

NCAP Discussion Group | 20 January 2021

Agenda:

1. Welcome and roll call
2. Update to SOI
3. Update on Study 2
4. Name Collision Outreach Efforts
5. .INTERNAL Case Study Review:
<https://docs.google.com/presentation/d/1qJB2GuYvEKfCdGwJ7wRPfVR4Ji8yODxFS09FzvmOTsg/edit#slide=id.p>
6. JAS Refresher
7. AOB

Table of Contents

Name Collision Outreach Efforts:..... 1

Global Advisors Study 3

.internal Case Study..... 3

Slide 1: Daily Query Volume4

Slide 2: Qtype Distribution4

Slide 3: Qtype Distribution of Strings5

Slide 4: Unique Daily Source IPs.....6

Slide 5: Geographical Distribution.....6

Slide 6: ASN Distribution6

Slide 7: ASN Distribution .mail/.internal.....7

Slide 8: Label Analysis8

Slide 9: Label Analysis 2.....9

Slide 10: .svc Segway and Kubernetes.....10

Slide 11:.....10

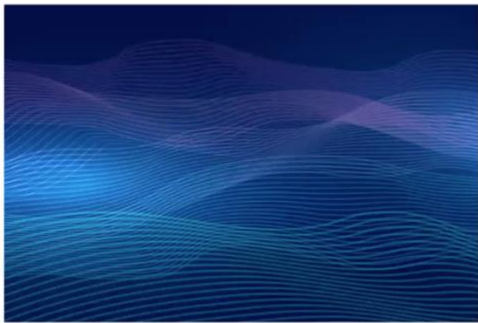
Slide 12: Root ASN Overlap and IP Growth.....11

Slide 13: Root ASN Overlap and IP Growth 211

Slide 14: SLD Overlap Analysis12

Name Collision Outreach Efforts:

started inclusion outreach endeavor looking to identify and remediate high query volume strings.



VERISIGN OUTREACH PROGRAM REMEDIATES BILLIONS OF NAME COLLISION QUERIES

JANUARY 15, 2021 • BY MATT THOMAS • DOMAIN NAMES

A name collision occurs when a user attempts to resolve a domain in one namespace, but it unexpectedly resolves in a different namespace. Name collision issues in the public global Domain Name System (DNS) cause billions of unnecessary and potentially unsafe DNS queries every day. A targeted outreach program that Verisign started in March 2020 has remediated one billion queries per day to the A and I root name servers, via 46

SUBSCRIBE TO VERISIGN BLOG

Enter your email:

SUBSCRIBE ME

DOMAIN NAME SEARCH

Use our NameStudio domain name generator:

SHOW ME

[Terms of Service](#)

RECENT POSTS

- SECURING THE DNS IN A POST-QUANTUM WORLD: NEW DNSSEC ALGORITHMS ON THE HORIZON
January 19, 2021
- VERISIGN OUTREACH PROGRAM REMEDIATES BILLIONS OF NAME COLLISION QUERIES
January 15, 2021
- NEWER CRYPTOGRAPHIC ADVANCES FOR THE DOMAIN NAME SYSTEM: NSEC3 AND TOKENIZED QUERIES
January 14, 2021

VERISIGN OUTREACH PROGRAM REMEDIATES BILLIONS OF NAME COLLISION QUERIES

JANUARY 15, 2021 • BY MATT THOMAS • DOMAIN NAMES

A name collision occurs when a user attempts to resolve a domain in one namespace, but it unexpectedly resolves in a different namespace. Name collision issues in the public global Domain Name System (DNS) cause billions of unnecessary and potentially unsafe DNS queries every day. A targeted outreach program that Verisign started in March 2020 has remediated one billion queries per day to the A and J root name servers, via 46 collision strings. After contacting several national internet service providers (ISPs), the outreach effort grew to include large search engines, social media companies, networking equipment manufacturers, national CERTs, security trust groups, commercial DNS providers, and financial institutions.

While this unilateral outreach effort resulted in significant and successful name collision remediation, it is broader DNS community engagement, education, and participation that offers the potential to address many of the remaining name collision problems. Verisign hopes its successes will encourage participation by other organizations in similar positions in the DNS community.

