DITL 2008
DNS Trace Collection

Duane Wessels
The Measurement Factory/CAIDA

OARC
June 4, 2008

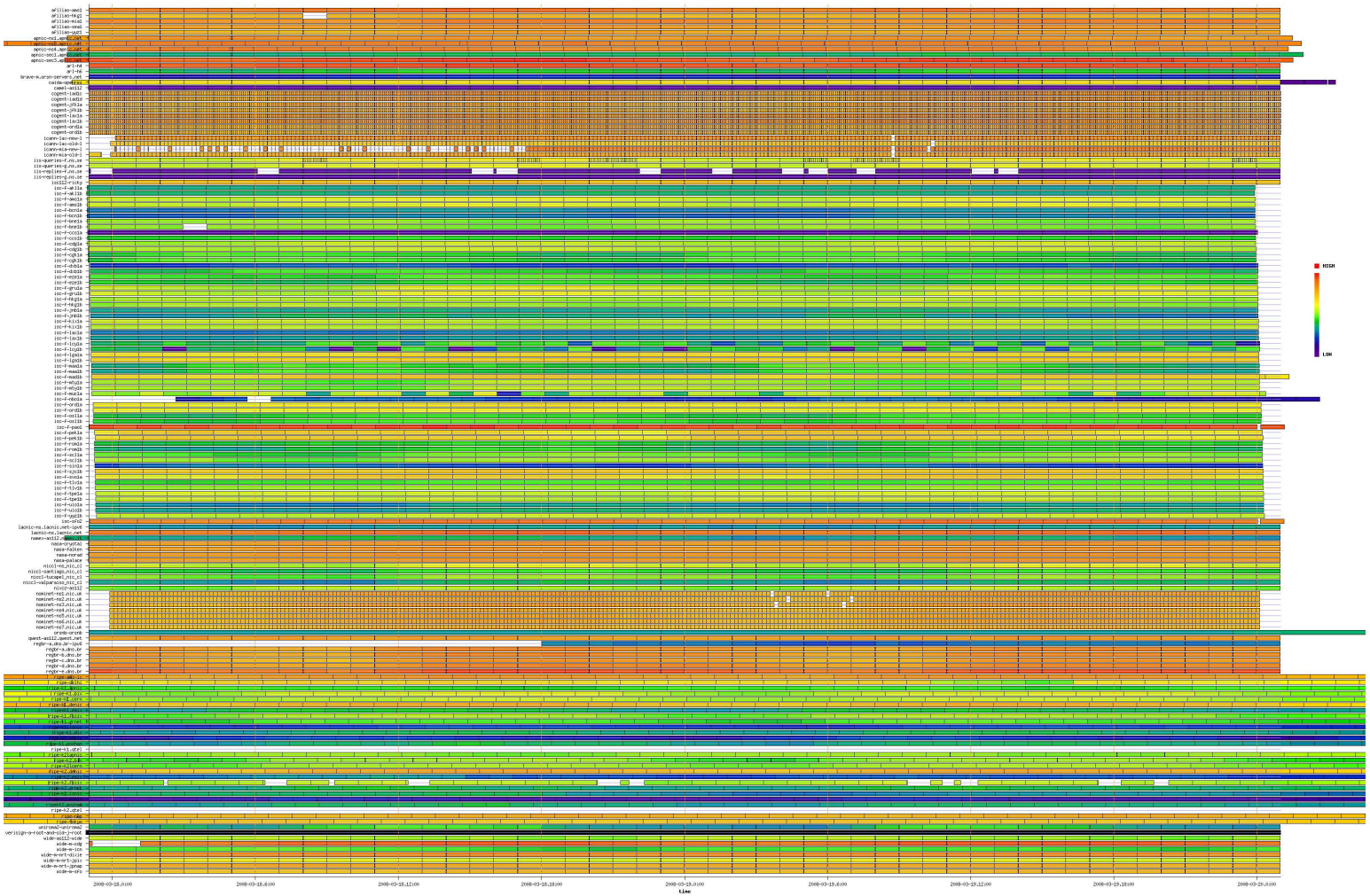
Day In The Life of the Internet

- An annual large-scale collection event coordinated by CAIDA and OARC.
- Goal is to collect (or at least catalog) a whole bunch of "Internet data" from a given 48-hour period.
- Focus on DNS traces, but also other types of data.
- 2008 is the third DITL.

