MAPPING INTERNET INTERDOMAIN CONGESTION

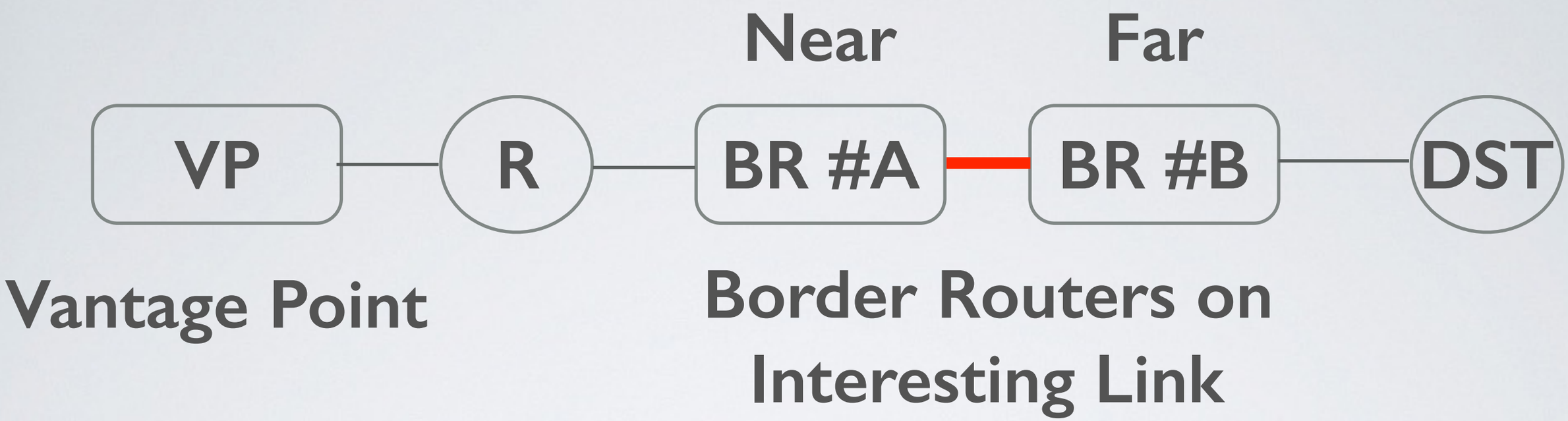
Amogh Dhamdhere, Bradley Huffaker, Young Hyun, Kc Claffy (CAIDA)
Matthew Luckie (Univ. of Waikato)
Alex Gamero-Garrido, Alex Snoeren (UCSD)
Steve Bauer, David Clark (MIT)

