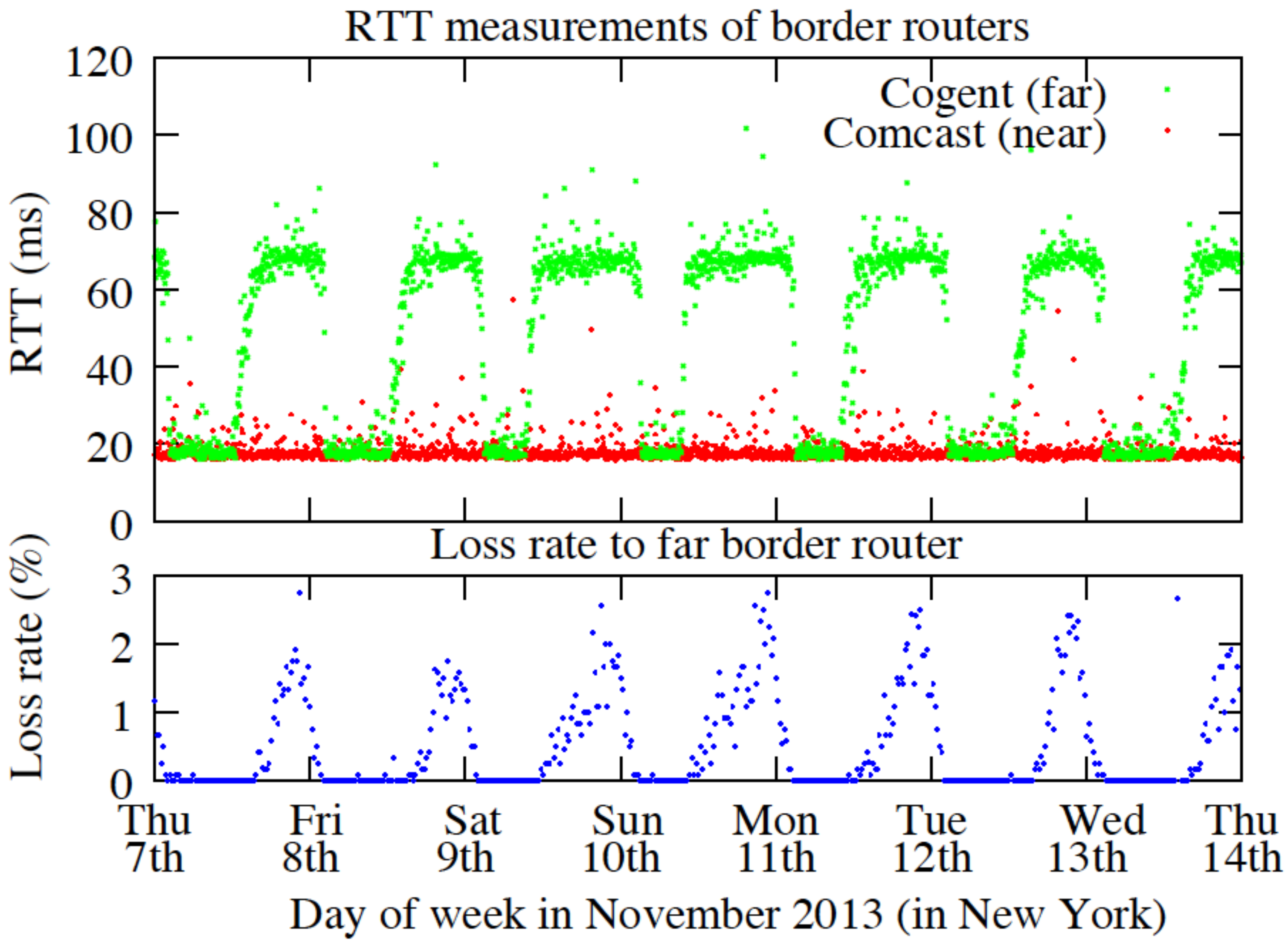


Frequently measure:

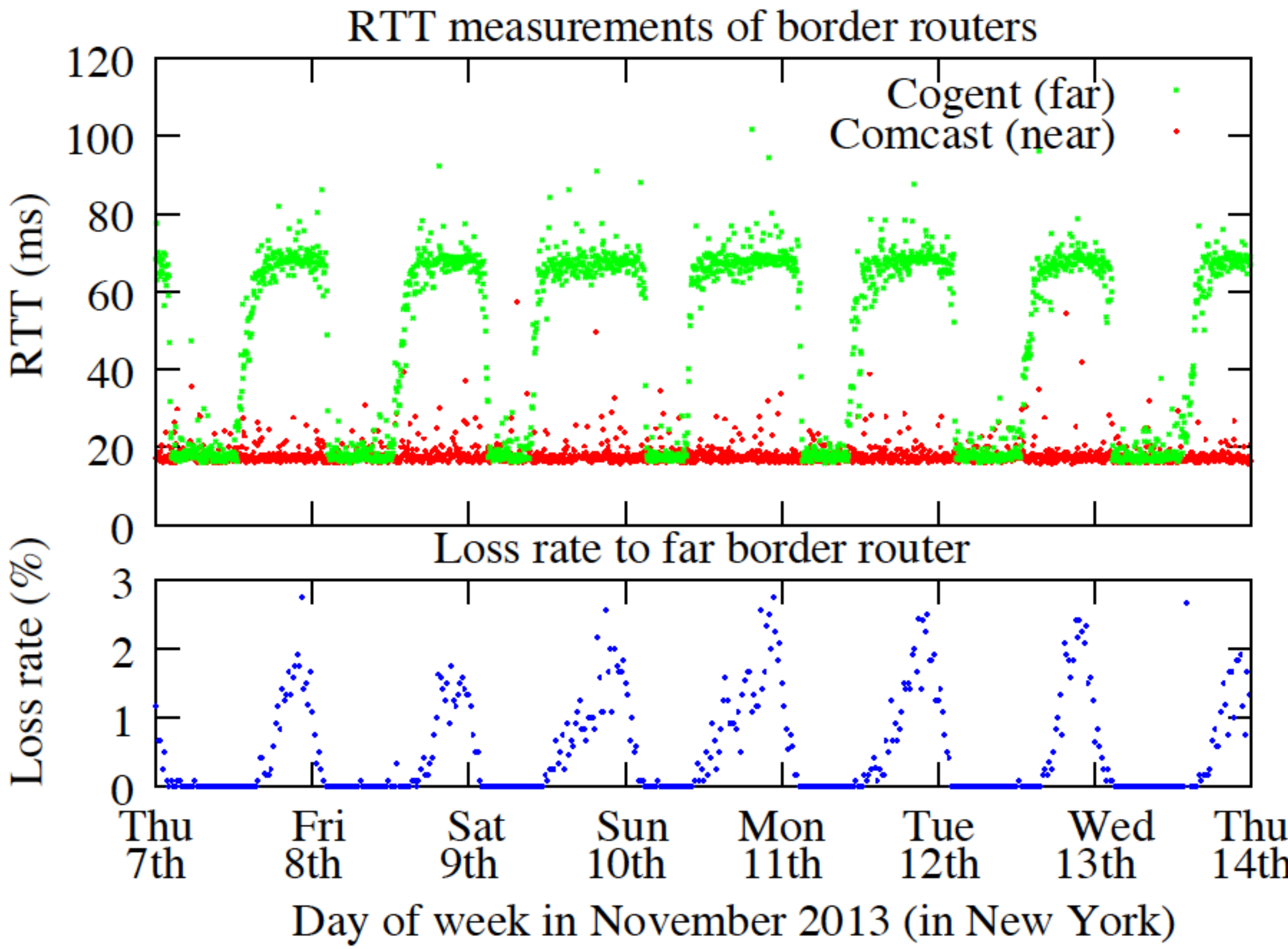


Send TTL-limited packets that expire in the "Near" and "Far" router

An Example (November 2013)



An Example (November 2013)



To infer diurnal pattern: FFT analysis of time series with frequency 1/day.

