Measuring the Leakage of Onion at the Root

A measurement of Tor's .onion pseudo-top-level domain
in the global domain name system

Aziz Mohaisen

Verisign Labs

(Joint work with Matt Thomas)

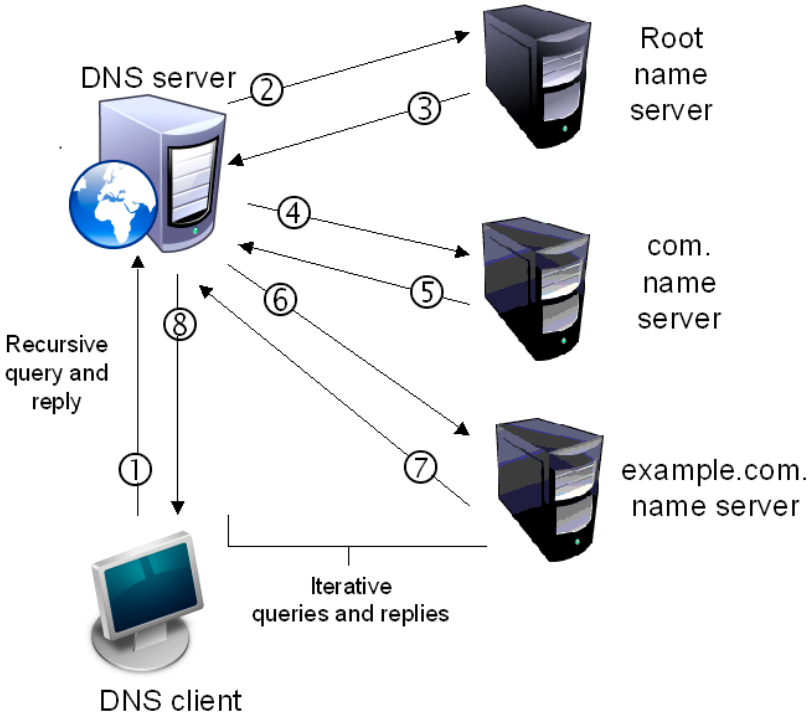
Agenda

- The global DNS and private namespace usage
- .Onion measurements from the A and J root servers
  - A Longitudinal Study of .Onion Traffic
  - Root Sampling Completeness (Representative?)
  - Volume and Diversity of Hidden Service Requests
  - Most Requested Hidden Services
  - ASN + Geo Diversity of Hidden Service Requests
  - Tor and the World: Event Correlation
- Trends from Day in the Life (DITL) of Internet
- Concluding Remarks and Future Work
- Q&A

The Global DNS and Private Namespace Usage

- The global DNS is a hierarchical system
- Currently there are 13 groups [A-M] of root servers
- Authoritative for TLDs such as “.com”, “net”, etc.

- Many installed systems utilize non-delegated TLDs for internal namespaces
  - E.g. “.corp” “.home”
- Queries to the roots for such TLDs result in NXDomains

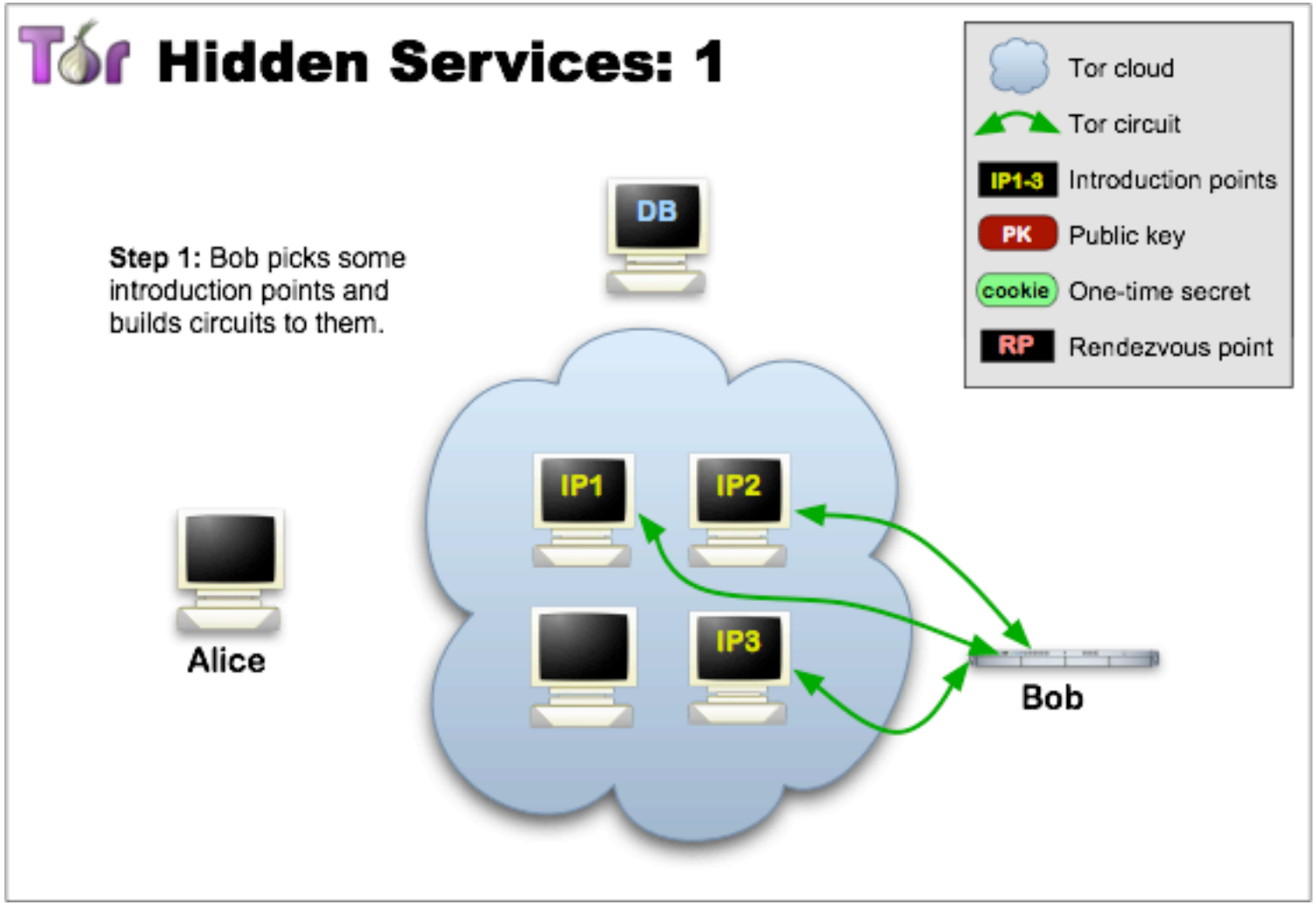


The Global DNS and Private Namespace Usage

- Delegation of new gTLDs spurred more critical studies of NXD requests at the roots.
  - Potential “Name Collision” Risks (see namecollisions.net).
  - Unintended leaked DNS queries may expose potentially sensitive private information and present additional threat vectors.

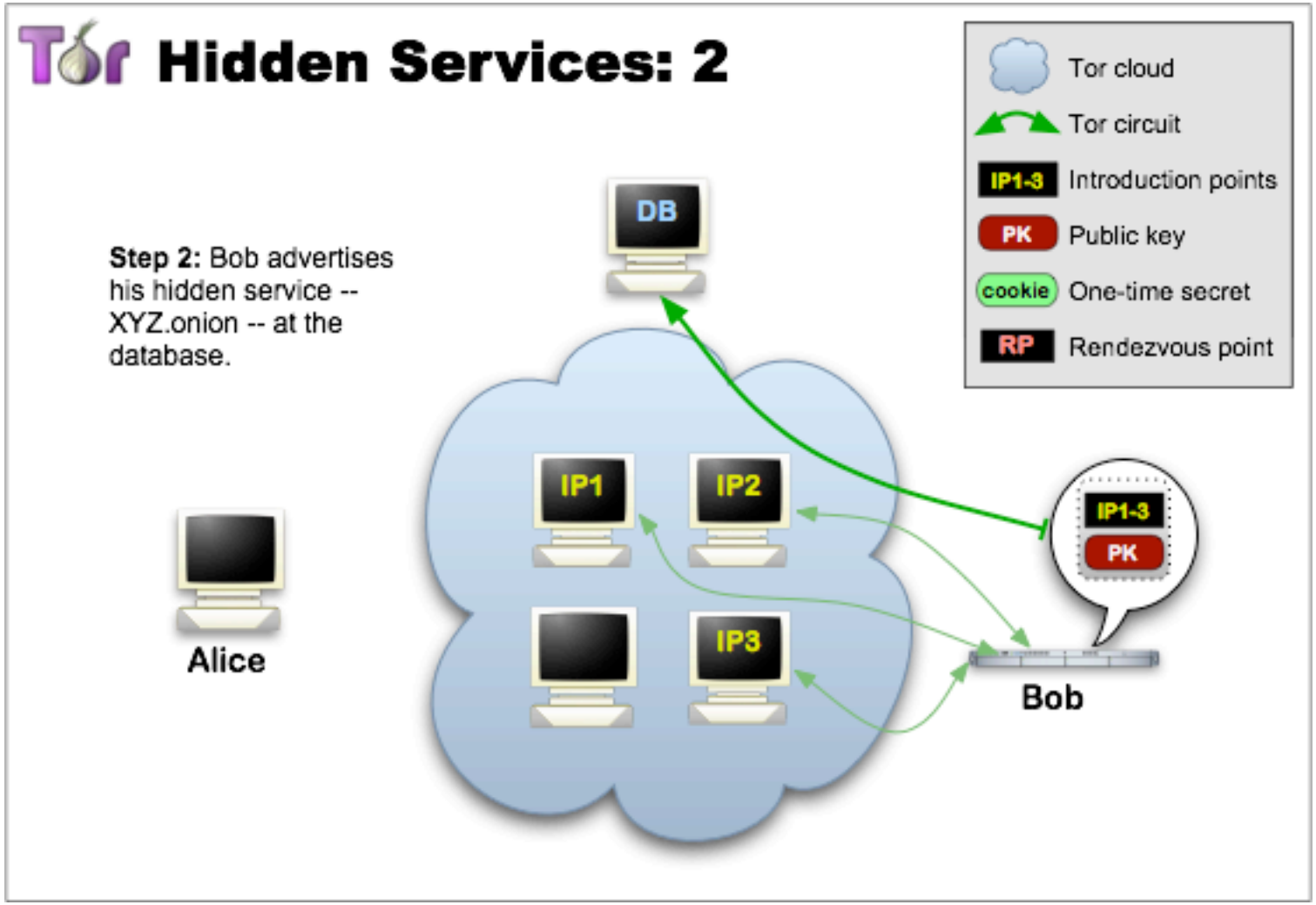
- Tor is a system that exploits the absence of a non-delegated TLD - .onion – for its Hidden Services
  - Tor is designed not to route .onion requests into the public DNS
  - It relies on the hidden service protocol for “torizing” requests.