amogh@caida.org

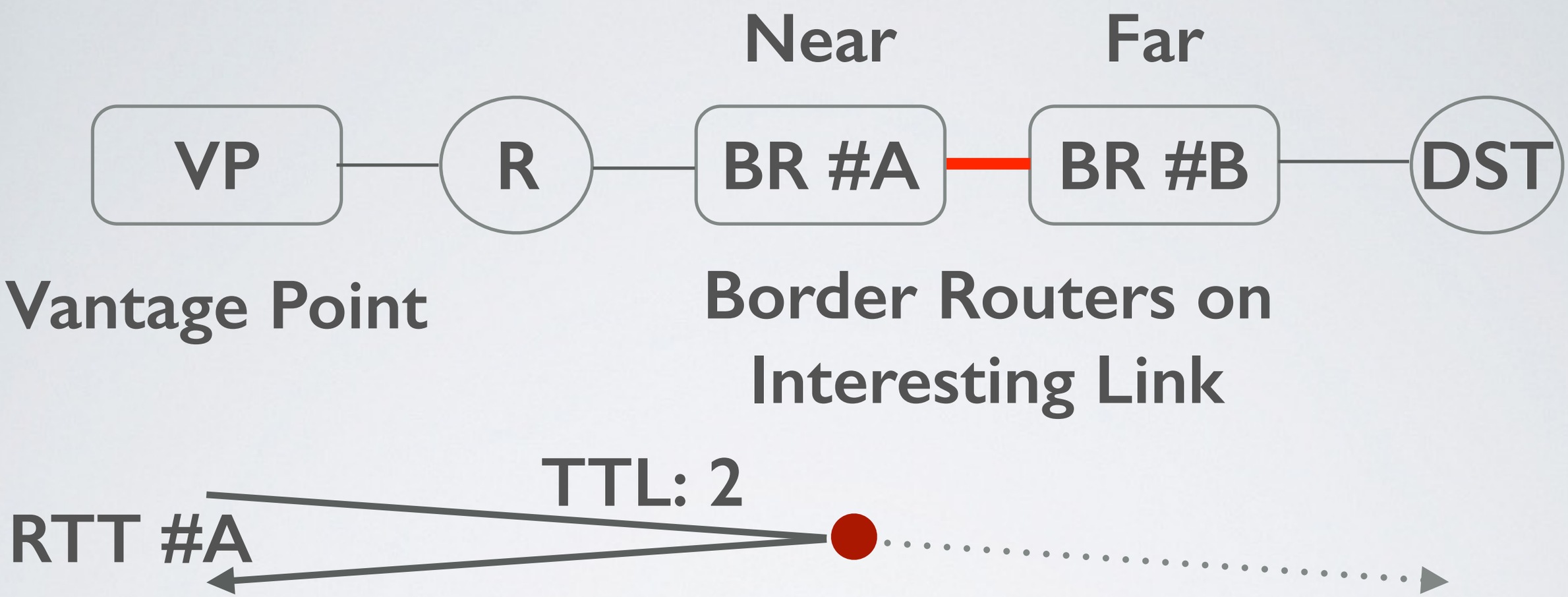
INTERDOMAIN CONGESTION

- 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

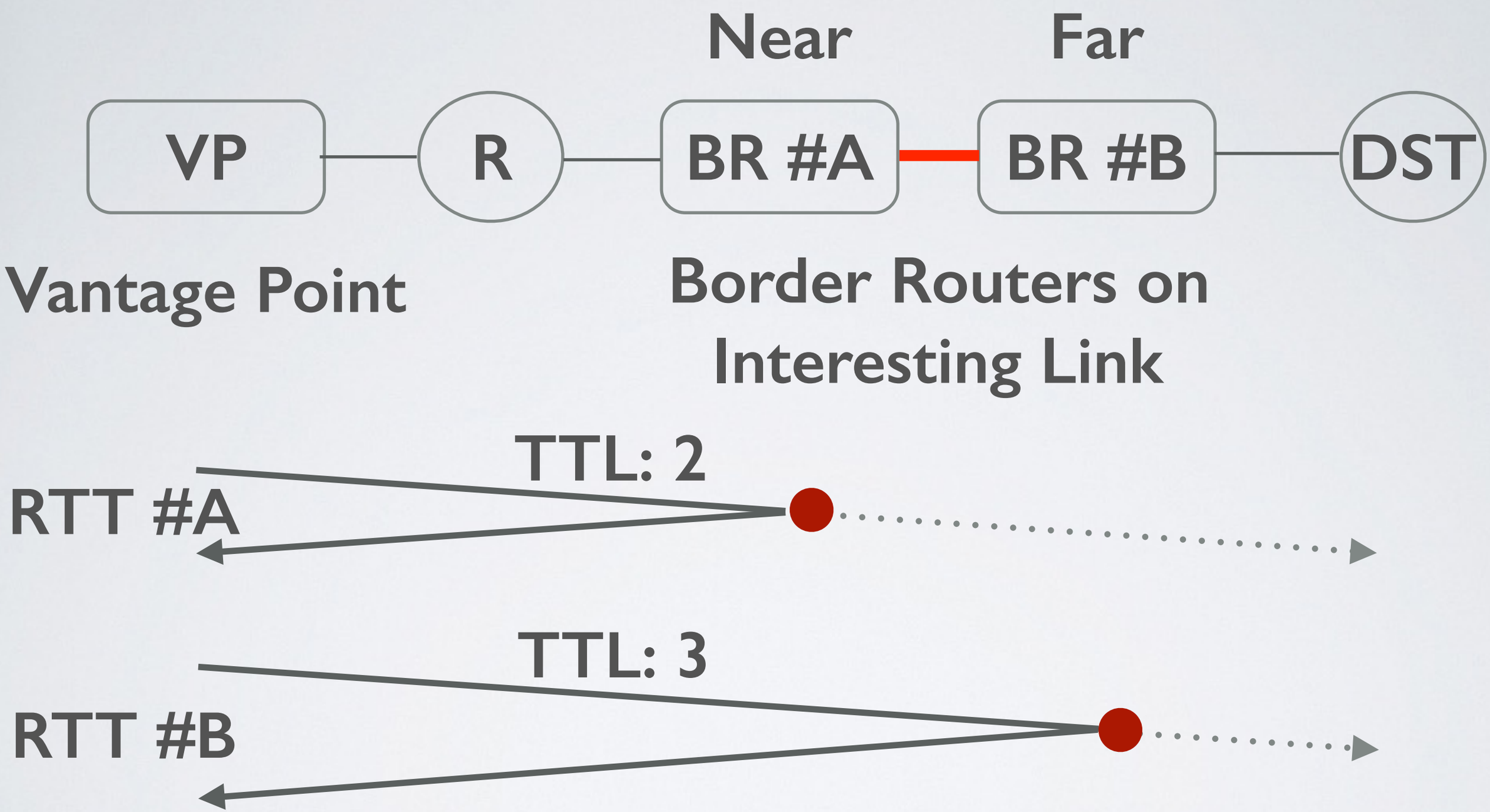
METHOD: TIME SERIES PING



