MAPPING INTERNET INTERDOMAIN CONGESTION

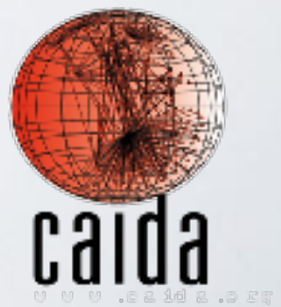
Amogh Dhamdhere, Bradley Huffaker, Young Hyun, Kc Claffy
(CAIDA/UCSD)
Matthew Luckie (Univ. of Waikato)
Alex Gamero-Garrido, Alex Snoeren (UCSD)
Steve Bauer, David Clark (MIT)

kc@caida.org

WHAT IS CAIDA?

Center for Applied Internet Data Analysis

– Founded by PI and Director k claffy
– Independent analysis and research group
– 20+ years experience in data collection, curation, and research
– Internet data collection tools, analysis, and data sharing
– Study security, stability, economics, performance of cyberinfrastructure
– located at the UC's San Diego Supercomputer Center
– supported by federal research funding (NSF, DHS), industry



OUTLINE OF TALK

- What is Internet cartography (mapping), why it is needed
- Transformations in Internet ecosystem that warrant regulatory (and cartographic) attention to interconnection
- NSF-funded UCSD/MIT collaboration to provide new techniques for mapping interconnection topology and performance
- Performance mapping: technique and validation
- Interconnection mapping: techniques and validation
- Some results and insights so far

INTERNET CARTOGRAPHY

why?

- best available data on Internet interconnection ecosystem is incomplete and of unknown accuracy
- no map of physical link locations, capacity, utilization, or interconnection arrangements
- hinders R&D, studies of robustness, resilience, economic sustainability

good reasons: technical, security, economic, market incentives

INTERNET CARTOGRAPHY

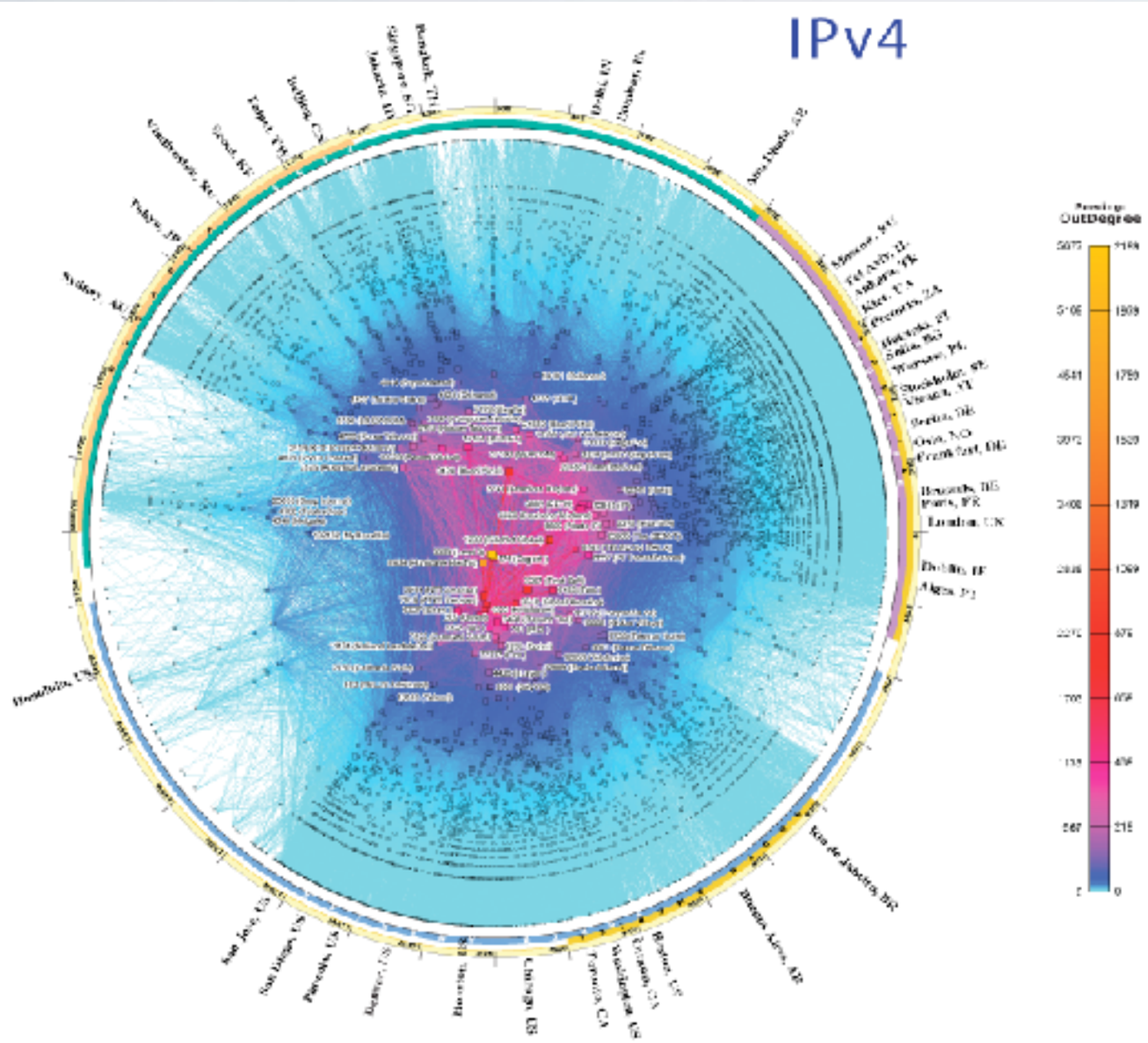
why?

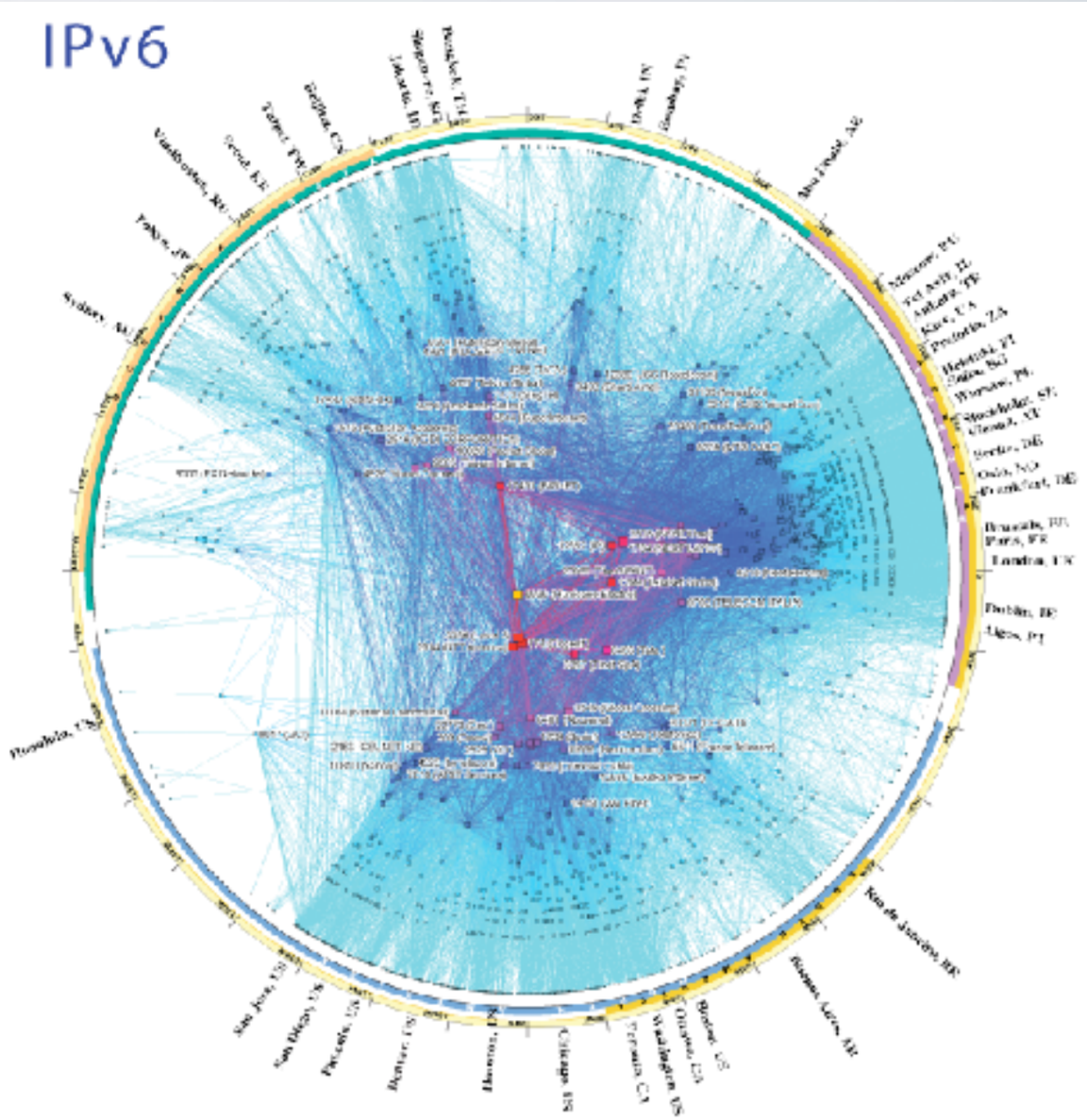
- capture and understand Internet security, resilience, performance characteristics
  - deployment of new technology, e.g., IPv6, DNSSEC
  - compliance with security best practices
  - performance issues, e.g., congestion, outages, hijacks
  - infer macroscopic properties

goals: support situational awareness, inform science, engineering, policy

2015 "AS CORE"