Limitations

- Asymmetric Routing
  - Reverse Traceroute [1] may unveil the reverse path (using IP options)
  - Both forward and backward path should be monitored; vantage points are needed at both ends

- Router Queuing Management
  - Measuring packets (ICMP packets) may be assigned to low priority queues
  - Random Early Detection (RED) before queue becomes full

- Router Ownership
  - It is not trivial to map an interface/router to a network; it requires analysis of massive amount of measurements (aliasing)

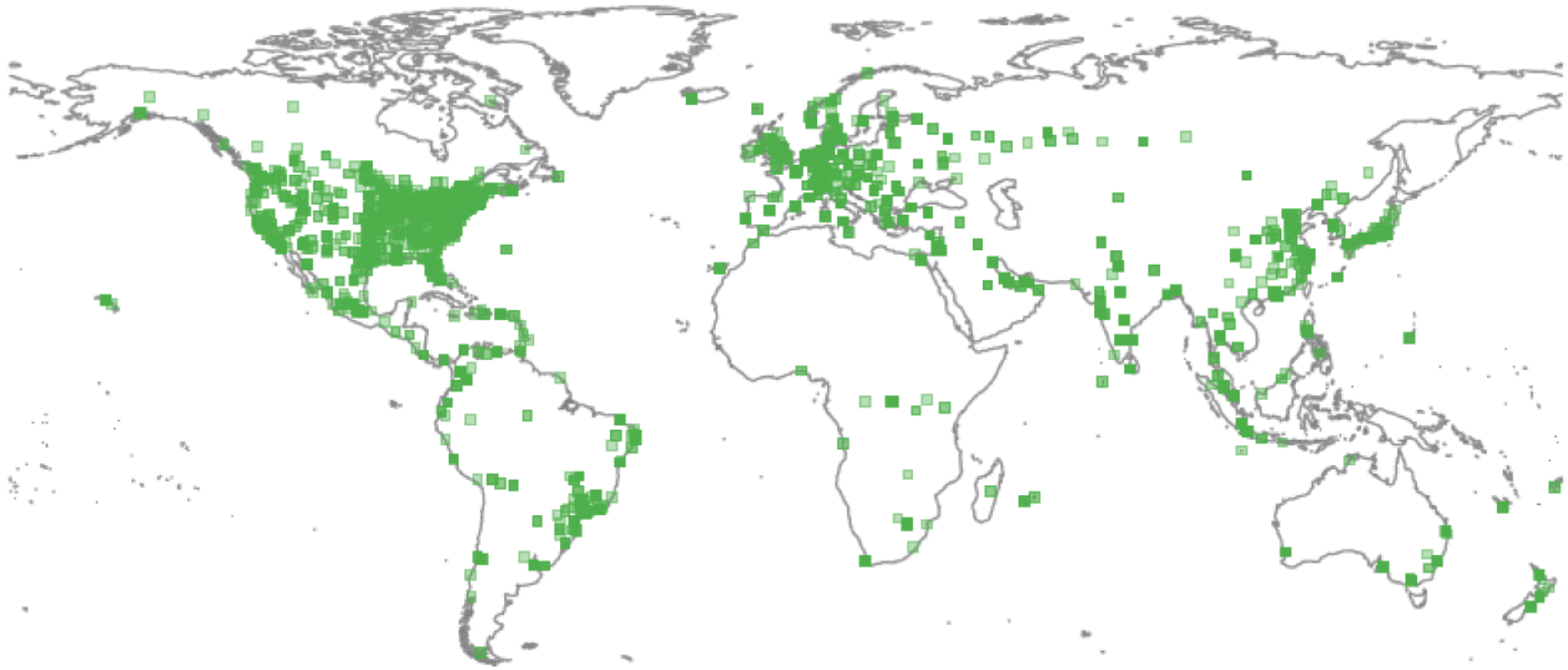
[1] "Reverse Traceroute" NSDI'10

Identifying and Measuring Points of Congestion

Part II: At Internet-wide Scale

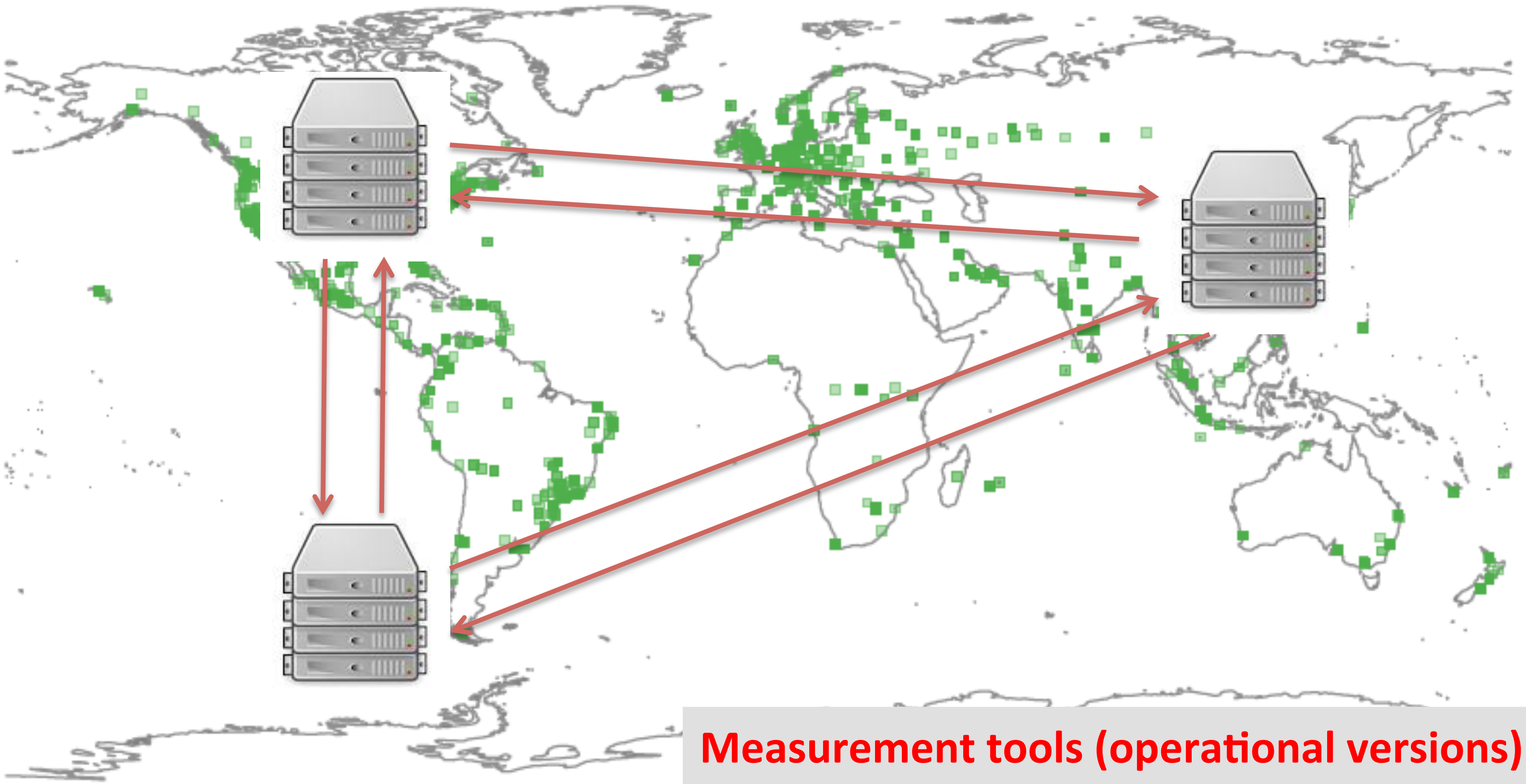
- "A Server-to-Server View of the Internet", Chandrasekaran et al., CoNEXT 2015