Verisign is proud to be the operator for two of the world's 13 authoritative root servers. Being a root server operator carries with it many operational responsibilities. Ensuring the security, stability and resiliency of the DNS requires proactive efforts so that attacks against the root name servers do not disrupt DNS resolution, as well as the monitoring of DNS resolution patterns for misconfigurations, signaling telemetry, and unexpected or unintended uses that, without closer collaboration, could have unforeseen consequences (e.g. Chromium's impact on root DNS traffic).

Monitoring may require various forms of responsible disclosure or notification to the underlying parties. Further, monitoring the root server system poses logistical challenges because any outreach and remediation programs must work at internet scale, and

- SECURING THE DNS IN A POST-QUANTUM WORLD: NEW DNSSEC ALGORITHMS ON THE HORIZON
January 19, 2021
- VERISIGN OUTREACH PROGRAM REMEDIATES BILLIONS OF NAME COLLISION QUERIES
January 15, 2021
- NEWER CRYPTOGRAPHIC ADVANCES FOR THE DOMAIN NAME SYSTEM: NSEC3 AND TOKENIZED QUERIES
January 14, 2021
- CRYPTOGRAPHIC TOOLS FOR NON-EXISTENCE IN THE DOMAIN NAME SYSTEM: NSEC AND NSEC3
January 13, 2021
- THE DOMAIN NAME SYSTEM: A CRYPTOGRAPHER'S PERSPECTIVE
January 8, 2021
- CHROMIUM'S REDUCTION OF ROOT DNS TRAFFIC
January 7, 2021

ARCHIVES

Select Month

TAGS

- .com .net .tv 中国 .id
- 品牌 Branding Cryptography
- Cybersecurity Cyberthreats DANE

Verisign: took A and J data and ranked top leaking tld strings by 2 factors: first was # of queries coming out of it; 2nd was looking at cumulative distribution function of # of ASNS for that string, look at what percentile the top 3 ASNS would be at becomes a gating mechanism to identify high query high concentrated strings of particular vendors. Look at Q names and IP addresses identify what kinds of systems were causing queries. Then did outreach to see if these entities had direct contests. Used DNS OARK. Sent emails to these people and people were open to working with us and they helped make many issues disappear. But some strings were associated with home networking equipment. Some tied to vendors. In these cases too time consuming to fix quickly.

Common causes: Suffix search list where sting was applied in a separate search list and appended on all queries.

Appendix A: Horizontal Study

Representative Regular Expressions across NXDOMAIN Responses



Global Advisors Study

Jeff Schmidt: refresher on analysis work done 5 yrs ago, on Jazz (??) horizontal and vertical study on DNS OARC data.

We did outreach to 200 firms. Meaningful engagement with 20 that resulted in improvements. Horizontal study effort to understand strings and query types across all name servers based on DNS OARC digital data. Across all queries that was in top level domain that was requested

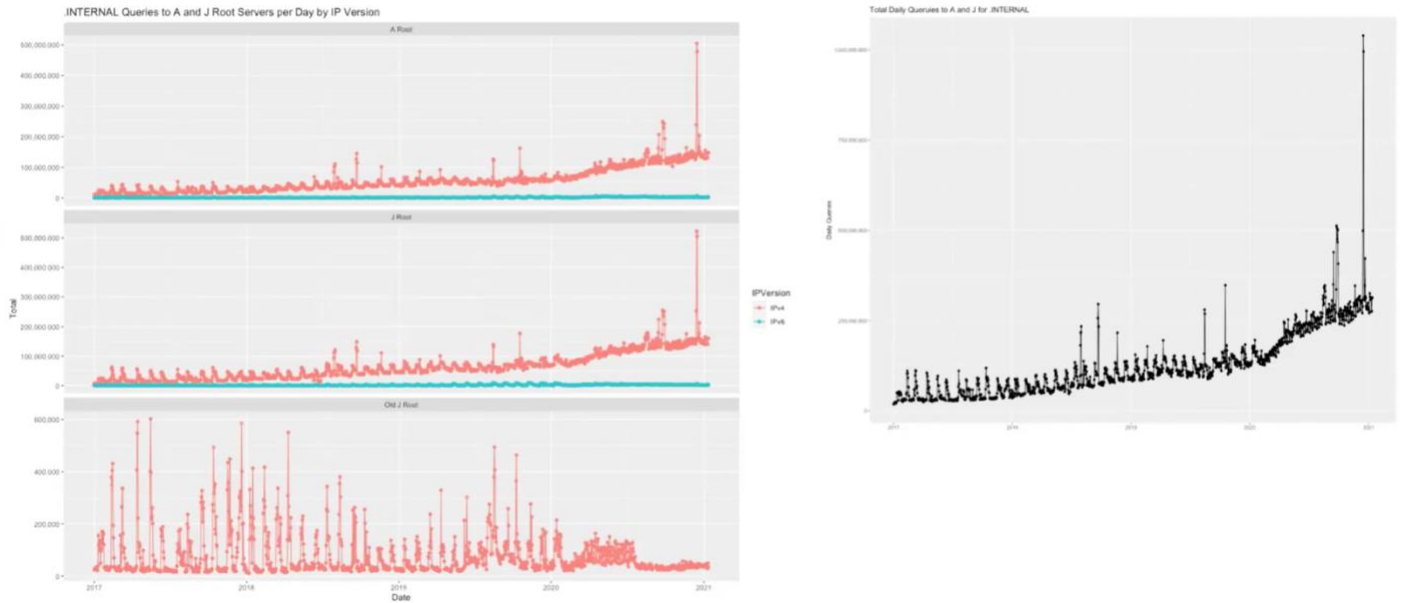
Some interesting things that stood out in the data:

What kind of software generates a query that starts with AD Route (strings that start with AD route and have some stuff at the end account for 1.3% of queries). Or software that generates a query that starts with compatibility additions or add ons.....this led us to discovering that Microsoft Active Directory has a big role- its queries start with them as DCS

.internal Case Study

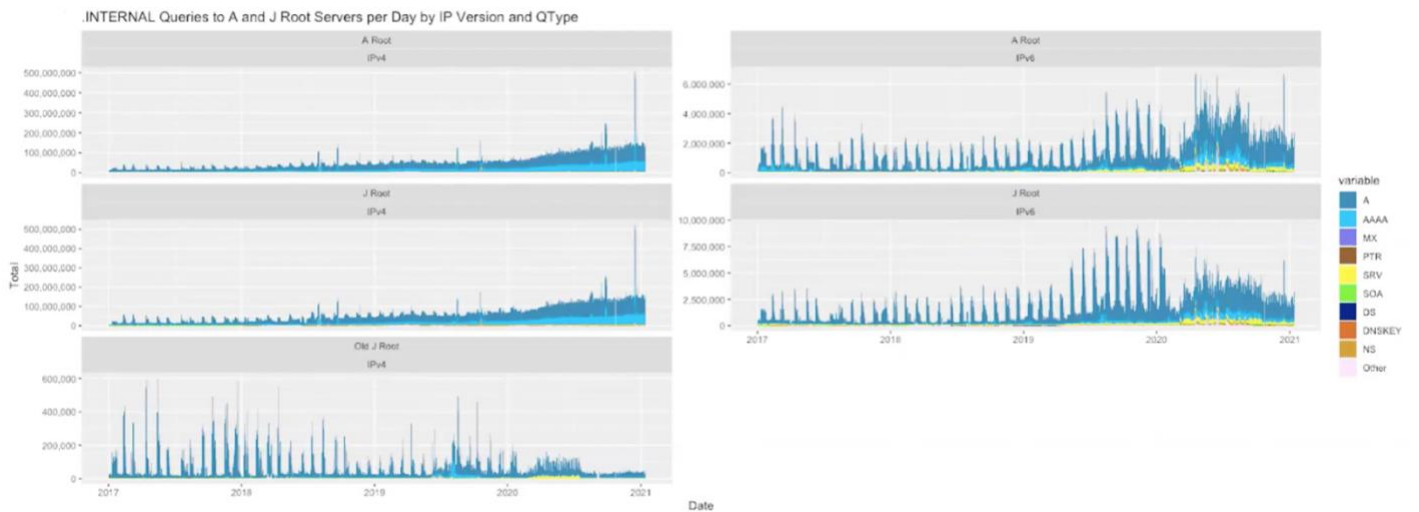
.internal case study

.INTERNAL Analysis :: Daily Query Volume

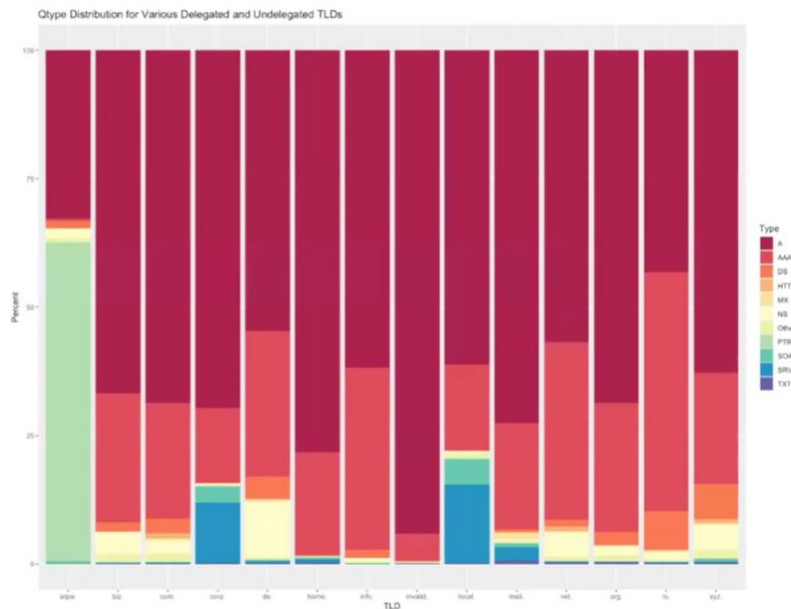


Graph on J, old J an dA route

.INTERNAL Analysis :: Qtype Distribution



.INTERNAL Analysis :: Qtype Dist. of Strings