A primer on Tor's Hidden Services



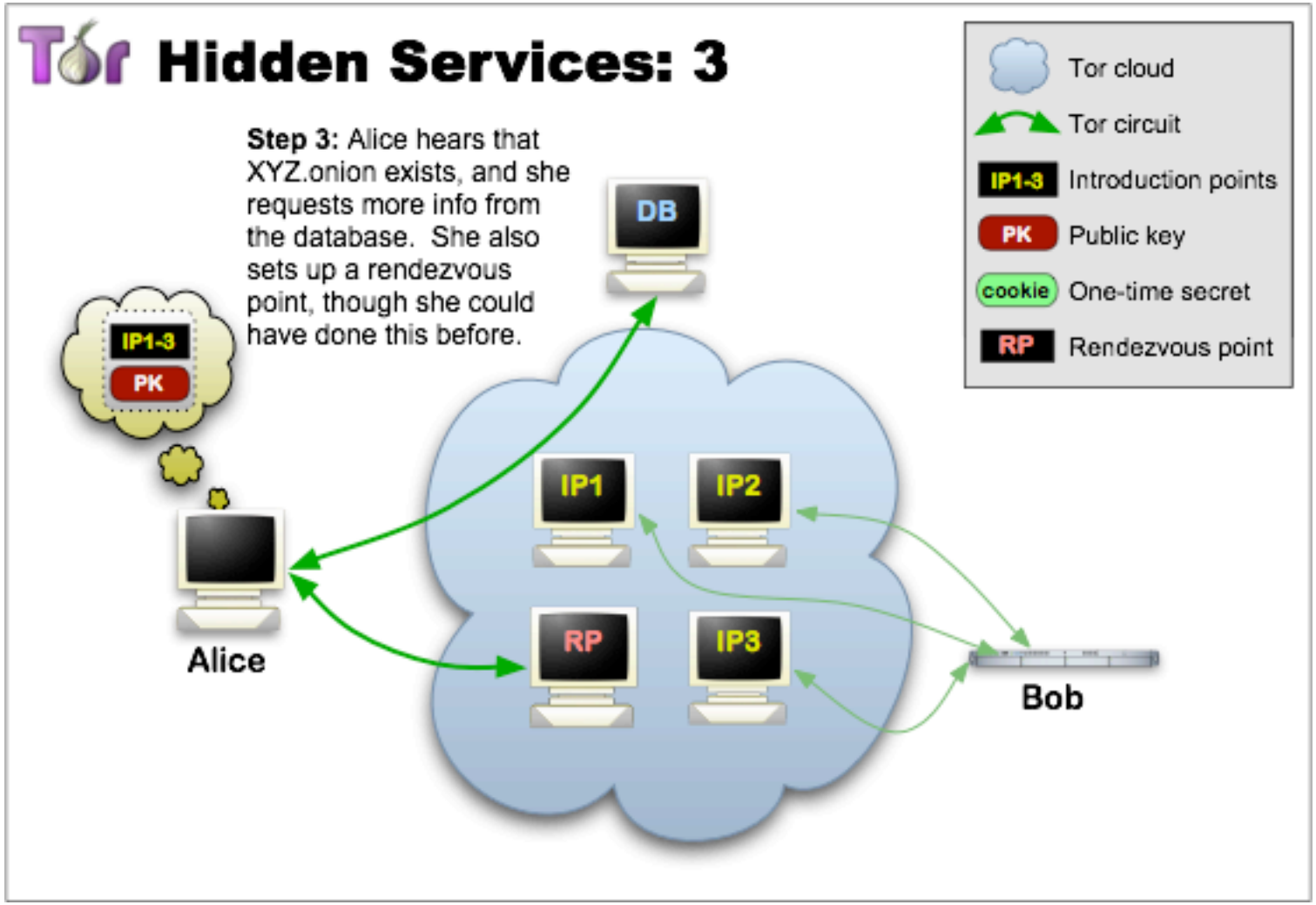
Source: torproject.org

A primer on Tor's Hidden Services, cont.



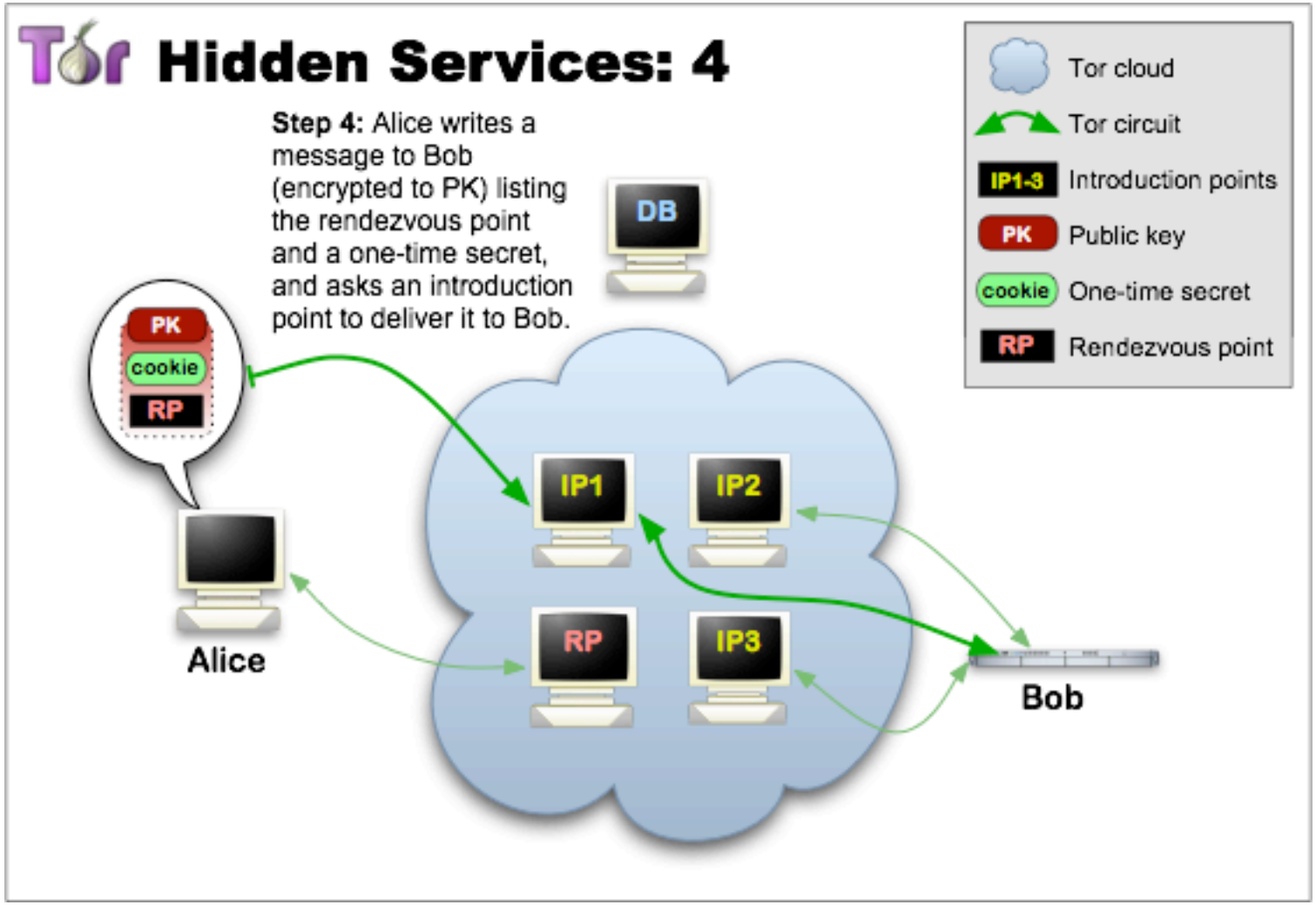
Source: torproject.org

A primer on Tor's Hidden Services, cont.



