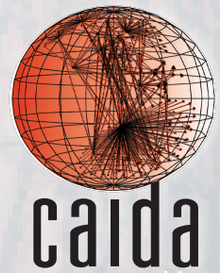
MAPPING INTERNET INTERDOMAIN CONGESTION

Amogh Dhamdhere, Matthew Luckie
Bradley Huffaker, Young Hyun, Kc Claffy (CAIDA)
Steve Bauer, David Clark (MIT)

amogh@caida.org



IN THE PRESS

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

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

BACKGROUND

- Modern peering (interconnection) disputes manifest as congested links
- Disputes among access, content, and transit providers
- 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

INTERDOMAIN CONGESTION

- We are developing methods to measure the 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**

- Part of a 3 year NSF-funded project on topology +congestion measurement

NSF PROJECT (CAIDA+MIT)

- Measurement-based characterization of the changing nature of the inter-domain Internet

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

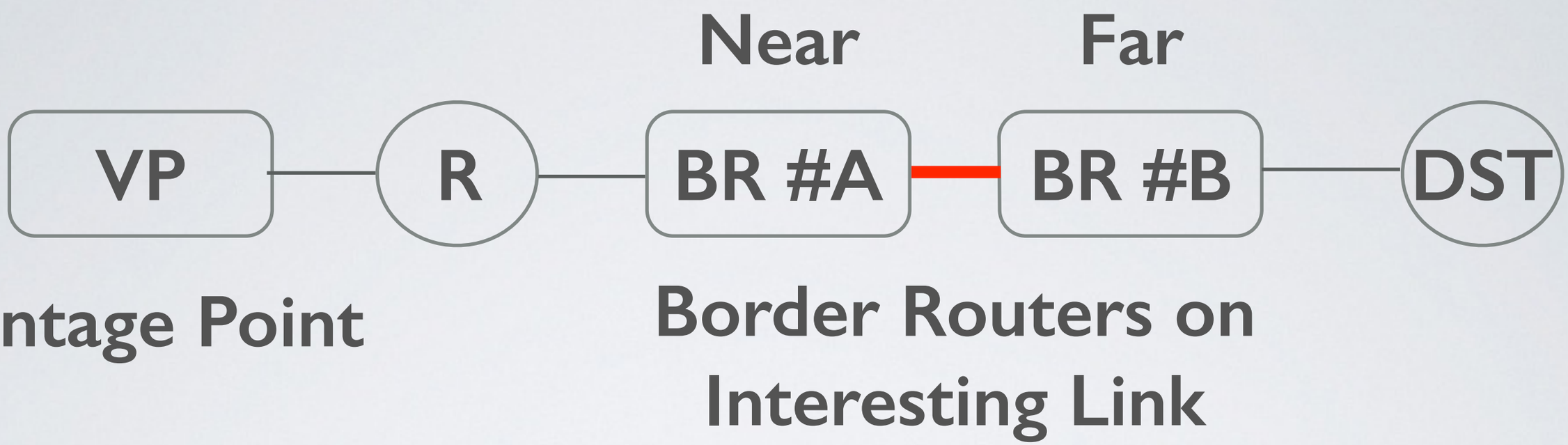
  - Growth of Content Delivery Networks (CDNs) and their ability to change traffic dynamics on short timescales