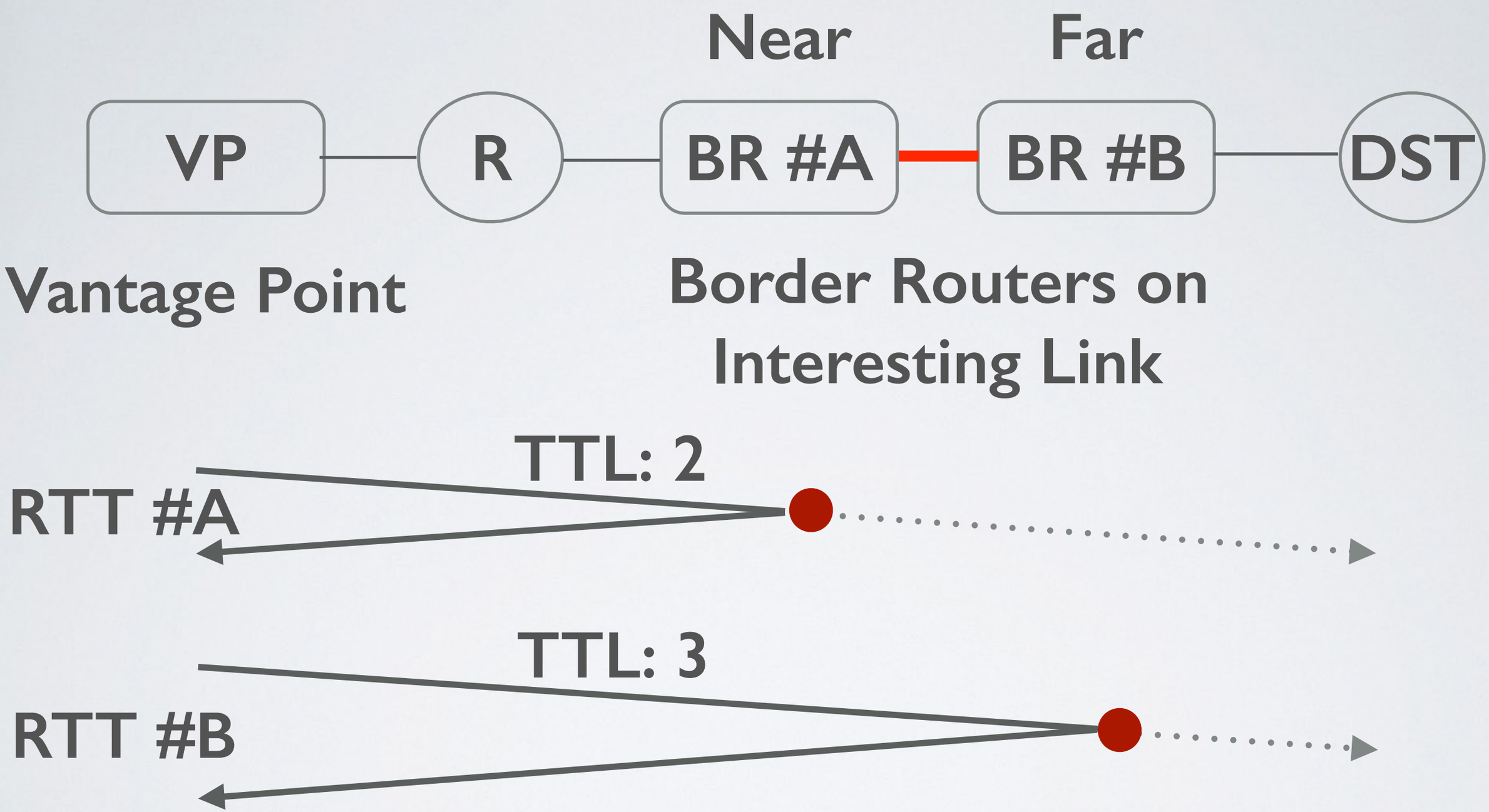
METHOD: TIME SERIES PING



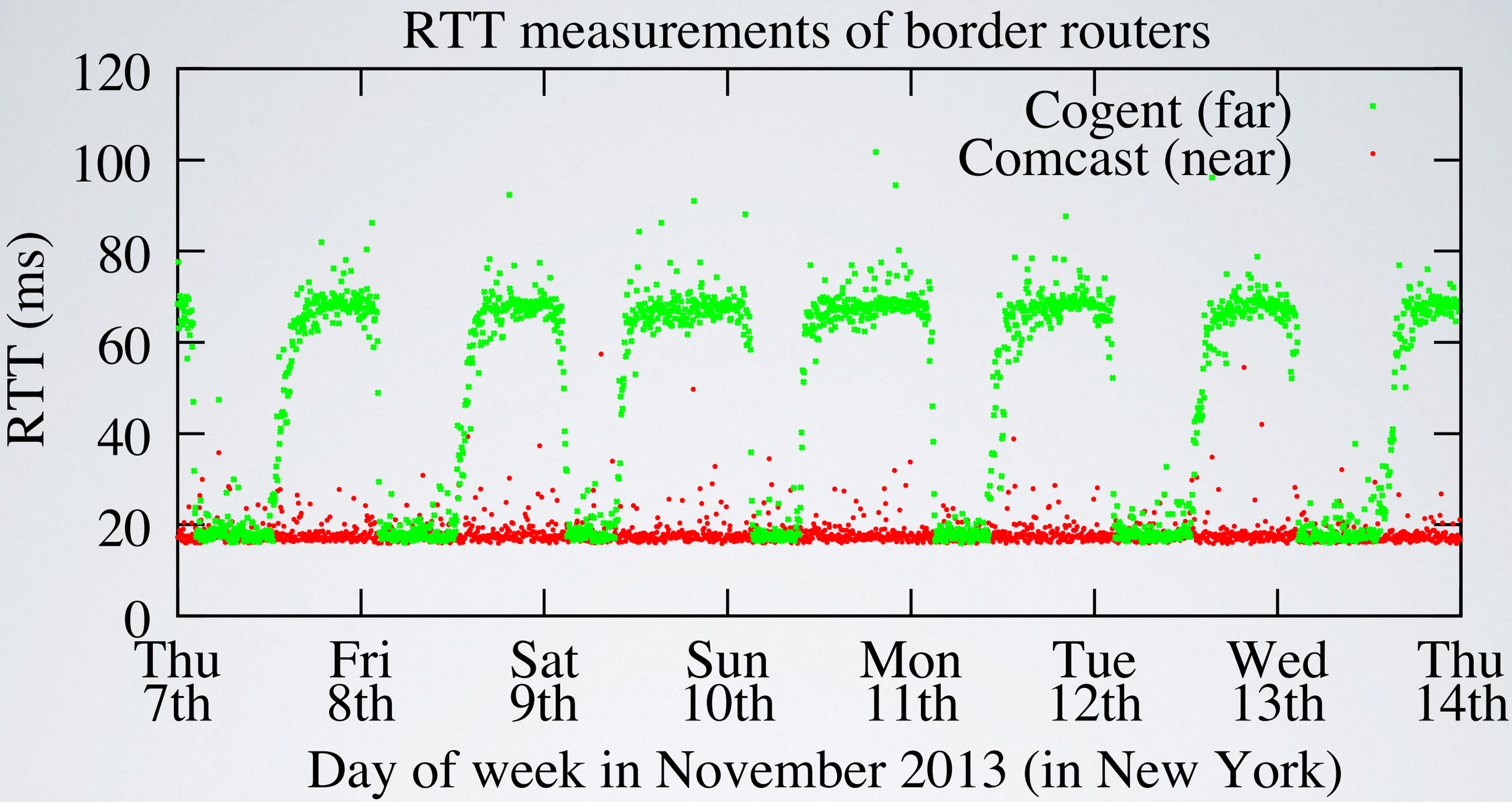
METHOD: TIME SERIES PING



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

