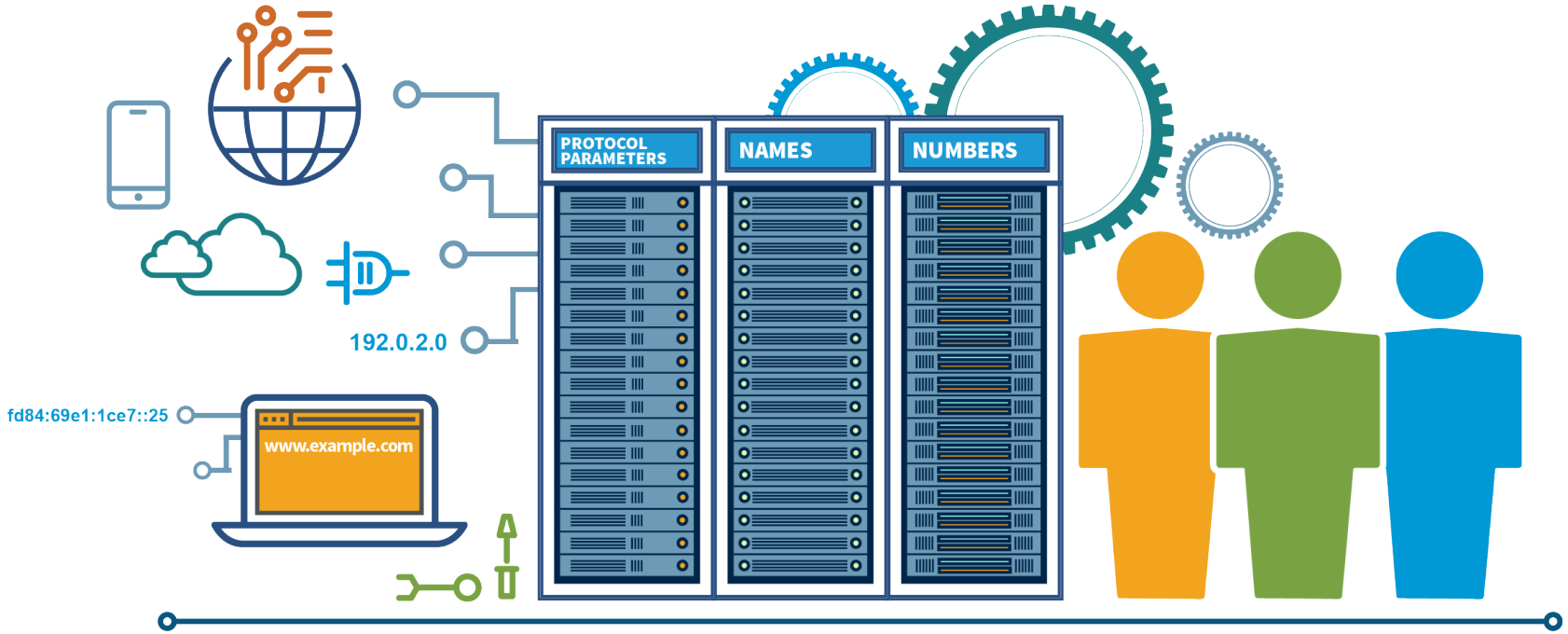
Identifier Technology Health Indicators (ITHI)

Yaovi Atohoun
Stakeholder Engagement & Operations Manager, Africa

UbuntuNet Connect 2018

19-20, November 2018; Zanzibar, Tanzania

Coordinating with our partners, we help make the Internet work.



What Are Internet Identifiers?

