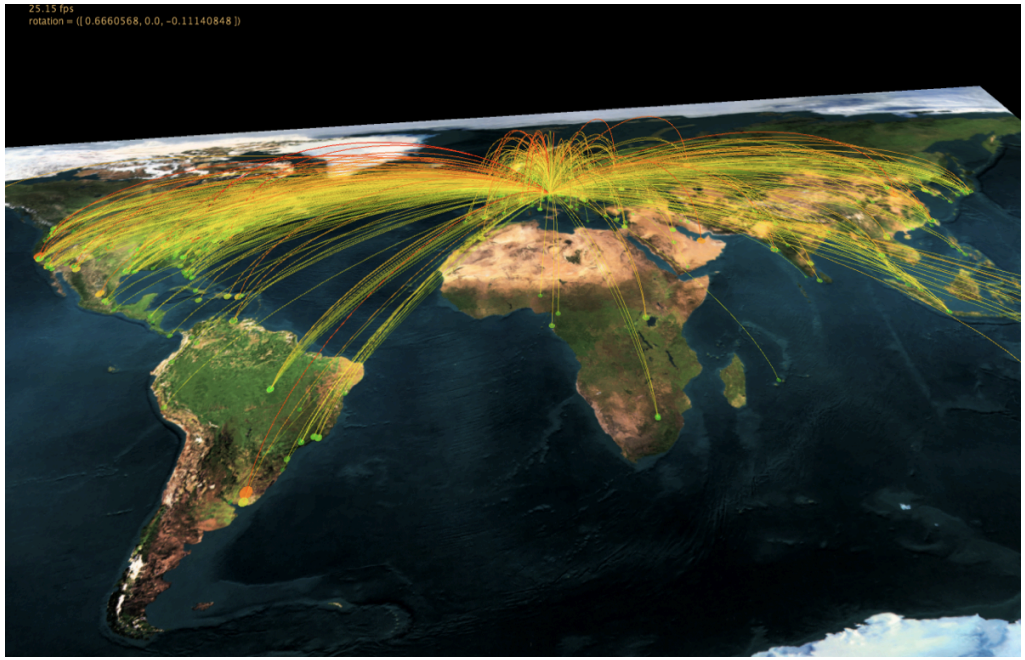
Passive DNS @ AcoNet / CERT.at ("pDNS") a.k.a. "DNS History project"

L. Aaron Kaplan kaplan@cert.at
Otmar Lendl lendl@cert.at

pDNS

- Idea in a nutshell:
  - Capture the **public DNS answer packet**
  - **at the recursor** (not the authoriative NS)
  - delete source IP, destination IP (🡐 privacy)
  - **timestamp the public DNS record** and finally
  - Store it in a DB
  - Provide a Query-Interface



- Invented in 2005 by Florian Weimer (BFK)