MAPPING INTERDOMAIN LINKS

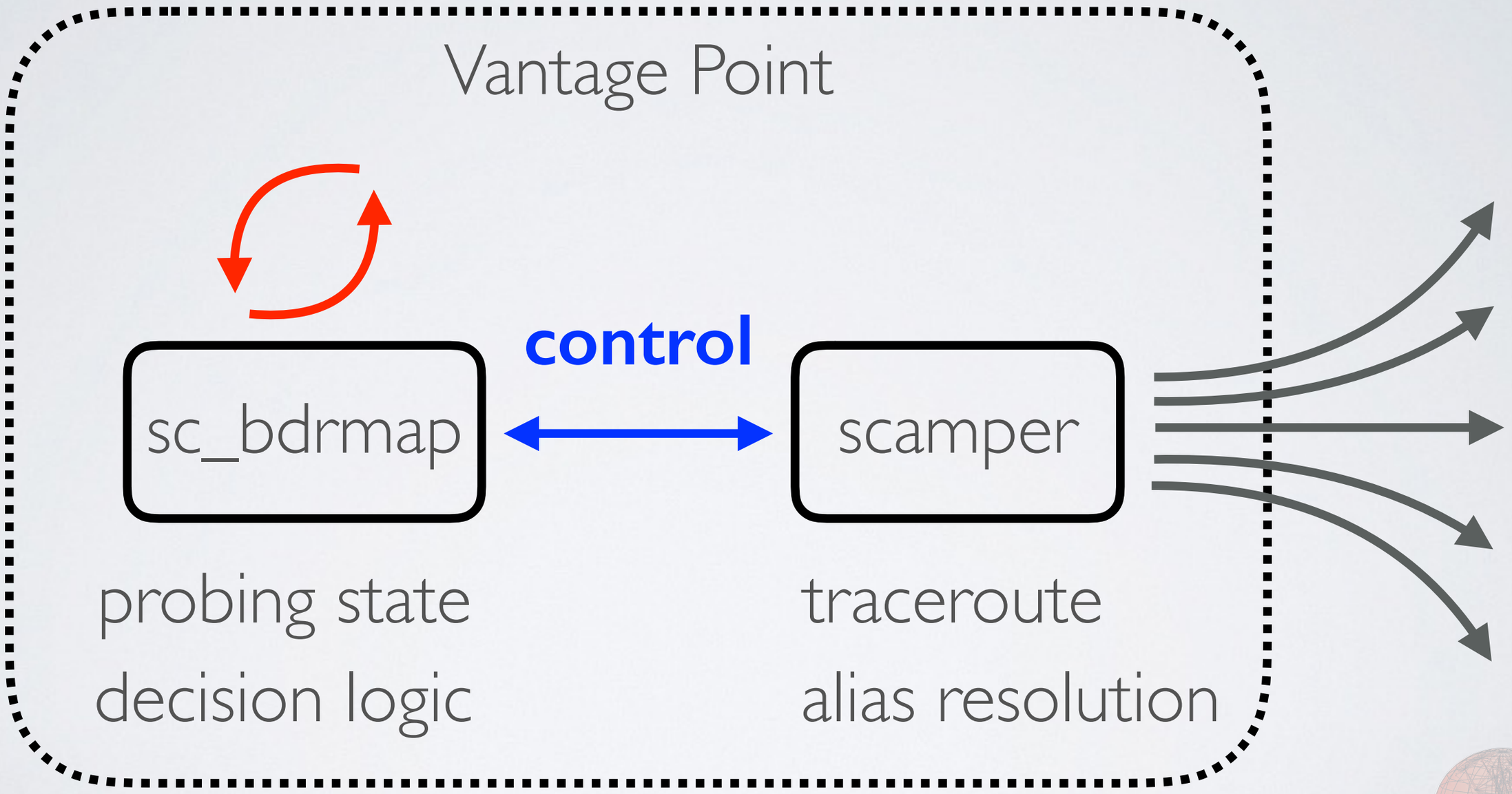
- Identifying interdomain router links and their owners is an unsolved problem in the Internet research community