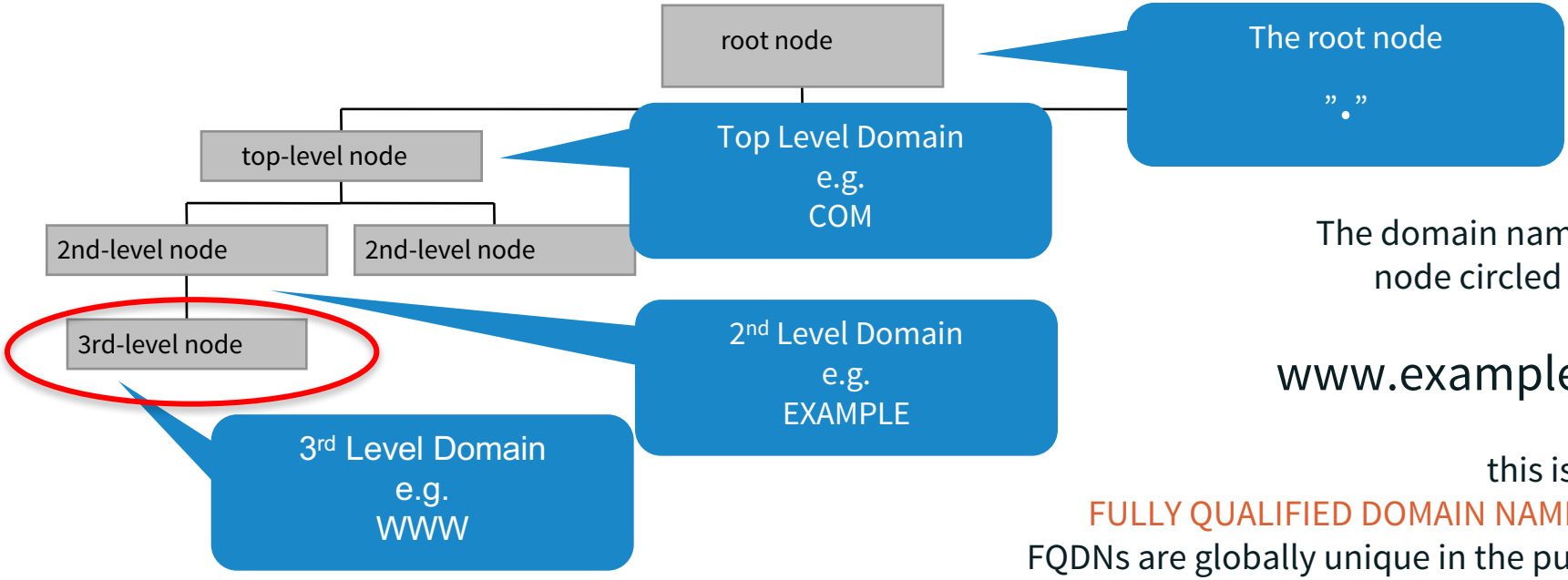
The Internet is a mesh of networks whose operators agree to communicate using predefined protocols (“TCP/IP”)

Networks use identifiers to name or number individual computers (hosts) so that these can communicate

- Medium Access Component (MAC) addresses identify the Internet’s doors or PO box numbers
- IP addresses identify the Internet’s streets house numbers
- Autonomous System Numbers identify the Internet’s neighborhoods
- Domain Names identify Internet hosts and services using a famliar language rather than IP addresses

Definitions: Labels and Domain Names

Each node in the DNS name space has a label
The domain name of a node is the list of the labels on the path from the node to the root of the DNS



The domain name for the node circled in RED is

www.example.com.

this is called a
FULLY QUALIFIED DOMAIN NAME (FQDN)
FQDNs are globally unique in the public DNS