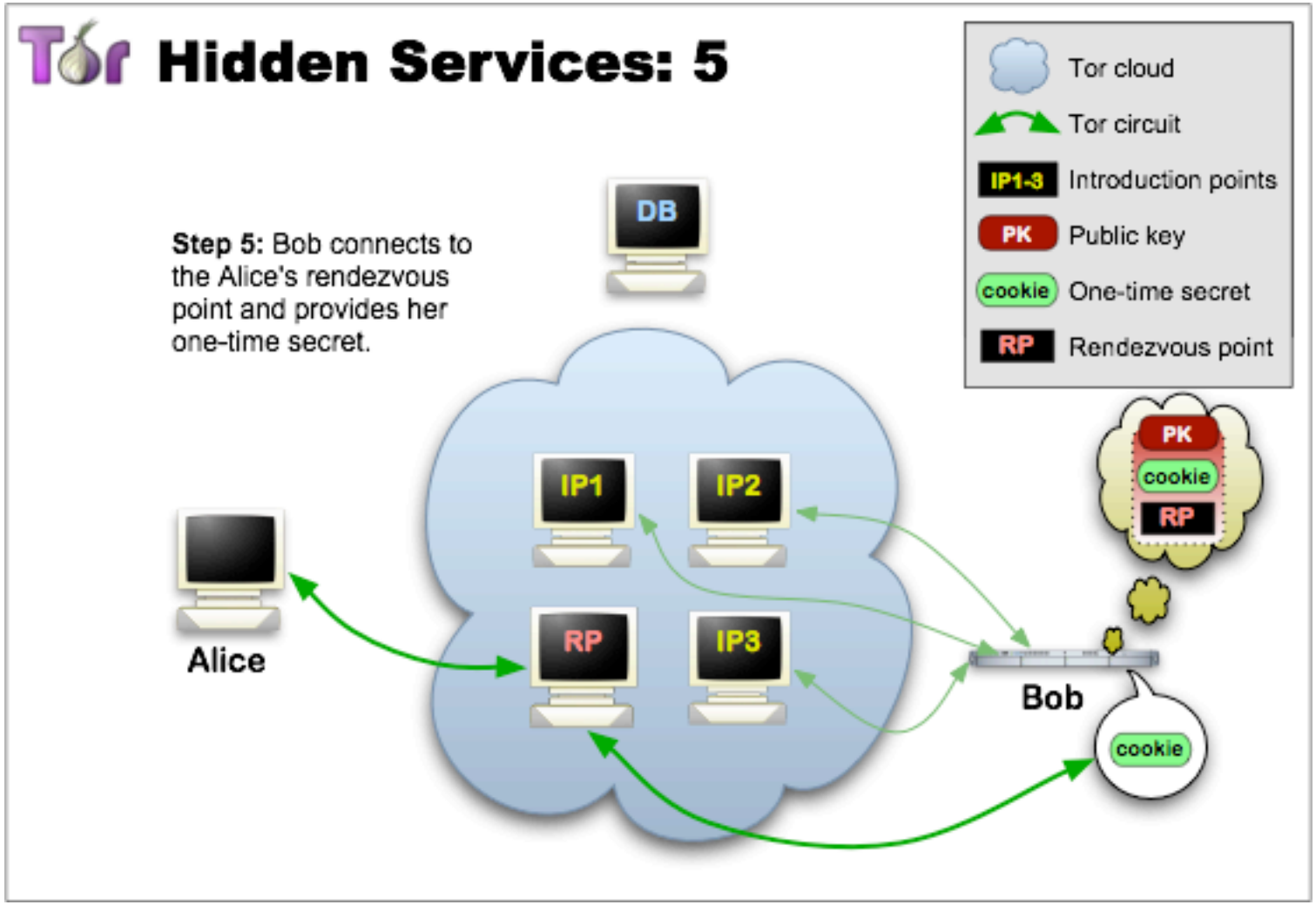
Source: torproject.org

A primer on Tor's Hidden Services, cont.



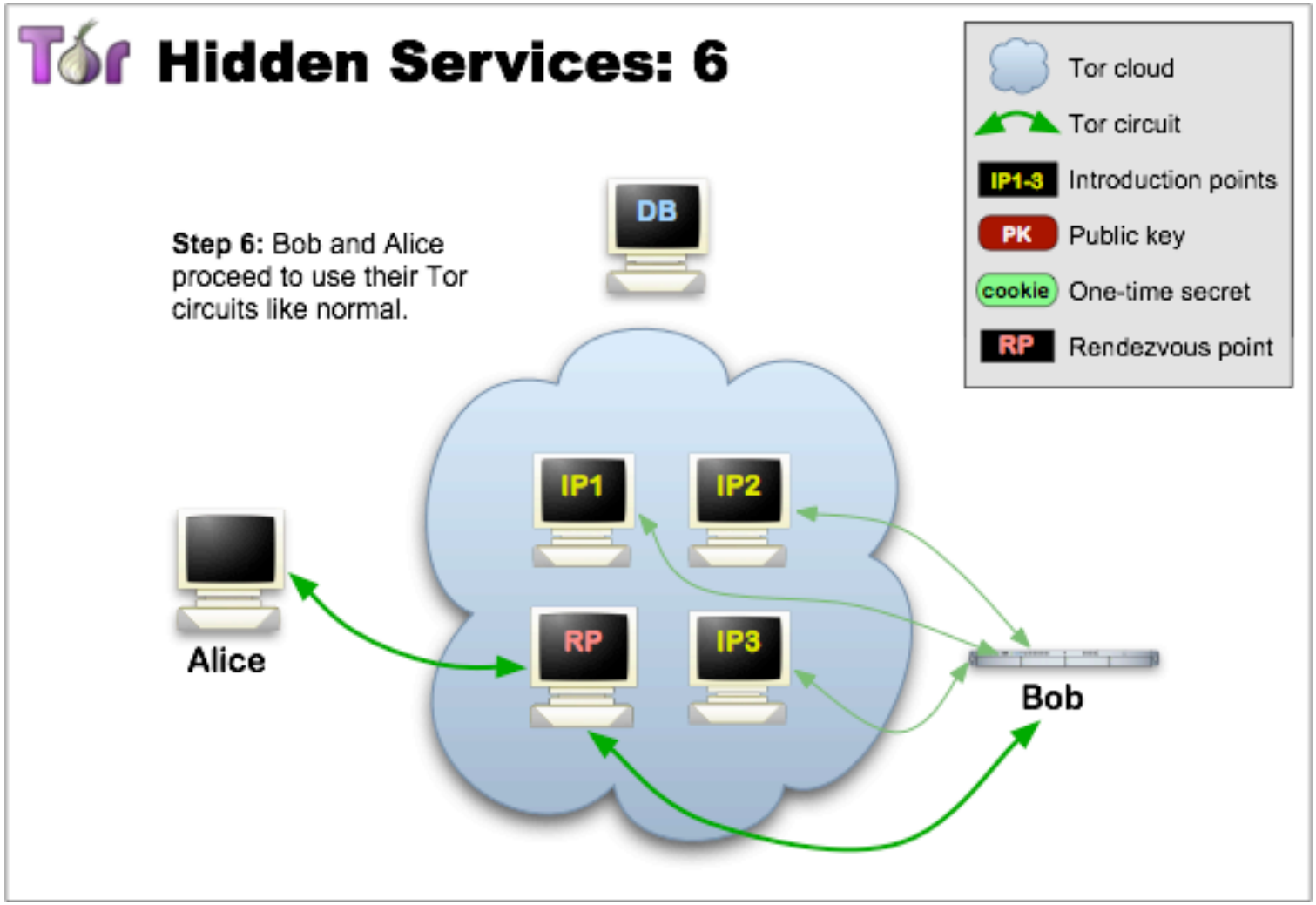
Source: torproject.org

A primer on Tor's Hidden Services, cont.



Source: torproject.org

A primer on Tor's Hidden Services, cont.