- We make a significant step forward with our tool to accurately map interdomain links (sc_bdrmap) for a given Vantage Point

  - Utilizes scamper's traceroute and alias resolution to produce graph of routers, annotated with pt2pt links and AS owners

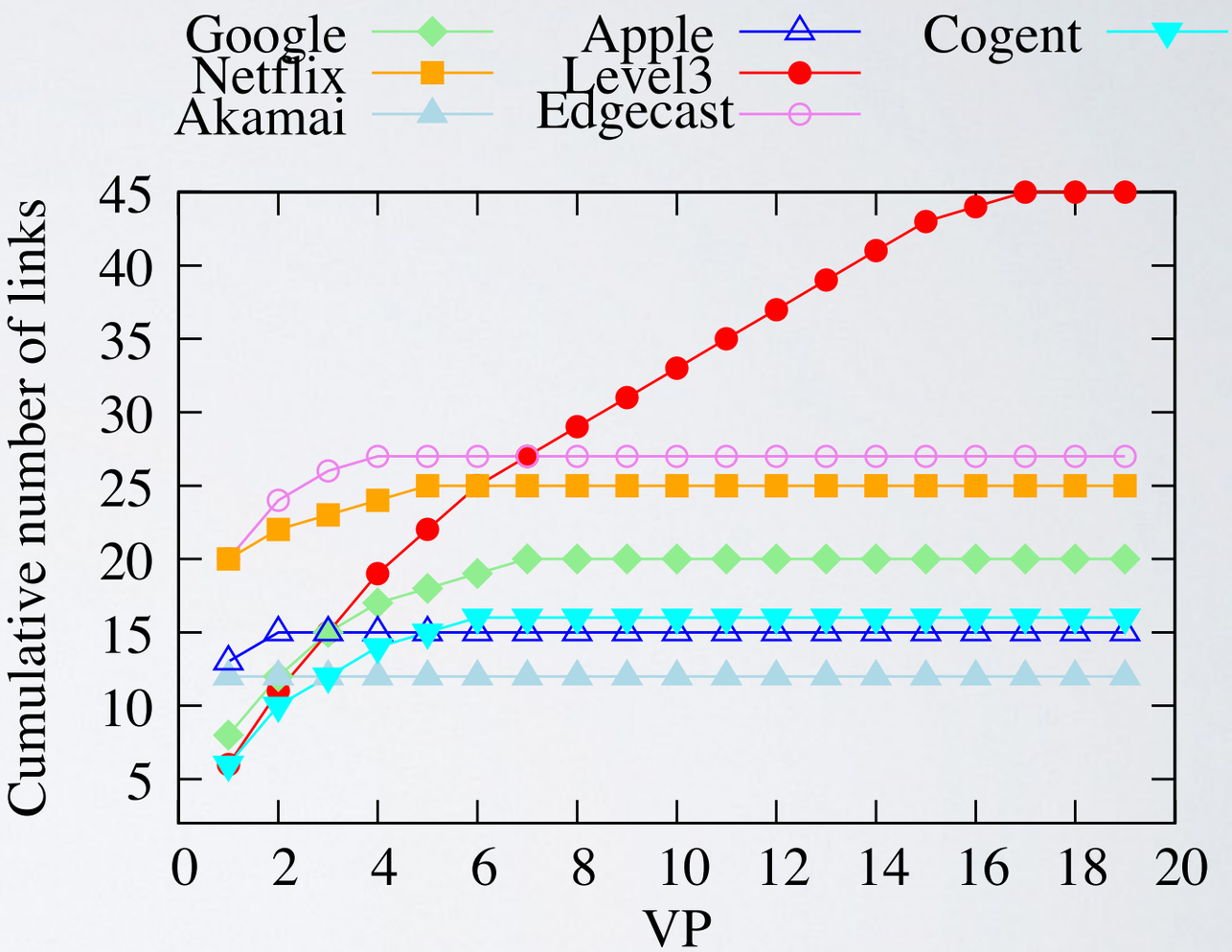
  - Resilient to traceroute artifacts (e.g., third-party IP addresses)

