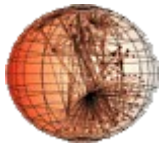
DNS root traffic: Analysis work to date ...

Nevil Brownlee, CAIDA



nevil@caida.org

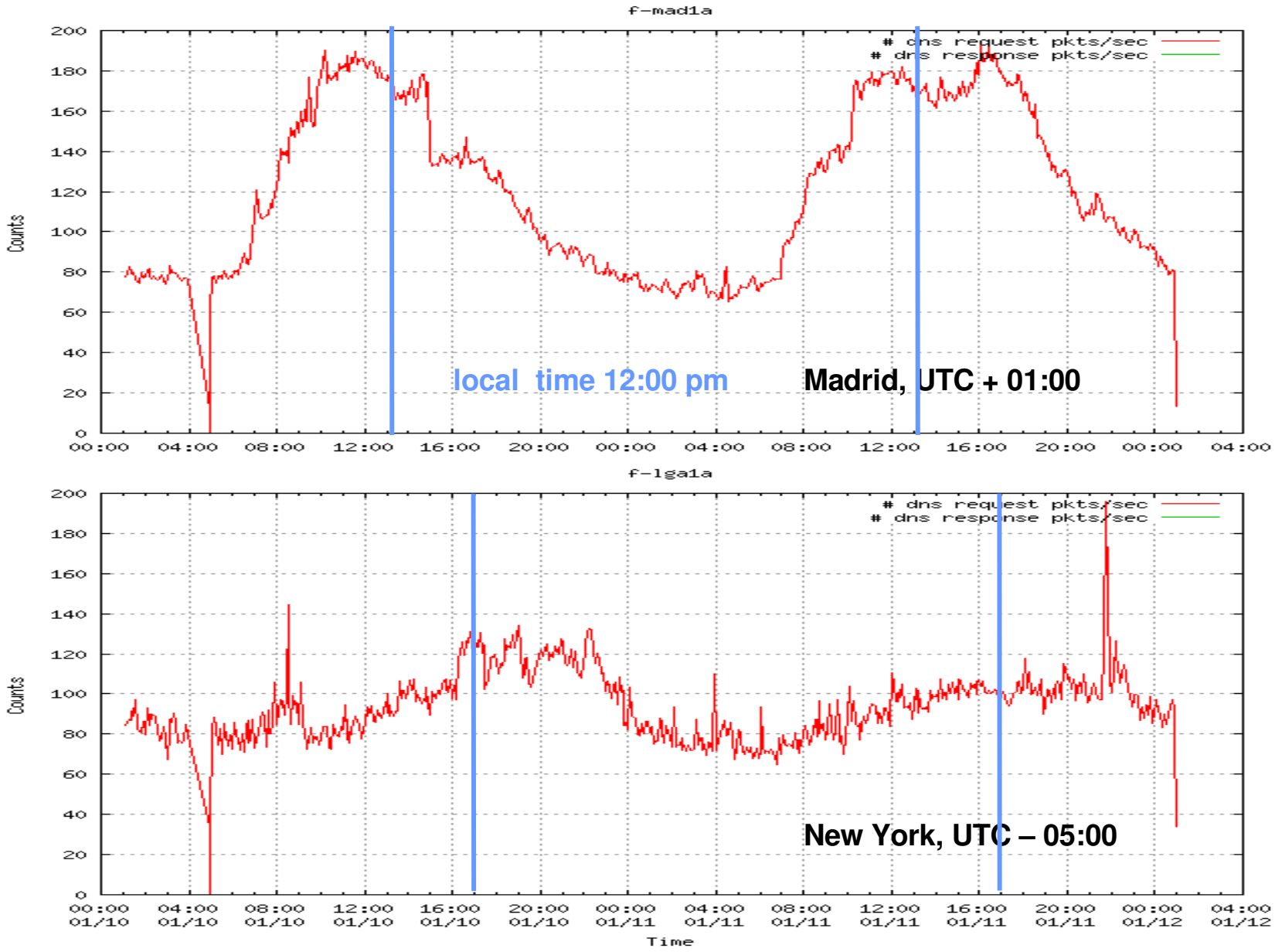
Outline

- **Analysis topics**
- **Some preliminary results**
- **Recommendations for future collections**