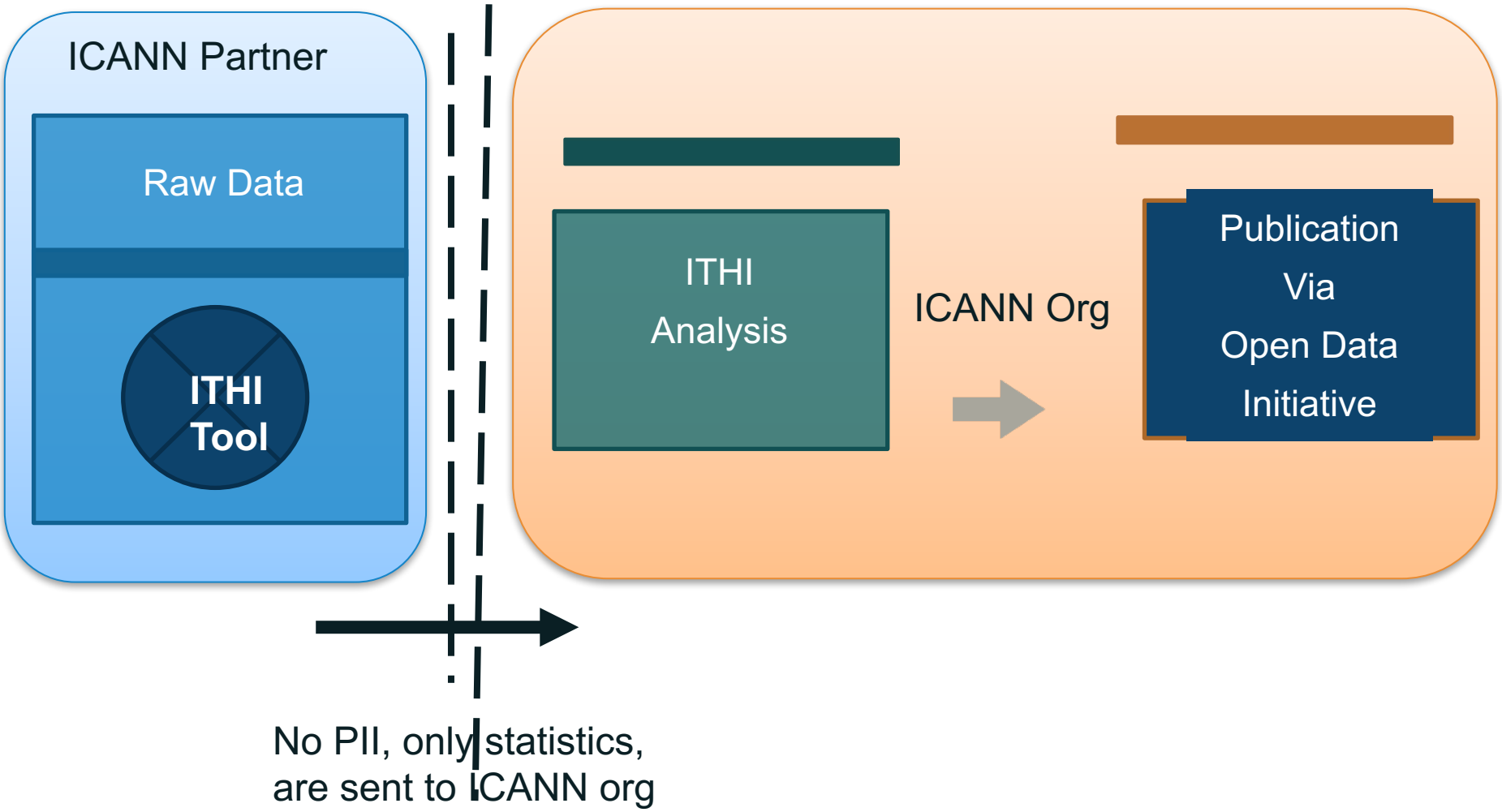
Operational elements of the DNS

- Authoritative Name Servers host zone data
  - The set of “DNS data” that the registrant publishes
- Recursive Name Resolvers (“resolvers”)
  - Systems that find answers to queries for DNS data
  - Caching resolvers find and store answers locally for “TTL” period of time
- Client or “stub” resolvers
  - Software in applications, mobile apps or operating systems that query the DNS and process responses

ITHI Principles of Operation

- Technical focus
- Problem areas → Metrics → Measurement
- Current value and trend over time
  - Automated process to collect & analyse data
- Measurement, not interpretation
- Extraction of statistics to avoid data privacy issues
- Open source tools & results

