VERISIGN®

Moving the Needle: One View on the Search for Impactful Research Projects

Eric Osterweil

It takes all kinds…

- As the title suggests, this talk is mostly just a look at my own view of research and ways in which it can be impactful
- My guess is that we *all* want our work to make an impact, and (at best) this talk is really just a ``for instance''
- My view is informed by where I sit:
  I'm a principal scientist at VeriSign, Inc., in the CSO office
  - I do a lot of DNS research (and some other stuff too)
  - I'll give Verisign's spiel in a bit

Level Set… Mr. Toad's Wild Ride

- I've stumbled into some interesting things, so I thought I'd share
- That said, it's a weird feeling to presume anyone wants to hear me gab about myself
- So, how about this: I'll talk about two different angles:
  - Things that have worked out
  - Things that haven't exactly gone 100% to plan?
- Hopefully this will at least entertain…

Grad School: Not for the Faint of Heart

- We all know how tough it can be, how many of us have ever asked ourselves,
  ``is it worth it?''
  - Obviously, the answer to that question is, necessarily, very personal
- The deep thinking we learn in grad school can be very synergistic with broad objectives in industrial labs
- One thing that I found helpful was focusing on the work that *I* found rewarding (obviously)
  - Many might say DNS is a simple / stale (boring?) protocol that (today) just sits in the background of mundane operations
  - I (for one) am at peace with that ☺

Outline

- A little about my background
- Getting settled at Verisign
- A look at the impact of some recent work
- High-order bits

VERISIGN

I started off in the industry

- After I finished my undergraduate work I was *done* with school
- I went off and got a job that paid me ``real'' *money*
- It was great, someone told me what to do, and I did it, and then I spent my money
  - This kind of left me wondering why anyone would want to be in industry

Going from industry to academia

- I recall thinking how exhausting it could be to try to explain important ideas to managers
- Sometimes, I was sure I was feeding them good ideas
- Though, sometimes the results were puzzling
- Luckily, I had a lot of and they sold me on

Working within corporate strategies isn't always bad

- True, in industry, I did feel confounded by seagulls
- I often wondered if I should have stayed in school

- BUT, there can be reason in madness with initiatives that are larger than one person's agenda(s)
- Indeed, I have *constantly* felt the rewards of my early industry experience

The looooooooong road (my thesis)

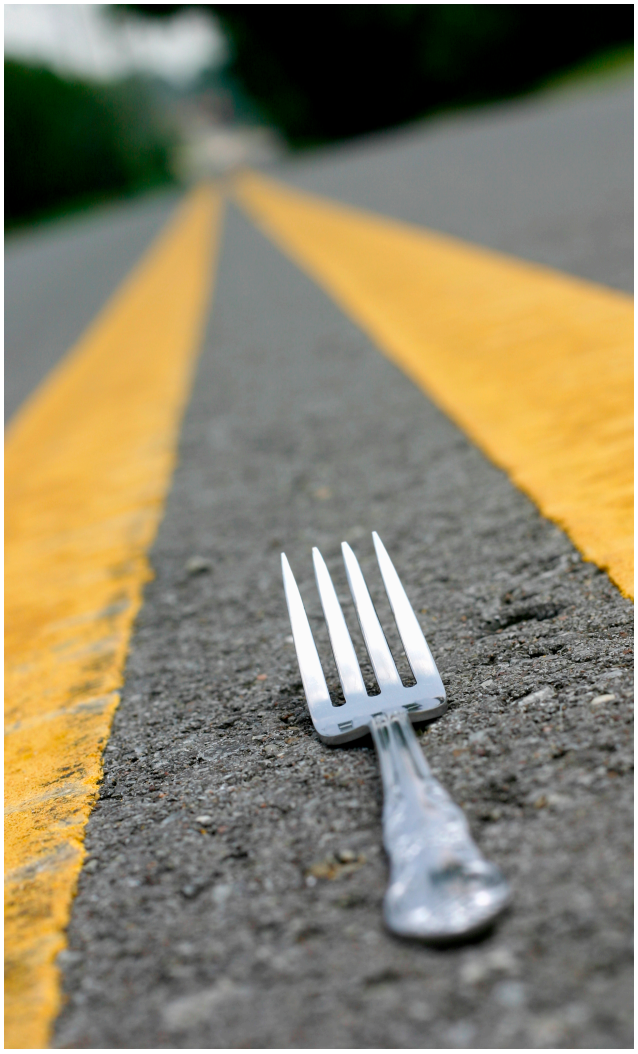
- I went to UCLA, and worked in Lixia' Zhang's lab
- DNSSEC newly standardized and perfectly… Boring
  - I will never forget words from my advisor: ``… monitoring and debugging is a detailed and tedious thing, but I believe there is some deep science one can find in the process…'' – Lixia Zhang '05
- So, I started a little Perl script, called SecSpider
  - It was quite a feat of engineering
  - It tracked something like 8 whole DNS zones!
- I wasn't *sure* it was going to ever go anywhere
  - I'm sure that I'm the only one who ever doubted their thesis early on