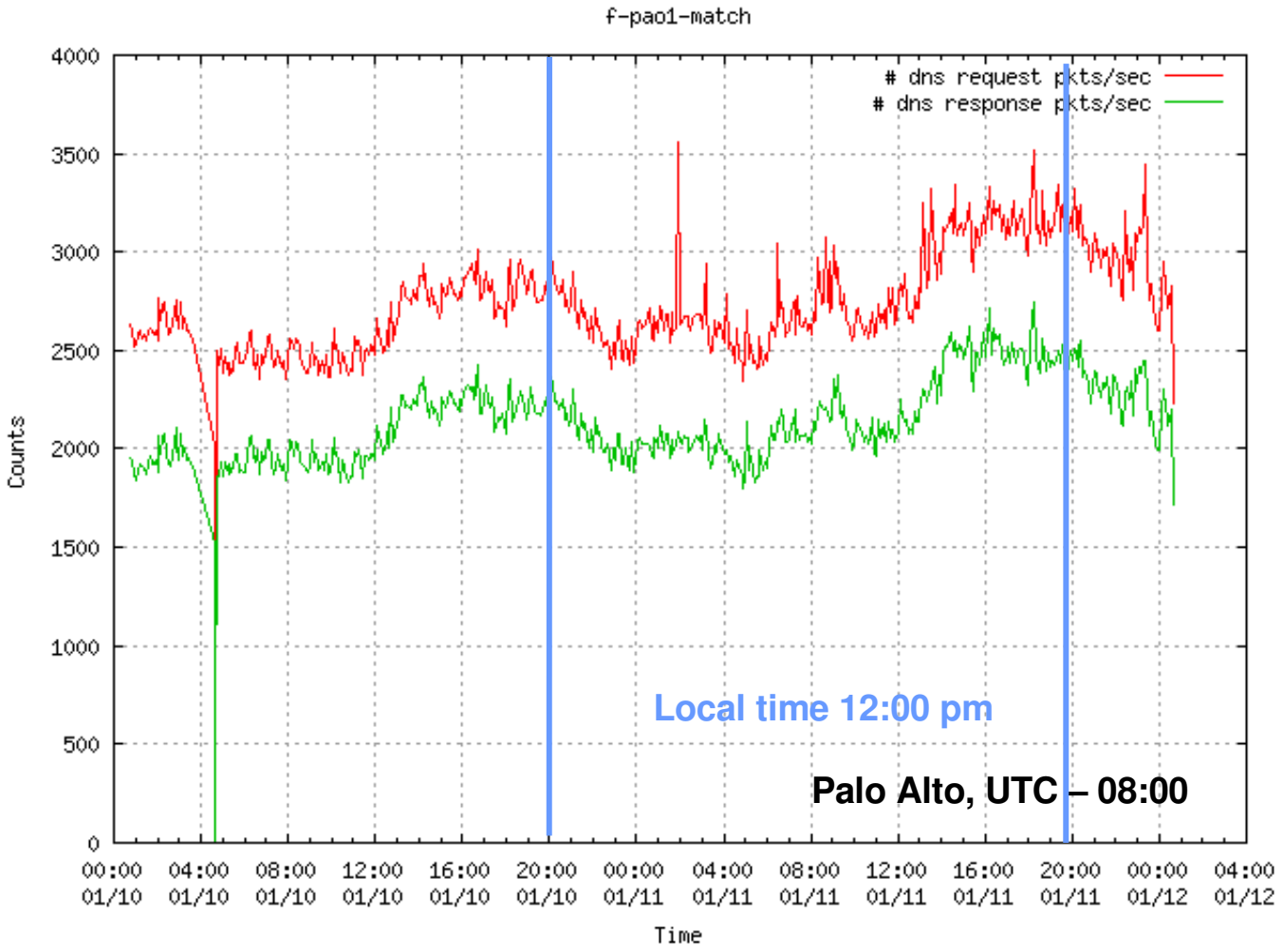
Analysis topics

- **Time-of-day usage difference**
- **Distribution of queries across anycast instances**
- **Distribution of response sizes and types**
- **Distribution of queries by gTLD and ccTLD**
- **Growth in and impact of DNSSEC**
- **Fraction of TCP request/response which are genuine DNS requests, not bogus**

- **Time-of-day usage difference**



- **Time-of-day usage difference**



Even the global instance shows slightly diurnal pattern!

(though we can not just match it with the local time)