ITHI: Process



8 Metrics and Data Sources

Metric	Name	Data Source
M1:	Inaccuracy of Whois Data	ICANN compliance dept.
M2:	Domain Name Abuse	ICANN's DAAR Project https://www.icann.org/octo-ssr/daar
M3:	DNS Root Traffic Analysis	Samples of DNS root traffic
M4:	DNS Recursive Server Analysis	Summaries of recursive resolvers traffic
M5:	DNS Resolver Behavior	APNIC
M6:	IANA registries for DNS parameters	Scan of recursive resolvers traffic
M7:	DNSSEC Deployment	Snapshots of DNS root zone
M8:	DNS TLD Traffic Analysis	Summaries of TLD traffic

Partners for measuring the DNS

- Existing Partners:
  - National University of La Plata (UNLP), Argentina,
  - University of Cape Coast, Ghana,
  - DNS Nawala, Indonesia, and
  - Kaznic, Kazakhstan (.KZ)
- Recruiting more partners:
  - Recursive resolvers
  - Authoritative servers

New: Software updates

- Open Source Software
  - Windows, Linux, Free BSD
- Packages available for Centos, Ubuntu
- Audit by NLLabs

Identifier Health Dashboard

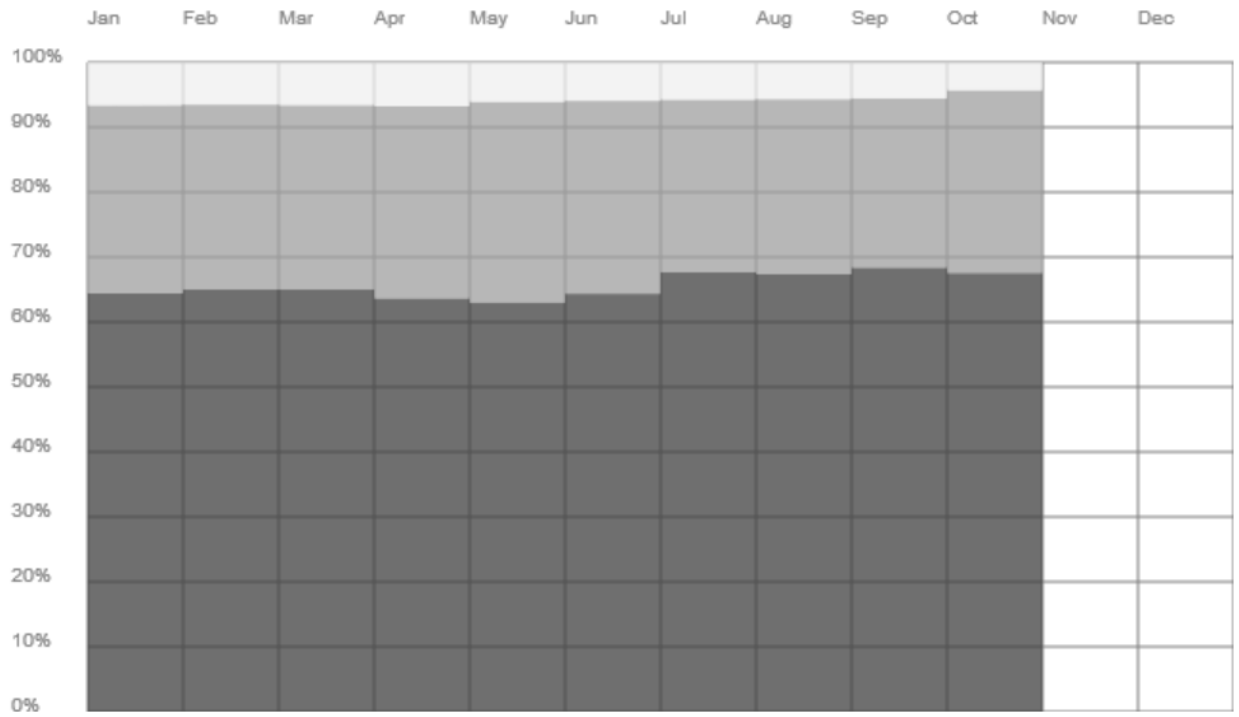
- Next slides show the proposed Health Dashboard
- Select specific data from rich metrics
- Detailed results available at https://ithi.research.icann.org/