  - Ability of large access ISPs to use market power

  - Congestion and choke points

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

METHOD: TIME SERIES PING

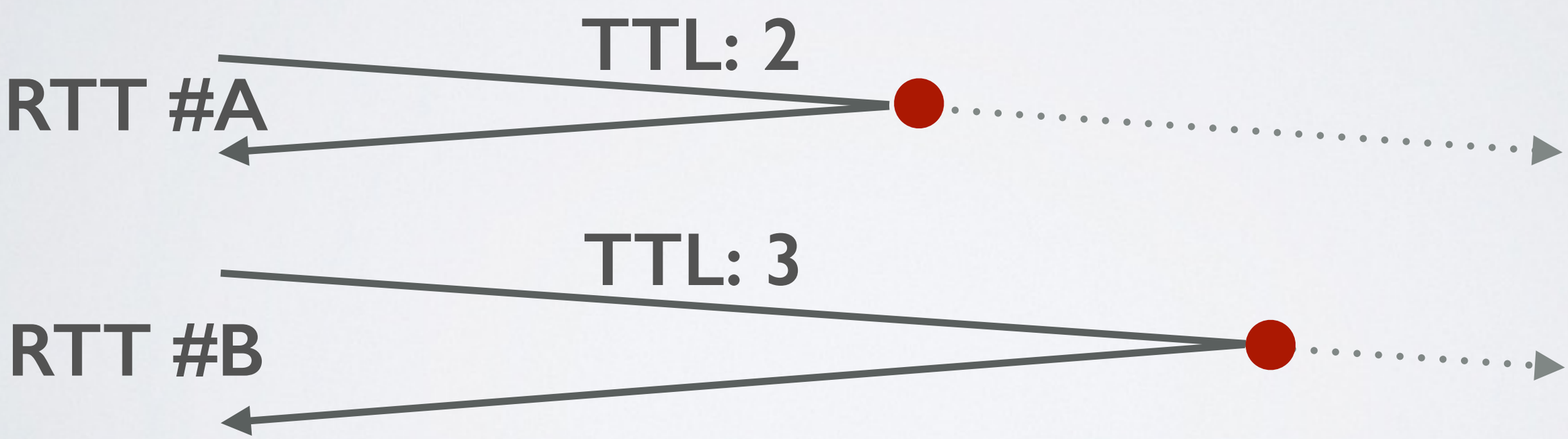


METHOD: TIME SERIES PING