- **DNS anycast analysis**

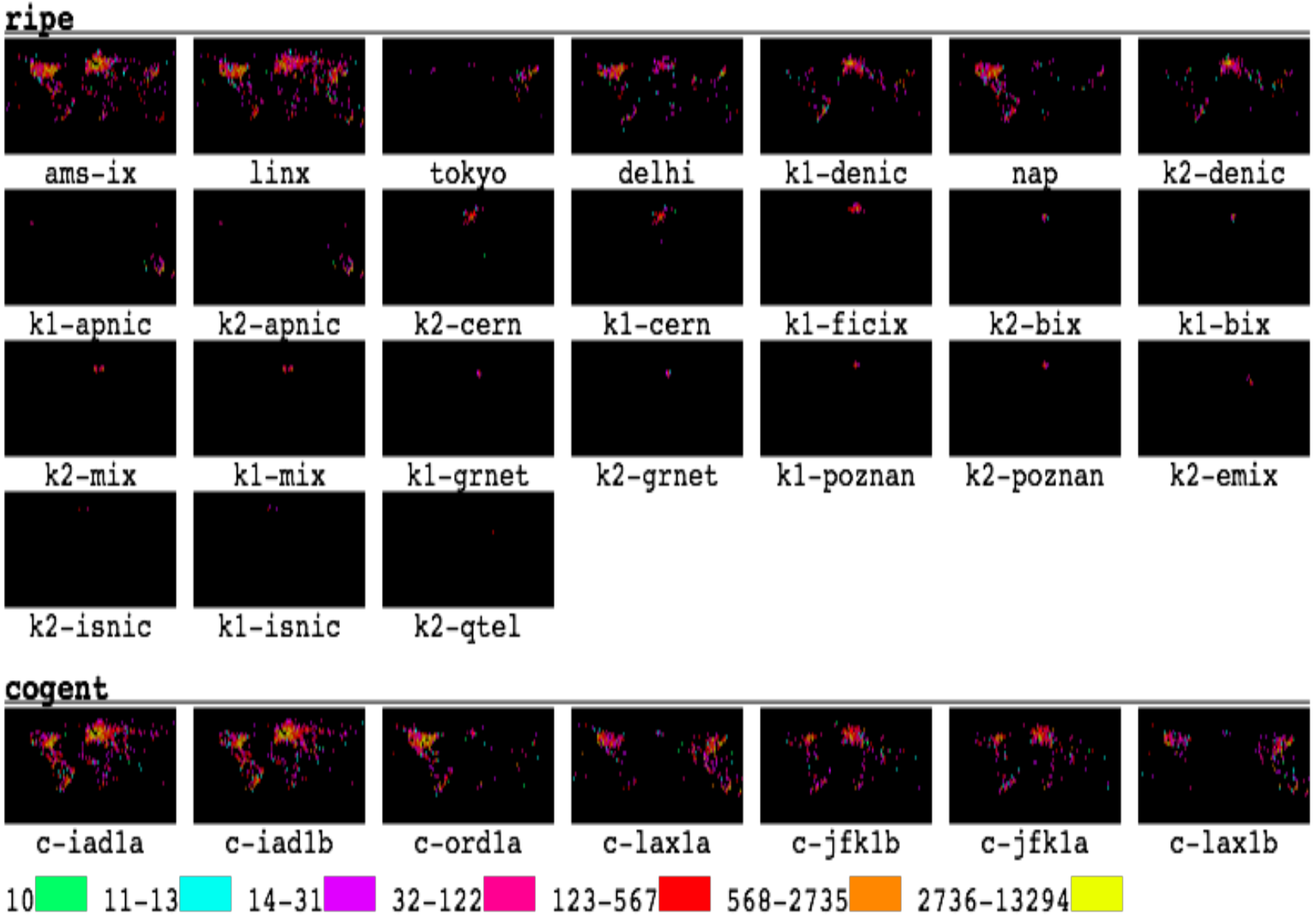
Datasets

- Date: 2006/01/10~11
- Member: Cogent (C root), RIPE (K root), ISC (F root)
- Geographic: Netacuity database for geographic mapping
- Topological: RouteViews BGP tables for ASs and prefixes (Jan 10, 2006)