VERISIGN

Eventually, however, SecSpider framed my thesis

- SecSpider continued to grow and actually discover interesting things
  - We found PMTU problems with DNSSEC in ~2006
  - We discovered replay vulnerabilities
  - etc…
- This ‘06 work centered around non-crypto based distributed key learning and verification
  - Led to a formal model called ***Public Data***, and a distributed system called ***Vantages***
- We still spent a lot of our time discussing this in operational venues and at the IETF
  - Splitting your focus is exhausting, but…

In retrospect…



- Sometimes it's hard to appreciate formative events
- It can be a lot of work:
  - Arguing on mail lists
  - Debating operators
  - Trying to gain deployment traction
- Tough choices: deciding between operational/standards debates or publications
- Indeed, it can be hard to see a fork in the road for what it is
- That is, until you hit the job market

VERISIGN

Operational presentations got me my job

- Turns out, arguing on NANOG and IETF mailing list can get you noticed
  - DNSSEC's Deliberately Unvalidatable Root Zone (DURZ) key
  - The "goodness" of this impact may be debatable, but it caught my attention: research can change the world around you
- If you like those venues, there's a handful of places (and a ton of need) for people with our skills
  - This includes the evolving set of industry research labs
- True, it's not all a bed of roses, but there can be a palpable impact to your work

VERISIGN

Industry labs: ymmv

- Everything is a tradeoff, and looking at what defines research agendas at an industry lab can be very useful when deciding if it's for you
- For example, there can be a difference between research that informs operations, and informing operations with your research
  - Industrial research can be quite interesting, and while a lot of the primitives are the same as we learn in grad school, it has some palpable differences
- The real question is, what makes *your* heart race?

Onto Verisign

- I was very used to being ``data poor'' in grad school
  - I'm sure that no one could possibly relate to that
- The allure of Verisign's data *alone* was almost too much to resist
  - We run .com/.net, and serves .edu / .gov / 2 of the 13 DNS root instances, and a bunch of other TLDs
  - We are the Root Zone Maintainer (RZM)
  - We see tens of billions of transactions / day
  - We run a DDoS protection service (VDPS)
  - We have a threat intelligence team (iDefense)
- i.e., Candy Land for a researcher

Working at Verisign has been eye opening

- Using our observation space, our team has:
  - Built and run a DNS reflector attack detection/remediation system (called Kraken)
  - Inter-domain routing security
  - Worked on new techniques, standards, and alliances around information sharing and reputation
  - Studied the attack trends of server-based DDoS attacks
  - Safeguarded the evolution of the top of the DNS namespace
- What has that all meant?
  - We've been using our data and research to quantify things that are not easily accessible elsewhere
    and
    put results in the hands of those who can act on our findings
- For the remainder, I'll dive into the DNS namespace
  - I can be bribed to talk about any of the other topics too