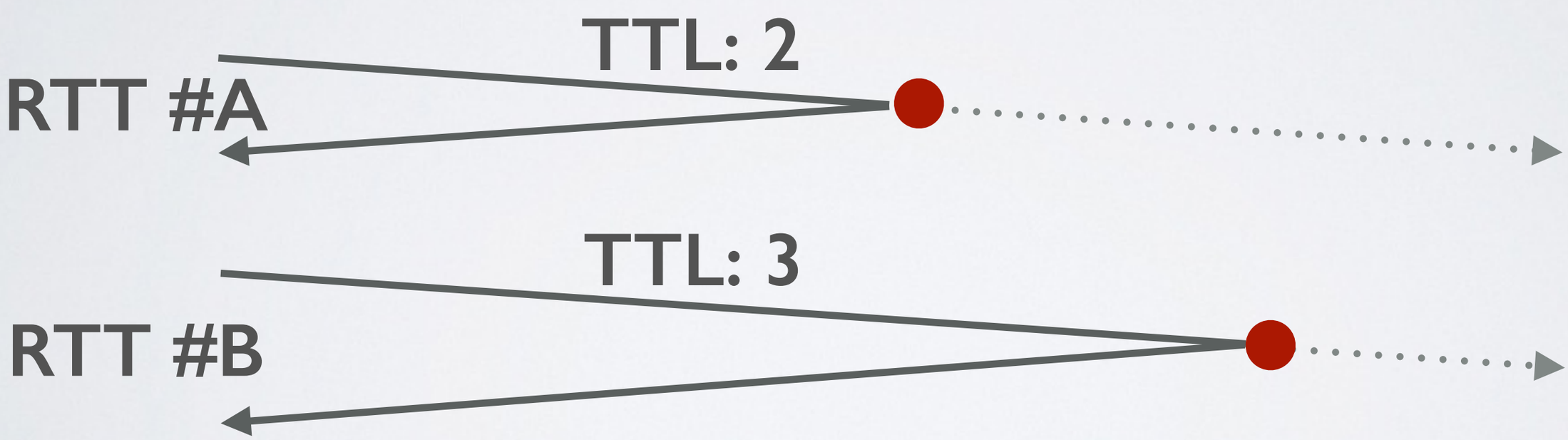
METHOD: TIME SERIES PING

Vantage Point

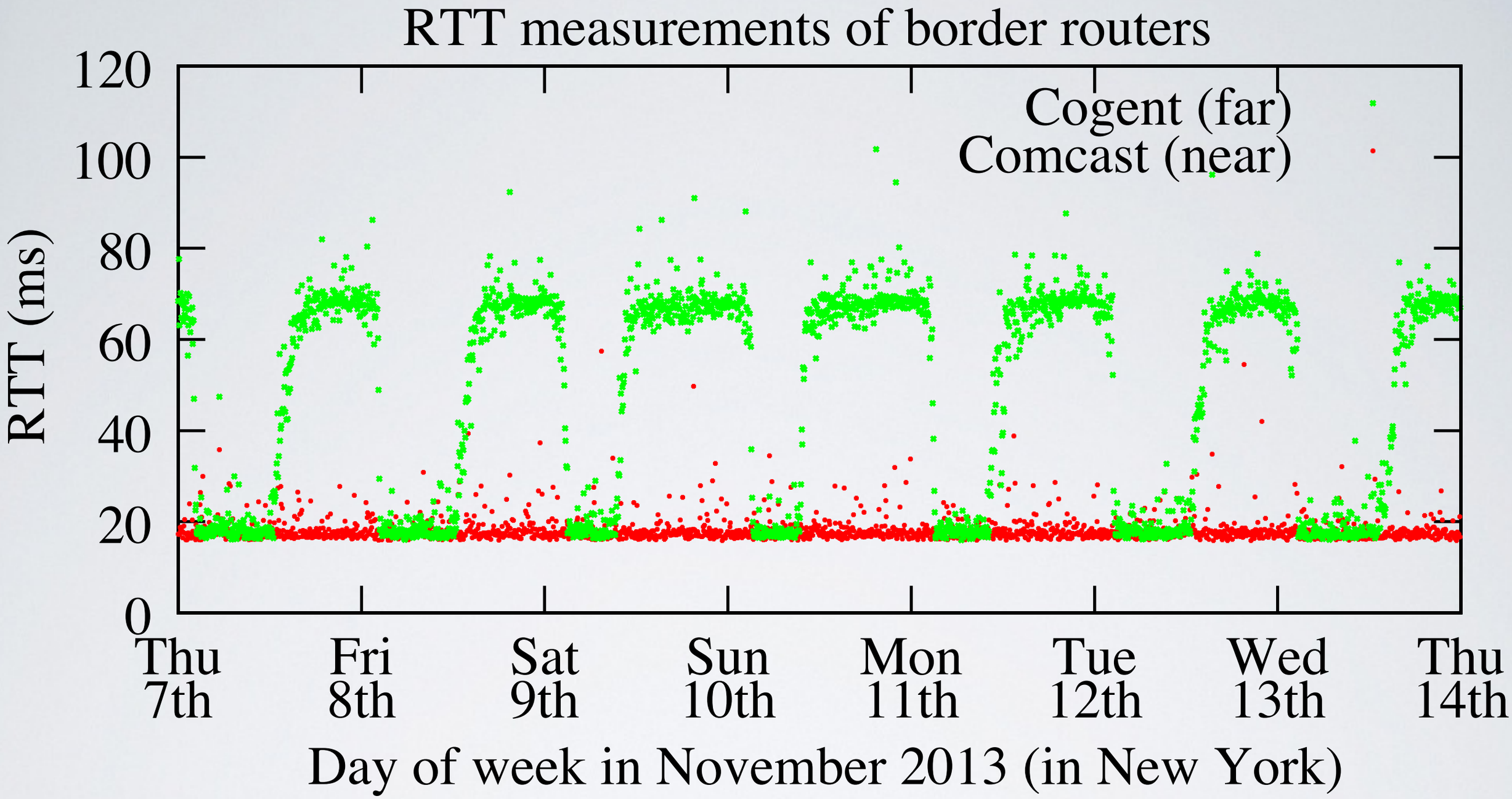
Border Routers on Interesting Link



METHOD: TIME SERIES PING



(repeat to obtain a time series)