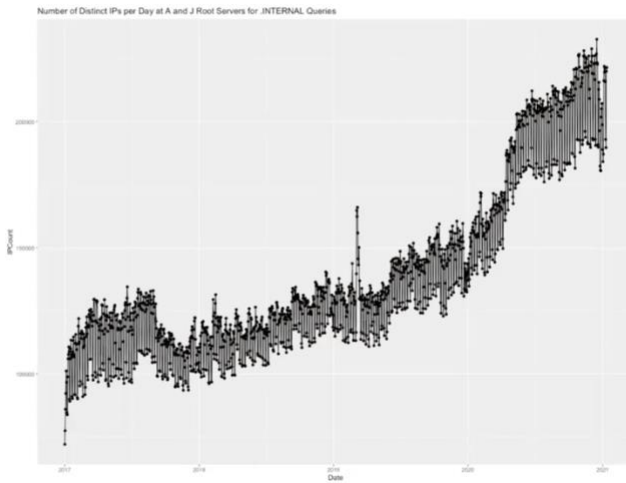


various delegated and non delegated strings on comparing their que types.

- Regular delegated seem to have a standardized mixture, but you will notice with .Corp, .home, .local and .mail as more of a Prevalence of SRT your service records, so maybe there is a little bit of a bias in some of these strings.
- Why as to why those strings are leaking in the first place. It's because they are DNS service discovery oriented type queries and just automatically getting sent out to find these new services and you have things like something search list depending to them and they're coming out.