VERISIGN

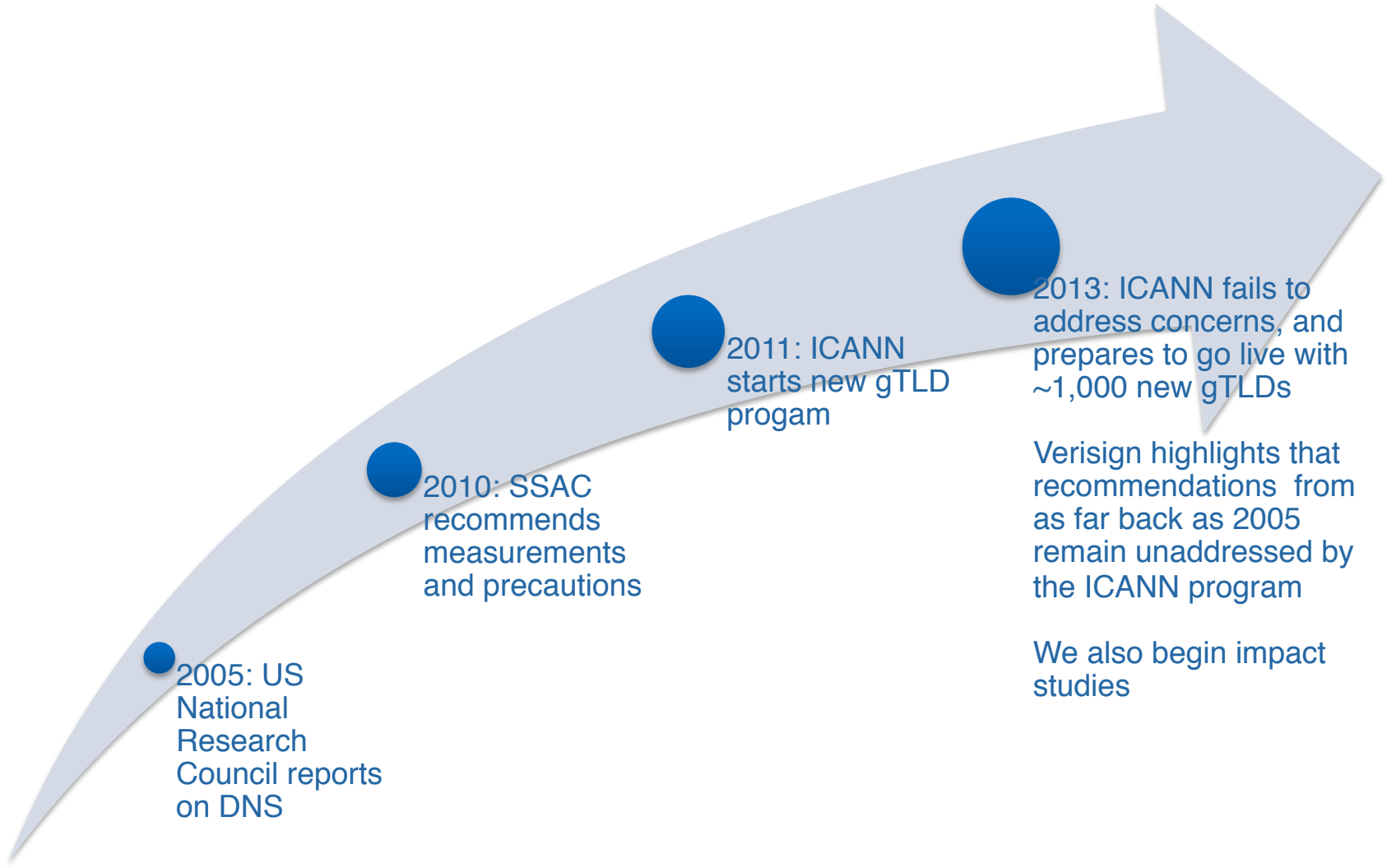
With great data comes great responsibility… or something like that