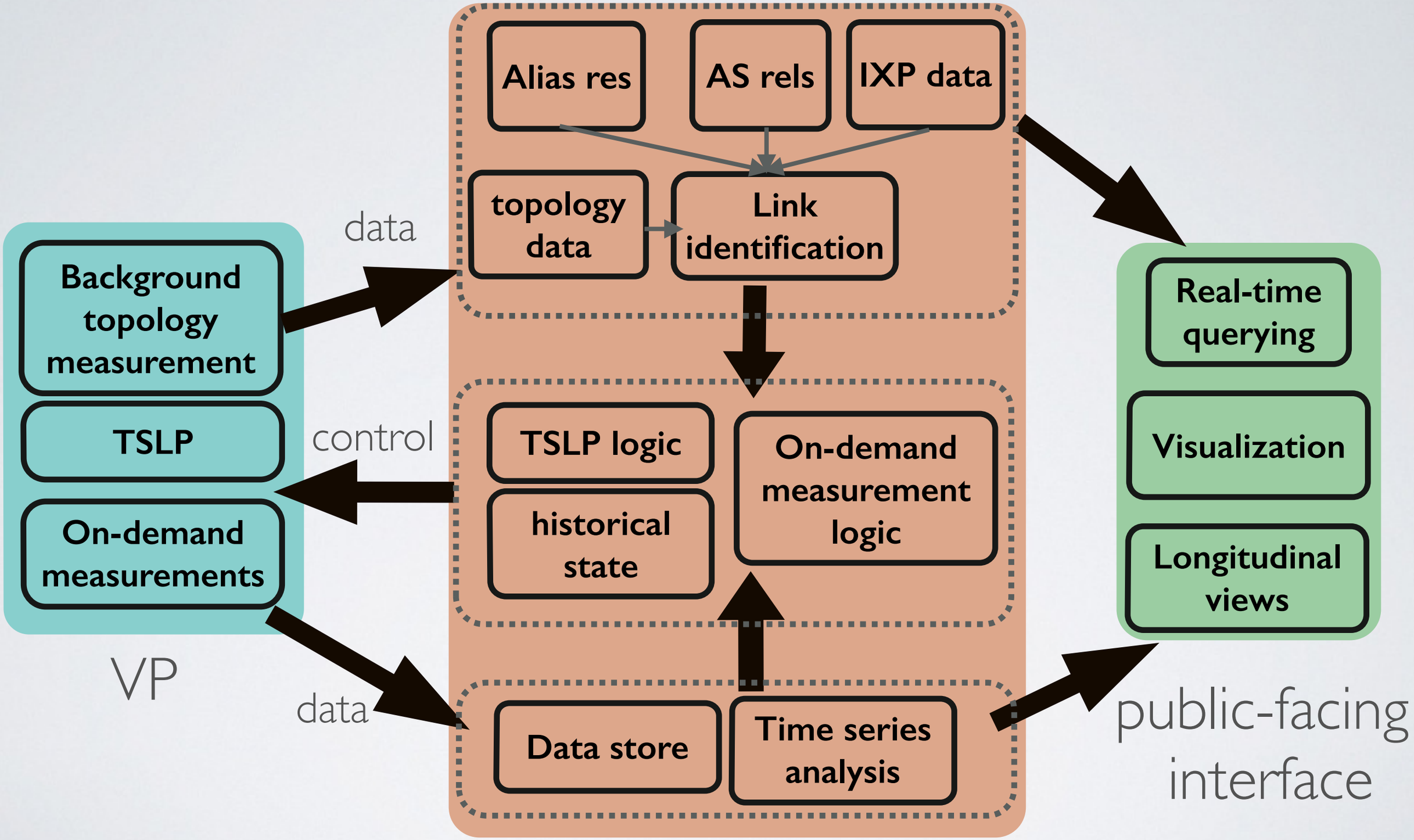


More congestion on weekend than weekdays. Monday 11th was Veterans Day

CHALLENGES

- Topology: Method requires us to know the location of border link between two networks; this is very difficult
- Adaptiveness: The network changes over time; our system needs to be adaptive to changes in the underlying topology and routing
- Data mining: Need scalable techniques to find patterns in the data that indicate congestion
- Validation: Difficult to get ground-truth; most peering agreements are covered by NDAs

MEASUREMENT SYSTEM



BACKEND SYSTEM

- Goal: adaptiveness — make sure we are probing an up-to-date state of the network
- Goal: detailed historical state — store enough information to be able to easily reconstruct topology, links/targets probed, destinations used per target, etc. at any given point in the past
- Goal: efficient data management and processing
- Goal: continuous monitoring — make sure various pieces of the system are up and running: VPs and TSLP probing, data collection/storage, data processing

ADAPTIVE PROBING

- Periodically (every 3 days) fetch background traces from VPs, perform link identification, produce probe list, push new probe list to monitor, produce various meta-data (e.g., reverse DNS lookups of all interdomain link IPs)

- All monitor state data stored in SQLite databases with an SQLalchemy framework for managing database updates

- Order of minutes to process traces, generate probing set, and produce all meta-data for each VP

- Updates for VPs can run in parallel — plan to use SDSC nodes for parallelism

DATA MANAGEMENT

- TSLP data is pulled from monitors and indexed into databases nightly. TSLP (near, far) time series produced weekly for all links probed from each VP during the previous week

- Level-shift detection run on all links from all VPs (thousands of links total). Time series data shipped over to MIT for FFT/ wavelet analysis (more on that from Steve)

- Eventually, some sort of time-series database to store time series data and feed the interactive front-end