Utilizing a Highly Distributed Platform



- Large-scale measurements utilizing 5,000+ server clusters (one server per cluster)
- 2,000+ locations: colocation facilities, IXPs, datacenters, residential networks, enterprise networks.
- 1,200+ networks

Utilizing a Highly Distributed Platform

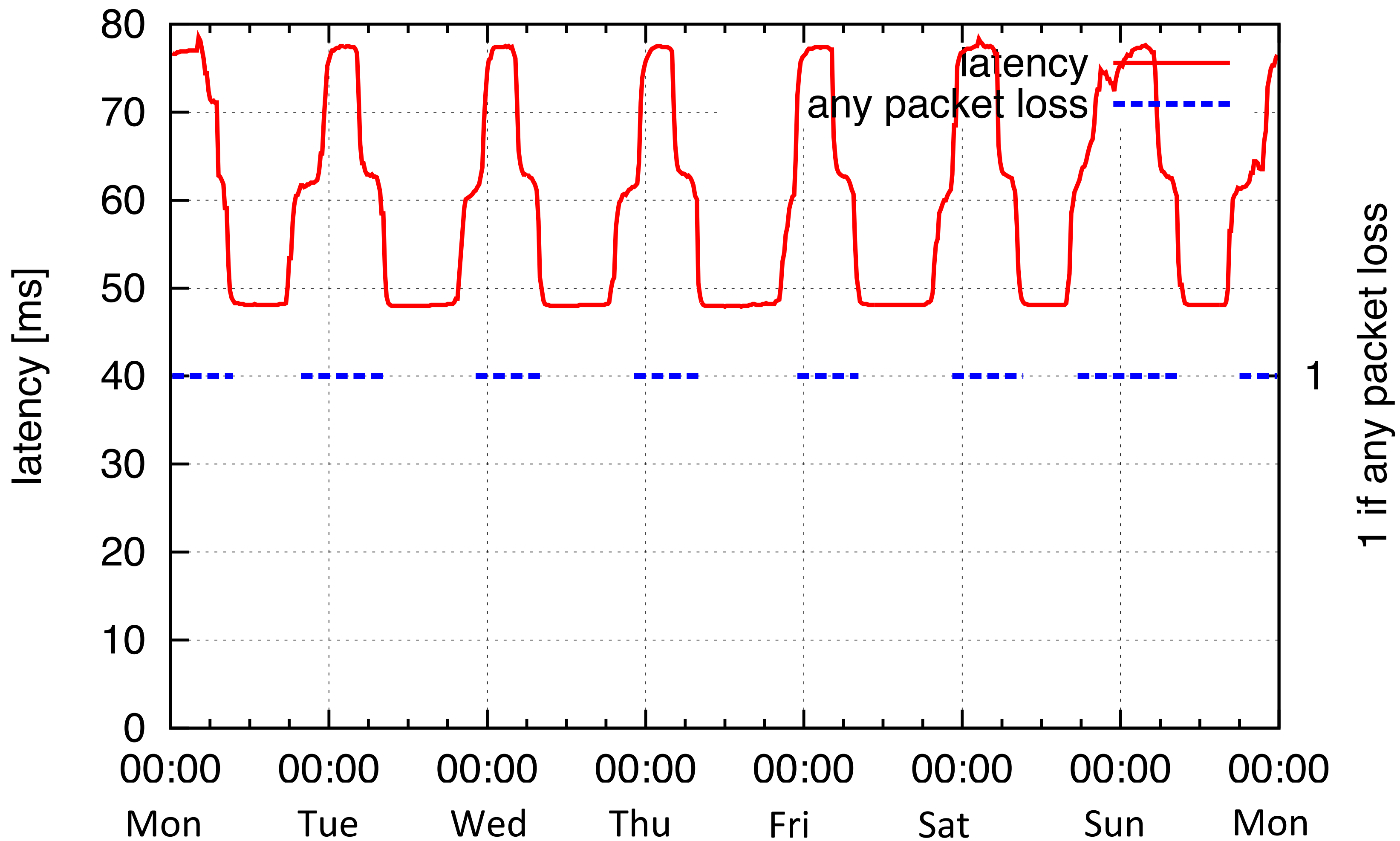


Measurement tools (operational versions):
- ping
- traceroute (Paris)

Methodology

Frequent Server-to-Server ping Measurements → Apply FFT to select candidate pairs with "congestion" → Perform traceroute campaigns → Infer the location of congestion

Bootstrap Phase: Server-to-Server Ping Measurements



Bootstrap Phase:
Server-to-Server Ping Measurements

- We collected and analyzed around 2 Million time series of pings

- Frequency: 1 sample per 15 minutes for 1 week

- The FFT analysis showed that around 6% are potential candidate pairs for congestion
  - Notice that routing may play a role
  - Notice that the increase of delay may not be always significant