- Moving the needle can be a function of being in the right place at the right time
- Our company has been studying the new gTLD program (an ICANN operation)
- We're all familiar with .com/.net/.mobi/etc.
  - These are called generic Top Level Domains (gTLDs)
- ICANN has been planning to introduce over 1,000 new gTLDs in the space of ~1 year
- We've been asking the simple question, ``is this safe?"
  - Specifically, we performed some measurements

The new gTLD program: risk vs. reward

- The new gTLD program offers a lot of positive opportunities
  - This is why Verisign has applied for over a dozen strings and is contracted to act as a backend registry for nearly 200 others
- But, the security and stability of the DNS is serious, and risk-taking at the DNS root has global implications
- ICANN's community has fractured: do applied-for gTLD strings pose risks? "Yes," "no," "yes, but…"
- *Evidence* and *measurements* are critical in clarifying what the appropriate level of caution should be

The timeline of ICANN's new gTLD program



VERISIGN

Impact: finding the right medium for the right audience

- Finding the right cadence for our research involved assessing who our target audience is
  - We issued technical reports and public comments in response to comment periods and RFIs
  - These have turned out to be very impactful mediums
- We primarily began using Technical Reports as stable references in our public comments
  - Our TRs #1130007 and #1130008, and our public comment about .cba have had an undeniable impact on the new gTLD program
  - http://techreports.verisinglabs.com
- We recognized existing concerns that had previously been raised, and then studied those to illustrate *risk*

VERISIGN

Some of the things we found

- Previous advice has urged for the measurement of name collisions (and other issues)
- We found that DNS name collisions, DNS search-list interactions, and Internal Name Certificates enable
  - Information leakage
  - DoS attack vectors
  - And Man in the Middle (MitM) attacks *(even against TLS)*
- Part of this problem is kind of an issue of eminent domain
  - Many DNS deployments have become dependent on the *absence* of many gTLDs

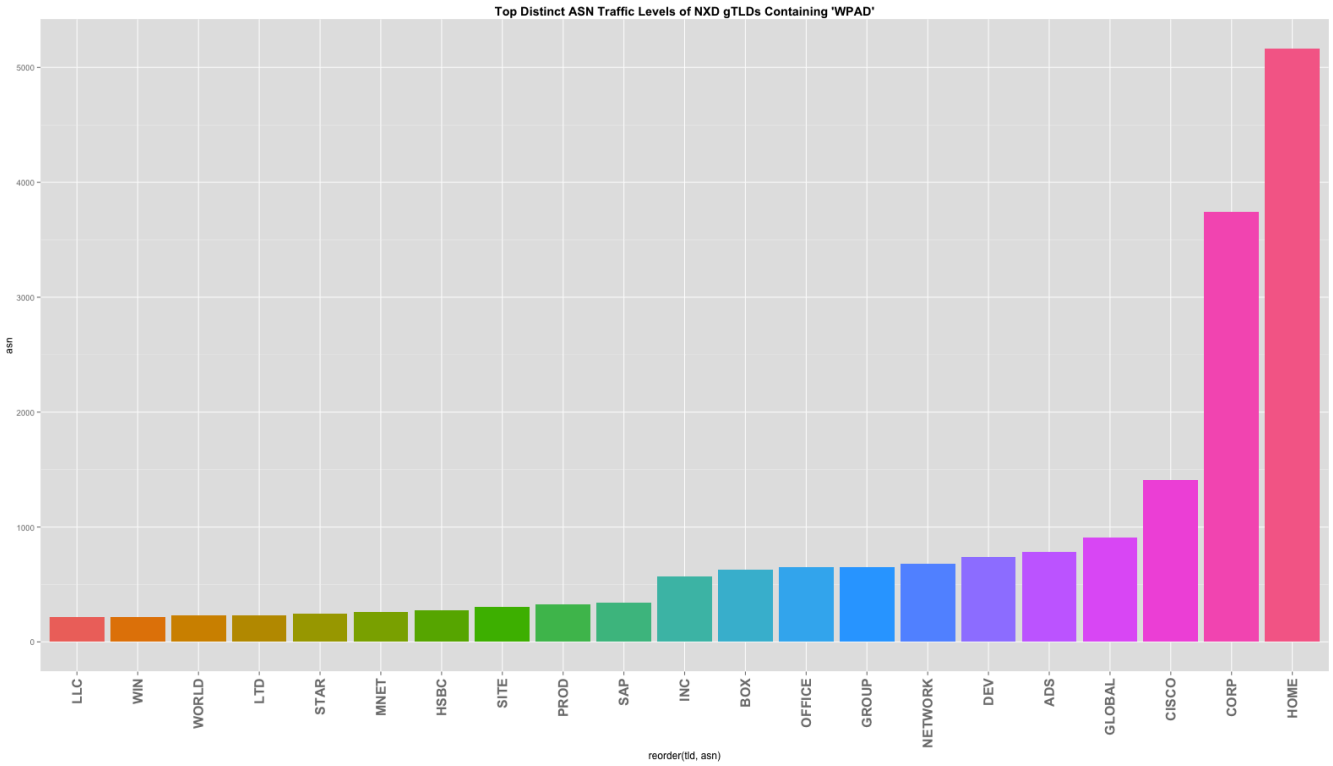
We assessed ``risk'' in both a broadly and in focused manners