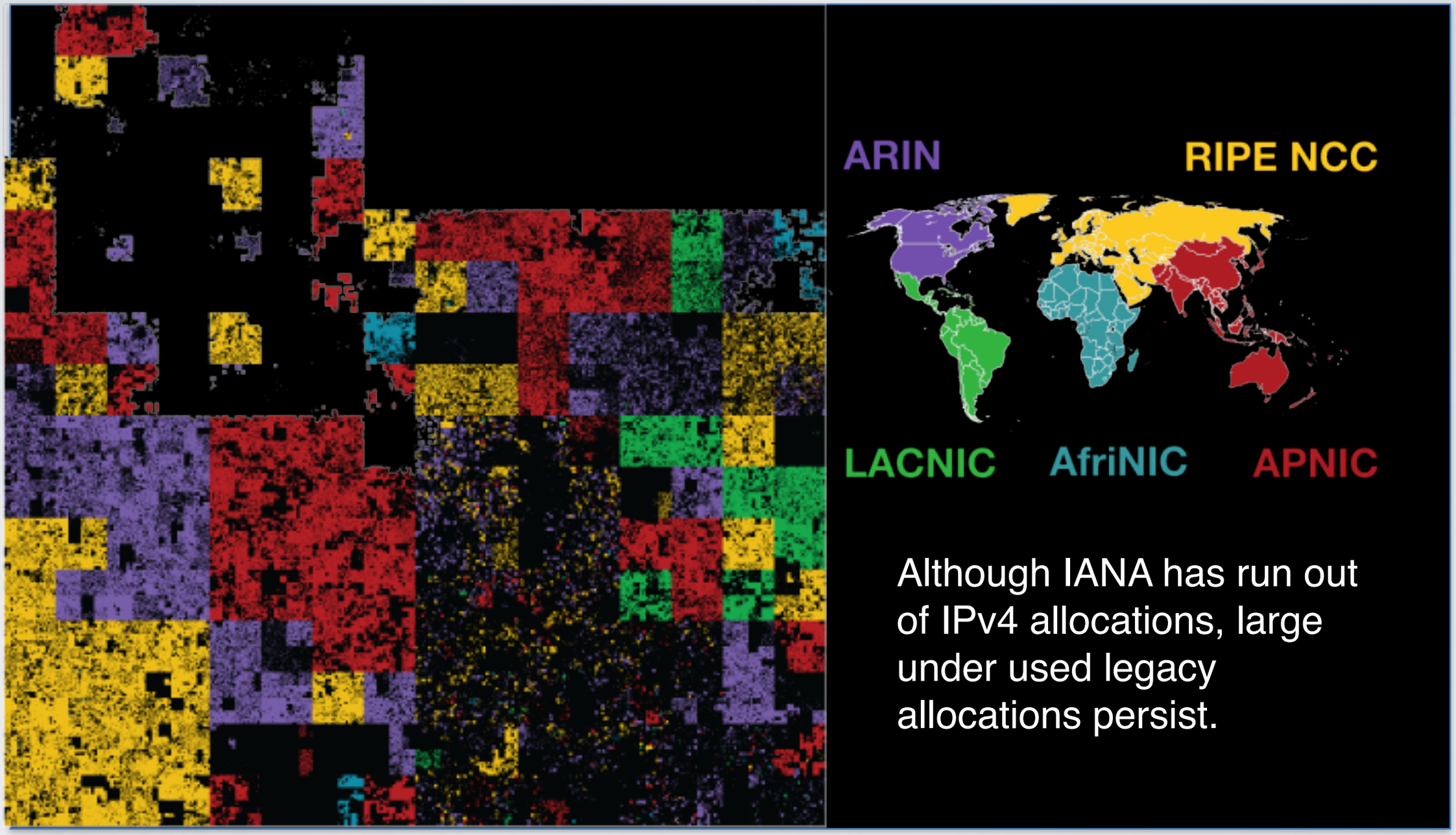
visualizing connectivity at ISP granularity





http://www.caida.org/research/topology/as_core_network/

2013 IPV4 CENSUS



http://www.caida.org/research/id-consumption/census-map/

TRUSTWORTHY CYBERSPACE?

(another angle)

France Telecom Accused Of Holding YouTube Videos Hostage Unless It Gets More Money

from the *more-peering-disputes* dept

'Peering' Into AOL-MSN Outage

Level 3 and Comcast Issue Statement

Jul 16, 2013

BROOMFIELD, Colo., July 16, 2013 – Level 3 and Comcast have resolved their prior interconnect dispute on mutually satisfactory terms. Details will not be released.

Confirmed: Comcast and Netflix have signed a paid peering agreement

by Stacey Higginbotham FEB. 23, 2014 - 9:27 AM P

Netflix packets being dropped every day because Verizon wants more money

Verizon wants to be paid by consumers and Cogent, but Cogent refuses to pay.

Cogent Gearing for Another Peering Battle

Verizon denies using net neutrality victory to sabotage Netflix, Amazon

BY **BRIAN FUNG** February 5 at 1:59 pm

Netflix still sucks on AT&T, and now AT&T plans to offer Netflix clone

AT&T partners with an investment firm to buy and launch streaming services.

caida

ISP COOPETITION

(historical context)

- TCP/IP architecture derived from a 60's project to build a "***cooperative network of time-shared computers***"

- Not designed for an adversarial environment except in the sense of resilience to disruption

- Not designed for competitive provisioning of services

- Not designed for sustained high-volume bitstreams

- Gives rise to vast set of compromises in security and trustworthiness

EXACERBATING TRENDS

- *Fundamental shifts in structure of ecosystem*

  - Increasing influence of Internet Exchange Points (IXPs) in facilitating co-location and inter domain connectivity

  - Concentration of content among few Content Delivery Networks (CDNs) who can change traffic dynamics on short timescales

- Ability of large access ISPs to use market power

- Resulting threat to stability: disputes over who invests in capital infrastructure to support traffic growth

ONE SUCH COMPROMISE

(performance degradation)

- ***Who upgrades the capacity between networks?***
  (Answer: Eventually, whatever is worked out under NDA.)

- Modern interconnection disputes among access, content, and transit providers manifest as congested links

- In the meantime, some content is carried over inadequate links between access and transit networks

- **Congestion on transit links affects everybody**, not just parties to the peering dispute

CSAIL

caida

INTERDISCPLINARY PROBLEM

(like most SATC problems)

- Economic basis for peering decisions lacks a discipline
- Commercial secrecy of all things related to peering inhibits development of discipline
- Internet architecture does not provide features to exchange money for tiered service across networks
- **Now a matter of regulatory interest**

CSAIL

caida

(this doesn't go on for too long before govt gets involved)

France Telecom Accused Of Holding YouTube Videos Hostage Unless It Gets More Money

from the *more-peering-disputes* dept

'Peering' Into AOL-MSN Outage

Level 3 and Comcast Issue Statement

Jul 16, 2013

BROOMFIELD, Colo., July 16, 2013 – Level 3 and Comcast have resolved their prior interconnect dispute on mutually satisfactory terms. Details will not be released.

Confirmed: Comcast and Netflix have signed a paid peering agreement

by Stacey Higginbotham FEB. 23, 2014 - 9:27 AM PDT

Netflix packets being dropped every day because Verizon wants more money

Verizon wants to be paid by consumers and Cogent, but Cogent refuses to pay.

Cogent Gearing for Another Peering Battle

Verizon denies using net neutrality victory to sabotage Netflix, Amazon

Netflix still sucks on AT&T, and now AT&T plans to offer Netflix clone

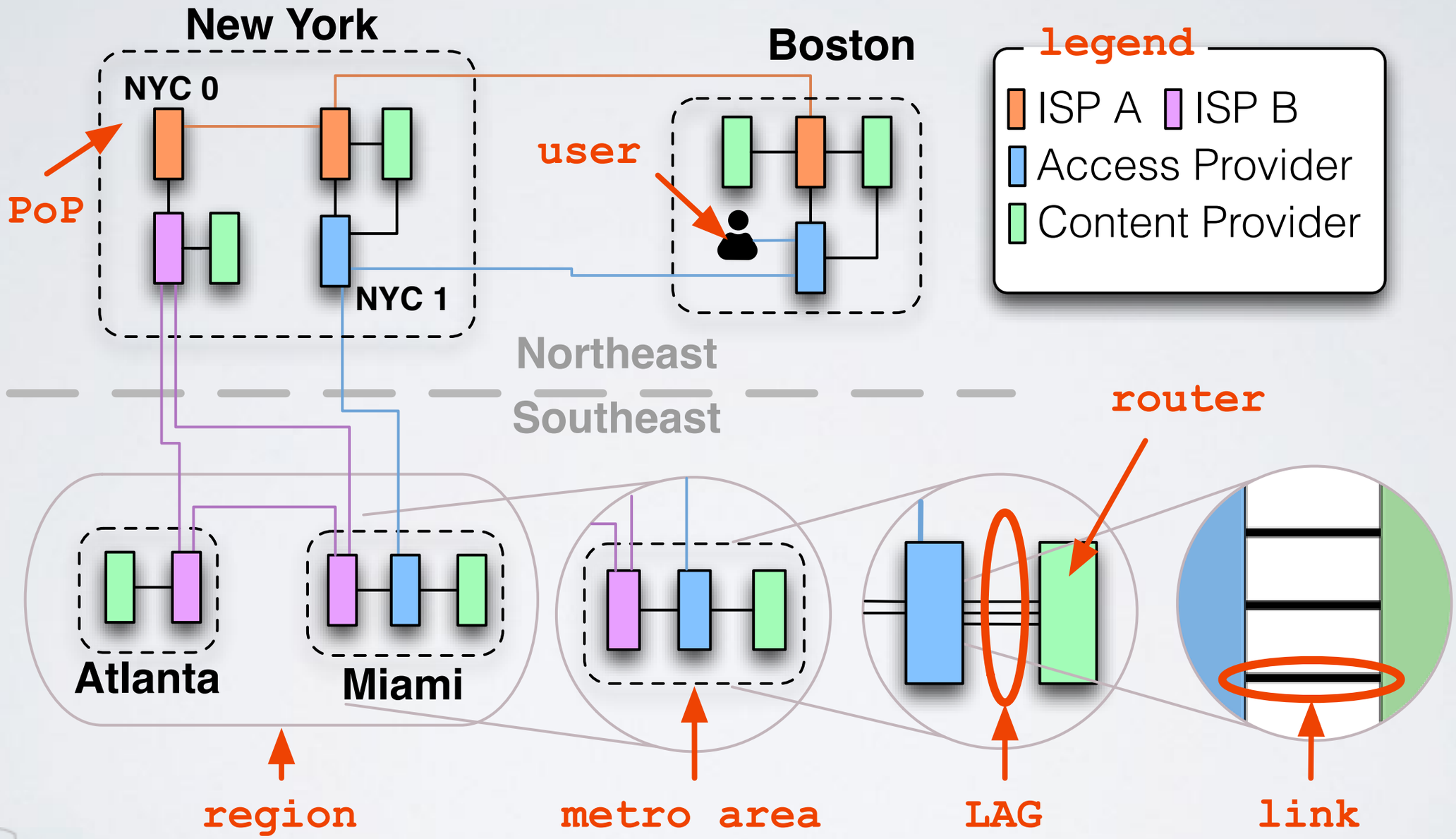
AT&T partners with an investment firm to buy and launch streaming services.