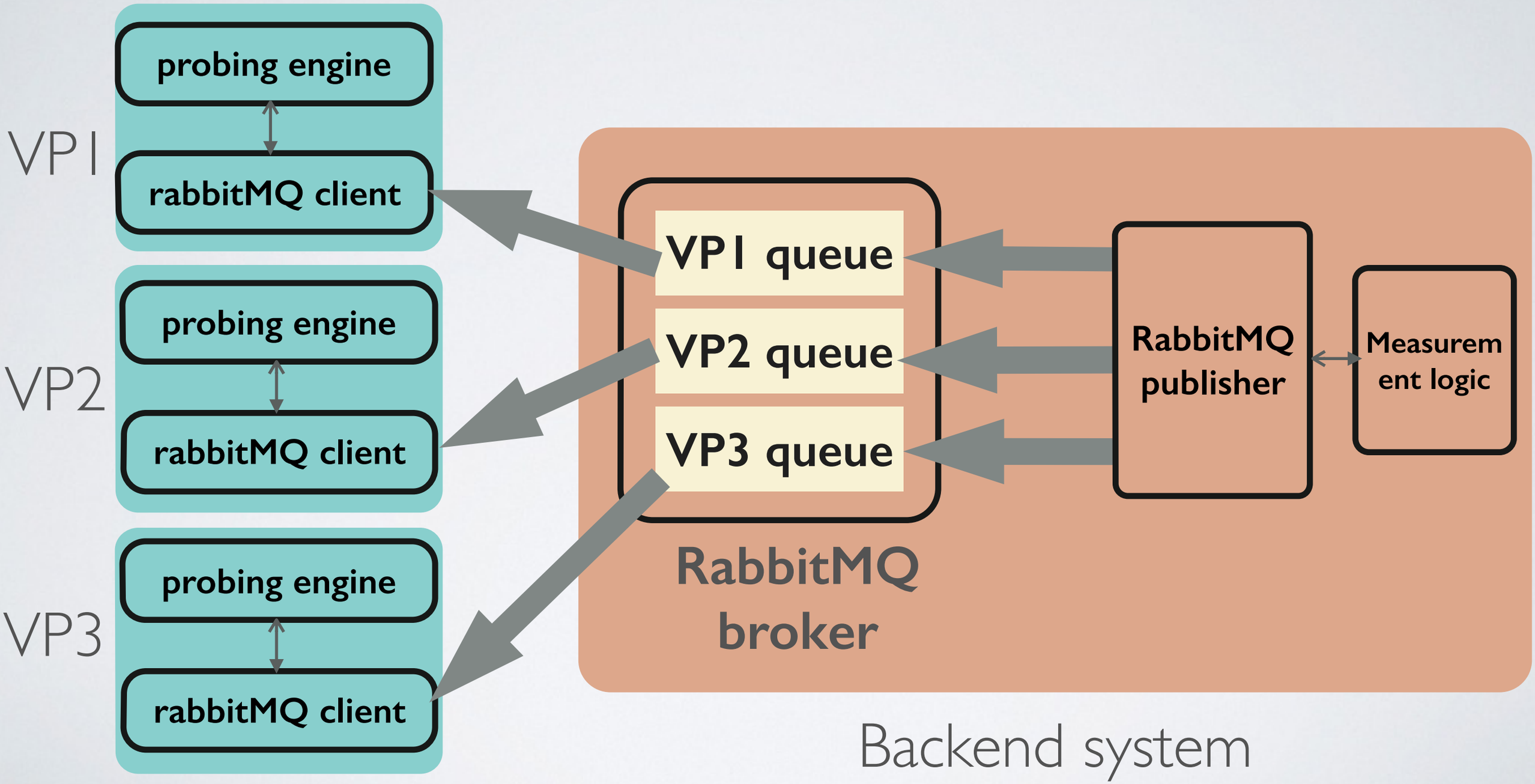
- Current system prototype is small-scale (~100GB of data per month). We anticipate ~1TB/month with large-scale deployment

ON-DEMAND MEASUREMENTS

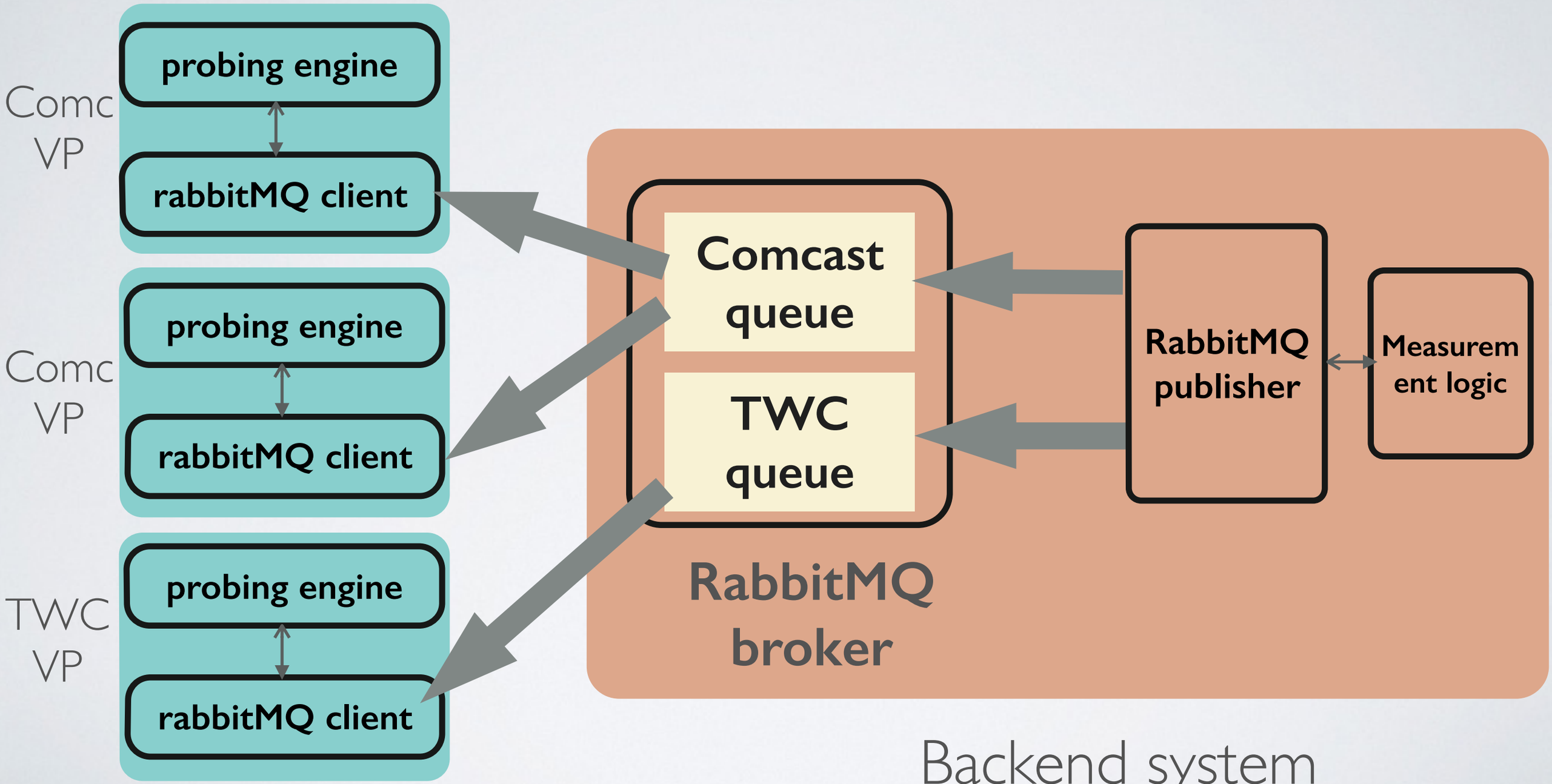
- We would like to use the results of time series analysis to trigger additional measurements from the VP — high-frequency probing to measure loss rate, or throughput measurements