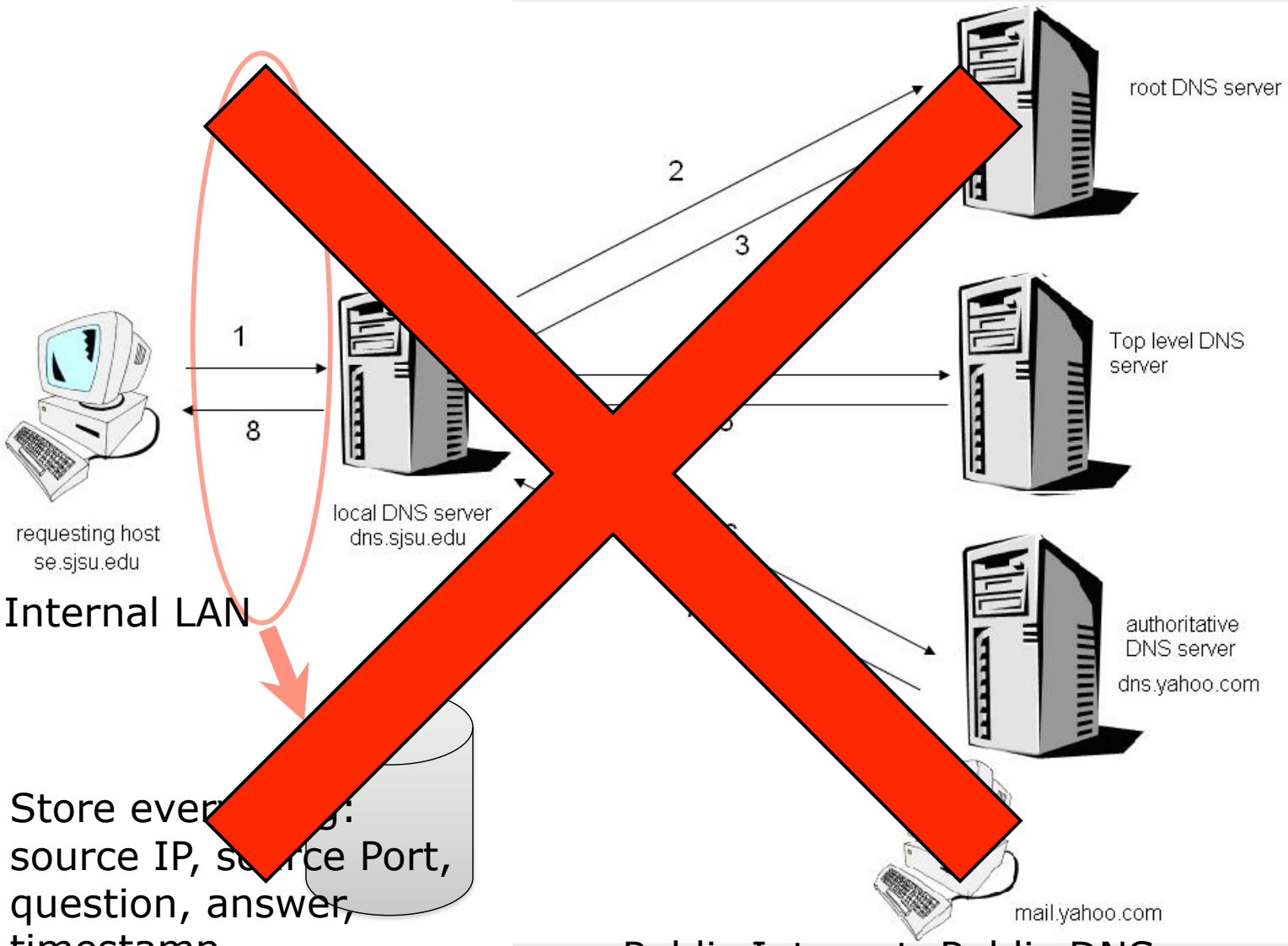
Two types of pDNS

- Pre-recursor passive DNS: the store-everything-that-you-can approach

  versus

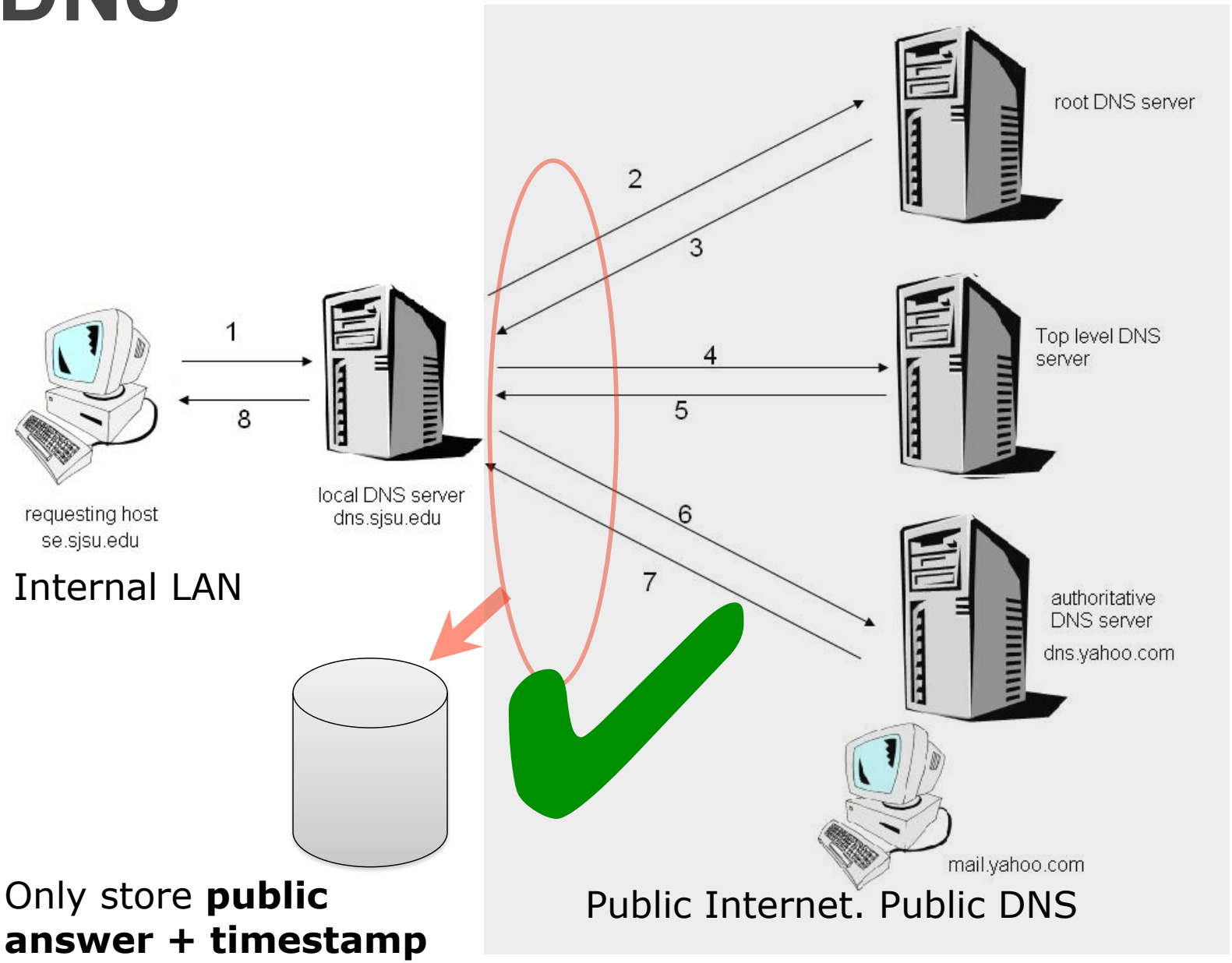
- Post-recursor passive DNS: only store what you really need.
  - Reduces volume
  - Respects privacy
  - Gets the benefit of caching of the recursor

PRE-recursor pDNS (Cisco)



Store ever[illegible]:
source IP, s[illegible]rce Port,
question, answer,
timestamp

BETTER: post-recursor pDNS



Only store **public answer + timestamp**

pDNS – the Data

CERT.at

rr-name: www.google.at
rr-type: A
rr-address: 173.194.35.184
seen-first: 2012-10-22 02:20:34
seen-last: 2014-03-02 20:10:42
count-requested: 40760