CSAIL

caida

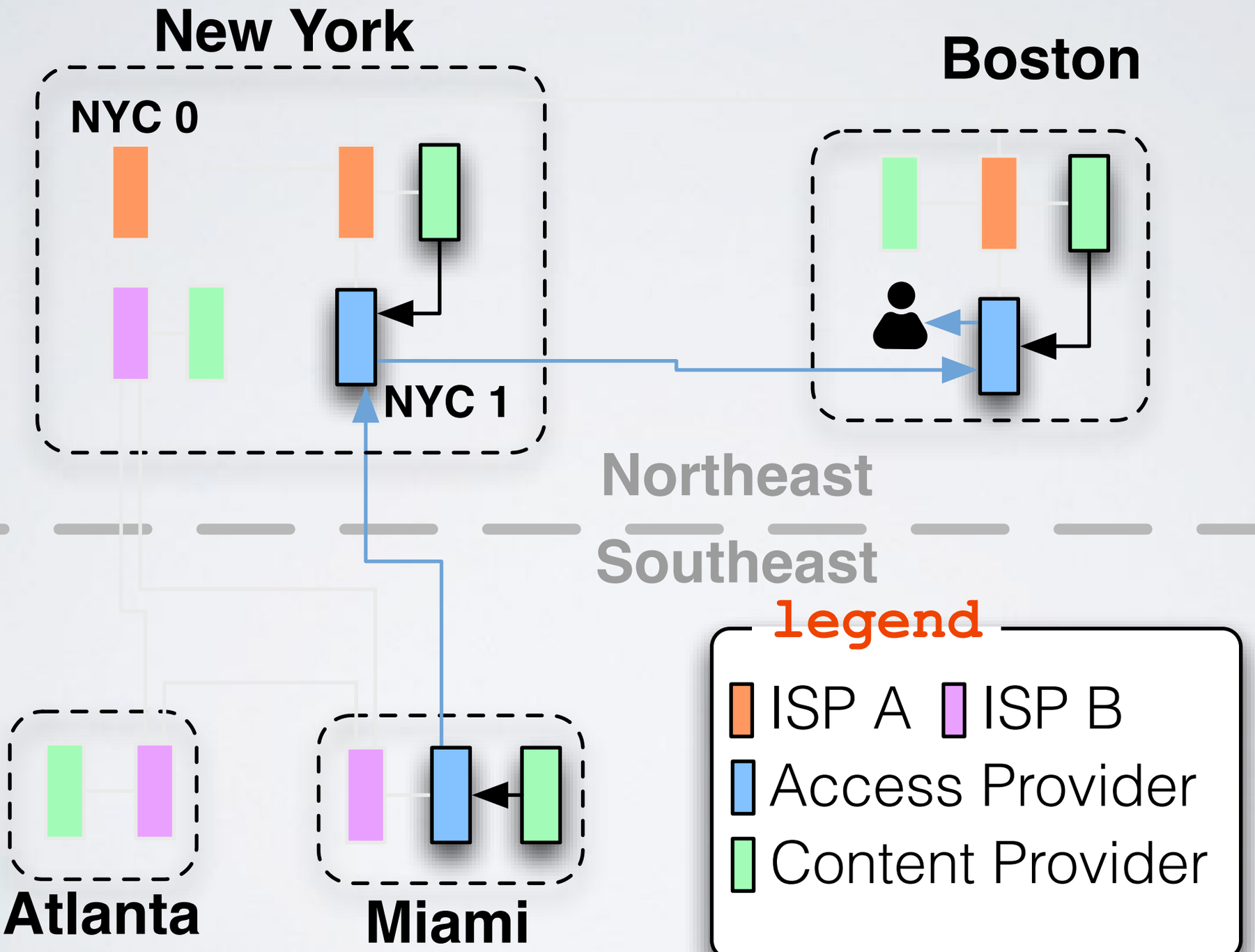
INTERCONNECT COMPLEXITY

(and this is extremely oversimplified)



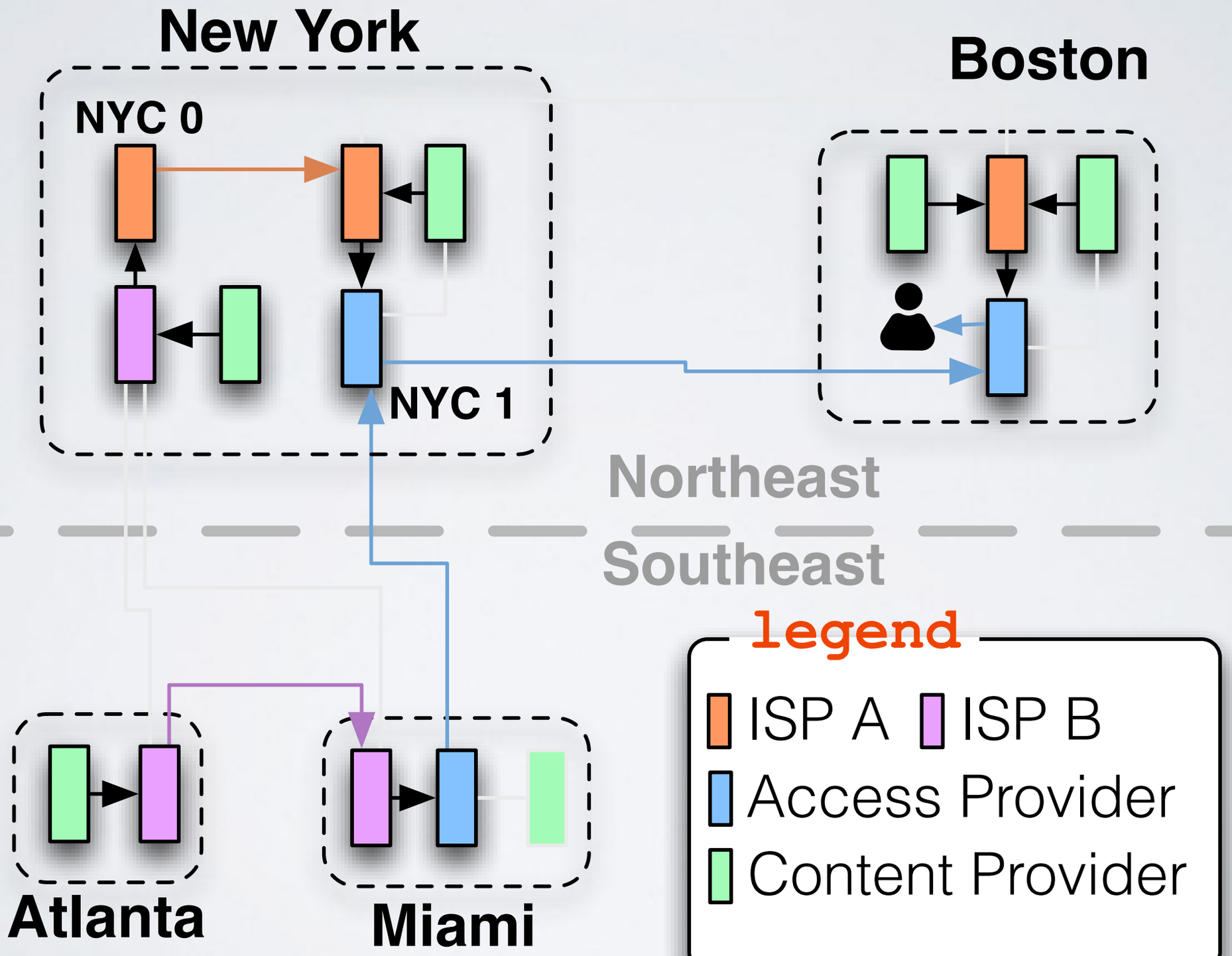
DIRECT CONNECTIONS

(increasing between large access and content providers)



INDIRECT CONNECTION

(because direct connections may be fully utilized)



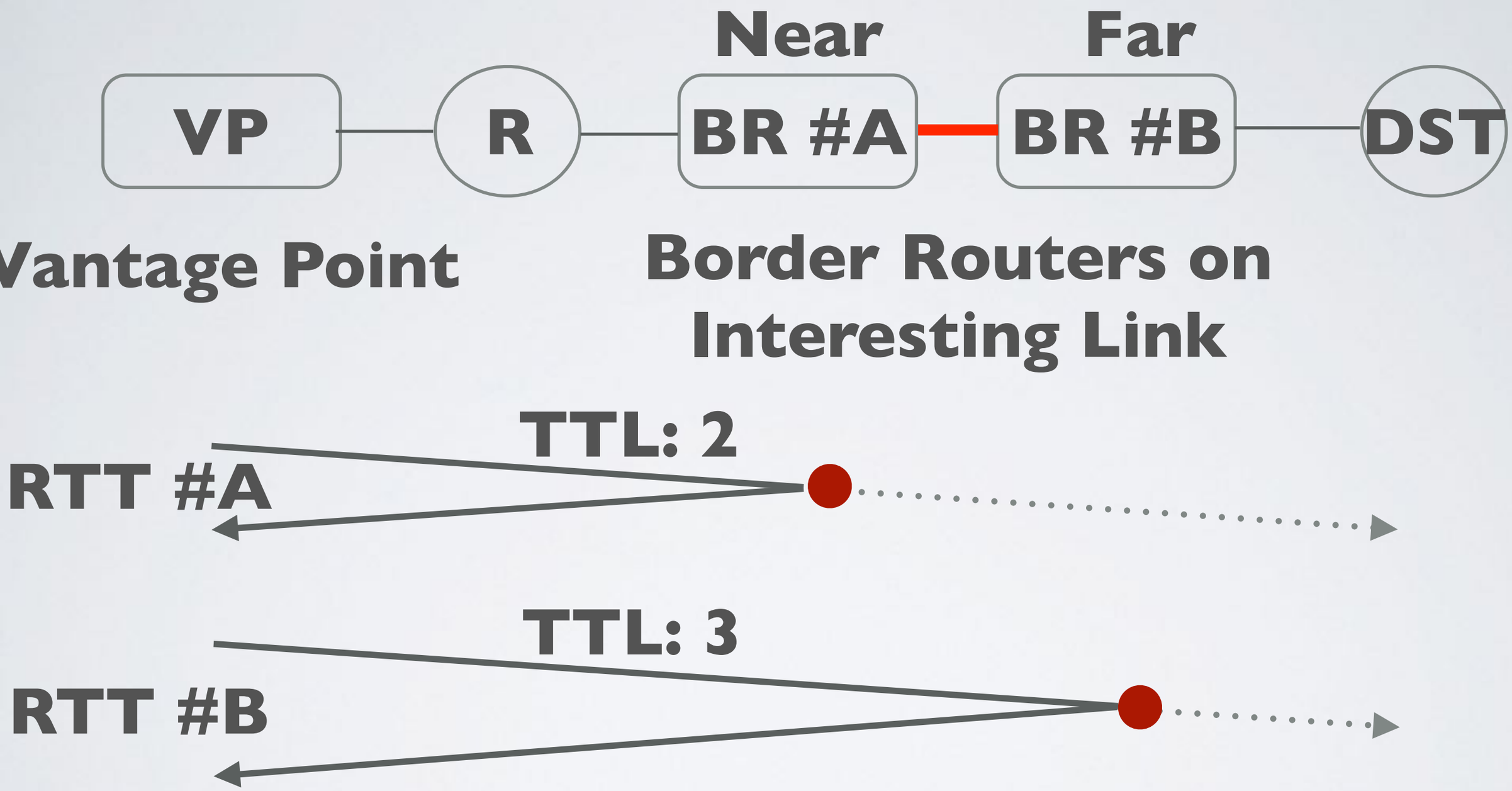
MAPPING INTERDOMAIN CONGESTION

NSF CNS-1414177 (UCSD) and NSF CNS-1413905 (MIT)