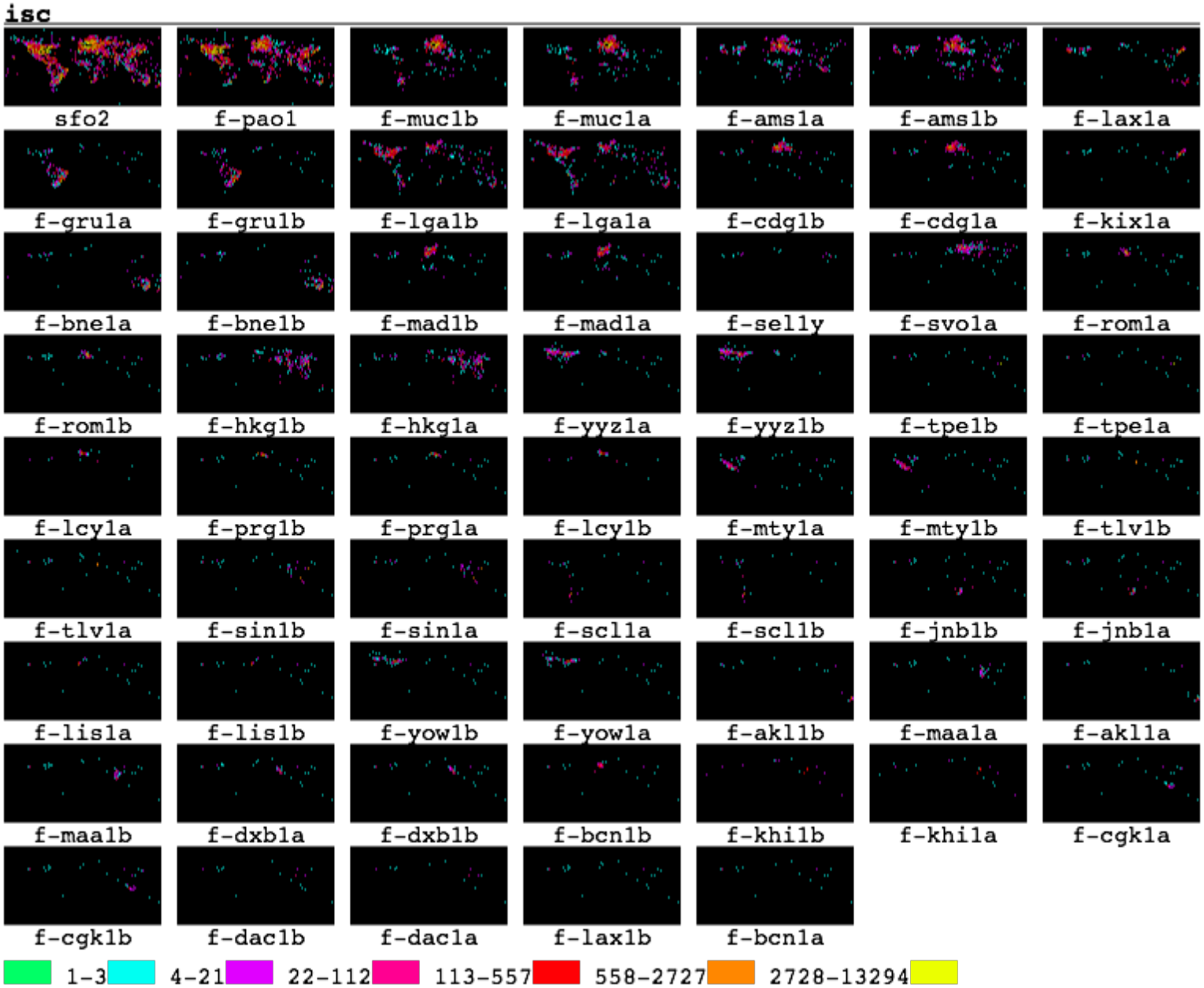
Scope

- Observed ASs: 19,237 (RouteViews: 21,883)
- Observed prefixes: 104,832 (RouteViews: 192,316)
- Observed IPs: 2,554,419

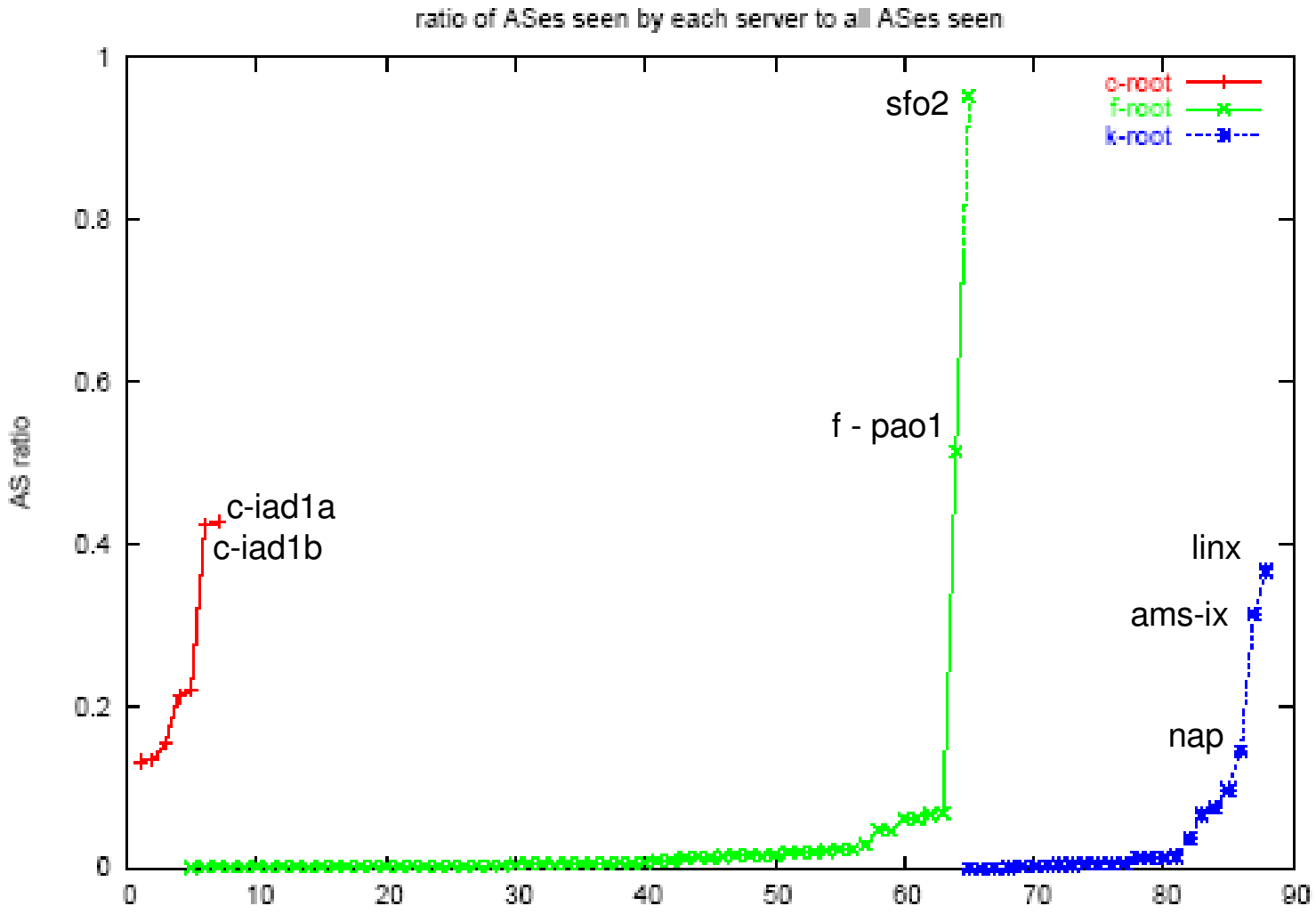
- **DNS anycast analysis – geographical distribution of clients**