Source: torproject.org

Tor Leaks DNS Queries…

WIRED GEAR SCIENCE ENTERTAINMENT BUSINESS SECURITY DESIGN OPINION MAGAZINE

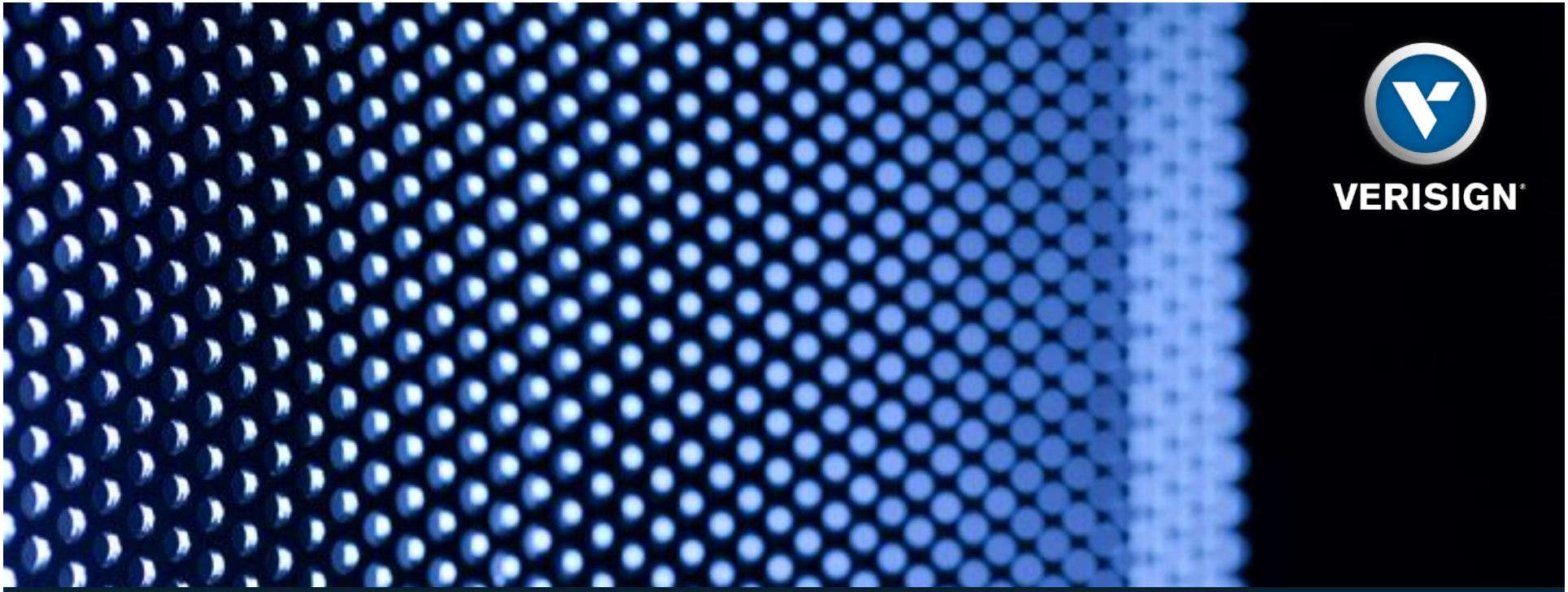
THREAT LEVEL

The Onion Router (TOR) is Leaky (Leeky)

BY RYAN SINGEL 10.19.06 | 4:34 PM | PERMALINK

Share 0 Tweet 0 g+1 0 in Share Pin it

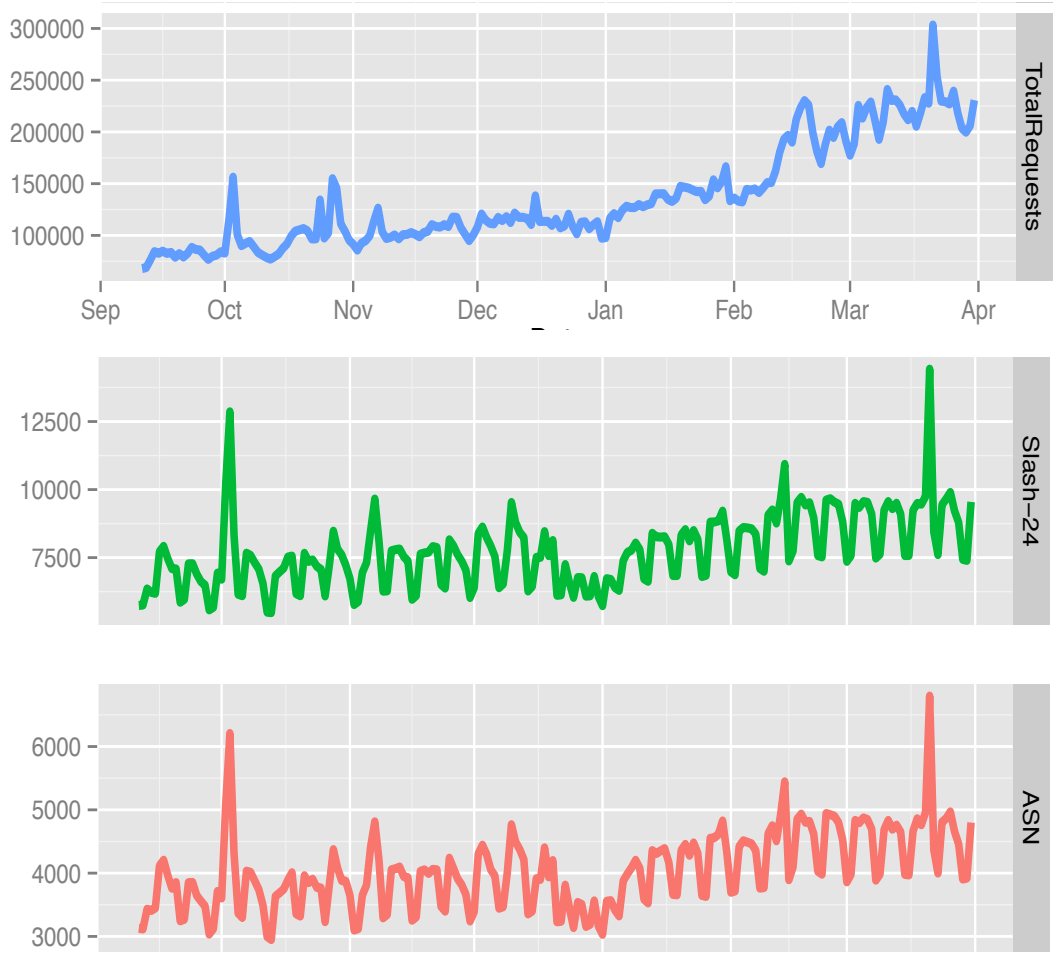
Source: wired.com



.Onion Measurements From A and J Root

General Overview

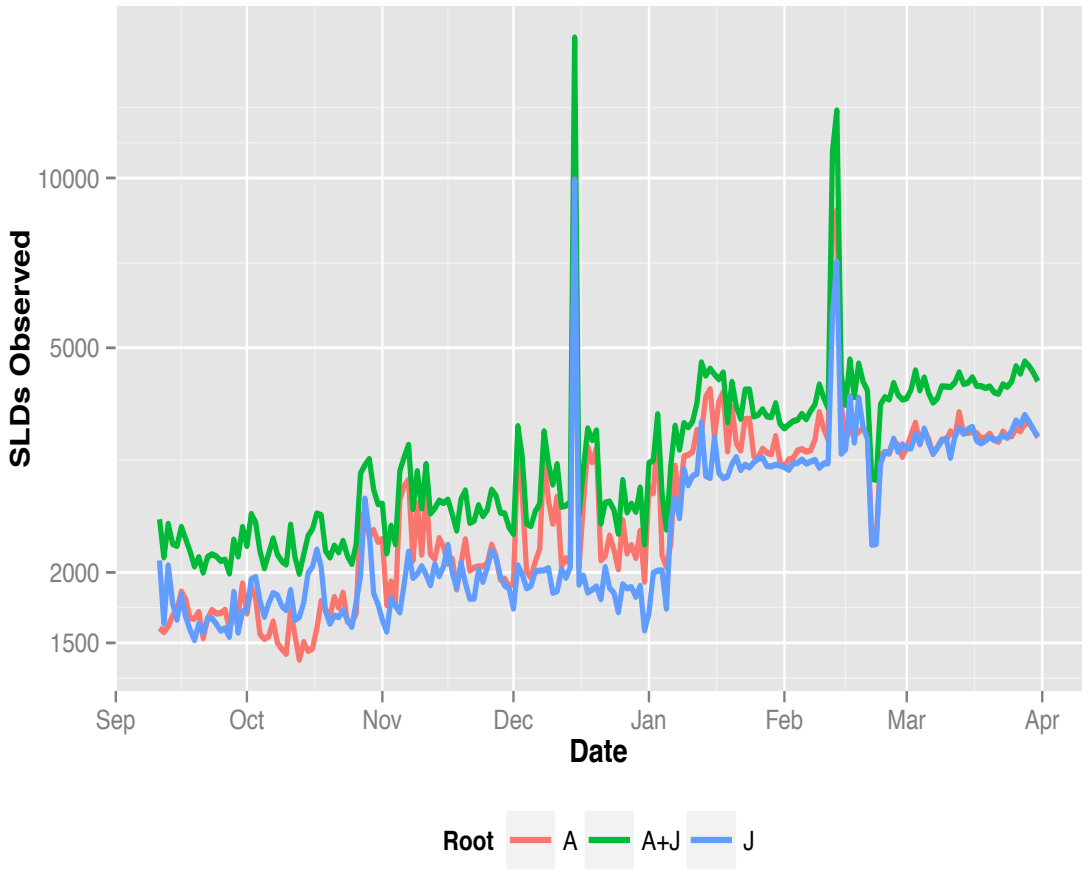
A Longitudinal Study of .Onion Traffic



- Six month capture from A & J Roots
  - *27.6M requests*
  - 81K SLDs
  - 172K IPs
  - *105K /24s*
  - *21K ASNs*