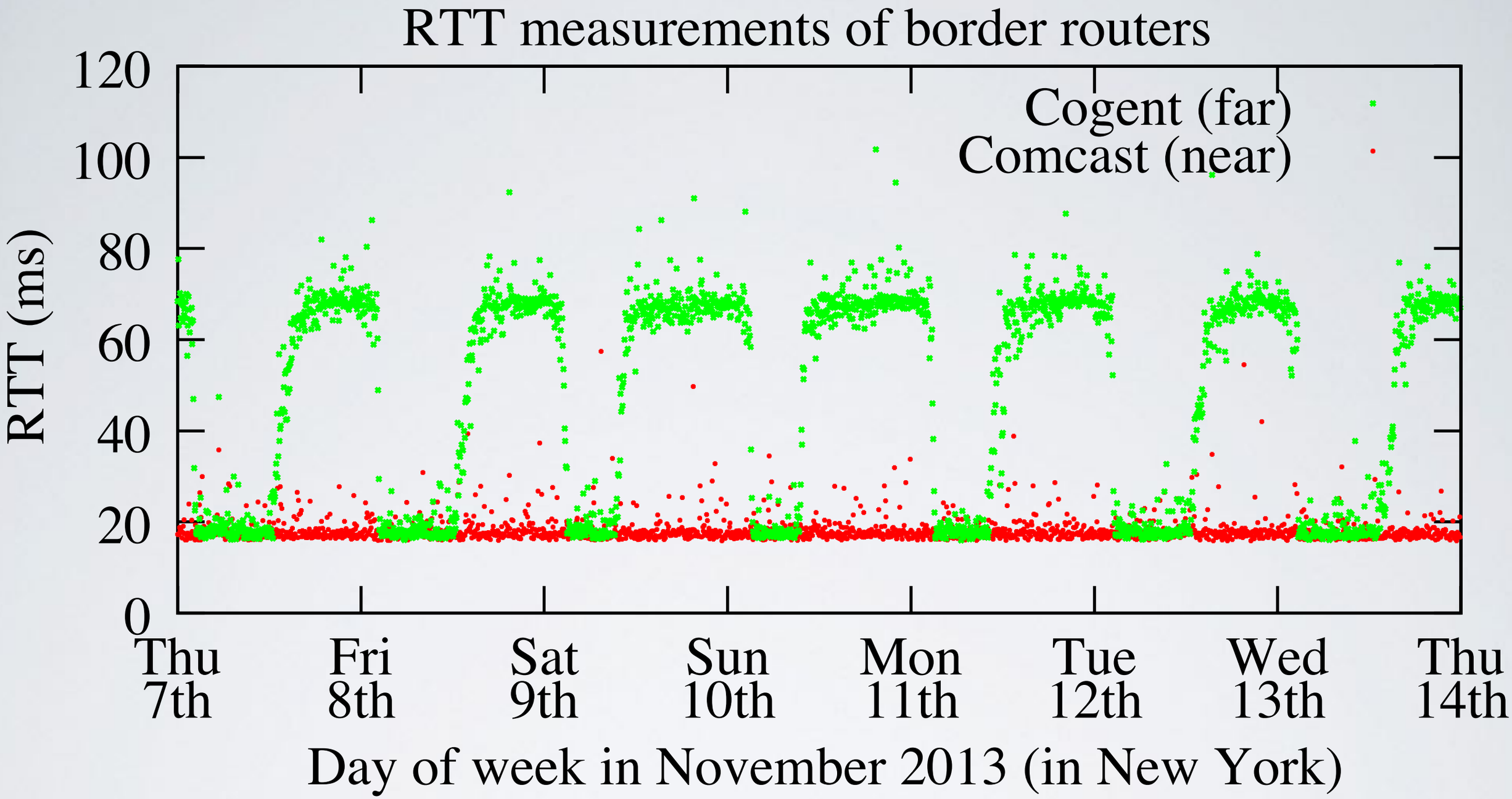
- Characterize changing nature of the inter-domain Internet
- Measure location and extent of interdomain congestion
- **Our goals (1) a system to monitor interdomain links and their congestion state, (2) a near real-time "congestion heat map" of the Internet, (3) increasing transparency, empirical grounding of debate**

"Mapping Interconnection in the Internet: Colocation, Connectivity, and Congestion" http://www.caida.org/funding/nets-congestion/

METHOD: TIME SERIES PING

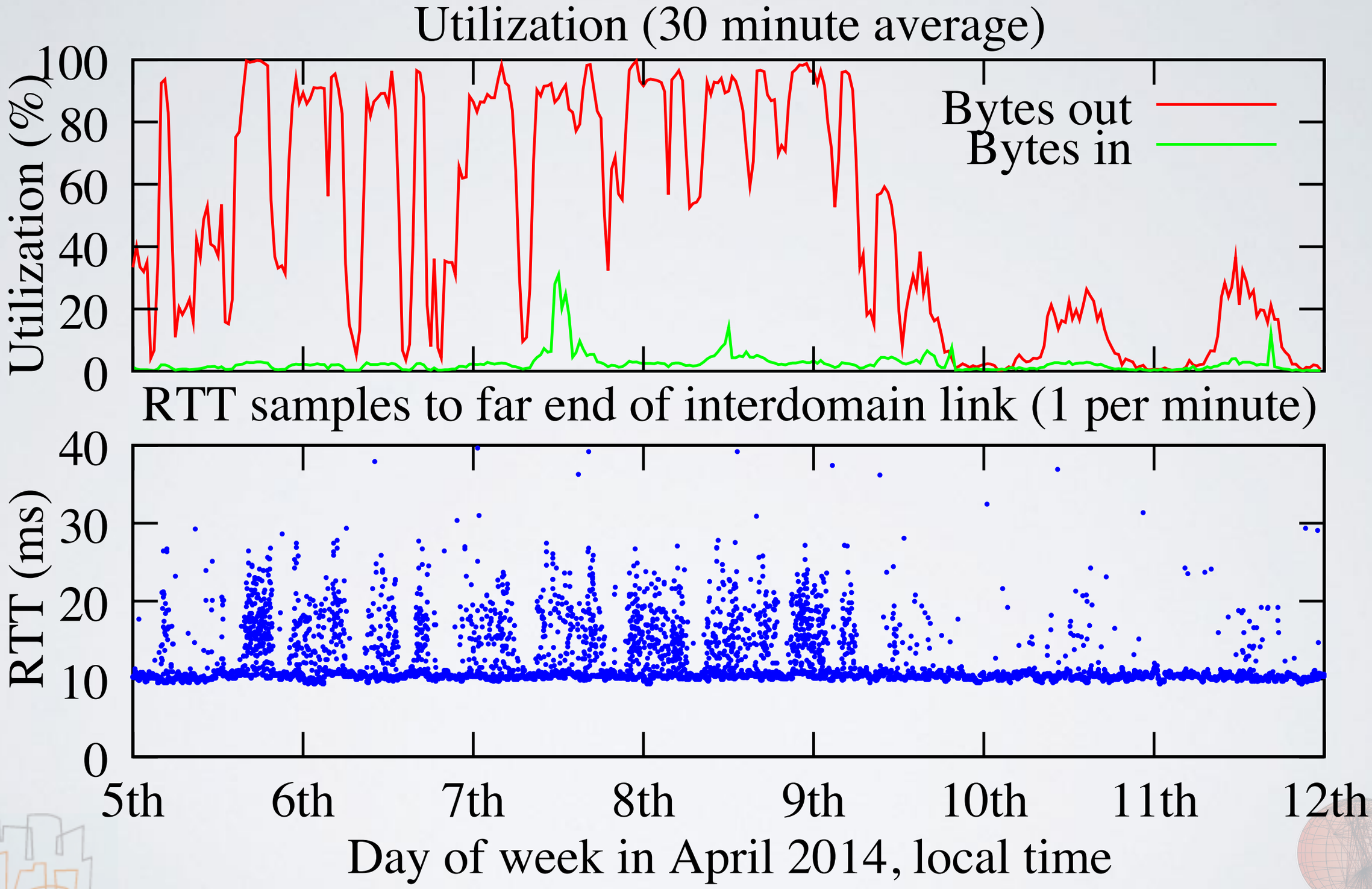


(repeat to obtain a time series)



More congestion on weekend than weekdays. Monday 11th was Veterans Day

VALIDATION: TRAFFIC VS TSLP



CSAIL

caida

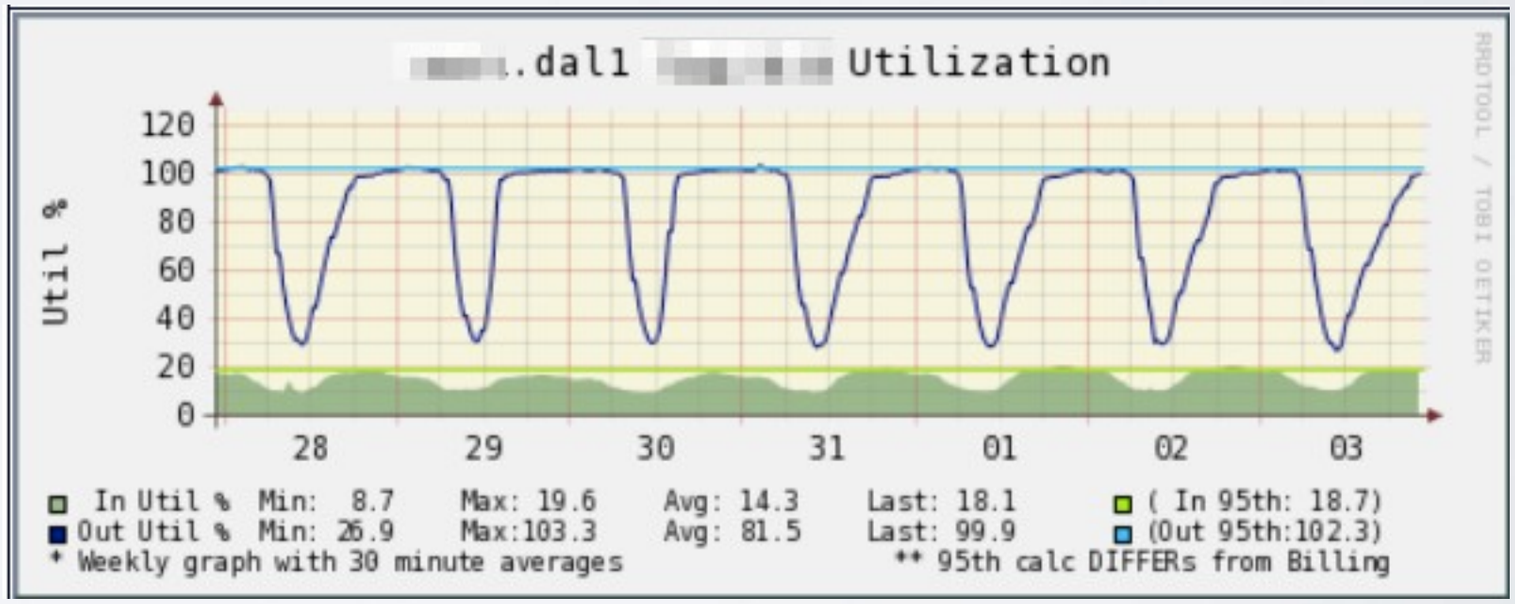
VALIDATION: WE GET LUCKY

(with concrete example)

- For links that show diurnal RTT pattern, how does pattern correlate with traffic data? But peering agreements contain NDA.