- In one of our TRs, we examined the broad "*spread*" of risk across all applied-for strings with a candidate *Risk Matrix*
  - We looked for measureable evidence of "what" the problem is
- Then, we followed that up a focused (per-impacted party) methodology
  - We add "who is impacted" to our "what's the problem" analysis
- We use network-level information (such as ASNs) and semantic information (namespaces) to identify impacted parties employing applied-for strings
  - How many namespaces are actually going to be impacted, and what might actually break for them
  - For example, why do we see DNS Service Discovery (SD) queries and virus scans that *seem* to be from residential apartment complexes?

For example, consider the MitM vector

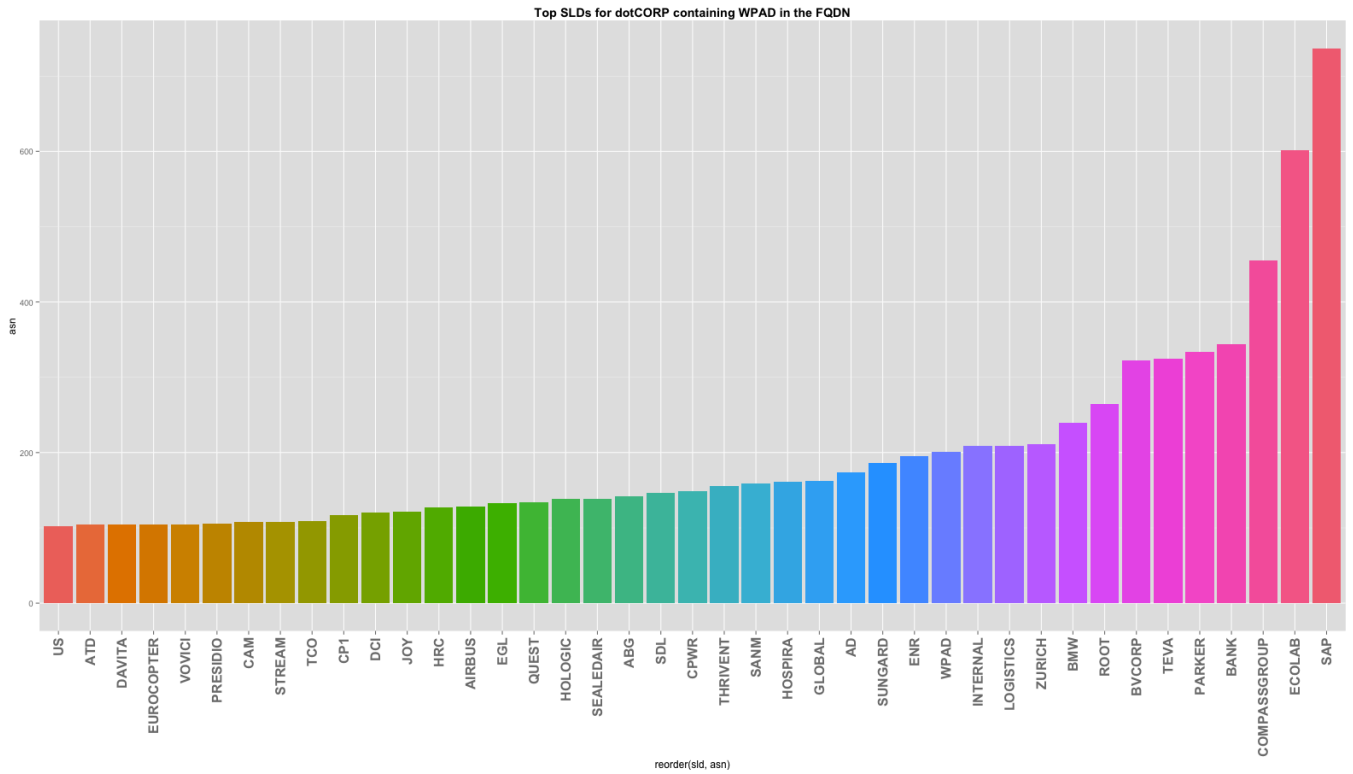
- Many many networks use internal Top Level Domains (iTLDs) to manage their internal networks
  - .corp, .home, .secure, .etc, etc.
  - Reduces local scoping, external dependencies, etc.
- BUT, there is a protocol called Web Proxy Auto-discovery Protocol (WPAD)
  - Browsers use this (today) to automatically determine if they should send all HTTP queries through a *learned* proxy
  - They issue wpad.<DHCP domain> to find this proxy
- Now, consider that Certification Authorities (CAs) will give you a cert for *any* DNS name that is not delegated
  - This includes iTLDs
- So… A new gTLD operator can answer iTLD queries and complete TLS handshakes

A closer look at who's using WPAD



- You can (almost) see, the TLDs using WPAD are varied, but disconcerting
  - Cisco, HSBC, SAP, box (this one takes some explaining), WIN
- High-order bit: a lot of namespace diversity

A closer look at who's using WPAD (2)



- Focusing in on just .corp, the Second Level Domains (SLDs) are interesting too
  - Airbus, BMW, Presidio, Eurocoptor, Teva, SAP, Bank, etc.
- High-order bit: even more namespace diversity

We illustrated this in our TR

- We also noted other MitM vectors (ISATAP, DNS-SD, search-list interactions, etc) in our TR #1130008
  - It's long, but maybe worth a read

- Next, rather than looking broadly, we looked more deeply at who was using these protocols
- To alleviate concerns, Commonwealth Bank of Australia claimed responsibility for .cba DNS queries

  "As the cause of the name collision is primarily from CBA… it is within the CBA realm of control…"