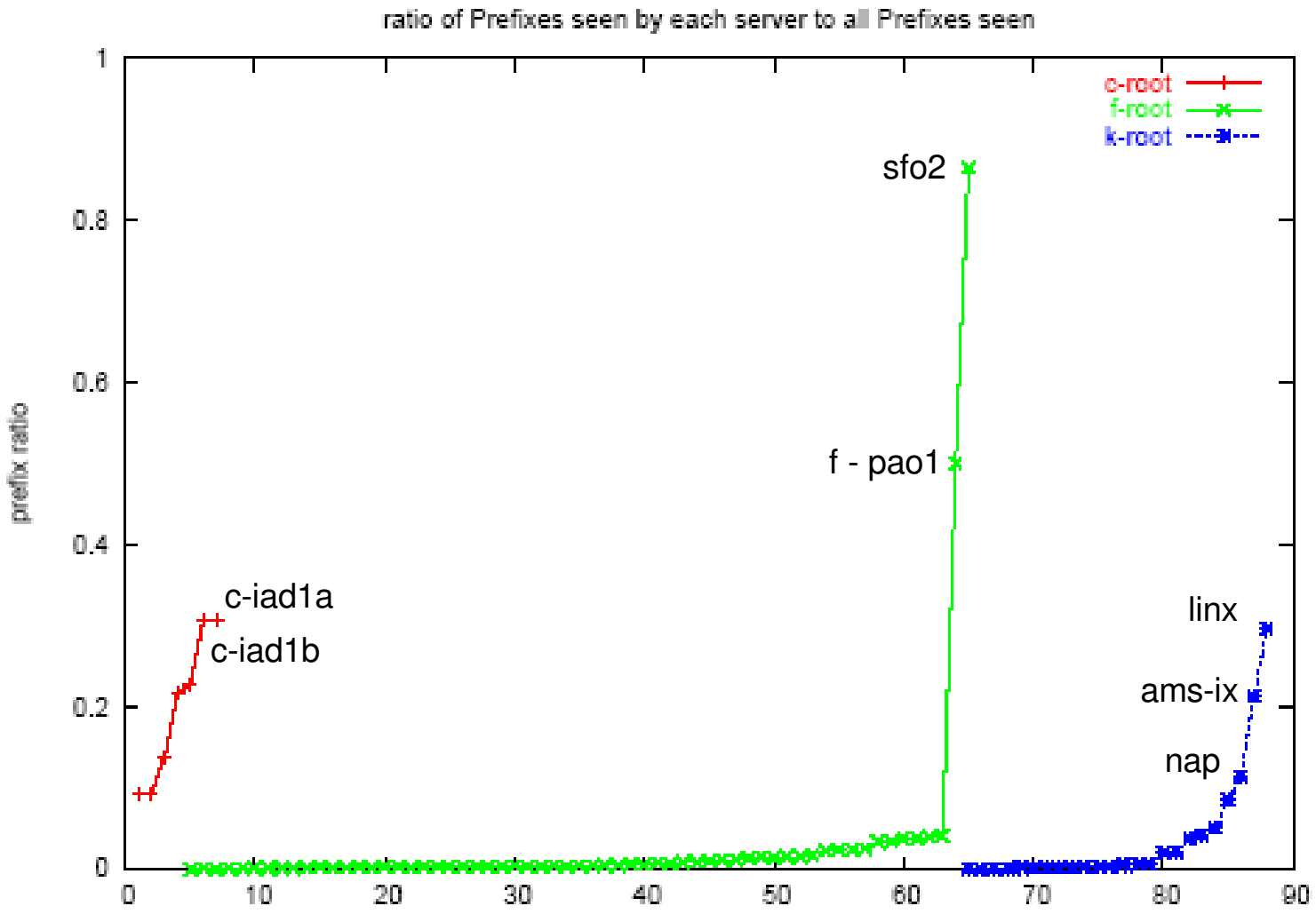
- **DNS anycast analysis – geographical distribution of clients**