- Closest to public data: Level3's blog "Observations of an Internet Middleman"

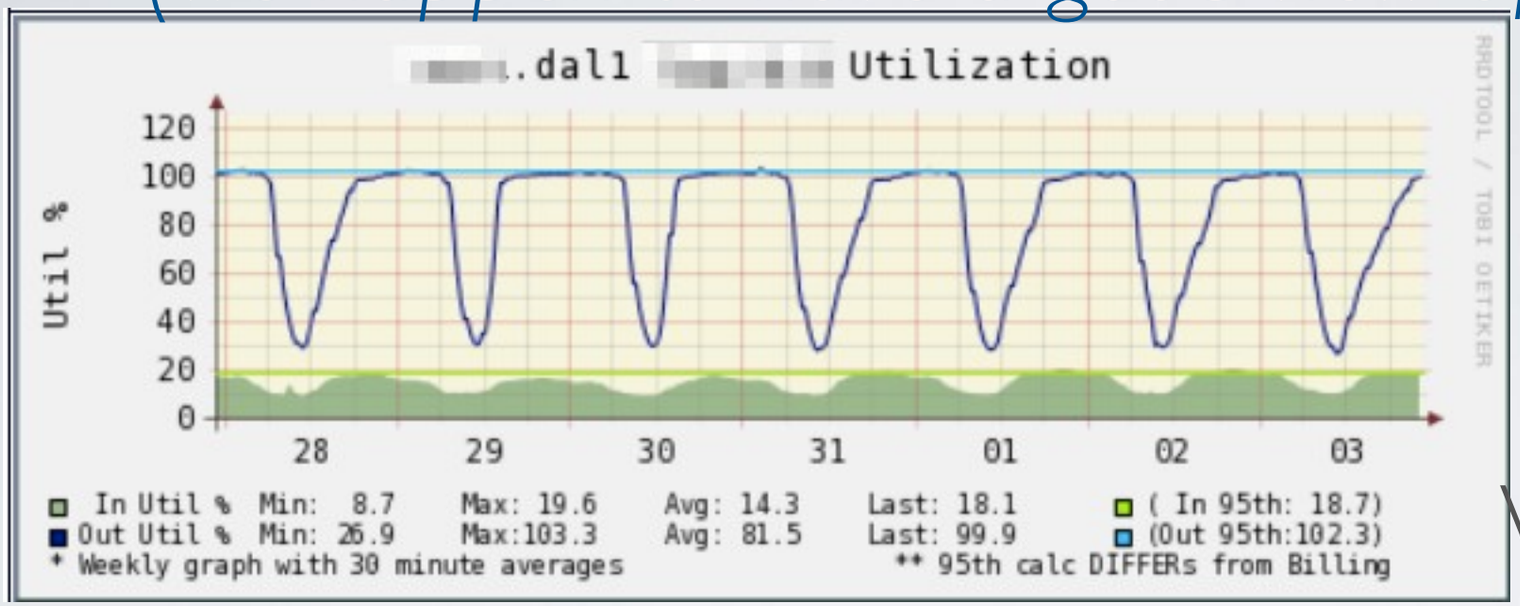
Anonymous Dallas Link

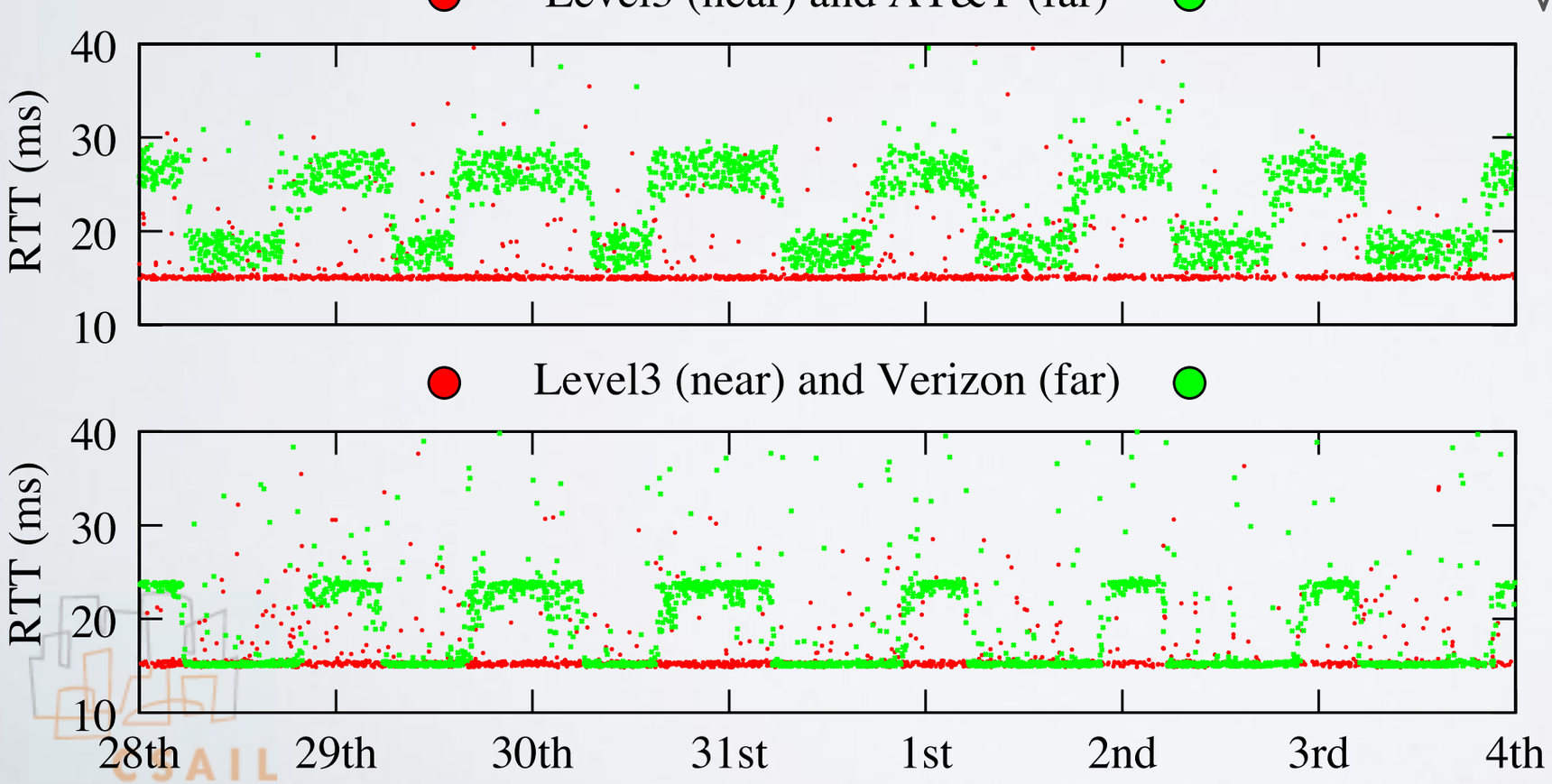


http://blog.level3.com/global-connectivity/observations-internet-middleman/

TSLP VALIDATION: LUCK

(we happen to have a good view of Level3-Dallas)





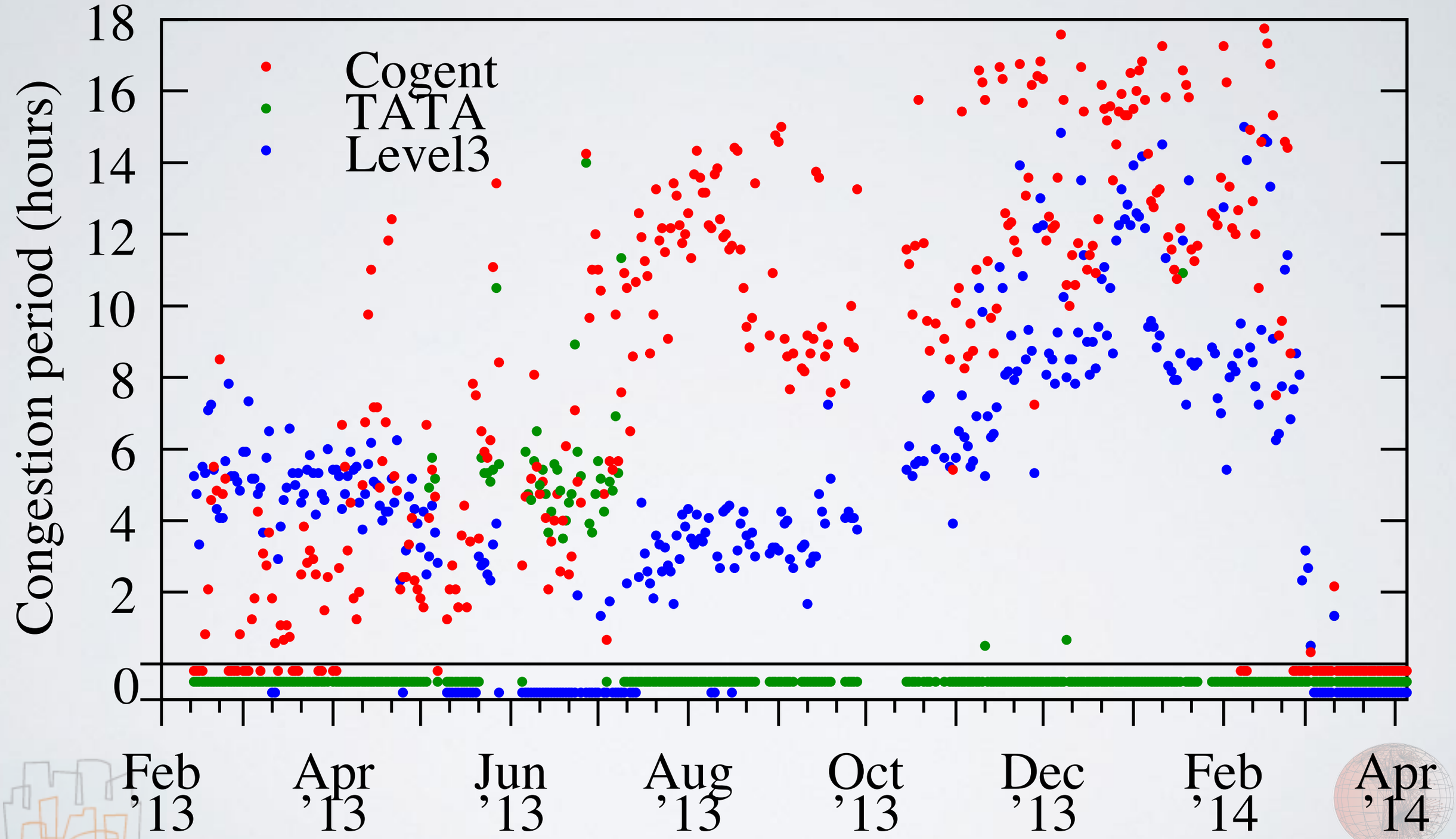
"Ground Truth"

We believe both links were congested with Level3 in Dallas

One looked like perfect fit (based on duration of level shift)

CONGESTION TRENDS

(three links of Comcast in Bay Area over time)



CSAIL

caida

TSLP CHALLENGES: AQM, WFQ

- AQM techniques such as **Random Early Detection** (RED) try to maintain a small average queue size by discarding packets before a queue becomes full

  - Hypothesize more gradual shift from low to high delay

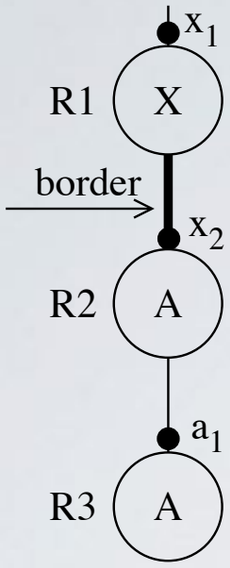
- Fair queueing techniques such as **Weighted Fair Queueing** (WFQ) place packets in different queues according to some property, such as the incoming port or sender.