  http://forum.icann.org/lists/comments-name-collision-05aug13/msg00004.html

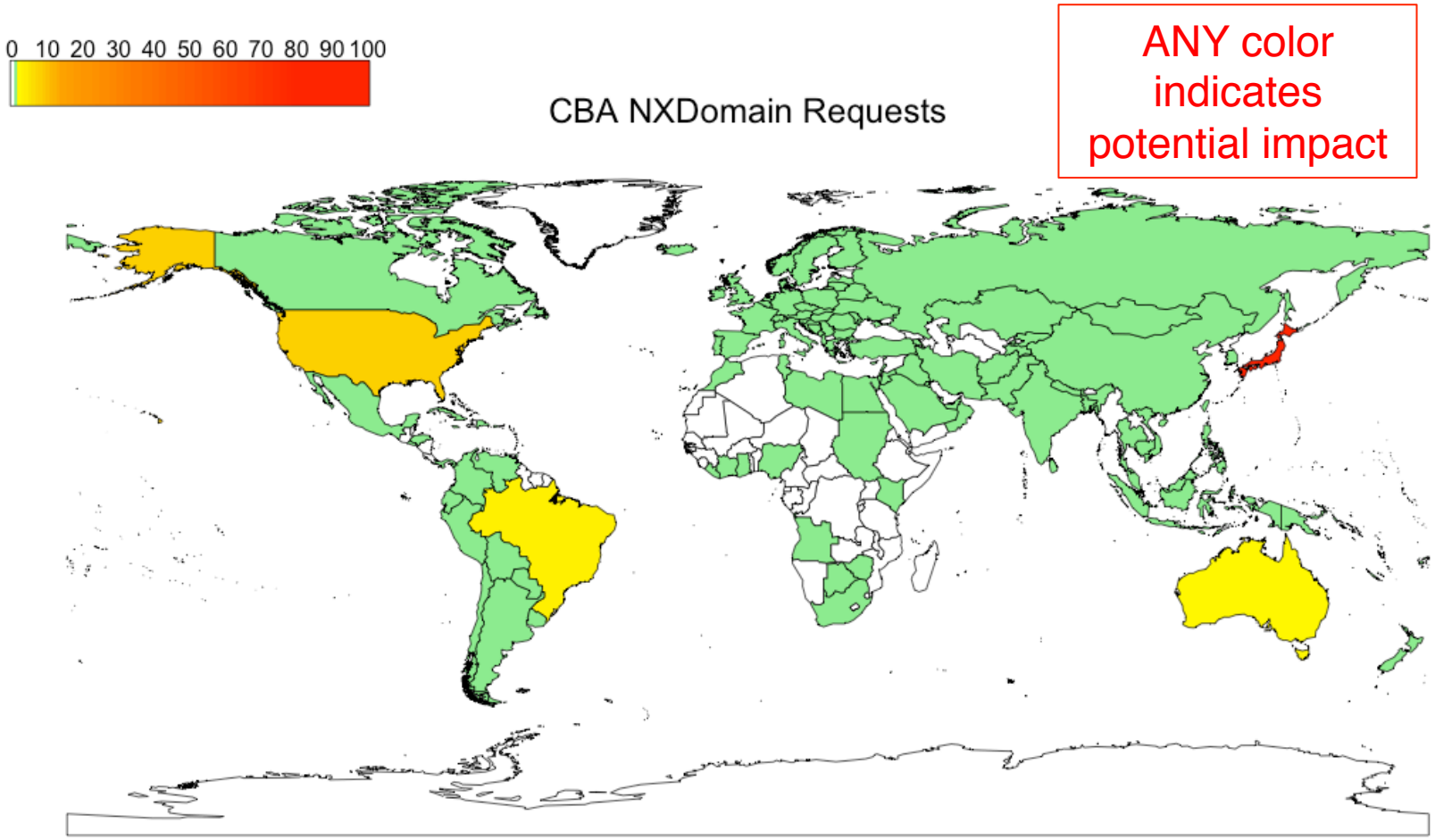
- Using our evolving methodology, we debunked this, and found another alarming trend

What network sources are making .cba queries

- 2,639 unique ASN's across 171 countries responsible for queries
  - 1,785 (~68%) ASNs made more than 1 query over the collection period
  - Top 20 ASN's account for ~90% of all queries
  - NTT-ME Corporation in Japan generates ~79% of all queries