- We require an agile, lightweight mechanism to distribute measurement tasks to the VPs. The scheme should be interoperable across different infrastructures

MESSAGE QUEUES



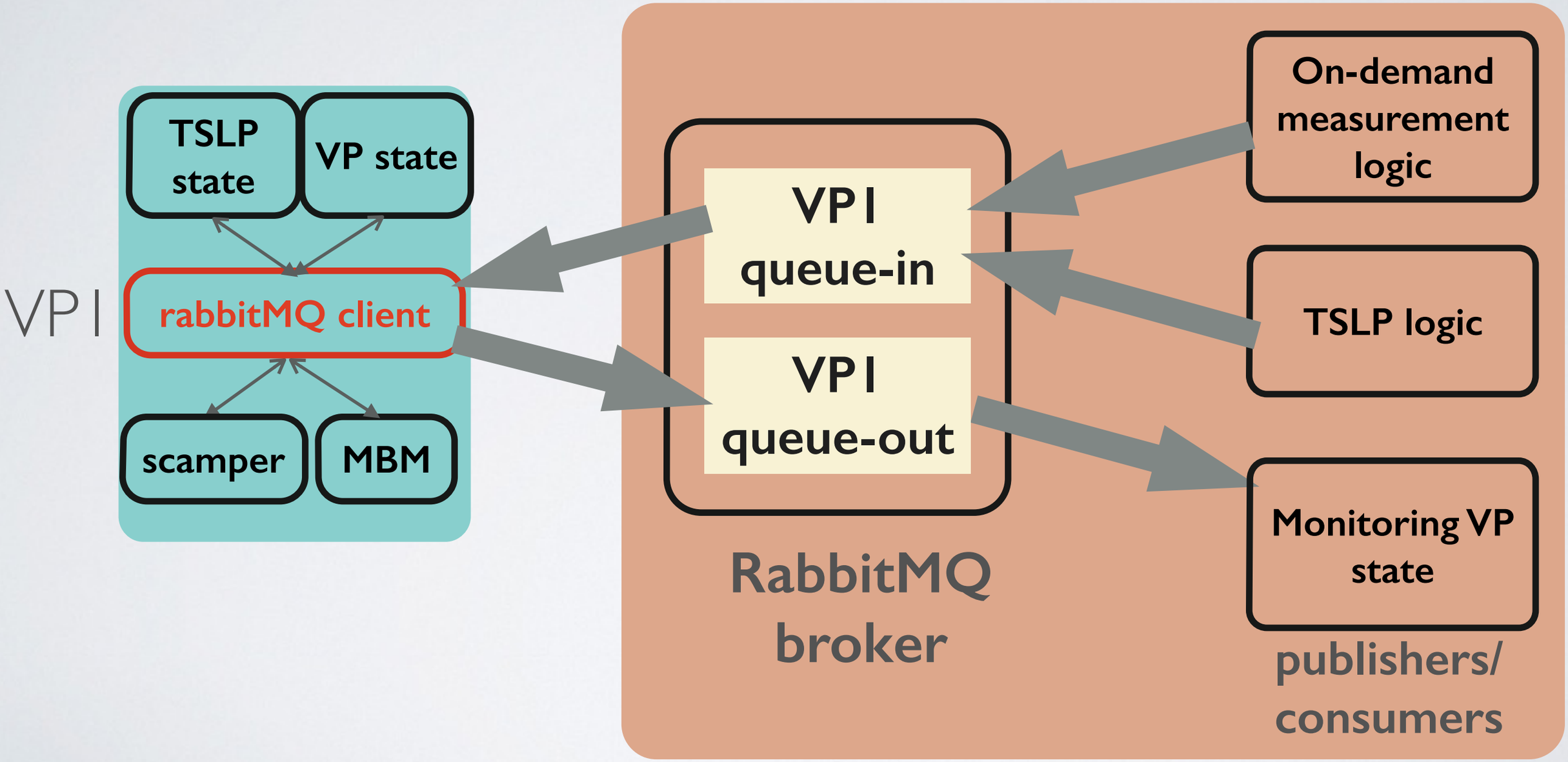
MESSAGE QUEUES



WHY MESSAGE QUEUES

- Scalable and lightweight
- Not tied to any specific architecture/infrastructure
- Many desirable features: reliability, high availability, security, flexible routing (point-to-point, broadcast), flexible message delivery semantics (exactly once/at most once)
- Wide range of usage scenarios: pushing on-demand measurement tasks to VPs, transferring probing lists and updates to VPs, notification of VP reboots, etc.

MESSAGE QUEUES: USE CASES



VP DEPLOYMENTS

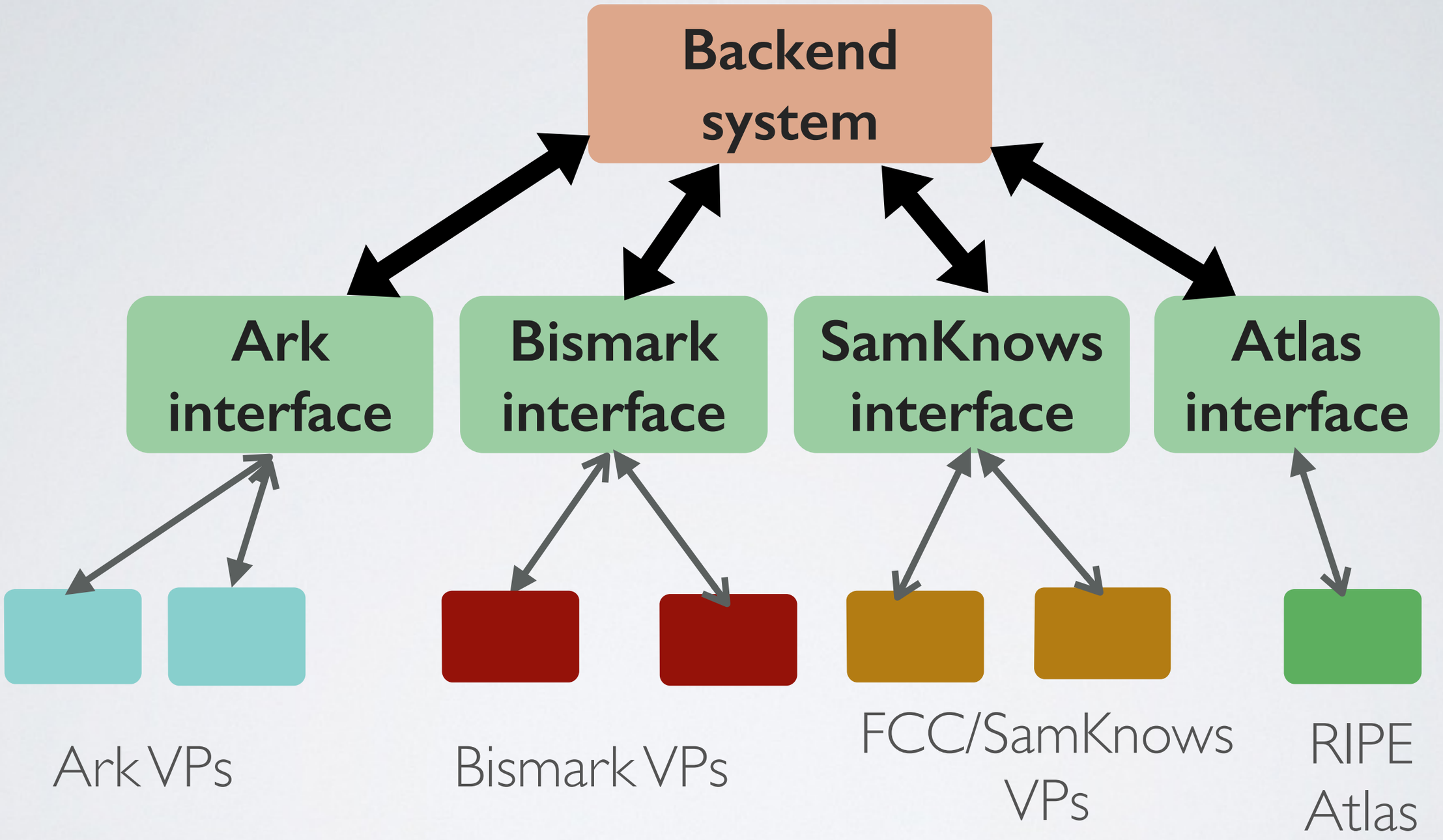
- Deployments in various access networks (and other network types, see http://www.caida.org/projects/ark/)