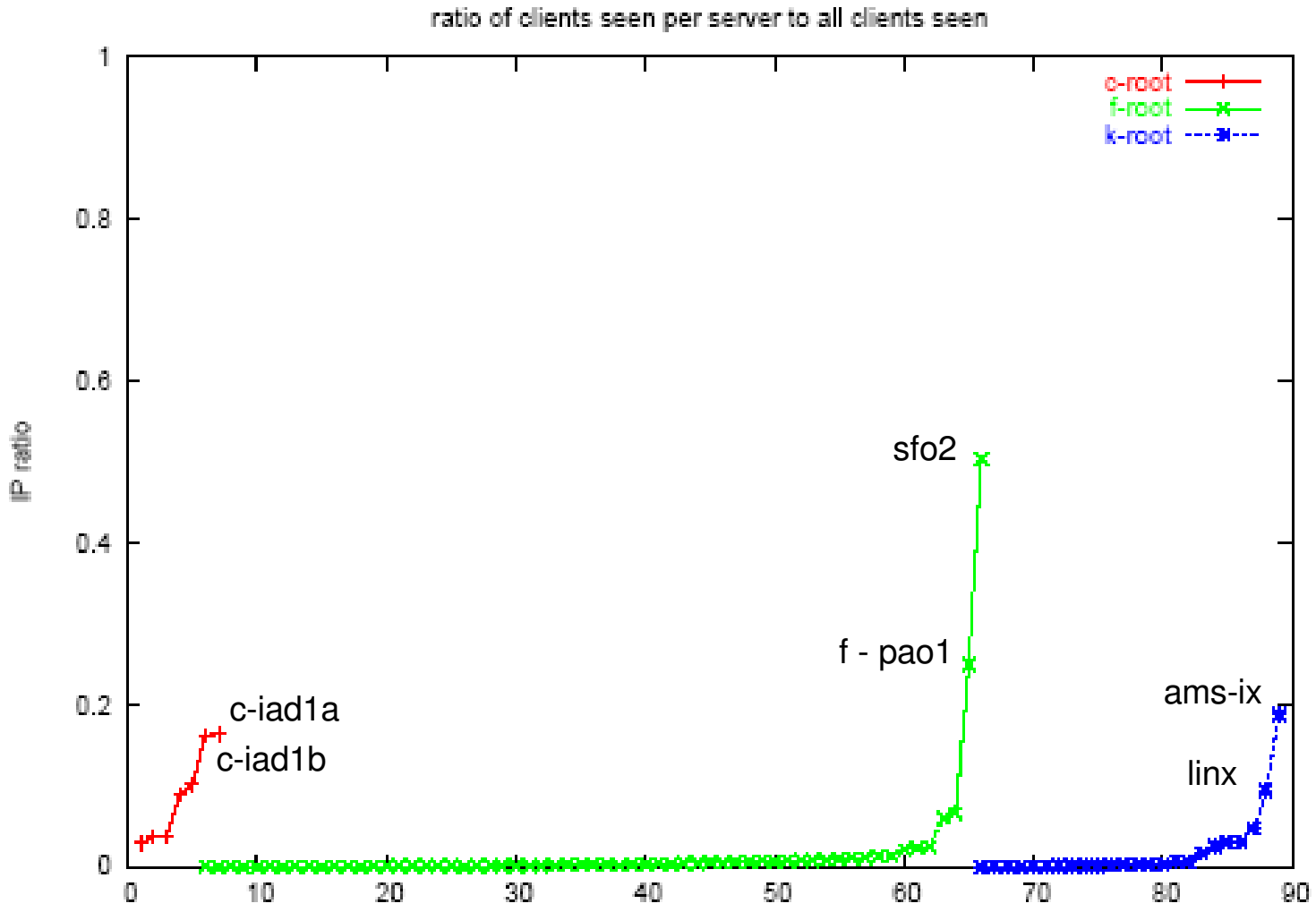
- **DNS anycast analysis – AS coverage per instance**
  ratio = ASs seen by instance / ASs seen by all