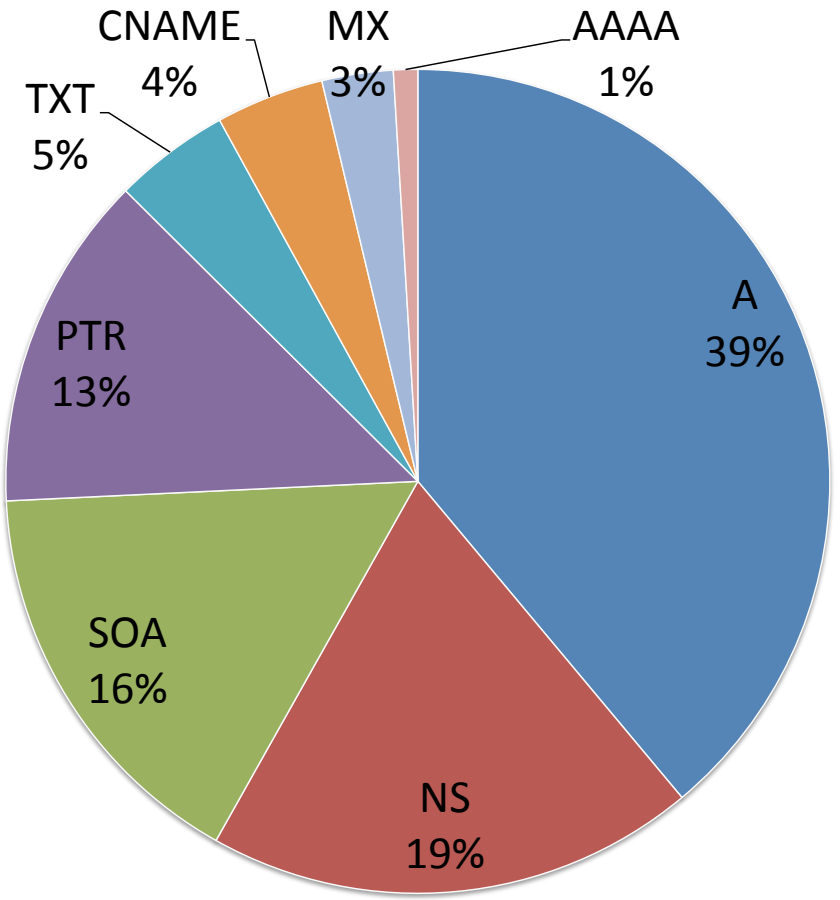
rr-name: www.google.at
rr-type: A
rr-address: 173.194.35.191
seen-first: 2012-10-22 02:20:34
seen-last: 2014-03-02 20:10:42
count-requested: 40760

rr-name: www.google.at
rr-type: A
rr-address: 173.194.35.183
seen-first: 2012-10-22 02:20:34
seen-last: 2014-03-02 20:10:42
count-requested: 40760

Public Data. Anyone on the Internet can query this. We do not know who asked that question

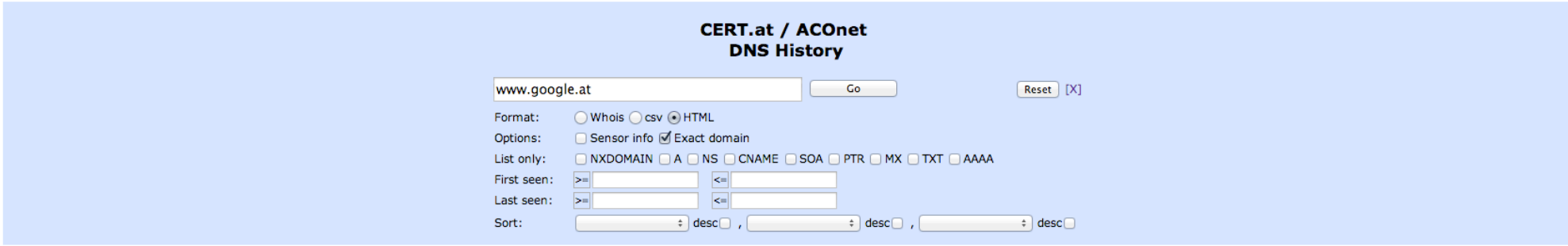
Our Dataset

- ~ 600 GByte of data
- ~ 2.5 billion rows
- ~ 520M Updates / day
- ~ 96% Last seen
- Lots of caching in RAM



- Based on PostgreSQL
- Using SSDs

Web-Interface for Queries




```
% CERT.at / ACOnet DNS replicator WHOIS server, version 2.0.
% (C) 2011 All rights reserved.
% Authors: L. Aaron Kaplan <kaplan AT cert.at>
%          Achim Adam      <achim.adam AT univie.ac.at>
%
% 419 elements, 0.1437s
```