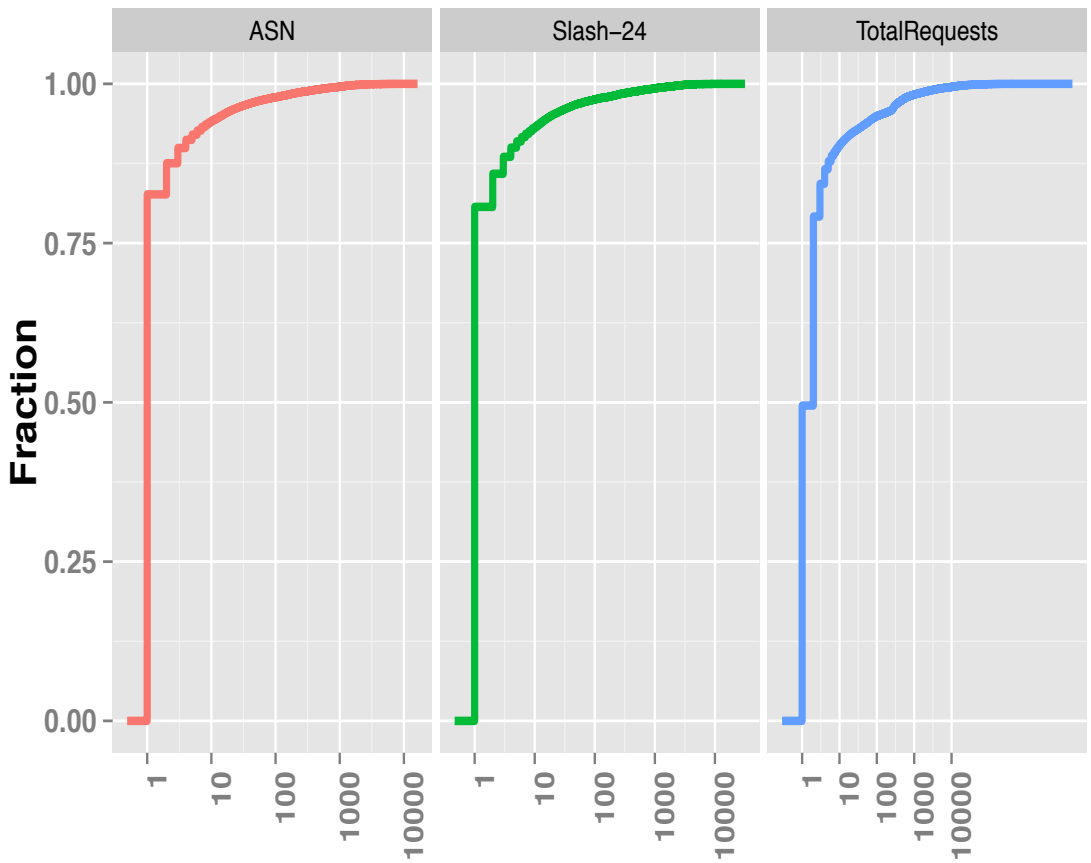
- Onion ranked 461

Root Sampling Completeness (Representative?)

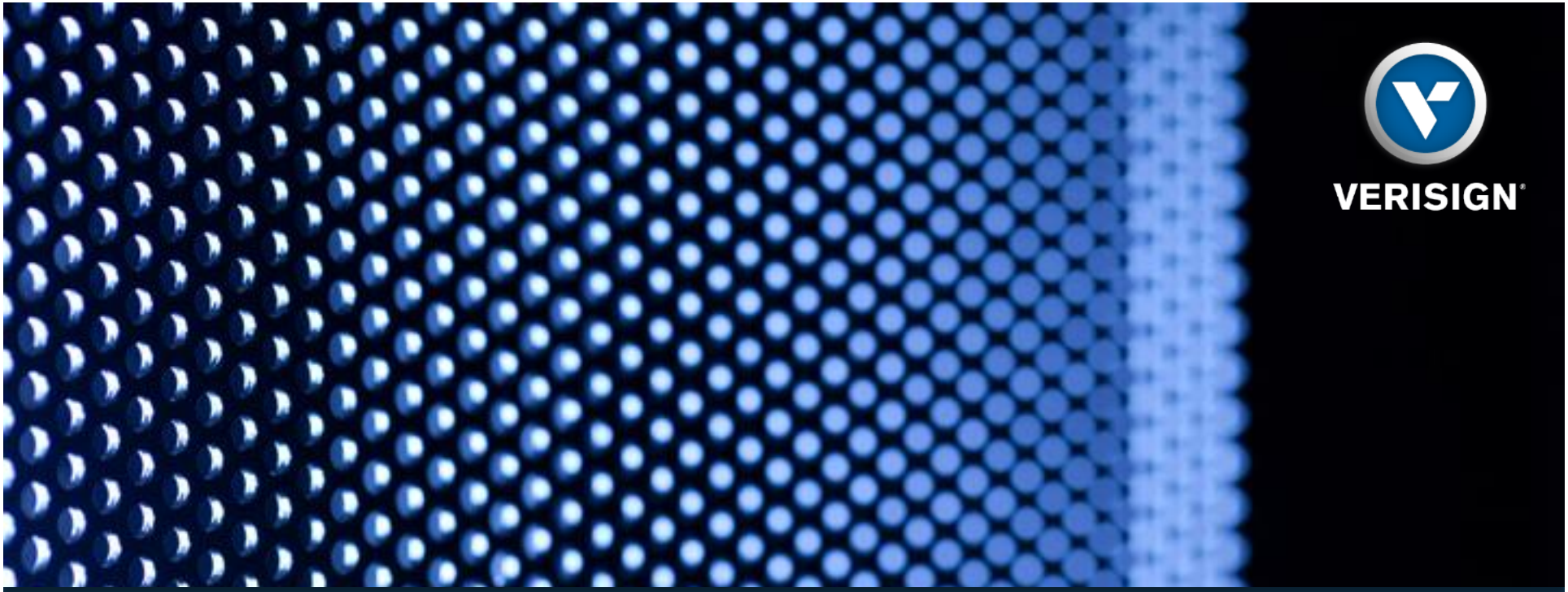


- A+J observe ~3300 SLDs/day
- Separately ~2500
  - 75% combined
- Prior SLD-root affinity reports suggests A&J observe 20% of total root traffic; confirmed by our study

Volume and Diversity of Hidden Service Requests



- Various measures of traffic to SLD distribution
  - 90% <= 10 requests
  - 95% <= 10 ASNs
- Very few SLDs with large and diverse traffic pattern



.Onion Measurements From A and J Root

Popular Services and Requesters

Most Requested Hidden Services

- Total of 81k SLDs!
  - Trackers
    - P2P systems
    - Tor Directory
    - Search Engines
    - Tor-related, etc
  - Deep web:
    - Silk road
    - Agora marketplace
    - (bitcoin)

Rank	Anonymized SLD	Type of Service	Traffic (%)
1	Z6---------43	Hidden Tracker	26.5
2	DK---------II	Silk Road	2.1
3	DP---------PC	TorDir	1.7
4	SI---------FK	Silk Road	1.4
5	3G---------4M	Search Engine	1.3
6	JH---------JX	Tor Mail	1.2
7	XM---------SL	Search Engine	1.1
8	AG---------WW	Agora Marketplace	1.1
9	FO---------UI	Bitcoin	0.9
10	TO---------NS	TorLinks	0.9