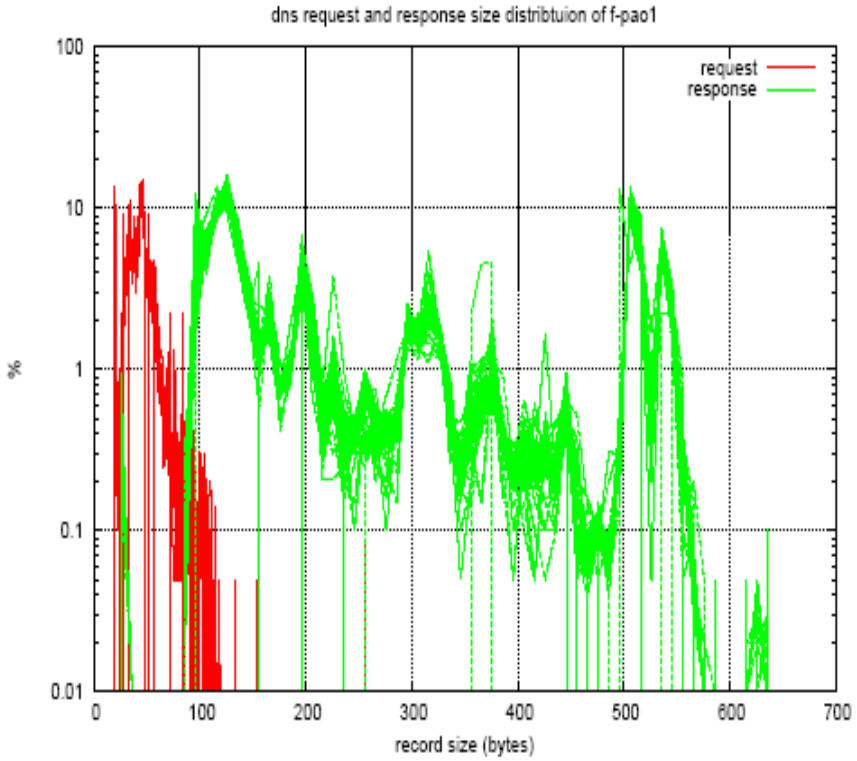
- **DNS anycast analysis – prefix coverage per instance**
  ratio = prefixes seen by instance / prefixes seen by all

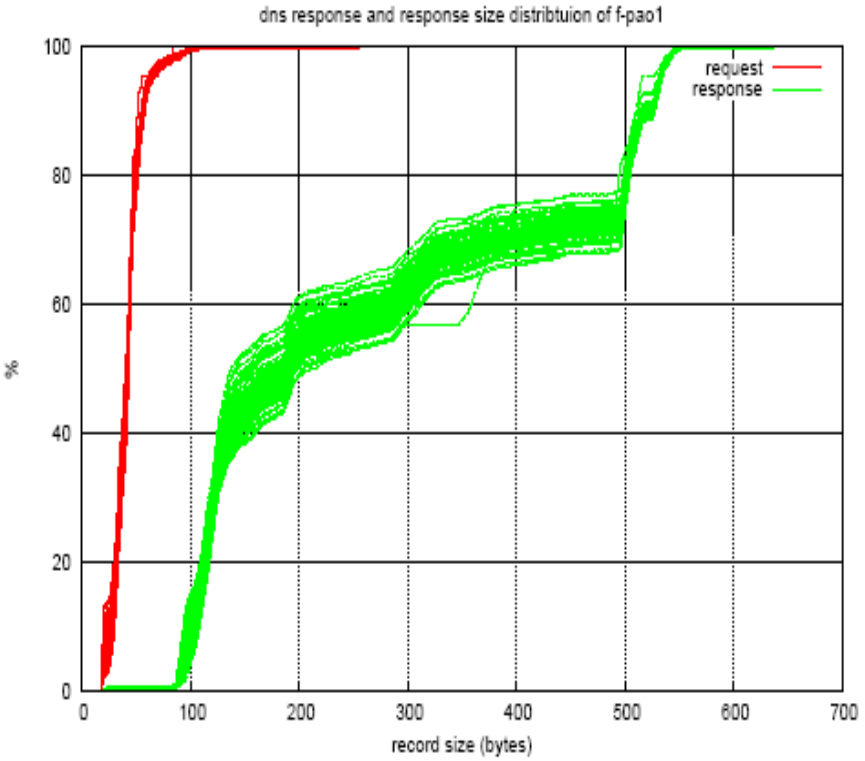


- **DNS anycast analysis – client (IP address) coverage per instance**
  ratio = clients seen by instance / clients seen by all



- **Size distribution of requests and responses**





pdf (y: log scale)

cdf (y: linear scale)