Slide 4: Unique Daily Source IPs

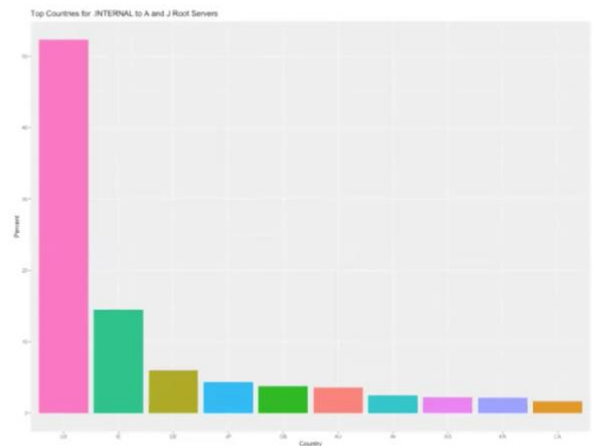
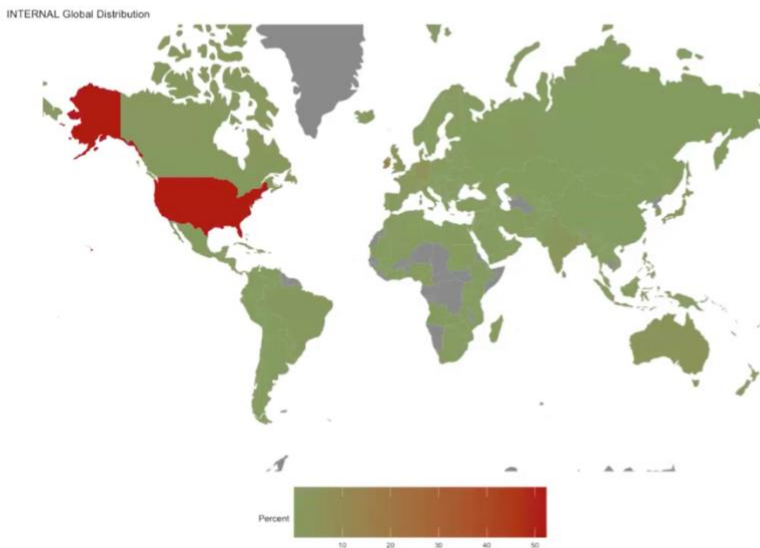
.INTERNAL Analysis :: Unique Daily Source IPs



Increase in distinct IP addresses - transient devices have been taken home due to quarantine, so .internal no longer resolving in enterprise resolution systems but going to residential ISPs

Slide 5: Geographical Distribution

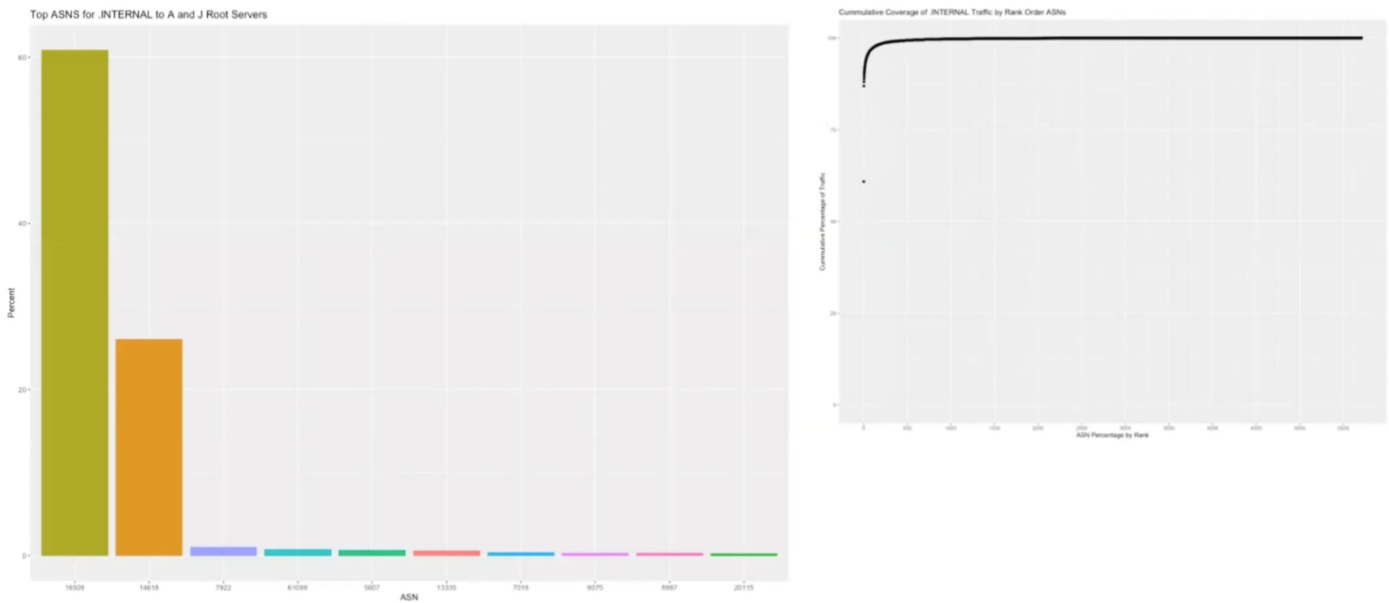
.INTERNAL Analysis :: Geographical Distribution