Answer + Timestamp

LEFT	RTYPE	RIGHT	FIRST-SEEN	LAST-SEEN	COUNT-SEEN
www.google.at	A	74.125.232.223	2012-09-21 06:05:39	2012-09-21 06:05:39	48
www.google.at	A	74.125.232.215	2012-09-21 06:05:39	2012-09-21 06:05:39	48
www.google.at	A	74.125.232.216	2012-09-21 06:05:39	2012-09-21 06:05:39	48
www.google.at	A	74.125.232.248	2012-09-21 12:39:39	2012-09-21 12:39:39	349
www.google.at	A	74.125.232.247	2012-09-21 12:39:39	2012-09-21 12:39:39	349
www.google.at	A	209.85.148.94	2012-09-11 17:27:31	2012-09-27 11:11:29	5
www.google.at	A	74.125.135.94	2012-09-10 13:06:35	2012-10-17 18:16:55	5
www.google.at	A	74.125.232.56	2012-11-22 18:40:04	2012-11-22 18:40:04	2
www.google.at	A	74.125.232.55	2012-11-22 18:40:04	2012-11-22 18:40:04	2
www.google.at	A	74.125.232.63	2012-11-22 18:40:04	2012-11-22 18:40:04	2
www.google.at	A	74.125.227.56	2012-11-22 18:40:31	2012-11-22 18:40:31	1
www.google.at	A	74.125.227.63	2012-11-22 18:40:31	2012-11-22 18:40:31	1
www.google.at	A	74.125.227.55	2012-11-22 18:40:31	2012-11-22 18:40:31	1
www.google.at	A	74.125.129.94	2012-11-22 18:40:36	2012-11-22 18:40:36	1
www.google.at	A	74.125.224.120	2012-11-22 18:40:41	2012-11-22 18:40:41	1
www.google.at	A	74.125.224.119	2012-11-22 18:40:41	2012-11-22 18:40:41	1
www.google.at	A	74.125.224.127	2012-11-22 18:40:41	2012-11-22 18:40:41	1
www.google.at	A	74.125.230.215	2012-10-23 08:26:04	2012-11-28 11:41:30	3
www.google.at	A	74.125.230.216	2012-10-23 08:26:04	2012-11-28 11:41:30	3
www.google.at	A	74.125.230.223	2012-10-23 08:26:04	2012-11-28 11:41:30	3

- Alternative UI: whois –h server <question>

Answer which Questions?

- Historic data
  - „What was the A record for a certain FQDN last year?“
- Inverse Lookups
  - „Which domains have A records that are in a given address range?“
- Generic reseach on bulk DNS data
  - T. Frosch, T. Holz: „Predentifier: Detecting Botnet C&C Domains From Passive DNS Data“

Example 1: C&C server

1. `hair3Choo8aibaaj.foo.ru` is a C&C server
2. Some PCs get infected
3. `hair3Choo8aibaaj.foo.ru` gets deleted
4. The CERT gets the information that `hair3Choo8aibaaj.foo.ru` was evil. Look at all connections / flows to that domain
5. But the domain got deleted. What to do?

6. Answer: look into the **DNS history** ➔ find the IP address ➔ look at netflows and find all infected PCs

Example 2: Is this a bullet proof hoster?

- Step 1: the netblock: 193.104.27.0/24. AS12604 / Kamushnoy Vladimir Vasulyovich - suspected BP host
- Step 2: ask pDNS:

```
rr-name: ns2.federalbankofnevada.com
rr-type: A
rr-address: 193.104.27.69
seen-first: 2010-02-17 09:57:25
seen-last: 2010-02-21 12:04:29
rr-name: pharmazoria.com
rr-type: A
rr-address: 193.104.27.164
seen-first: 2009-12-03 17:16:39
seen-last: 2009-12-30 12:33:43
rr-name: www.genericmedsusa.com
rr-type: A
rr-address: 193.104.27.162
seen-first: 2009-12-16 16:04:07
seen-last: 2009-12-21 11:47:22
```

- Here we found 500+ entries!
- Many very shady records
- Strong indication that this hoster is a Bullet Proof Hoster

Example 3: suspicious domains in my netblock

- Step 1: create a list of known good domains in my network range
- Step 2: ask pDNS for my network range:
- Step 3: make a diff. Find domains which point to your IP range, but you are not aware that they were there!