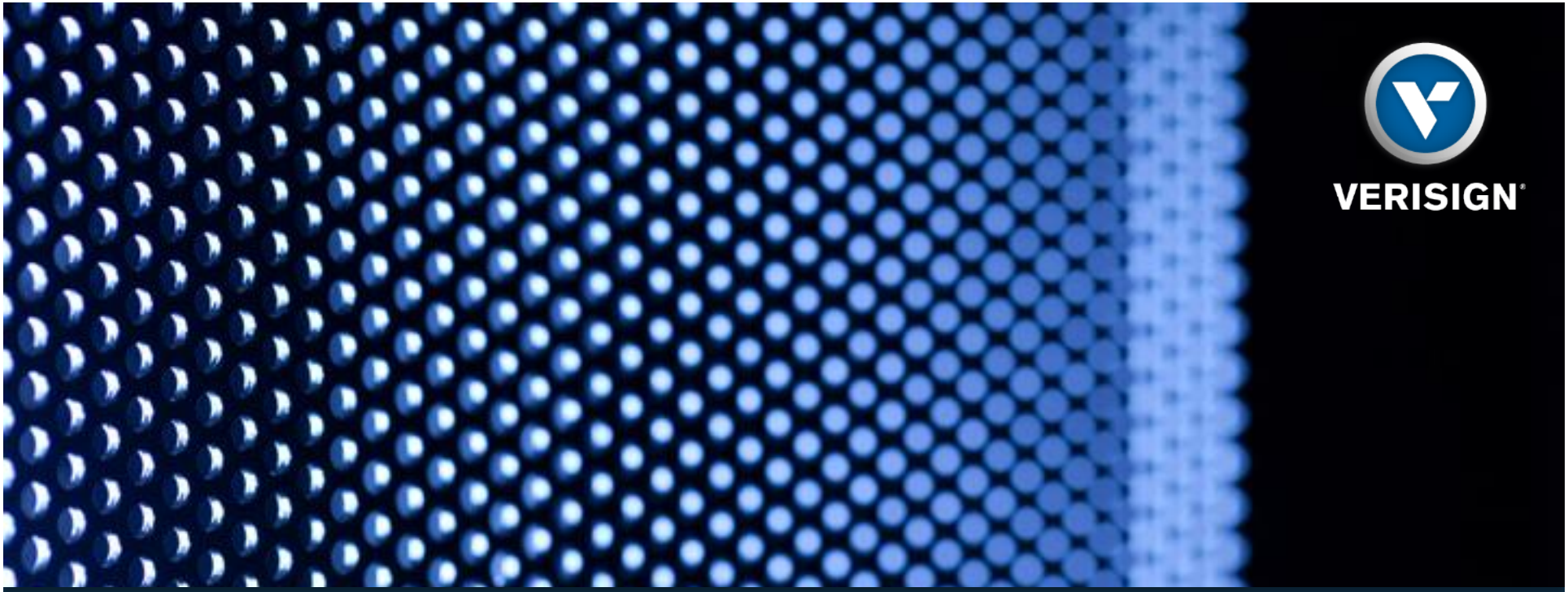
- 26.5% of all .Onion traffic to one Hidden Service
- Long tail distribution over remaining Hidden Services
- Top 10 SLDs account for 38% of all .Onion traffic

ASN + Geo Diversity of Hidden Service Requests

Country Code	Requests	% Traffic
US	9878093	35.7
RU	2213691	8.0
DE	1482075	5.3
BR	1258468	4.5
CN	996130	3.6
GB	984059	3.5
KR	980656	3.5
PL	918948	3.3
CA	785184	2.8
FR	670103	2.4
AU	510745	1.8
NL	454441	1.6
ES	448171	1.6
IE	425469	1.5
IT	423550	1.5
AR	387594	1.4
MX	363389	1.3
IN	295122	1.0

Autonomous System	Requests	%Traffic
AS15169	2267250	8.2
AS7922	1222955	4.4
AS7018	654680	2.3
AS36692	571609	2.0
AS30607	561349	2.0
AS4766	560739	2.0
AS701	512989	1.8
AS7132	447528	1.6
AS22773	400657	1.4
AS6830	392233	1.4
AS20115	342716	1.2
AS3786	326885	1.1
AS28573	309751	1.1
AS5617	290577	1.0
AS3356	290160	1.0
AS7738	284726	1.0
AS22773	273845	0.9
AS4134	258832	0.9

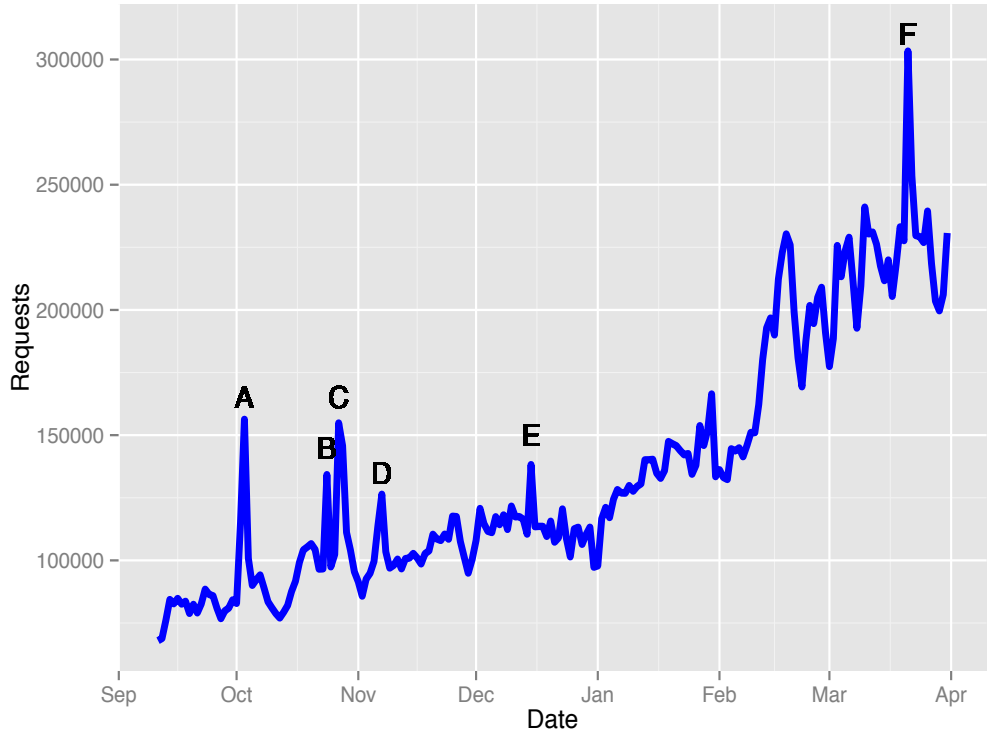
- Geo distribution of requests differs from that reported by Tor
  - USA (↑13.4). Germany (↓8.8), France (↓6.2) and Spain (↓4.4).
- Large percentage of traffic issued from public DNS services
  - Google (AS15169), OpenDNS (AS36692)