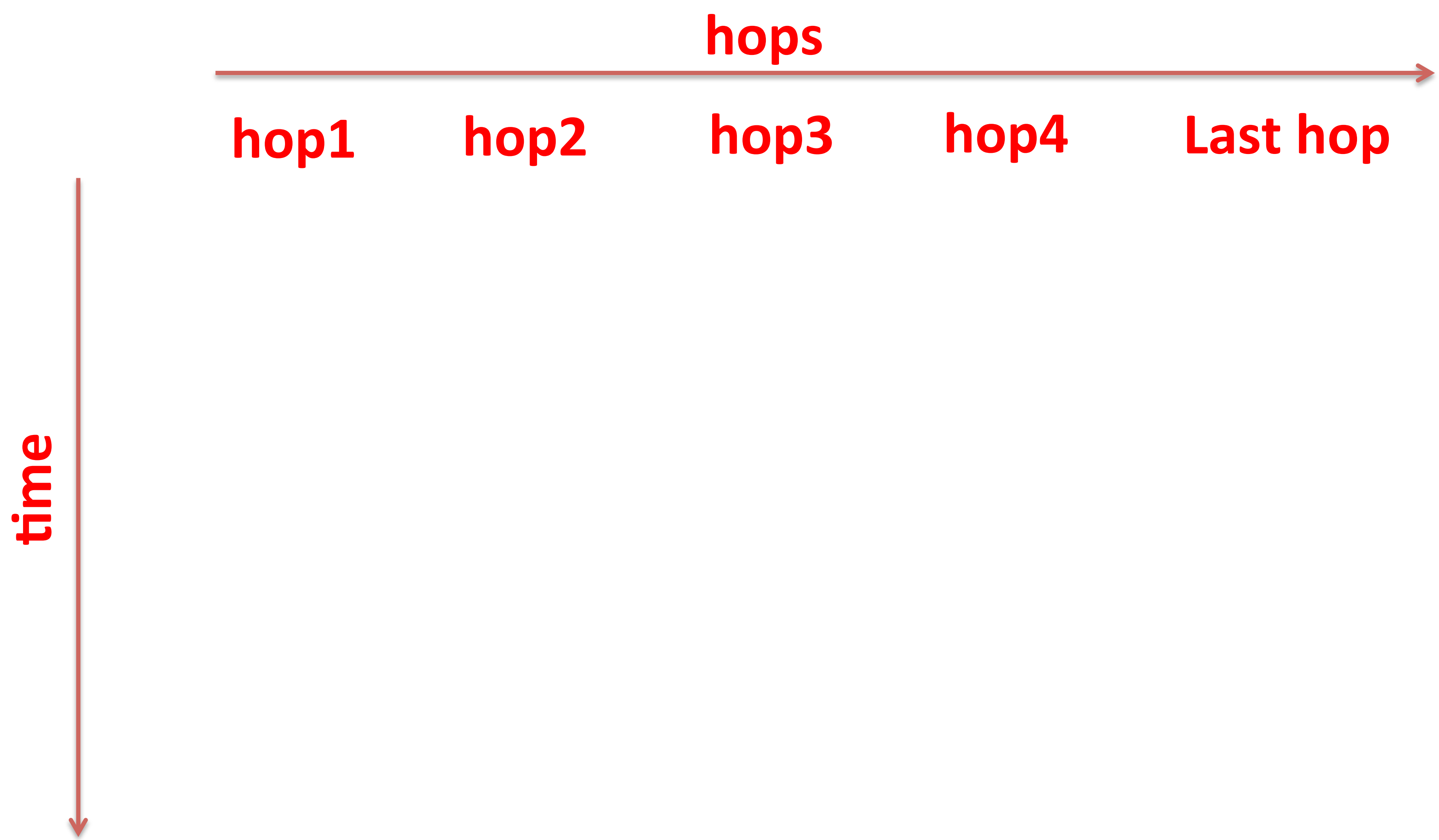
Server-to-Server Traceroute Measurements

- Unfortunately with ping measurements is not possible to locate where the congestion occurs.

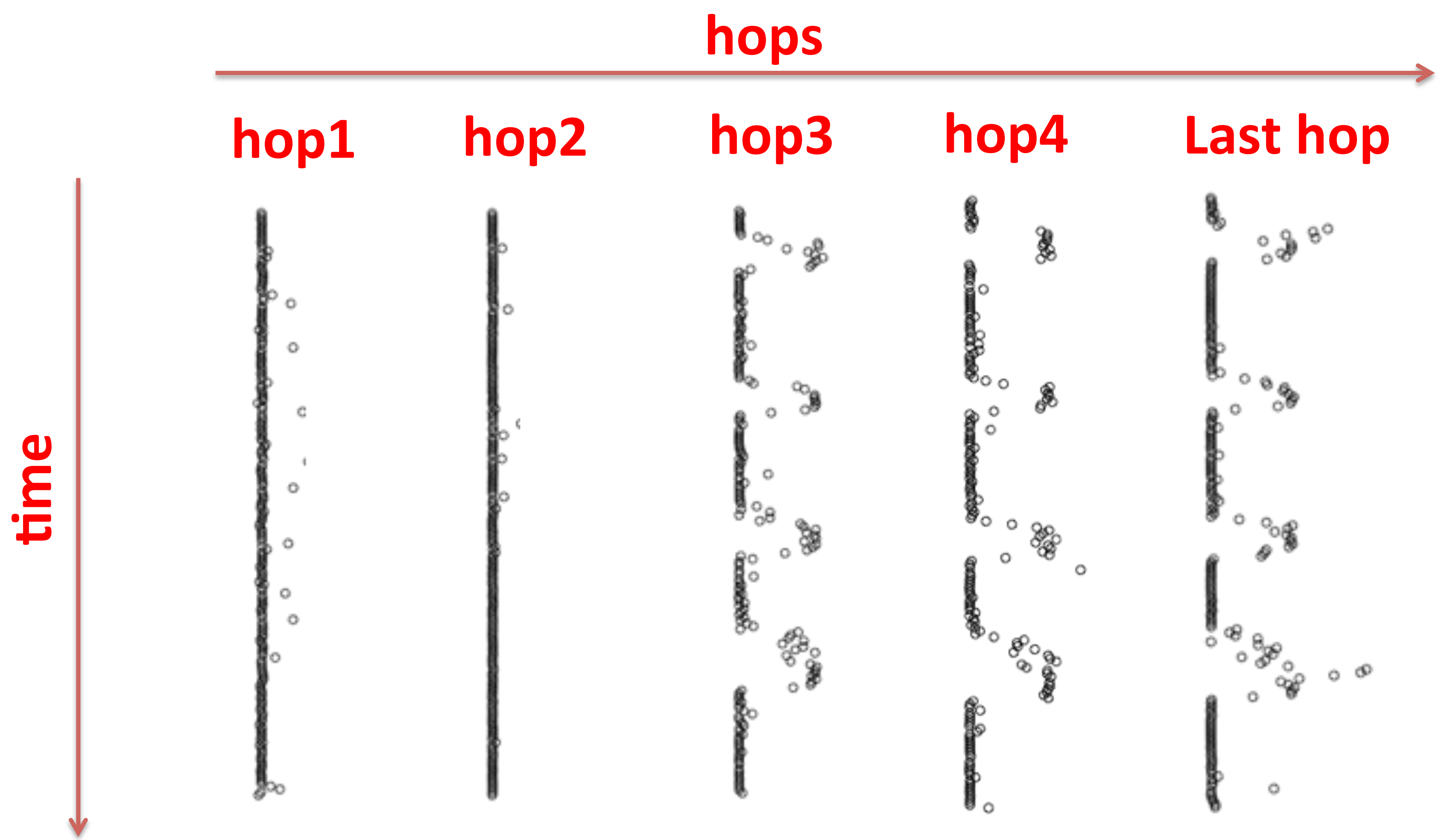
- We perform server-to-server traceroute measurements in both directions, for around 100K pairs

- Measurements span two weeks with frequency 1 traceroute every 30 mins.