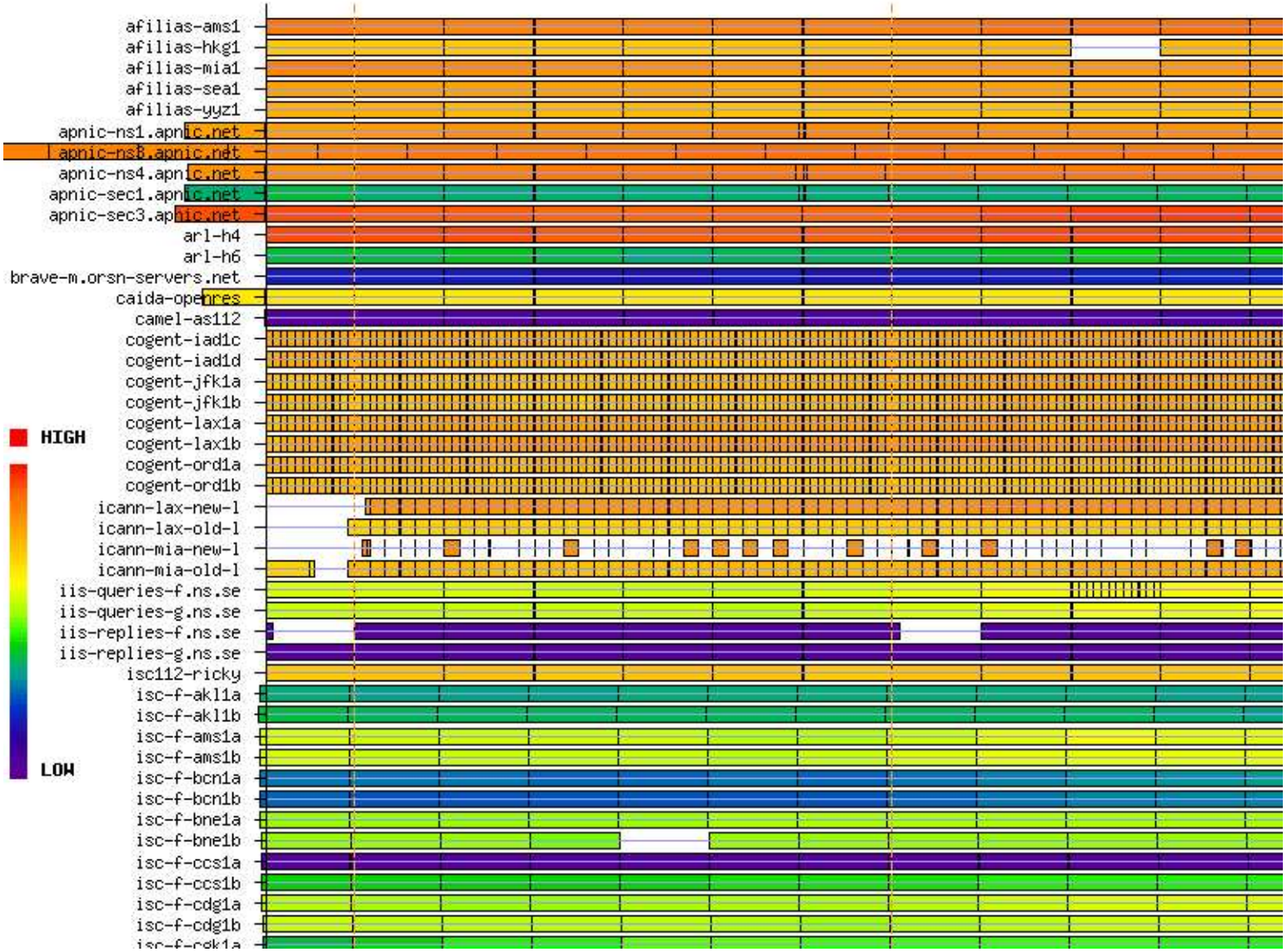
DNS Traces

	2006	2007	2008
Roots	C,E,F,K	C,F,K,M	A,C,E,F,H,K,L,M
ORSNs		B,M	B,M
AS112s		WIDE	Qwest, nix.cz, Camel, namex.it, ISC, WIDE Afilias
TLDs			uk, br, cl, se, org
RIRs			LACNIC, APNIC
Old-Roots			J,L
Resolvers			Uniroma2.it, SIE
Files	9,547	9,910	21,319
GBytes	229	742	1942