SC_BDRMAP OPERATING MODEL



BORDER MAPPING

- Running sc_bdrmap from ~60 monitors in diverse networks
- Obtained validation from a few networks; method shows good accuracy (we need more validation)
- Some findings: Comcast-Level3 connected with 45 router-level links!
  - Need VPs close to those locations to find these links



ADAPTIVE PROBING

- TSLP probing depends on up-to-date view of network
- Run bdrmap continuously on VPs
- Periodically (every day) fetch bdrmap data from VP, 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)
- Order of minutes to process traces, generate probing set, and produce all meta-data for each VP

DATA!

- We probe the near and far interface of every interdomain link every 5 minutes
- TSLP data is pulled from monitors and goes into a time series database
- Not re-inventing the wheel: use InfluxDB as the time-series database. InfluxDB provides an SQL-like query interface
- InfluxDB + Grafana: Interactive querying plus visualization
- Bonus: updates in near real-time!

Grafana - Access - Content

beamer.caida.org:3000/dashboard/db/access-content-interconnections

Access - Content interconn...

Zoom Out | Last 7 days | Refresh every 15m

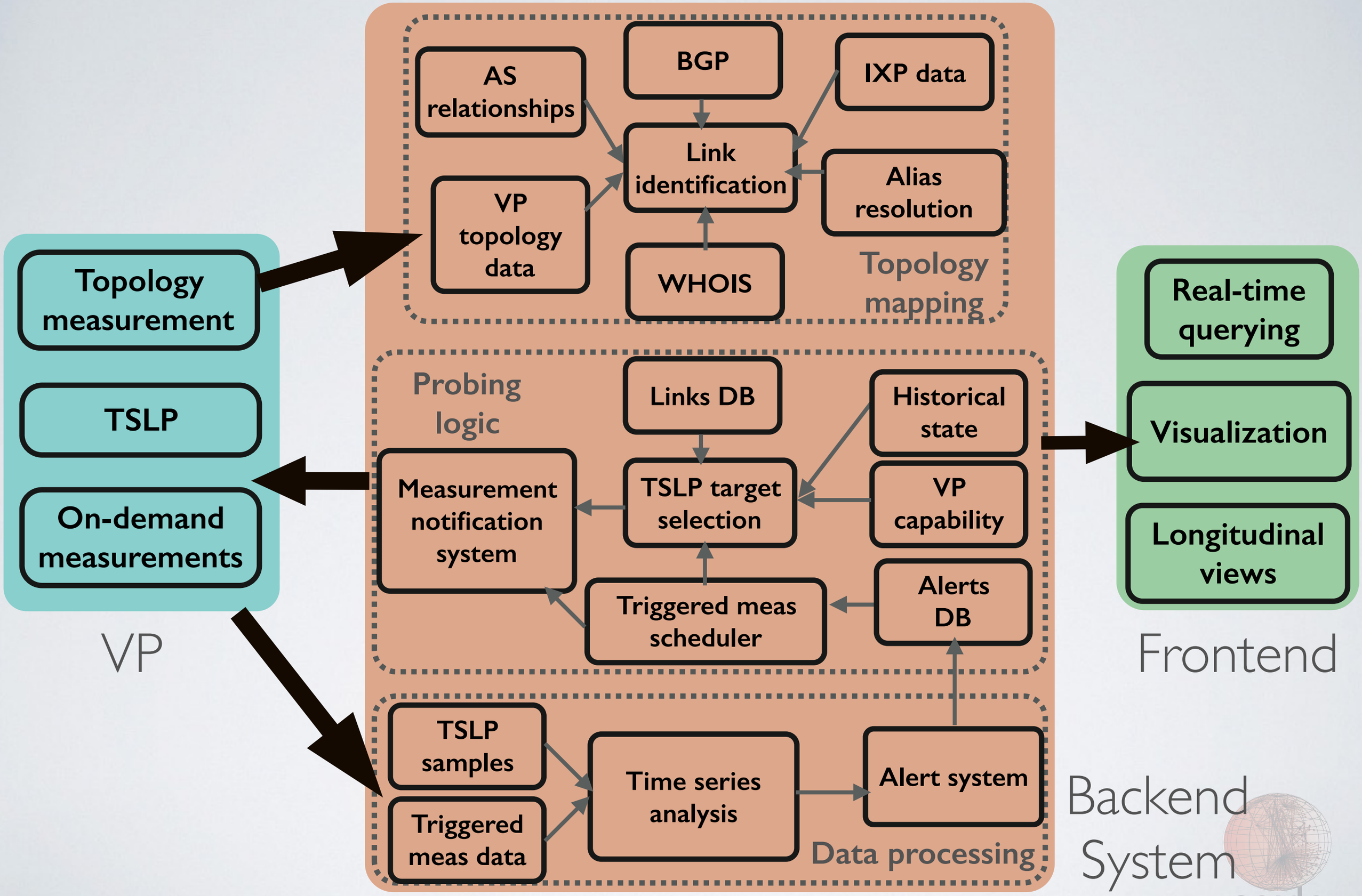
caida

DATA EXPLORATION

- Demo
- (let's hope this works)