Locating Congestion Points

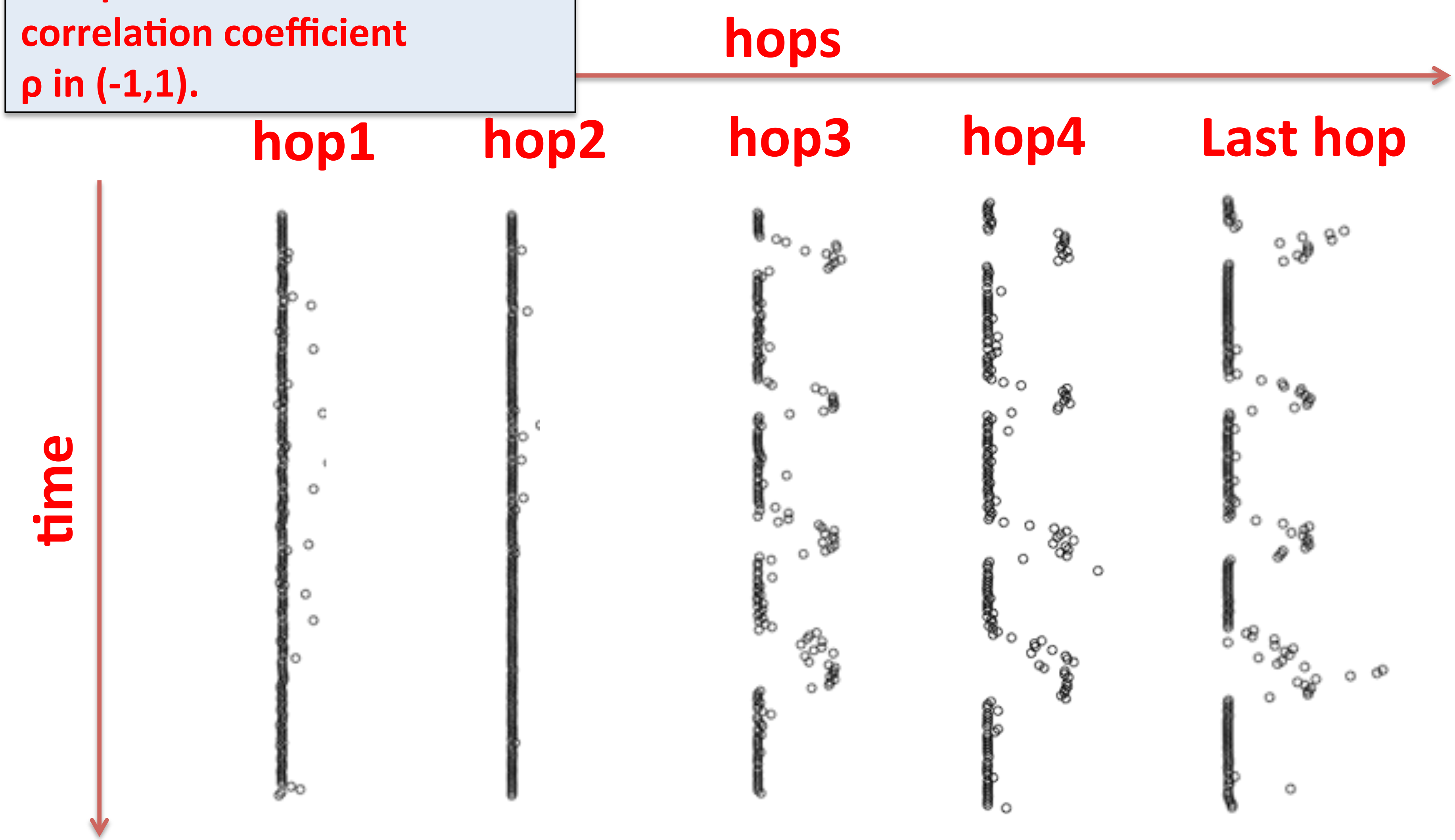


Locating Congestion Points



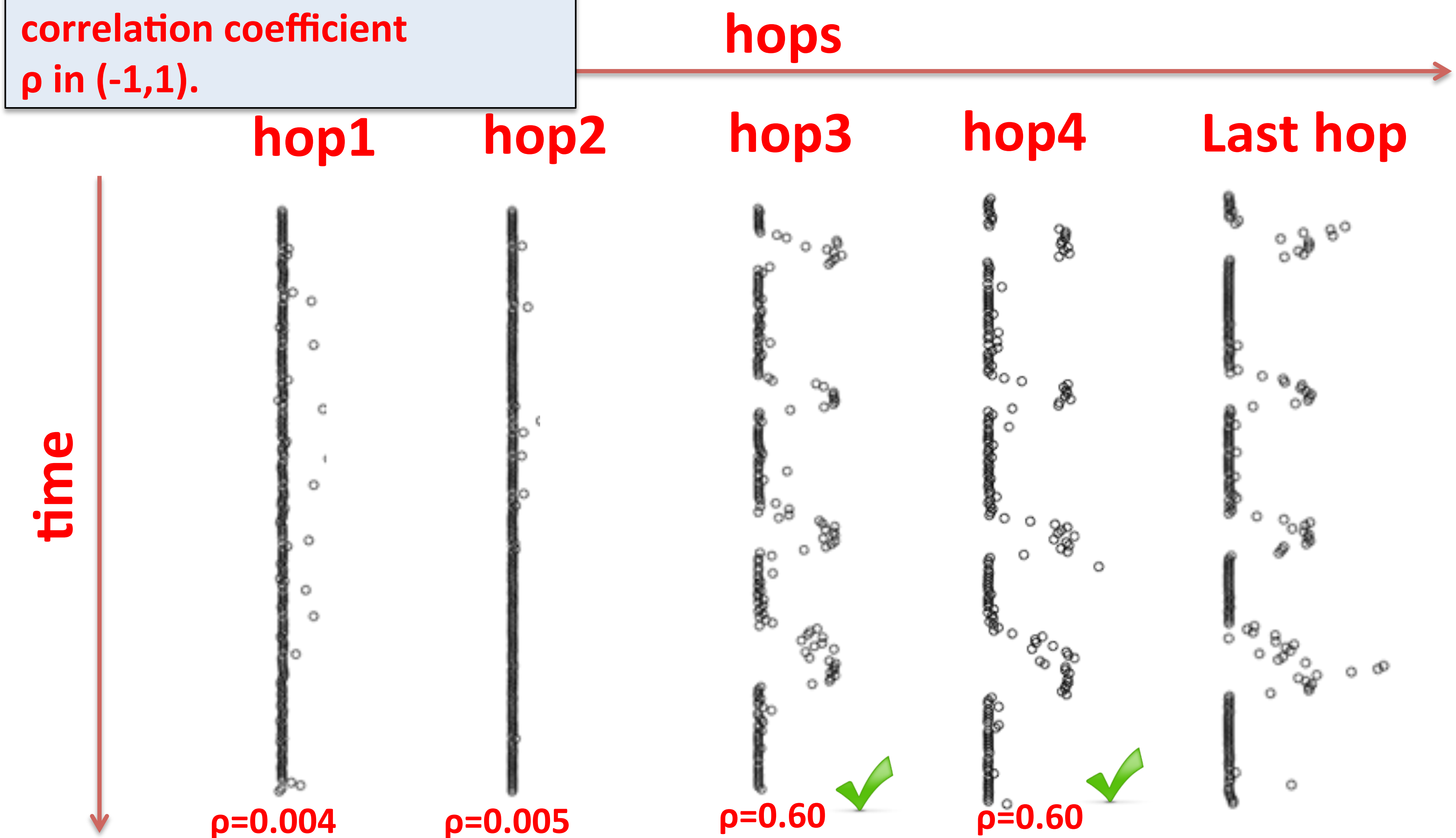
Locating Congestion Points

To locate the congested link:
Compute the Pearson correlation coefficient
ρ in (-1,1).



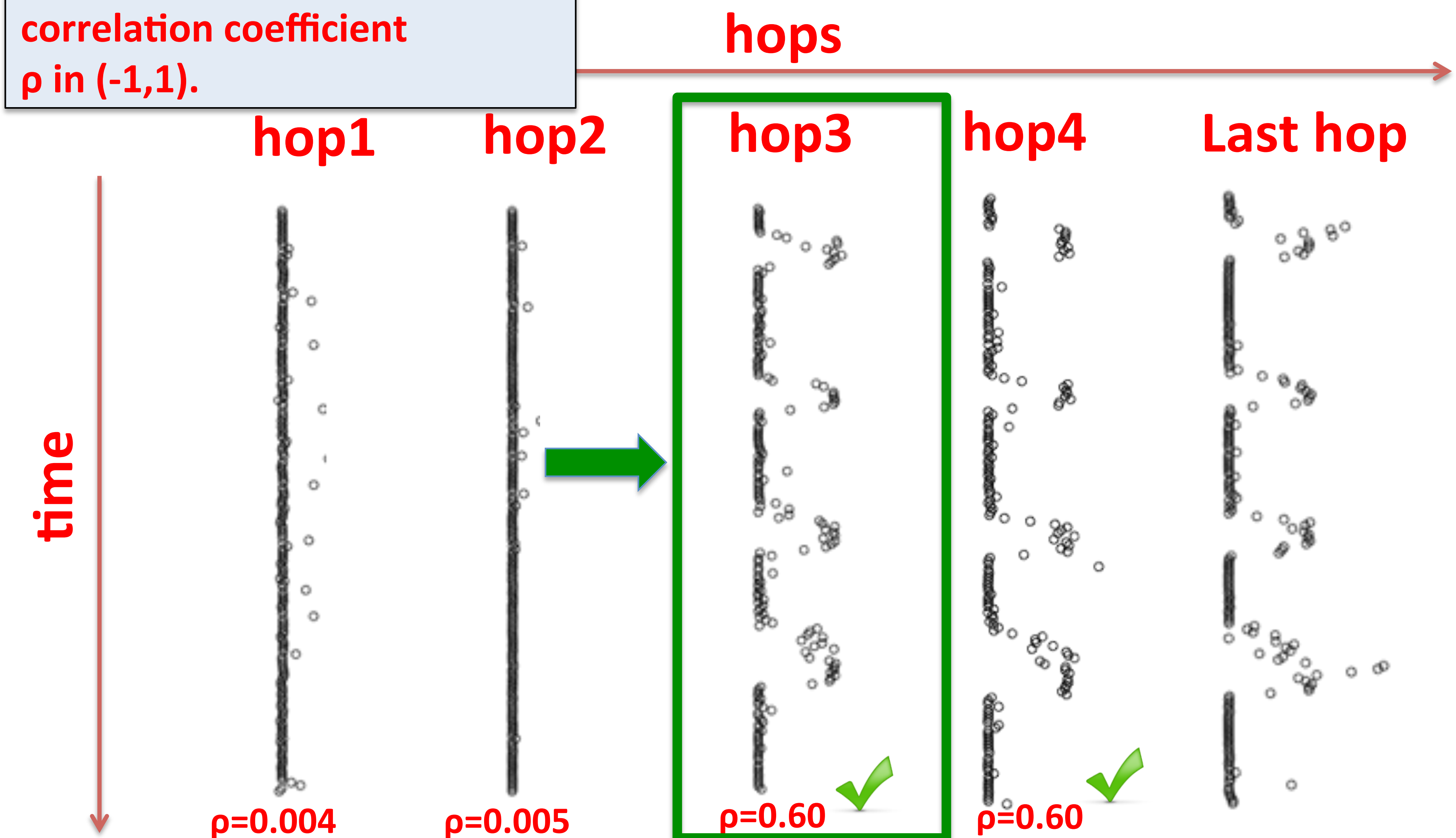
Locating Congestion Points

To locate the congested link: Compute the Pearson correlation coefficient ρ in (-1,1).



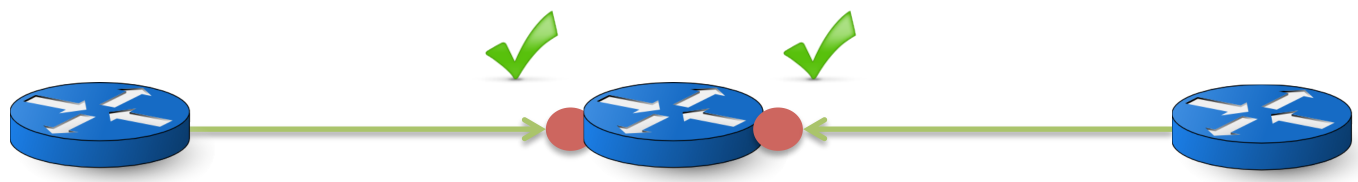
Locating Congestion Points

To locate the congested link:
Compute the Pearson correlation coefficient
ρ in (-1,1).