- Currently 19 monitors running TSP measurements

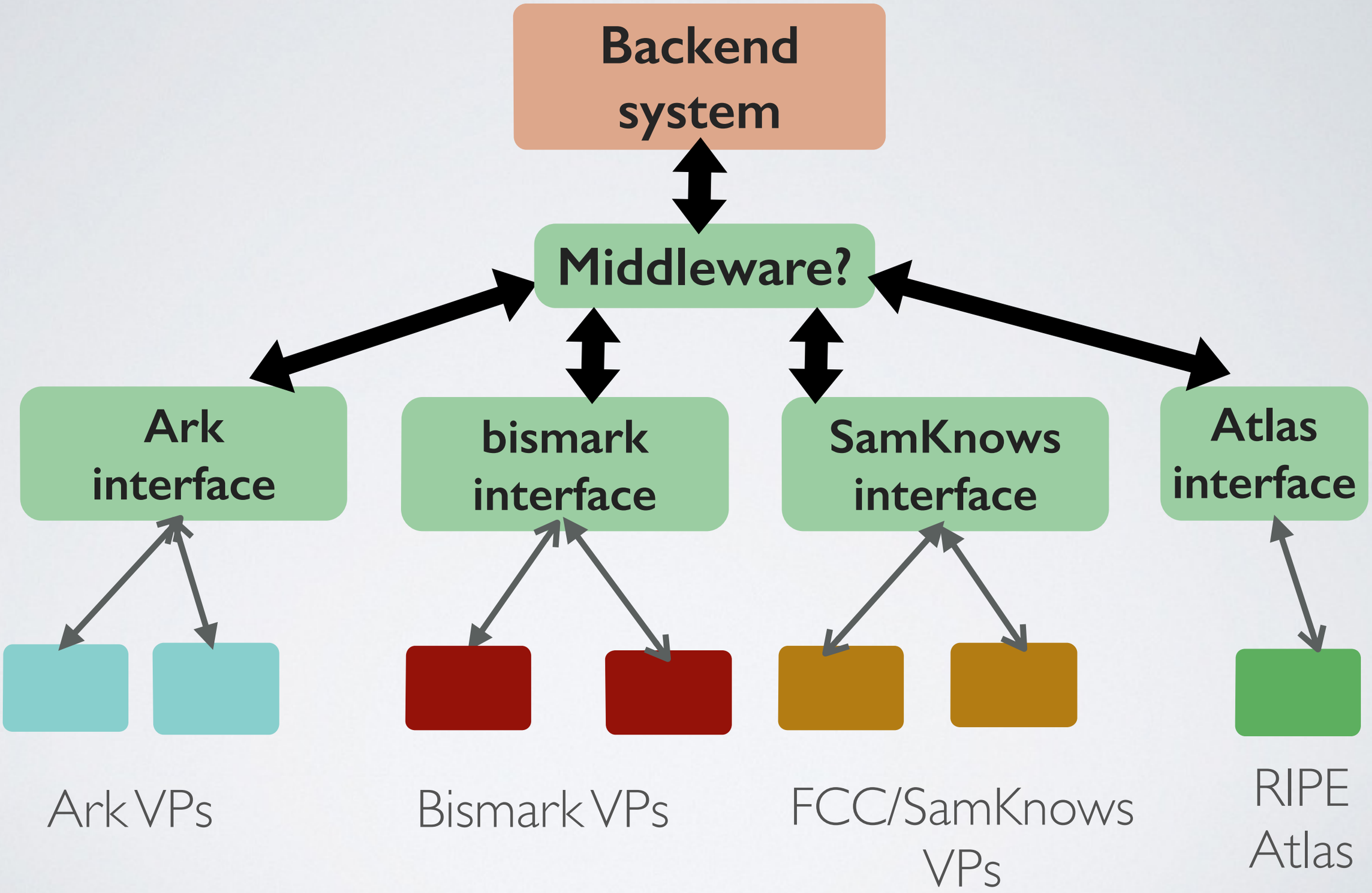
- We continue to deploy Ark nodes using Raspberry Pi hardware in homes of our friends (or friends of friends)

- **Goal: deploy our experiments on other platforms: Bismark, FCC-Samknows (hundreds of vantage points)**

TSLP ON OTHER INFRASTRUCTURES



TSLP ON OTHER INFRASTRUCTURES



SUMMARY

- We are working on various pieces of the congestion measurement system
- There's a lot of work to be done
- Would like to discuss:
  - Running our measurements on other infrastructures
  - Better/more scalable/more efficient backend and data management systems

THANKS!

amogh@caida.org