US, then Ireland for .internal Different geo distribution then .mail

Slide 6: ASN Distribution

.INTERNAL Analysis :: ASN Distribution



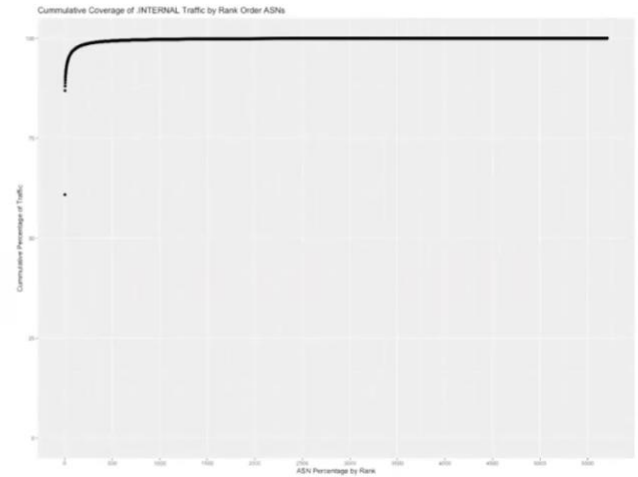
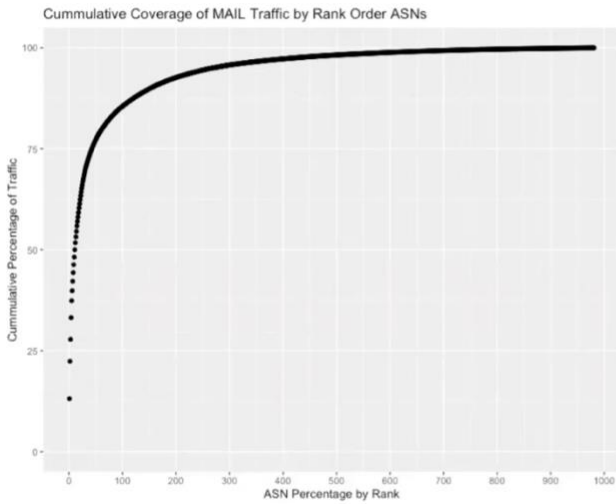
85% or 86% of all of the queries for .internal are coming from 2 autonomous systems that are actually the same company. It's just they have two different ASes. So there is one company out there that is responsible for upwards of 80 some percent of the leakage of .internal

graph on the right is the cumulative distribution of those as isn't the percentage of traffic that they're leaking out.

Have been in contact with the "company" mentioned and working with them to fix it.

Slide 7: ASN Distribution .mail/.internal

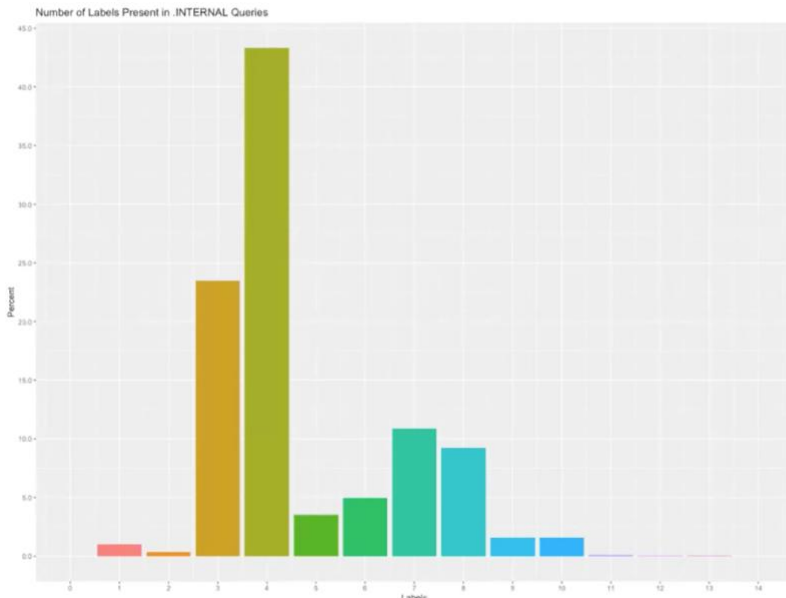
.INTERNAL Analysis :: ASN Distribution MAIL/INTERNAL