  - TSLP cannot infer link congestion when our packets are

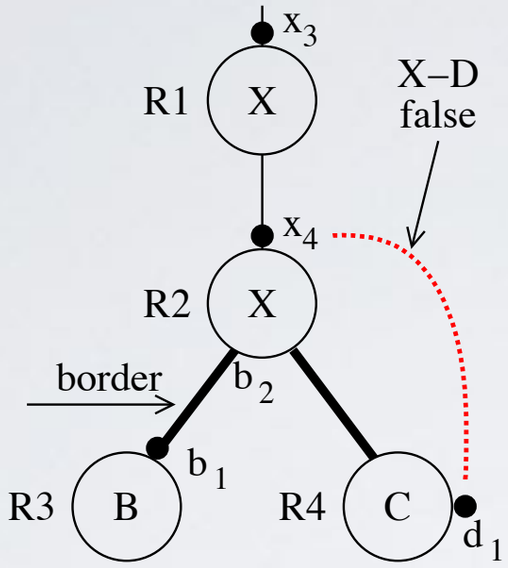
UNSOLVED CHALLENGE: MAPPING NETWORK BORDERS

- TCP/IP architecture has no notion of network boundaries (though there are hundreds of thousands of them!), nor mechanism to identify a complete router
- Canonical topology measurement tool (traceroute) & router-level inferences are a series of clever hacks and heuristics
- Operator address assignment and router implementation practices limit accuracy of mapping IP address to owner
- Traceroute sample bias reduces accuracy of router inference
- Operator interest in privacy inhibits transparent naming

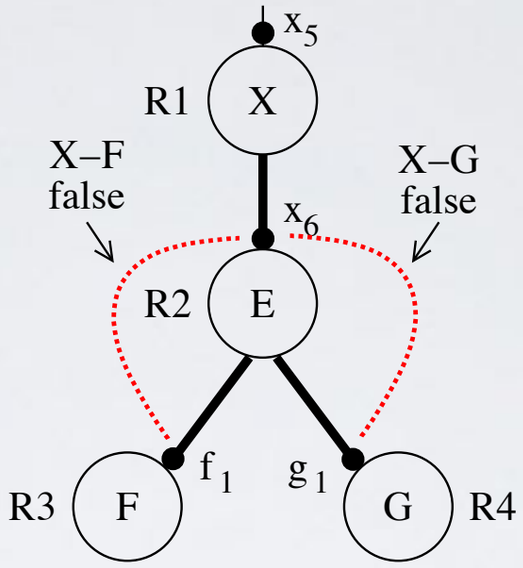
CHALLENGE: ROUTER OWNERSHIP



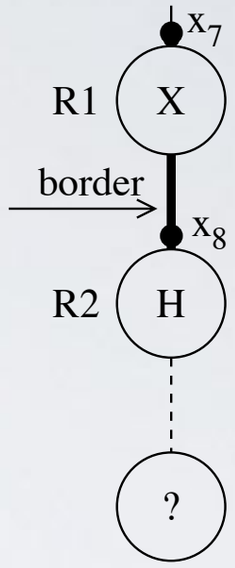
(a) unclear which link is the interdomain link



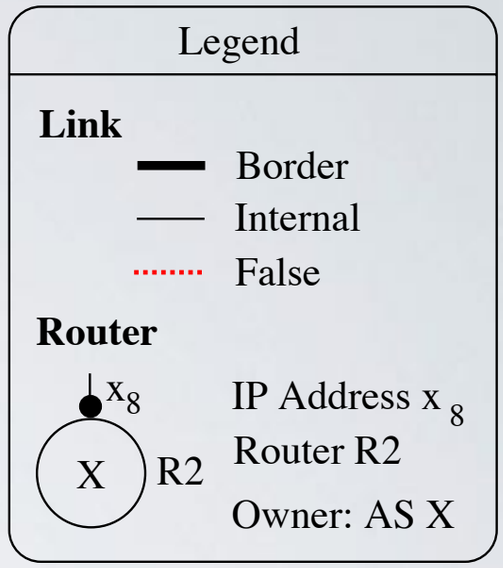
(b) third party addresses confuse border identification



(c) 1–hop in neighbor network



(d) no responding destination



- More details on the difficulty of inferring boundaries: "*bdrmap*: Inference of borders between IP networks" (just one hop upstream…) IMC2016

- Related work: Chen *et al.* — Where the Sidewalk Ends: Extending the Internet AS Graph Using Traceroutes From P2P Users. CoNEXT 2009.

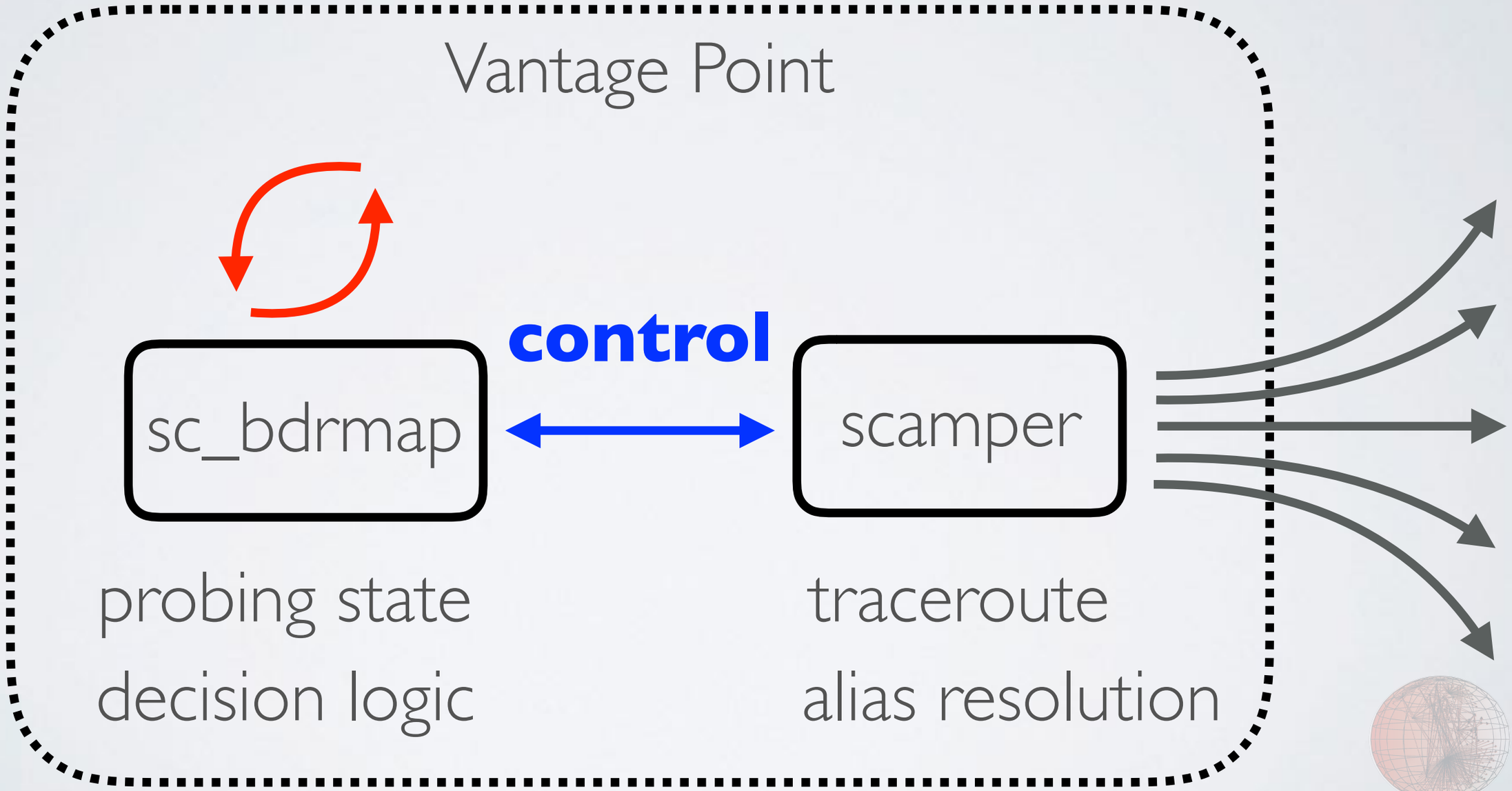
CSAIL

caida

SIGNIFICANT STEP FORWARD

sc_bdrmap

- Produces graph of routers, annotated with links and AS owners

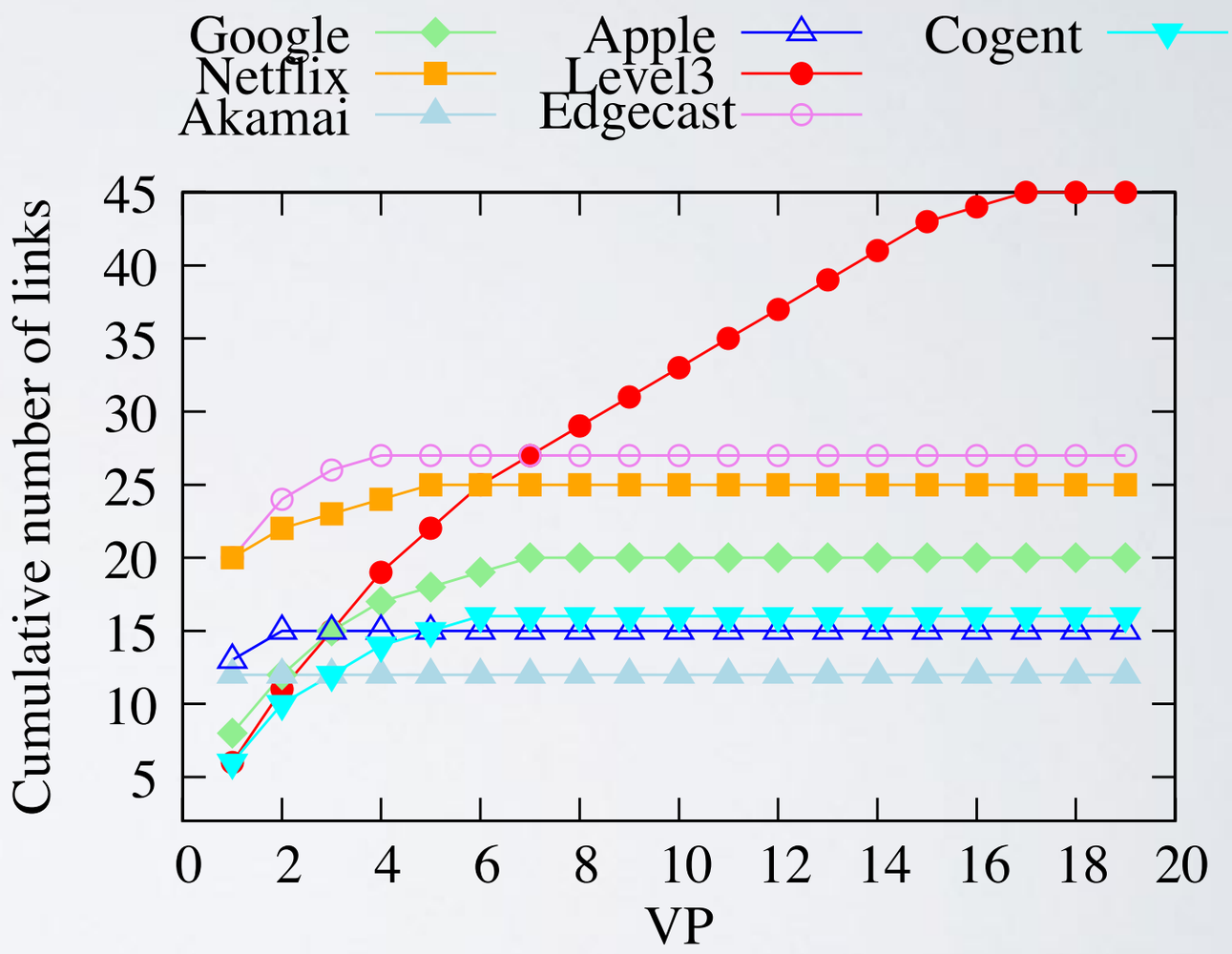


CSAIL

caida

BORDER MAPPING

- Running process from ~60 monitors in diverse networks
- Validation succeeded from a few networks (need more)
- Rich fabric: One network was connected to Level3 with 45 router-level links!
  - Need VPs close to those locations to find these links