Raw Data Received



Coverage Chart Zoom



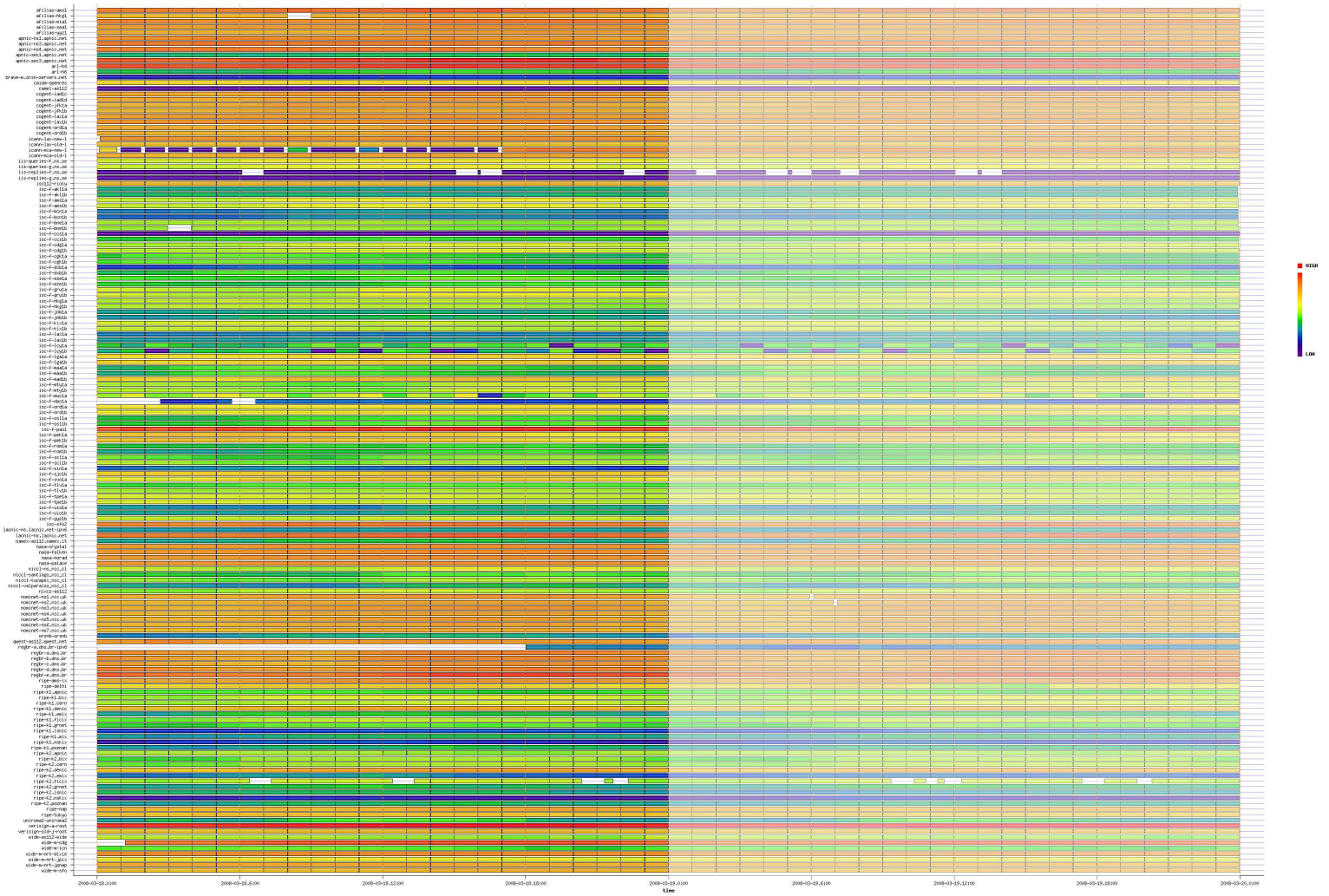
Problems

- Collector clocks not NTP-synced.
  - Please don't fix your clocks mid-collection.
- *dnscap* ignores TCP.
- *dnscap* and *tcpdump* with port-based filters ignore (UDP) fragments.
- Confusion about "node-id" leads to filename collision and data loss.
- *dnscap* script should have the option to **not** remove files after uploading.
- AF_INET6 has a different value on Linux and BSD. Files written by *dnscap* on Linux were not read properly on BSD.
- "Afilias" has only one "f."

Clean Data Set

- Corrected packet timestamps for incorrect clocks where known.
- Made all files 1-hour long.
- Separated Verisign's A-root and old-J-root.
- Separated v4 and v6 files for LACNIC.
- Truncated some files at the first sign of corruption.
- Fixed the AF_INET6 on Linux problem.
- Stripped VLAN tagging headers.
- Ensured that all packets are in chronological order.

Selected the latter 24 hours as the Analysis Period



More Info

- Raw and clean DITL traces available through DNS-OARC.
- http://www.caida.org/projects/ditl/summary-2008-03/
  - Describes additional, non-DNS data collected for DITL'08.

The End