Identifier Dashboard: Four Specific Health Indicators

After 9 month of data collection, we can propose four key indicators:

1) Root Traffic Characterization
2) Resolver Concentration
3) DNS SEC deployment
4) Name Leakage

1) Root Traffic Characterization



Metric	Current Value	Average Value
M3.1 (% No Such Domain queries)	68%	65%
M3.2 (% cacheable queries)	28%	28%
Core (100% - M3.1 - M3.2)	4%	6%

2) Resolver Concentration (proposal)

- Motivated by concerns of current and potential future concentration of the DNS Resolver market
- One data-point: July 2018
  - Resolvers aggregated by AS number
  - Top resolver market share was 13%
  - 25 resolvers account for 50% of eyeballs
  - 460 resolvers to account for 90% of eyeballs
  - 28,000 resolvers seen in the eyeball study
- Three Data sources to be combined:
  - Root (M3), TLDs (M8), APNIC (M5)
  - Will require some development

3) DNSSEC Deployment (Jul 2018)

Origin	Metric definition	Value
M7.1	%TLD signed	**91%**
M7.2	%CCTLD signed	**51%**
M3.5	% of resolvers that set DO bit in queries to the root	**82%**
M5.3.2	% of resolvers that set the DO bit in queries to APNIC test	**84%**
M5.3.1	% of users using resolvers that set the DO bit in queries	**92%**
M4.5	% of stub clients setting the DO bit in queries to recursive	**1%**
M5.4.1	% of users using resolvers that perform DNSSEC validation	**51%**
M5.4.2	% of resolvers that perform DNSSEC validation	**25%**

- *Recursive resolvers appear ready (DO bit)*
  - *But this may be due to software default setup*
  - *Only 25% actually perform DNSSEC validation*
- *Stub resolvers (clients) rely on recursive resolvers for validation*

4) Name Leakage: Root and Recursives

Origin	Value	% Leaks
M3.3.1	LOCAL	6.0%
M3.3.2	HOME	4.8%
M3.3.2	IP	1.1%
M3.3.2	INTERNAL	1.1%
M3.3.2	LAN	0.8%
M3.3.2	LOCALDOMAIN	0.6%
M3.3.2	DHCP HOST	0.5%
M3.3.1	INVALID	0.5%
M3.3.2	DHCP	0.5%
M3.3.1	LOCALHOST	0.5%
M3.3.2	CORP	0.3%
M3.3.2	DLINK	0.3%
M3.3.2	GATEWAY	0.3%
M3.3.2	DLINKROUTER	0.3%
M3.3.2	BELKIN	0.2%
M3.3.2	HOMESTATION	0.2%
M3.3.2	OPENSTACKLOCAL	0.2%
	Other names	81.8%

Origin	Value	% Leaks
M4.2	LOCAL	1.5 %
M4.3	UNIFI	1.5 %
M4.3	HOME	1.4 %
M4.3	TOTOLINK	0.7 %
M4.2	LOCALHOST	0.5 %
M4.3	LAN	0.4 %
M4.3	LOCALDOMAIN	0.3 %
M4.2	INVALID	0.01%
	Other names	94 %

In theory, root and recursive should be similar.
In practice, they are not:

- *Local effects dominate recursive traffic*
- *Small number of recursive resolvers in study*
- *No corporate resolvers in study*

Large fraction of leaks is not explained by "frequent names":

- *Host names of local machines (recursive),*
- *Wi-Fi routers (recursive, root),*
- *Made up names (root traffic).*

.MAIL is found, but way down the list

Contact us!

- Alain Durand <alain.durand@icann.org>
- Christian Huitema <huitema@huitema.net>

- https://ithi.research.icann.org

One World, One Internet

Visit us at **icann.org**

@icann

linkedin/company/icann

facebook.com/icannorg

slideshare/icannpresentations

youtube.com/icannnews

soundcloud/icann

flickr.com/icann

instagram.com/icannorg