CSAIL

caida

BORDER MAPPING

OTHER CHALLENGES

- Adaptiveness to changes in underlying topology and routing
- Scalable techniques to find signals of congestion
- Efficient data management and processing, alerts, reactive measurements
- Store enough information to be able to reconstruct topology, links/targets probed, at any given point in the past
- Validation: Difficult to get ground-truth; most peering agreements are covered by NDAs

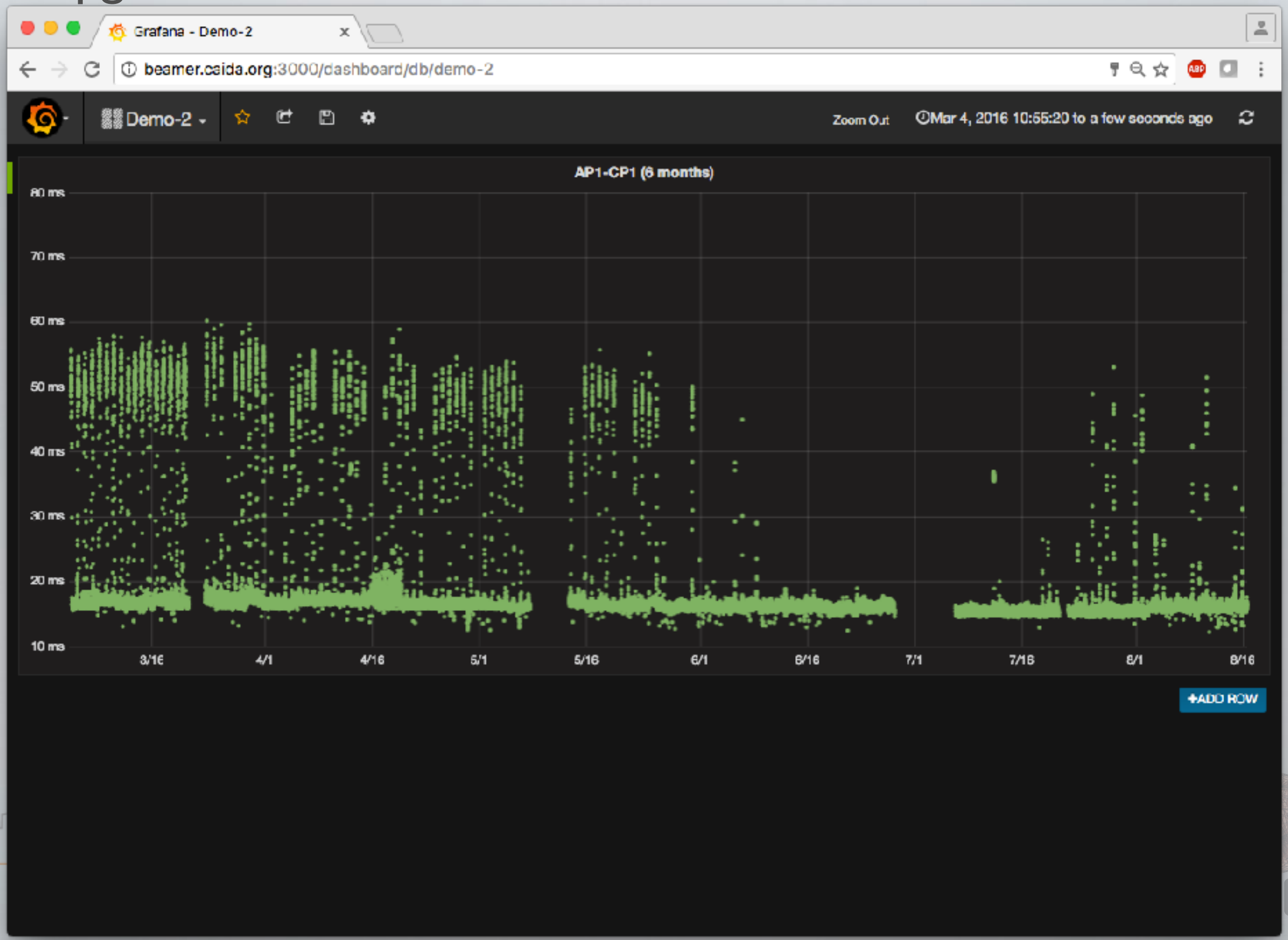
DATA!

- We probe the near and far interface of every interdomain link every 5 minutes
- TSLP data pulled from monitors and put into database
- Use InfluxDB: SQL-like query interface
- InfluxDB + Grafana: Interactive querying plus visualization
- Bonus: updates in near real-time!

DATA EXPLORATION

- Screenshots of System in Action
- Demonstration of your favorite ISP/content-provider pair after the talk

Upgraded Link from Access Provider to Content Provider?