REACTIVE MEASUREMENTS

- Use the results of time series analysis to trigger additional measurements from the VP — high-frequency probing to measure loss rate, throughput measurements, or application-specific QoE measurements
- We have built an agile, lightweight reactive measurement system to distribute measurement tasks to the VPs. The scheme is interoperable across different infrastructures
- In progress: time series analysis to feed into reactive measurement system


AS relationships
BGP
IXP data
Link identification
Alias resolution
VP topology data
WHOIS
Topology mapping
Topology measurement
TSLP
On-demand measurements
VP
Probing logic
Links DB
Historical state
Measurement notification system
TSLP target selection
VP capability
Triggered meas scheduler
Alerts DB
Real-time querying
Visualization
Longitudinal views
Frontend
TSLP samples
Triggered meas data
Time series analysis
Alert system
Data processing
Backend System
caida

VP DEPLOYMENTS

- 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 continue to deploy Ark nodes using Raspberry Pi hardware in homes
- **Goal: deploy our experiments on other platforms: BISmark, FCC-Samknows (thousands of vantage points)**

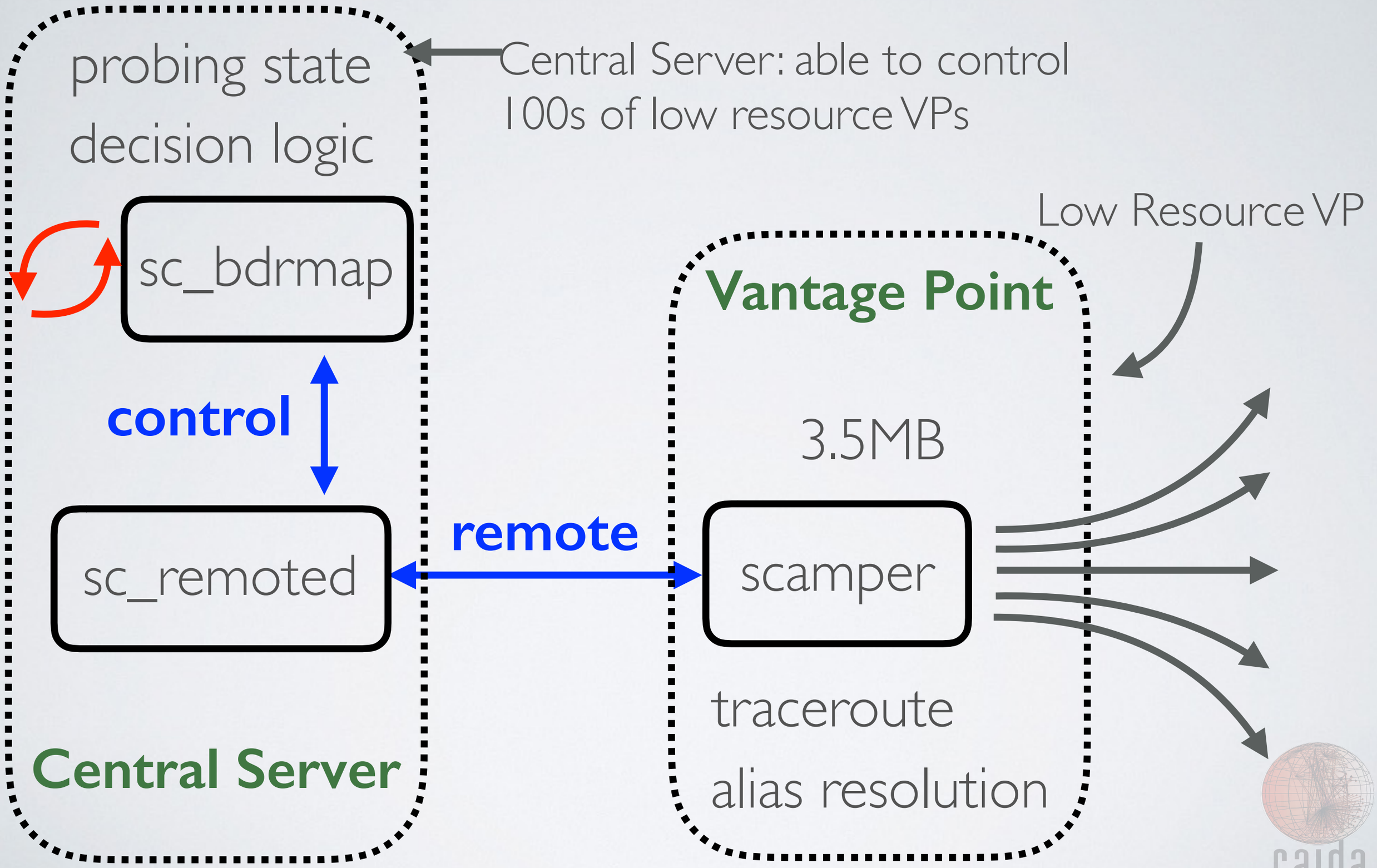
TSLP ON CONSTRAINED DEVICES

- 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)