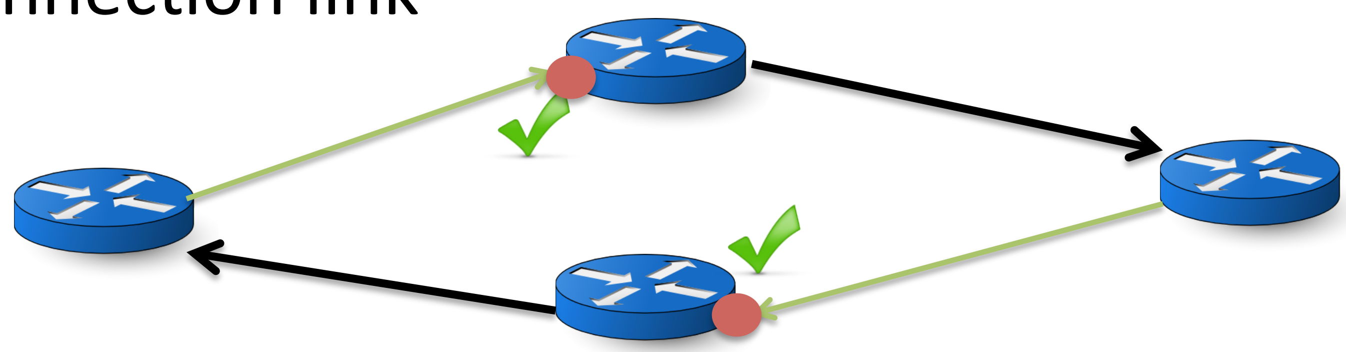


Locating Congestion Points

- Symmetric Routing: Forward and reverse infer the same router



- Asymmetric Routing: We can only argue internal/ interconnection link



Some Observations

- We investigated 310K links; we inferred around 3,000 links with persistent congestion

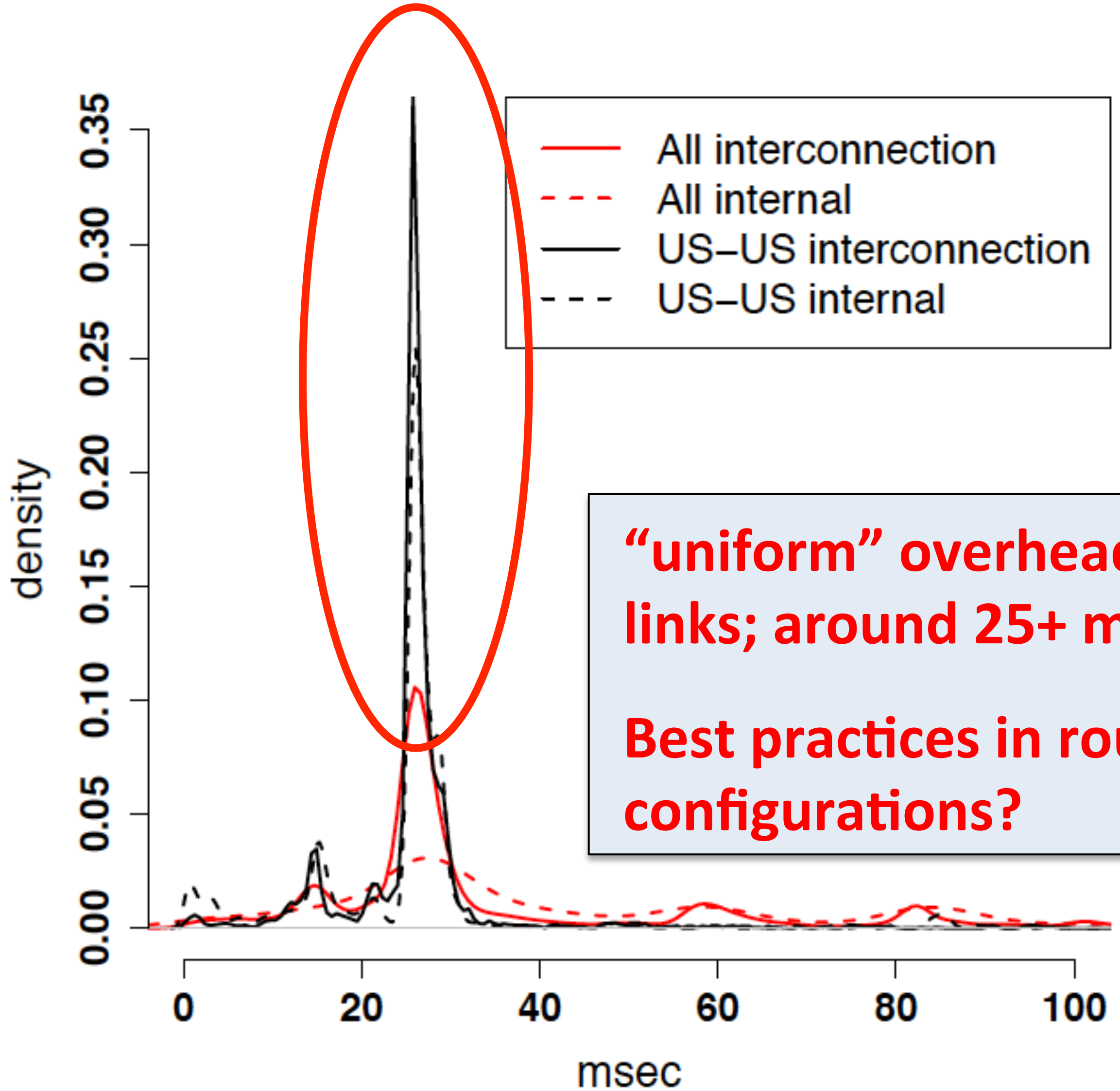
- Both internal and interconnection links were congested

- But, interconnection links were inferred from a large number of traceroutes, in some cases by >300 probes.

- Both customer-provider and peer-peer interconnections were congested

- Public peering links (at IXPs) were less congested than private interconnects.