.Onion Measurements From A and J Root

Event Correlation

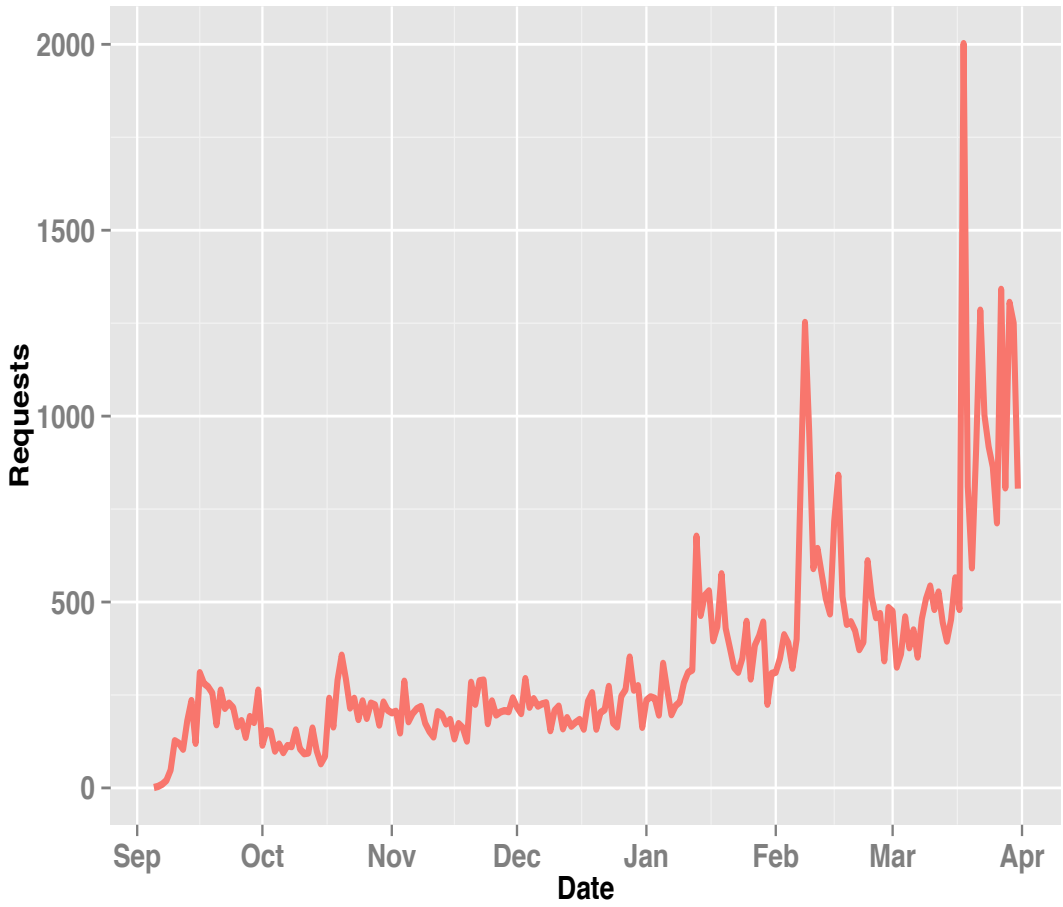
Tor and the World: Event Correlation



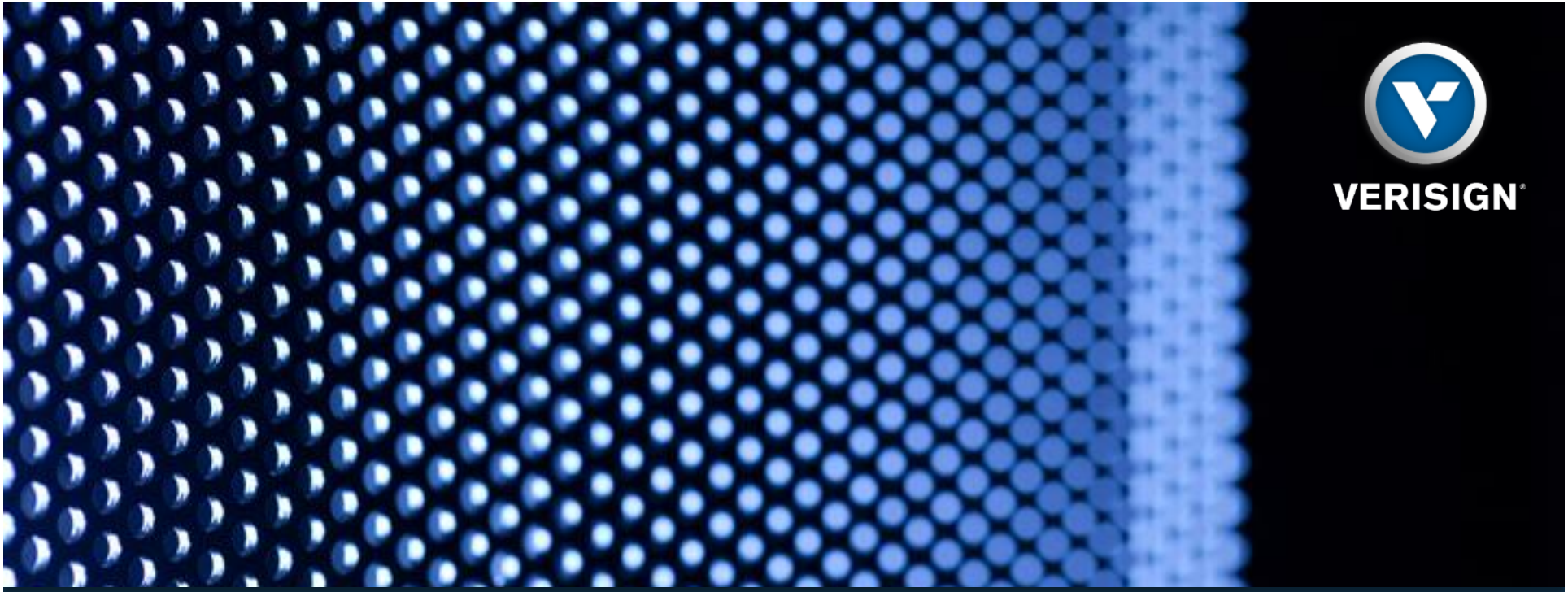
Event	Date	Requests	Event
A	10/03/13	156312	Silk Road Shutdown [13]
B	10/24/13	134236	TorATM Traffic Spike [14]
C	10/27/13	154855	URL Posted on Reddit [15]
D	11/07/13	126398	New Silk Road URL [16]
E	12/15/13	138231	Pirate Bay URL Posted [17]
F	03/21/14	303347	Multiple URLs Posted on Reddit [18]

- Several spikes in Onion Traffic can be correlated to specific hidden services and reported events.
- Many spikes coincide with postings of Onion URLs on popular websites

Tor and the World: Event Correlation, cont.



- Reported events within Turkey during elections
- Measure onion requests from Turkish IP addresses



.Onion Measurements From DITL

An Overview

DITL Dataset

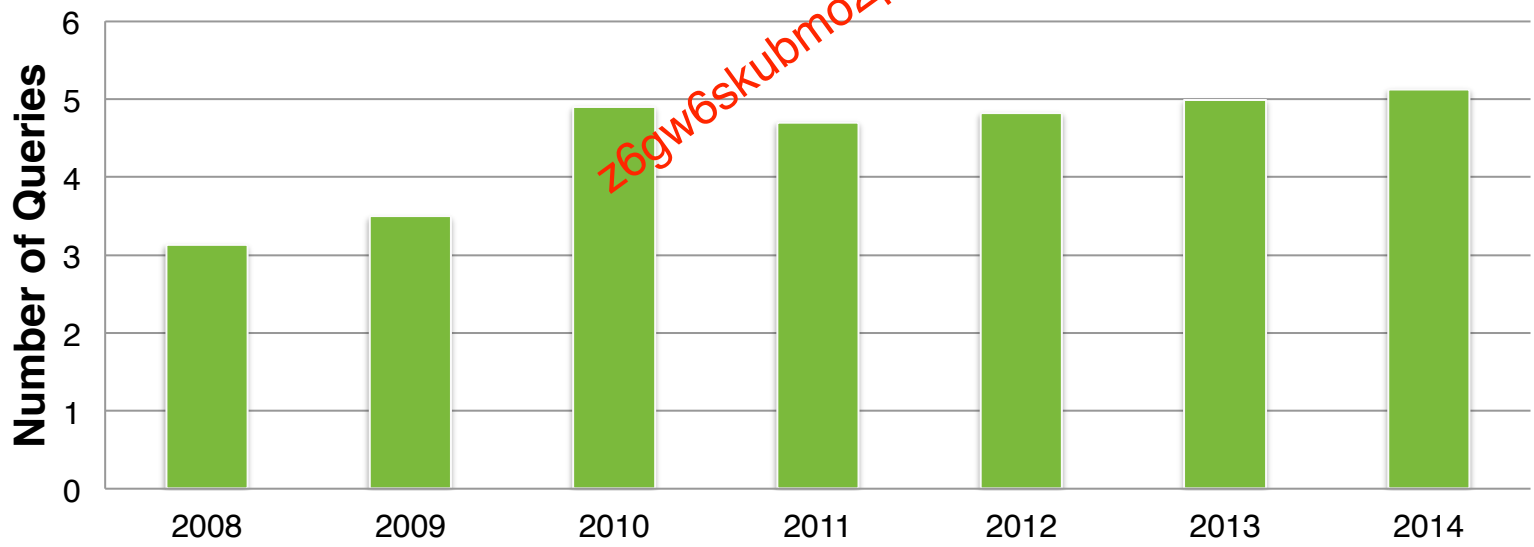
- DITL provides archival data
  - Simultaneous measurement effort from roots and name servers; two days a year; data managed by DNS-OARC.
  - Covers 7 years, from 2008 to 2014 (1-2 days per year)
  - From all root servers ('A' through 'M'); 6,850,728 .onion queries
  - Originated from 5,324,412 IP address, over 336,273 /24 addresses
  - Queried 18,330 .onion SLDs the total of 7 years

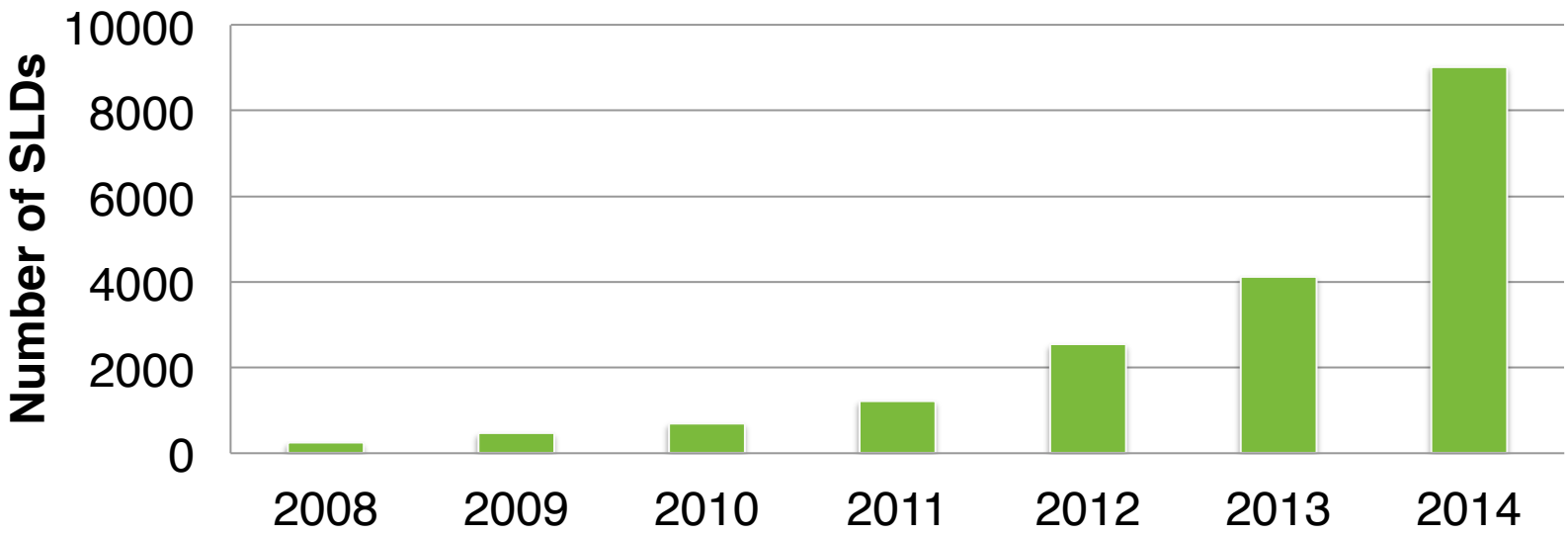
Year	# roots	Root servers	Total queries
2008	7	(A,C,F,H,K,L,M)	3,710
2009	8	(A,C,E,F,H,K,L,M)	13,343
2010	13	ALL	2,371,869
2011	11	All except B and G	691,385
2012	10	All except B, D, and G	693,524
2013	11	All except B and G	1,371,650
2014	9	All except B, D, G, and L	1,705,247

DITL Dataset, cont.