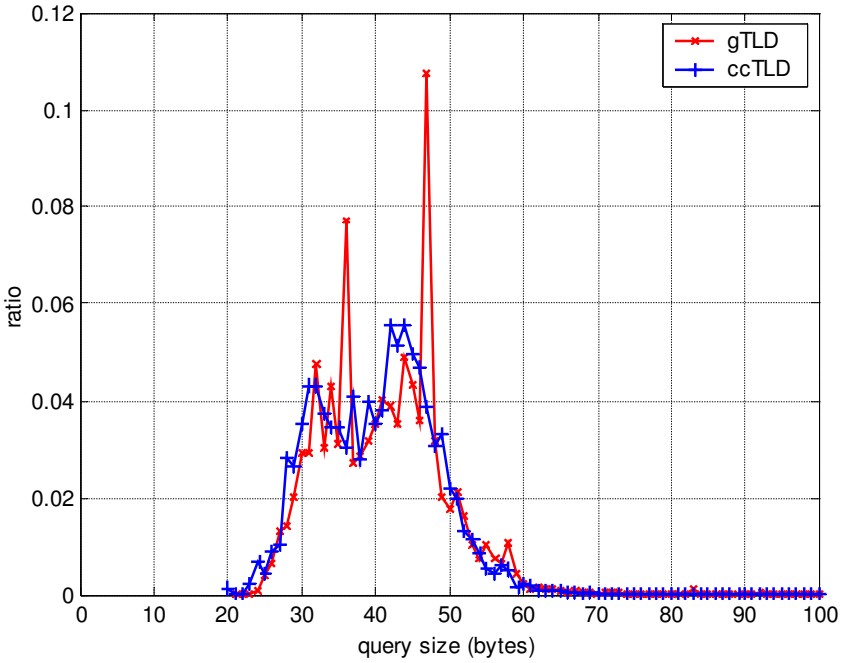
Request and response size distributions every 5m at Palo Alto

-- 80% of requests are shorter than 50B,

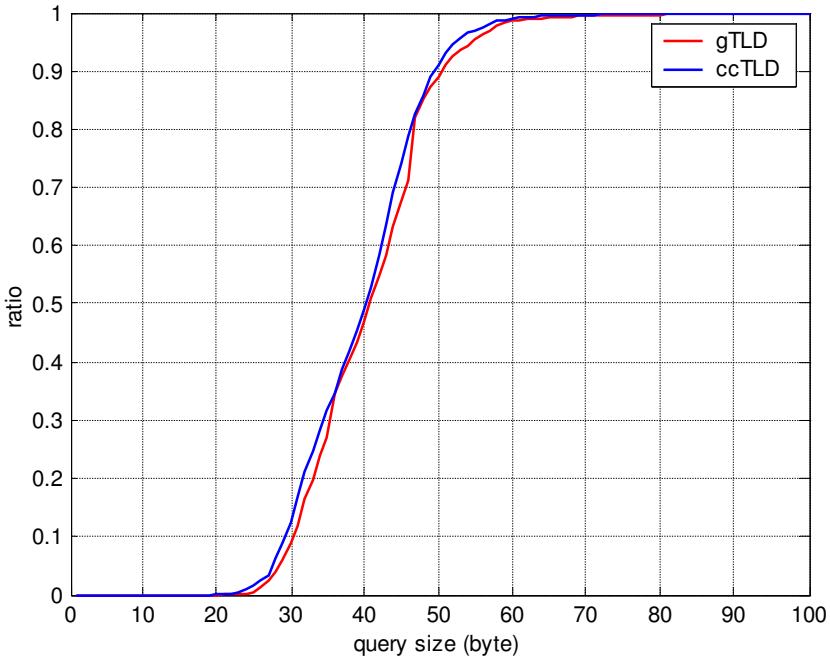
-- 80% of responses are shorter than 500B

- **Size distribution of UDP queries for gTLD and ccTLD, over all the OARC datasets**

The ratios of the query sizes larger than 100 bytes are tiny, not shown here…



pdf



cdf

	# of queries	min size (B)	max size (B)	avg. size (B)
gTLD	2.08G	21	1279 (faked?)	41
ccTLD	0.98G	20	1279 (faked?)	40

Overall, 90% of the queries are shorter than 50 bytes.

- **Fraction of genuine TCP DNS traffic**
  - We saw few genuine TCP requests
  - Most of TCP "requests" (to port 53) are bogus (syn, fin, ack, … no payload)
  - Further analysis needed
- **Growth in and impact of DNSSEC**
  - More analysis is needed. For example, what % of queries/responses include DNSSEC-related RRs?
  - Only a few instances collected response data

Recommendations for future collections

- **We want as much data as possible**
  - **Both TCP and UDP for port 53**
  - **Both queries and responses**
  - **From all root and gTLD instances**
  - **Over 48 hour long, mid-week preferred**
  - **Synchronized to UTC time**
- **Why so greedy?**
  - **TCP accounts for a small fraction of DNS traffic, but may well increase**
  - **Responses require significant storage, but are necessary to answer questions about DNSSEC usage**
  - **Differences between 'global' and 'local' instances**
  - **Average daily traffic, diurnal patterns, anomalies within a day**

Thanks

- **To my colleagues at CAIDA (who did most of the actual work)**
  - **Brad Huffaker**
  - **Ziqian Liu**
  - **Marina Fomenkova**

- **And to the OARC team, ISC and root operators (who collected the traces)**

- **Questions?**