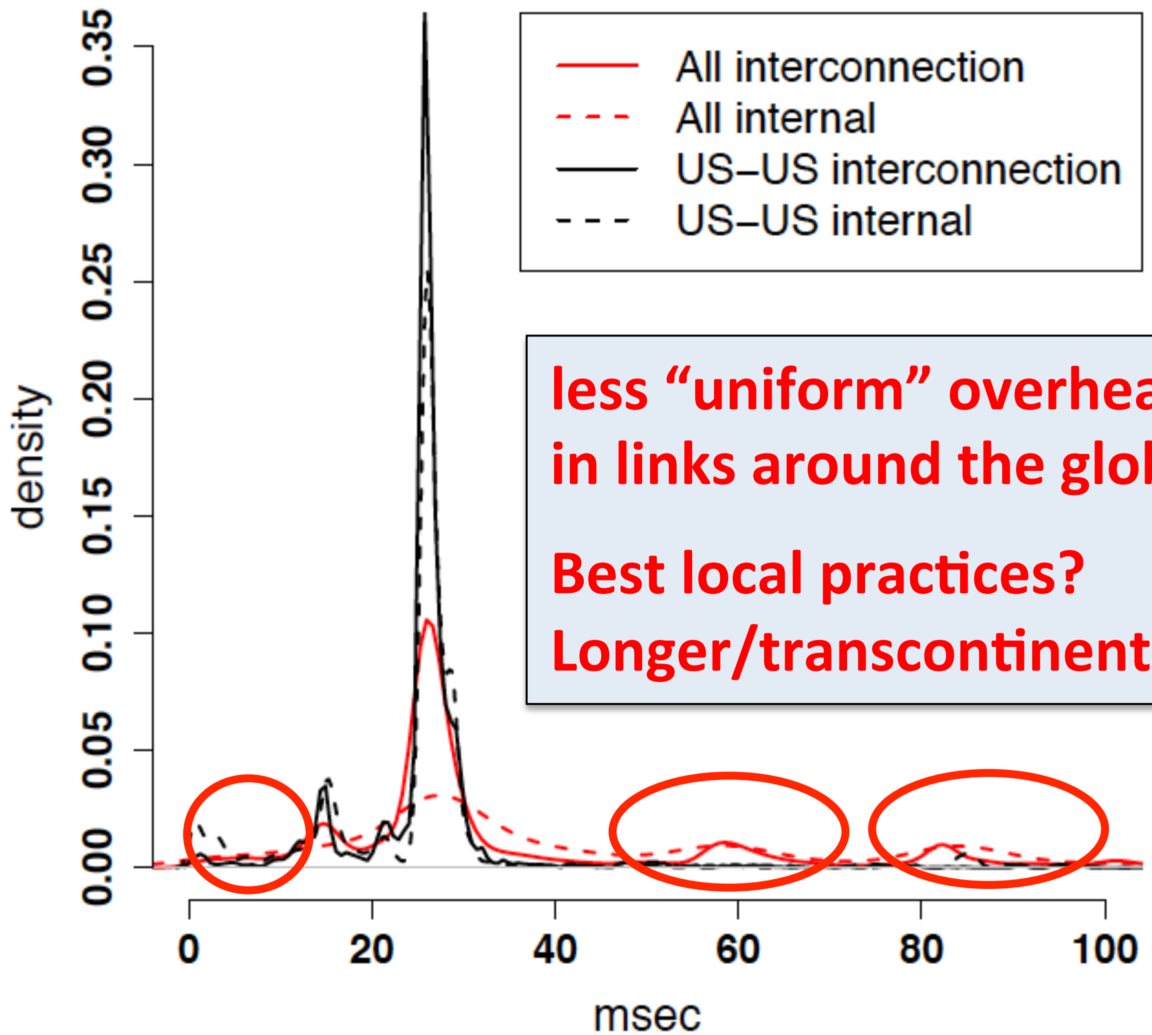
What is the Overhead of Congestion?



"uniform" overhead in US-US links; around 25+ msec

Best practices in router configurations?

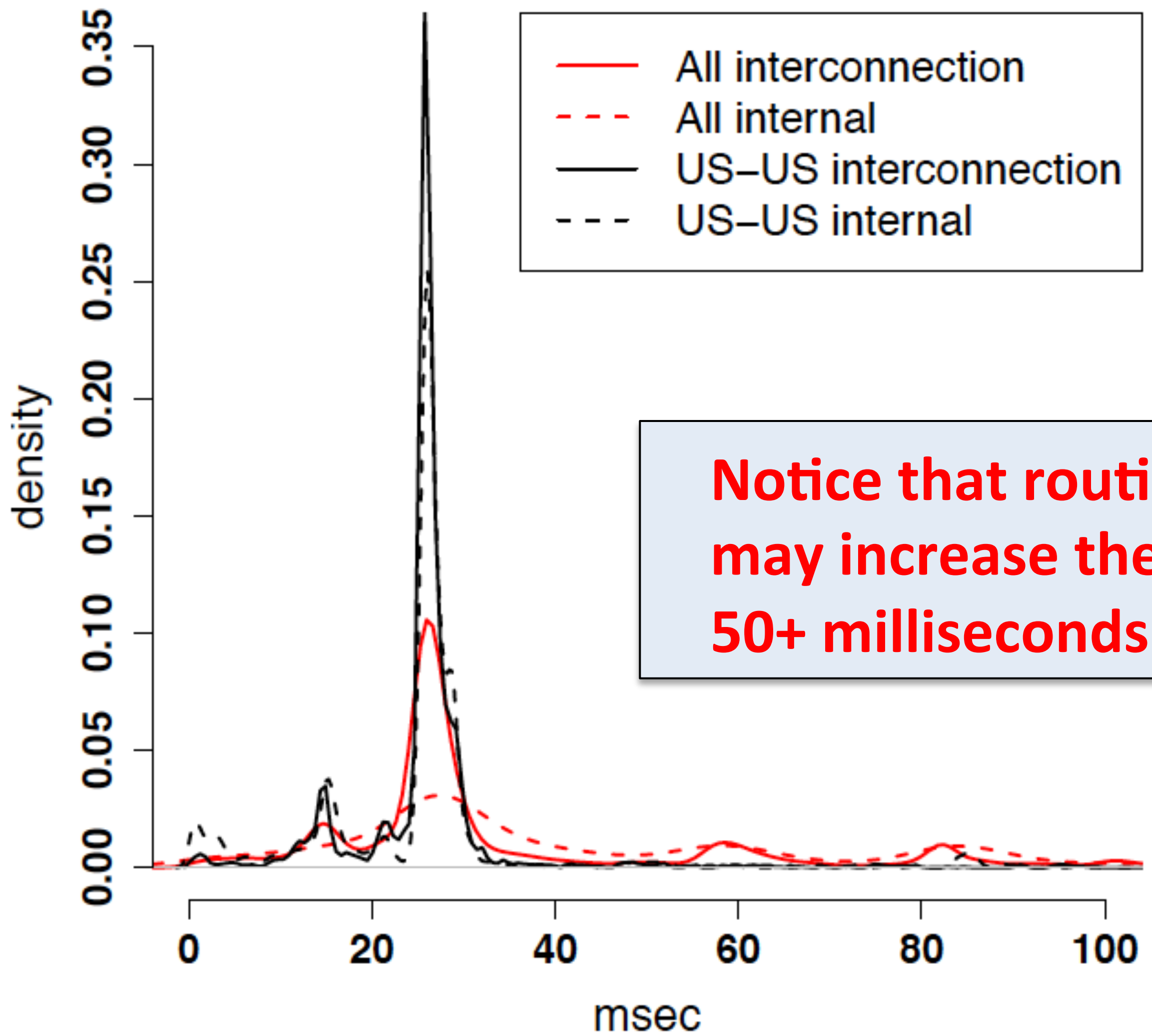
What is the Overhead of Congestion?



less “uniform” overhead
in links around the globe

Best local practices?
Longer/transcontinental distances?

What is the Overhead of Congestion?



Notice that routing changes may increase the delay by 50+ milliseconds.

Summary

- As we rely on a smooth operation of the Internet, any disruption such as congestion, has a negative impact on user experience and productivity

- We presented techniques to measure congestion and localize it to a link or a network

- Our large-scale study shows that congestion is not the norm, but in some paths it contributes to the end-to-end delay

Next Steps

- Continue to measure the Internet and seek for points of congestion

- Improve our techniques and deal with "black box" behavior; we welcome your help!

- Scale up our analysis

- Make an Internet "heat map" of congestion publicly available

Thank you!