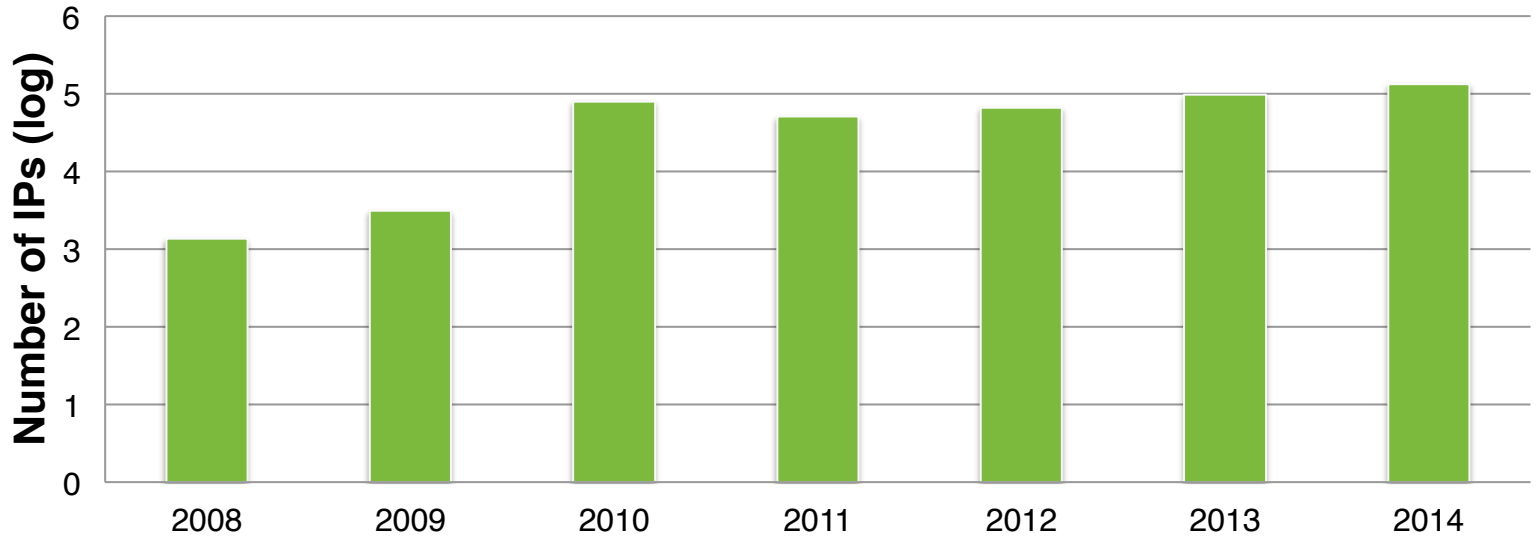
Root	Organization	# queries	# years	Traffic (%)
A-root	Verisign	515,107	7	7.52
B-root	USC-ISI	97,119	1	1.42
C-root	Cogent	723,152	7	10.56
D-root	UMD	205,403	3	3.0
E-root	NASA	151,014	6	2.2
F-root	Internet Sys	763,663	7	11.15
G-root	Defense Info Sys	72,232	1	1.05
H-root	US Army	360,490	7	5.26
I-root	Netnod	975,579	5	14.24
J-root	Verisign	842,361	5	12.3
K-root	RIPE	733,951	7	10.71
L-root	ICANN	649,648	6	9.48
M-root	WIDE	761,009	7	11.11

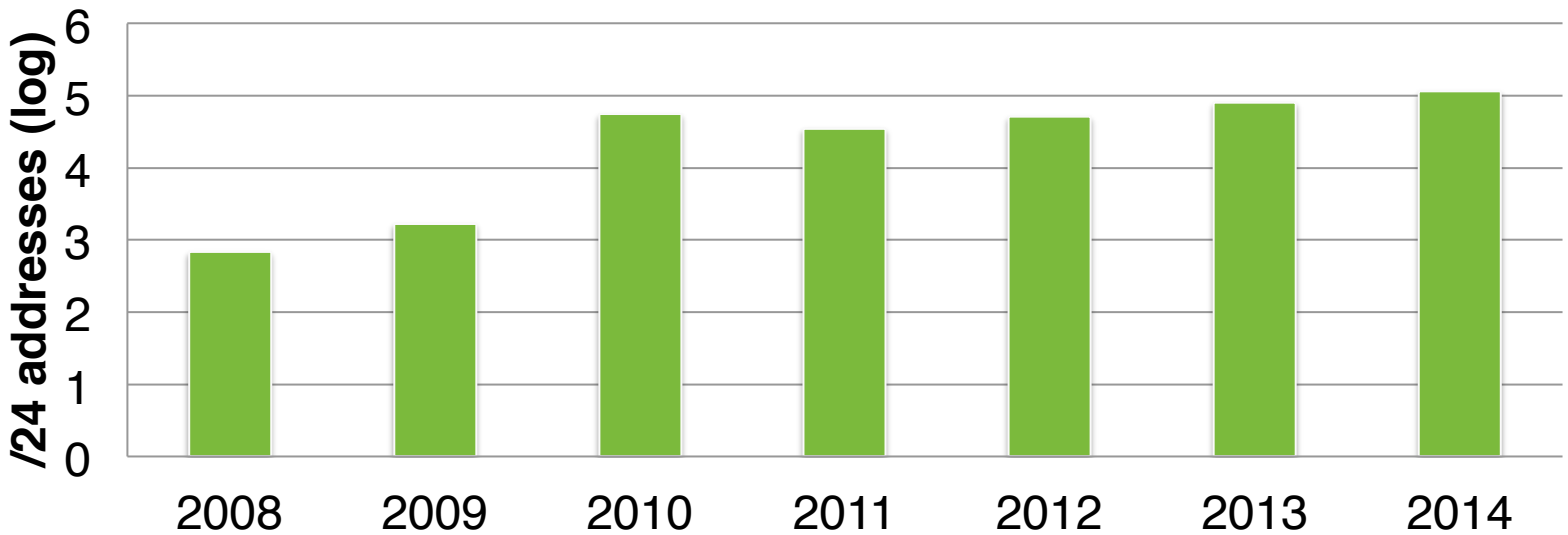
DITL Dataset, cont.





DITL Dataset, cont.





Potential Causes of Leakage and Remedies

- “Ignorance of the crowd”
  - Hypothetical scenarios support that from analogous contexts.
- Search list processing
- Browser prefetching
  - A problem with other collision-related incidents.
- Malware: Chewbacca, 64-bit variants of Zeus, etc
  - As suggested in many reports.
- Bundle misconfiguration
- Potential remedies :
  - Enabling blocking at the stub and recursive resolvers.
  - Automated configuration.
  - User notification for further actions.

Potential Implications

- The potential implications depend on who is querying
  - *Individual user IP:* most severe; clear identification and potential privacy threat to the individual users.
  - *Recursive DNS resolvers:* outbound queries are aggregated, and less threat to privacy. Incentives may prevent recursive from sharing such information.
    - *ISP resolver:* outbound queries are aggregated. Incentives may guard user privacy, even when ISP sees individual user IP.
    - *Open resolvers:* outbound queries are aggregated, but some threat to privacy. Open resolver's incentives are unclear.
    - DPRIVE (DNS PRIVate Exchange) and privacy enhanced resolution are two potential ways to address an observer and remedy risk

Concluding Remarks and Future Work

- We measured a sample of .Onion DNS requests to A+J
  - Examined unique characteristics of these requests longitudinally
    - Network and Geographical
  - Increased traffic spikes correlated with specific events
    - URL postings, Censorship
- Certain causes of leaked DNS queries remains unknown
  - Misconfiguration, search lists, typos, poor user understanding, etc.
- We plan to continue the examination of the leaked queries
  - Other non-delegated privacy TLDs (i2p, exit, etc).
  - Malware (by name/family)

powered by

VERISIGN™