TSLP ON CONSTRAINED DEVICES

- We extended scamper to support interactive remote control
  - resource constrained device still runs scamper (3.5M RSS)
  - scamper connects to more powerful remote system to receive instruction
  - border mapping and TSLP can run on the remote system and use sc_remoted to communicate with scamper instance

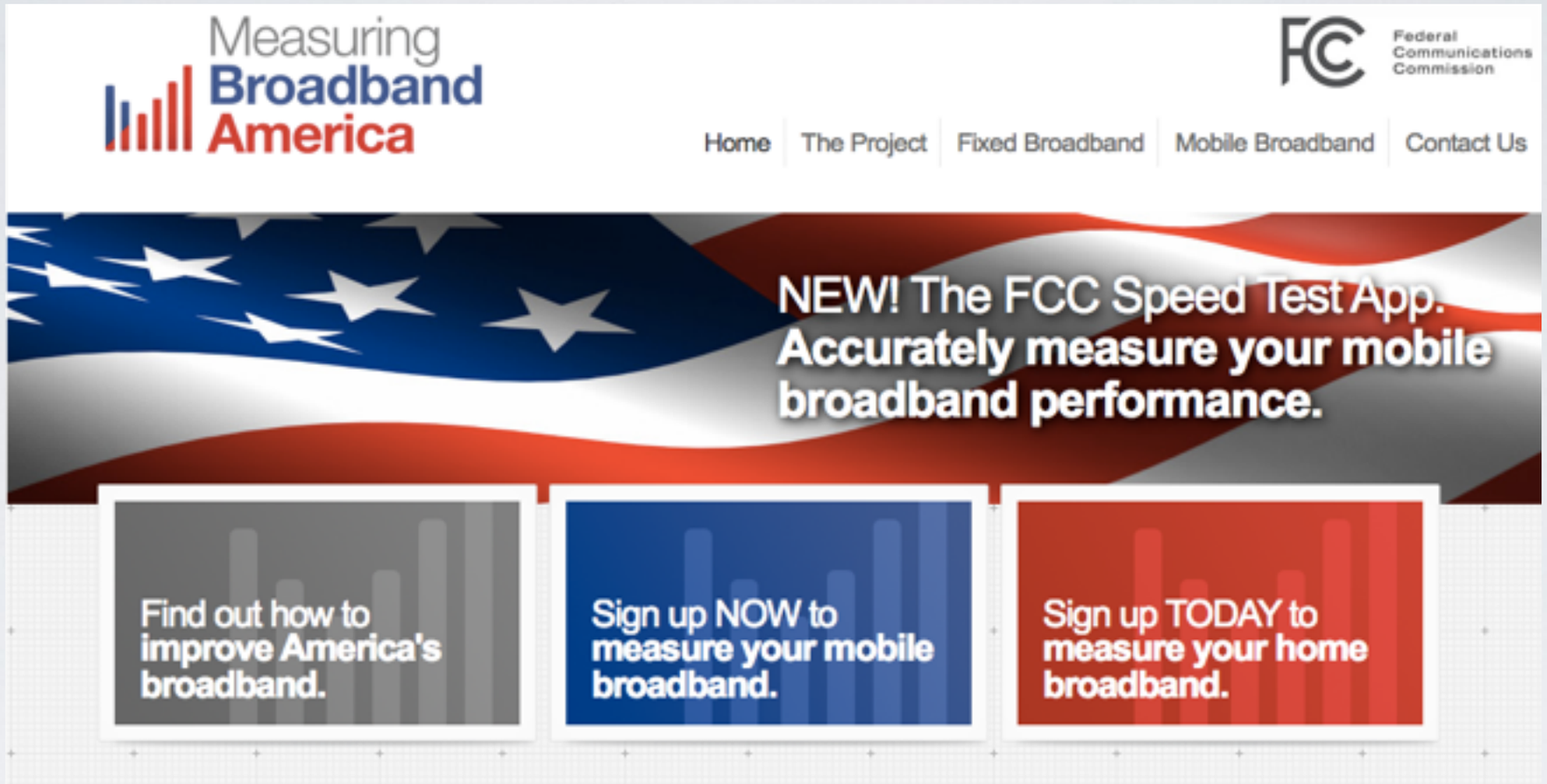
SC_REMOTED


probing state
decision logic
Central Server: able to control
100s of low resource VPs
sc_bdrmap
control
sc_remoted
remote
Central Server
Low Resource VP
Vantage Point
3.5MB
scamper
traceroute
alias resolution

BISMARK TRIAL

- Deployed border mapping and TSLP on ~15 Bismark OpenWRT routers (Thanks to the Project Bismark team!)
- Results are promising so far
- Varying degrees of reliability: router reboots and address changes interrupt border mapping and TSLP processes
- Enhancements to scamper remote operation mode should make things more robust

FCC-SAMKNOWS



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

THANKS!

amogh@caida.org