Top ASN's	Query Count
AS9595 Corporation	396,742
AS15169.	16,806
AS7018.	7,179
AS8075	5,629
AS7922.	2,746
AS30607	2,737
AS16880.	2,247
AS27882.	2,184
AS28573.	1,916
AS4804	1,796
AS1221	1,711
AS577	1,629
AS4230	1,507
AS6830.	1,434
AS45867	1,388
AS36692	1,326
AS7132	1,192
AS22773.	1,130
AS3356	1,119
AS71	1,052

Heat map of query sources for .cba



Generalized impact statement

- Roughly 80% of all queries seen for .cba are for namespaces conducting Bonjour and other DNS-SD protocols
  - Of the 65 namespaces:
    - 49 are based in Japan
    - 5 are based in Brazil
    - 2 are based in Canada
    - Only 1 sees activity from Australia and that namespace has the most diverse set of queries.
- Bonjour is a DNS Service Discovery protocol for network services like:
  - printers, Apple TV, etc…
- Also enables ***smart home*** automation technologies like:
  - thermostats, remote and physical access systems, energy management, alarms, etc.., e.g.,:

http://www.marvell.com/smart-energy/assets/Marvell-Smart-Energy-Platform-Brief.pdf

DNS-SD/Bonjour-based namespaces

- 65 different namespaces seen making Bonjour queries

Example Namespaces
parkside-kamagaya
makuharibaytown-mirama3
maehara-b2
winstown-inage45
makuharibaymirama-ru
ilink-ichikawa-03
wt-inagekaigan3
w-g-c-2
yachiyo-pc12
gp-sonnou-4
w-g-c-1
Winsinage2
ilink-ichikawa-04
ver-ichi6
Toyoshiki2
ilink-ichikawa-02
takanedai-9
g-n-2
ichikawaminami-3
a-takane5

Parkside Kamagaya – Rental property in Chiba, JP (parkside-kamagaya)

Makuhari Baytown High-rise in Chiba (makuharibaytown-mirama3 and makuharibaymirama-ru)

I-Link Ichikawa The Towers East (ilink-ichikawa-03)

Why would this be alarming?

"Bonjour, also known as zero-configuration networking, enables automatic discovery of devices and services on a local network…" https://developer.apple.com/bonjour/

- DNS-SD queries leak from the local namespace when a machine thinks it exists in a "zone" that does not exist in global DNS
  - 106,881 (21.2%) of all CBA queries represent explicit DNS-SD related queries
- Anyone could potentially answer DNS-SD queries
  - So, when a Bonjour or other DNS-SD client asks where its printers are, *who* is really answering?
- Some have dismissed this as FUD,
- But some ISPs have not!
  - Also, a consortium of Power companies issued a comment underscoring the dangers this could pose to live power grids
  - http://forum.icann.org/lists/comments-name-collision-05aug13/msg00058.html

VERISIGN

So, what does that all mean?

- Our studies found evidence of new classes of risks
  - For example, smart spaces in Japan could have service discovery queries for remapped for alarms, or SIP communications could be impacted in homes in Germany, etc.
- One of the challenges is that *since* DNS is so pedestrian, it's transparently woven into our daily lives, with complicates systemic effects
  - ``The most profound technologies are those that disappear'' – Mark Weiser
  - Even TLS won't be safe anymore!
- We have begun doing more qualitative impact studies, and are preparing for conscientious disclosure

High-order bits

- What we learn in school is *very* valuable, and very needed
  - Even mundane things like DNS need deep thinkers!
  - There are lots of ways for our work to make a difference
- The ops communities are in desperate need of deep of the skills we have
  - It is not always clear to me that this is obvious to everyone
  - That can be source of pain, or the hallmark of opportunity!
- Our schooling ought to be a standing challenge to each of us: find where our passion is needed, and apply it there!
  - Sometimes, the most banal topics can be truly impactful
  - Timing is so important, but being the right person, at the right place, at the right time can be deeply rewarding

VERISIGN®