- We could offer this as a service. Anyone interested?

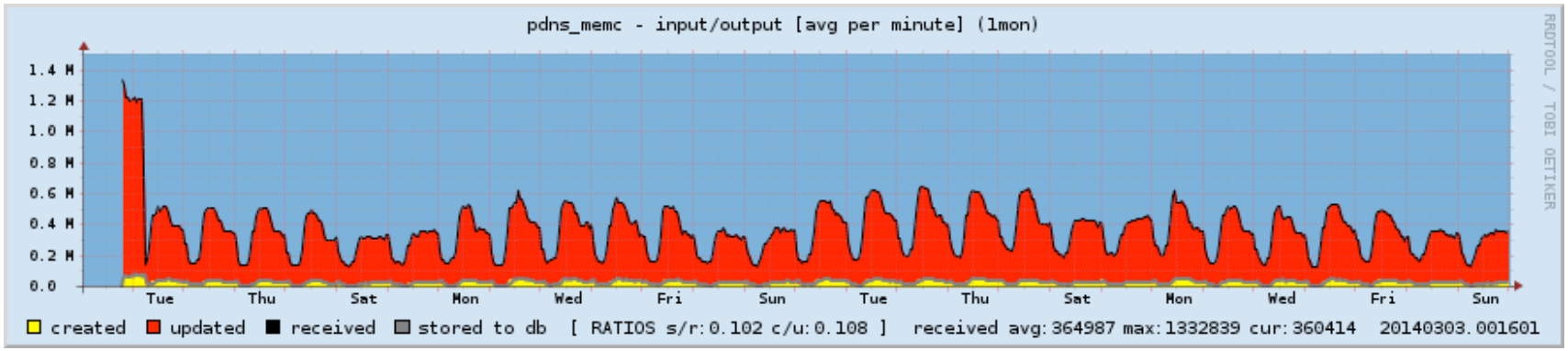
Example 4: Egypt goes offline

- Egypt goes offline, this poses a research question:
  - What other domains are offline because all of their NS are in egyptian IP space?
  - ➔ we can find out

DATA VOLUME AND TUNING

Data volume

- Creating a pDNS server is easy **BUT** does it ~~blend~~ scale?
- As of 2014: ~ 2.5 Billion entries in a Postgres DB
- Number of DNS answers/minute coming in:



- It was not trivial and took a lot of time to tune the system to be able to handle so much data in PostgreSQL + 100GBytes of RAM + many SSDs.

Data Volume (2)

- Largest pDNS Servers that we are aware of:
  1. ISC/Farsight
  2. BFK
  3. CERT.at/Aconet

Current topics with pDNS

- Multiple implementations (ISC/Farsight, BFK, CERT.at/Aconet, …)
- Aim: Make them interoperable

`Passive DNS - Common Output Format`

`draft-dulaunoy-kaplan-passive-dns-cof-02`

- Submitted to the DNSOP WG @ IETF
- Supported by FIRST.org

Participation

- Access to our DB is limited to:
  - Specific reason + signed MoU: researcher or legitimate IT Security (CERTs)
  - Contributors of data
    - Run a sensor
    - Feed in the data, mix it up further with other sensors (mixing is good)
    - The more diverse the user-base of the sensors is, the better the overall data quality

Summary

- It is possible to keep a DNS history while at the same time preserving privacy

- Applications:
  - Research
  - IT Security
  - Monitoring / Alerting of suspicious domains
  - ... your idea?...

kaplan@cert.at, lendl@cert.at

THANK YOU!