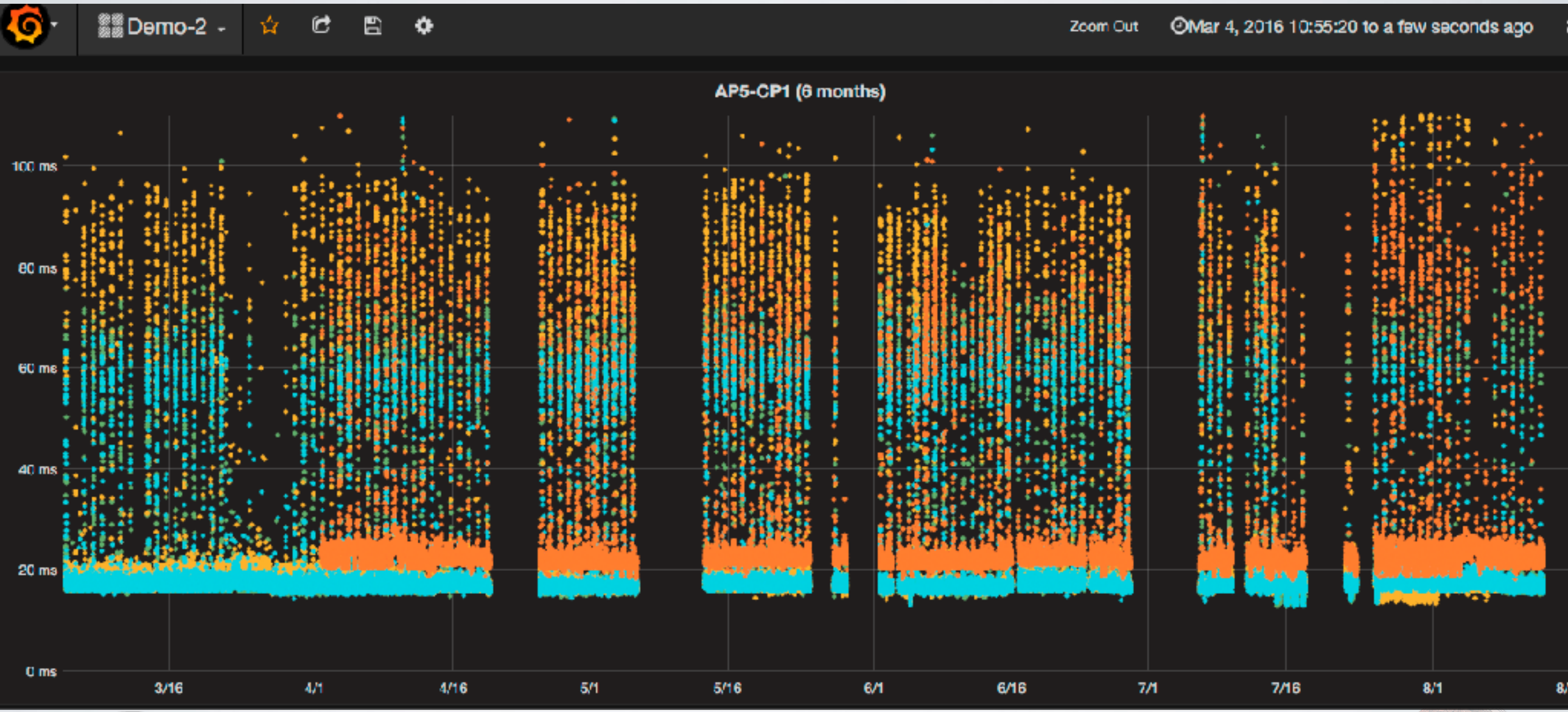


Two Access Providers in Italy to Large Content Provider

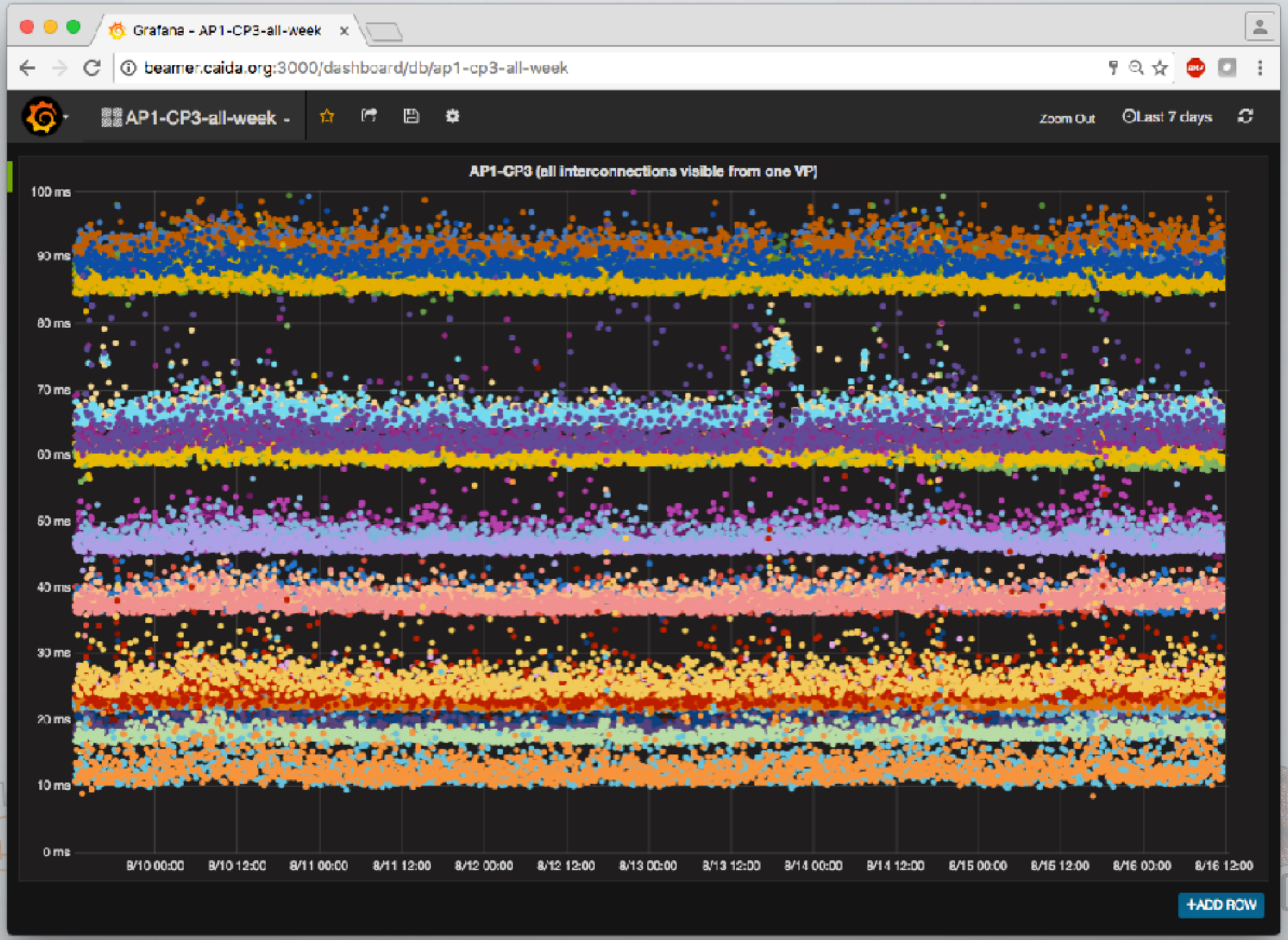
CSAIL

caida

Congested links from Access Provider to Content Provider



All links from an access provider to a content provider



Select Interconnections from Top 3 Cps to Top 6 APs

CONGESTION AND QOE

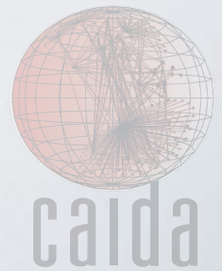
(quality of experience)

- User QoE could be impacted by various network-level performance metrics - delay, delay jitter, loss rate, available bandwidth, stability of available-bandwidth….

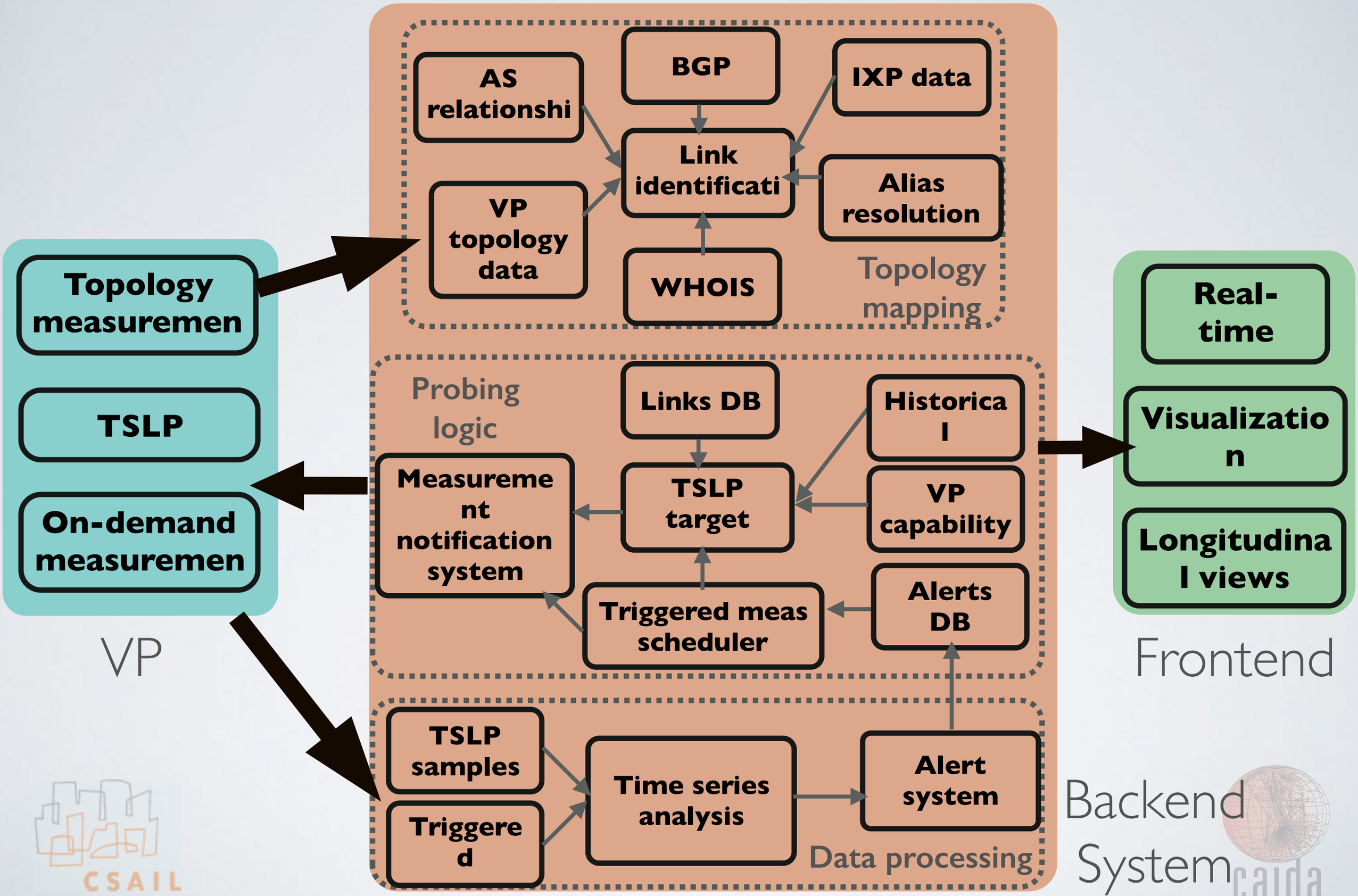
- A congested peering link increases the delay for flows crossing that link (full buffer adds latency)

- A congested peering link has a non-negligible loss rate

- A congested peering link has low (almost zero?) available bandwidth.



MEASUREMENT SYSTEM



VP DEPLOYMENTS

Archipelago

- Deployments in various access networks (and other network types, see http://www.caida.org/projects/ark/)

- Currently 60+ monitors running TSLP measurements: 39 networks, 26 countries

- We to deploy Ark nodes using Raspberry Pi hardware in homes

- **Goal: deploy our experiments on other platforms: BISmark, FCC-MBA (thousands of vantage points, but not enough CPU/memory to do mapping)**

CSAIL

caida

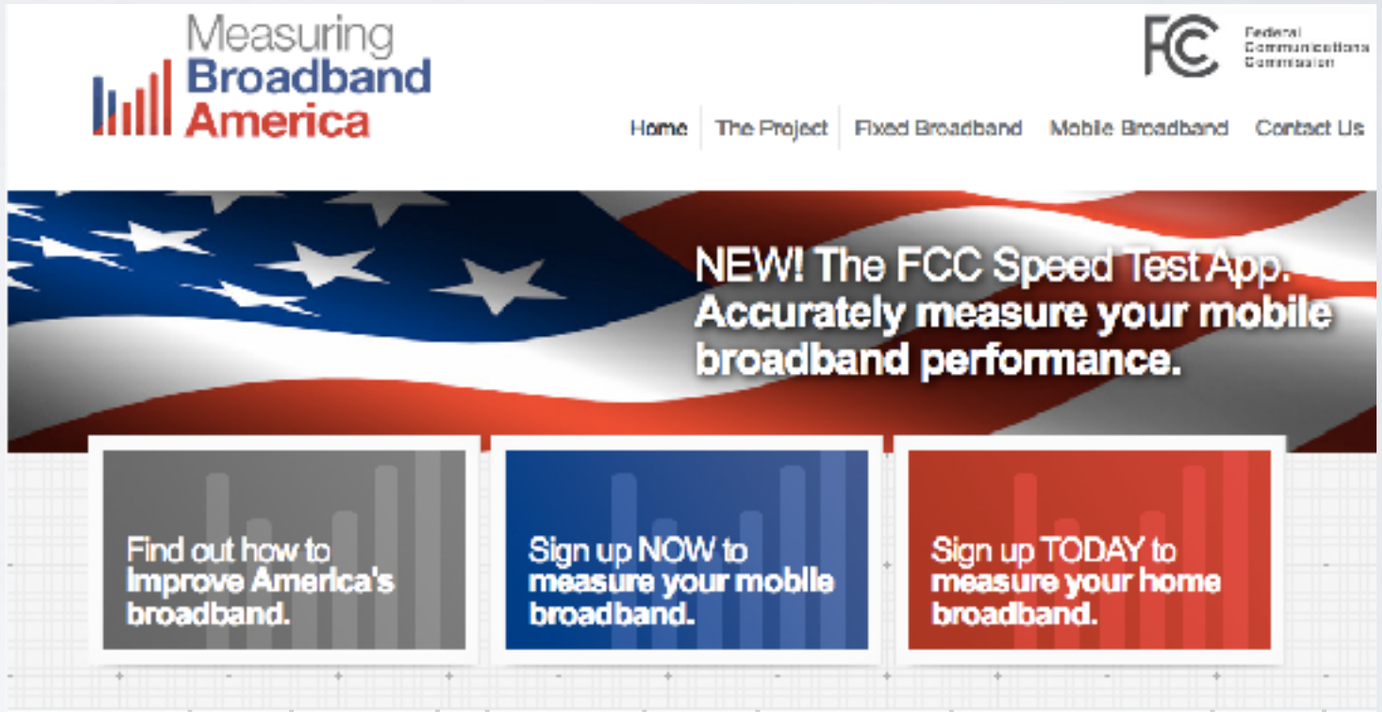
TSLP ON CONSTRAINED DEVICES

- TSLP and sc_bdrmap are CPU, memory, and disk intensive
  - TSLP data can grow large owing to sampling frequency
  - sc_bdrmap requires ~150MB of memory and some CPU to maintain mappings, probing state
- Some measurement devices (SamKnows, BISmark) are resource constrained
  - 64-128M memory
  - 400Mhz MIPS CPUs
  - No significant storage (small flash w/ limited write cycles)

FCC-SAMKNOWS

with any luck..

- ~10,000 home routers operated by SamKnows as part of the FCC's Measuring Broadband America program
- Next target for TSLP



- Our deployment on Bismark should make it easier to deploy on SamKnows (right??)
- Issues: Several different router hardware and firmware versions, much larger scale

PARTING THOUGHTS

- Longitudinal measurement of infrastructure yields insights into technical as well as economic and market dynamics that hinder stability and performance of the ecosystem

- Lots of exciting problems in computer science and engineering, and ability to re-use open source software components to build system to solve them

- High impact project: triggered FCC selection of team as Independent Measurement Expert for AT&T/DirecTV merger (see NANOG Feb 2016 keynote and FCC filings linked from caida.org)

THANKS!

kc@caida.org