Fewer # of sources responsible for much more of the traffic

Slide 8: Label Analysis

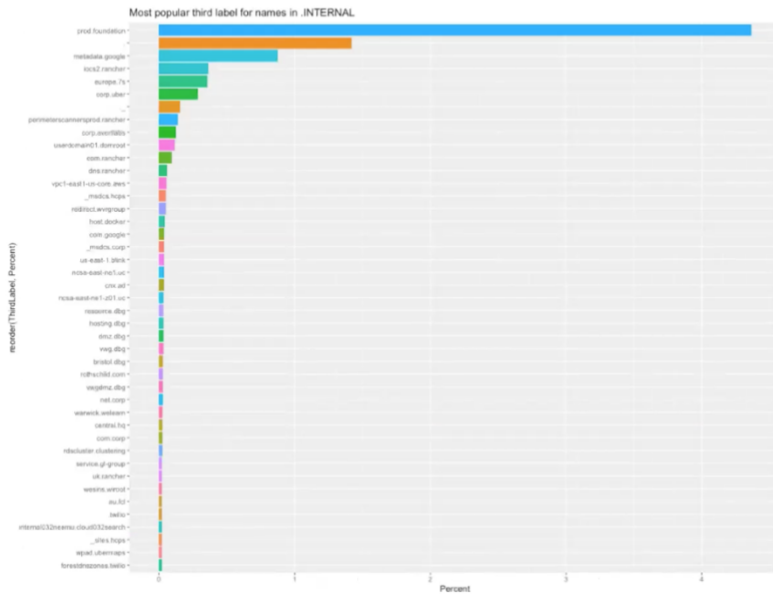
.INTERNAL Analysis :: Label Analysis



SLD	Percent	ThirdLabel	Percent
1: compute	64.06637157	1: us-west-2.compute	17.1589977
2: ec2	24.25419474	2: us-east-2.compute	16.4330502
3: foundation	3.29748196	3: eu-west-1.compute	12.5938315
4:	1.07270648	4: eu-central-1.compute	6.3346503
5: compute-1	0.87439469	5: ap-northeast-1.compute	5.6828674
6: google	0.75785944	6: prod.foundation	4.3639820
7: rancher	0.71297776	7: ap-southeast-2.compute	3.6556092
8: clustering	0.55616635	8: eu-west-2.compute	2.9668113
9: hcps	0.39810000	9: ap-southeast-1.compute	2.8879647
10: corp	0.27232804	10: ap-south-1.compute	2.5913326
11: 7s	0.26998527	11: ap-northeast-2.compute	2.5346497
12: uber	0.23709058	12: com.ec2	2.4392407
13: ubermaps	0.18868883	13: sa-east-1.compute	2.2049954
14: dbg	0.16952733	14: internal.ec2	1.7763651
15: ghndnet	0.15295878	15: us-east-1.compute	1.7595327
16: domroot	0.13477335	16: ca-central-1.compute	1.4508265
17: -	0.11673546	17:	1.4198201
18: wvrgroup	0.10461611	18: us-west-1.compute	1.4166473
19: westada	0.10071690	19: eu-west-3.compute	1.3051468
20: avertlabs	0.09441954	20: metadata.google	0.8742713
21: cmw	0.06675903	21: eu-north-1.compute	0.8657445
22: aws	0.06358604	22: ap-east-1.compute	0.6080187
23: uc	0.06308703	23: us-gov-west-1.compute	0.4586415
24: ad	0.06271632	24: iocs2.rancher	0.3656085
25: maxim-ic	0.05971856	25: europe.7s	0.3569946
26: efi	0.05687003	26: net.ec2	0.3532283
27: fcl	0.05675254	27: me-south-1.compute	0.3512589
28: wge	0.05359981	28: af-south-1.compute	0.3455081
29: domain	0.05233912	29: de.ec2	0.3376713
30: blink	0.05020264	30: eu-south-1.compute	0.3026175

the graph on the left is looking at the percentage of the number of labels contained within the key name. vs .mail with 60% having only 1 label, .internal has more context in terms of labels and content within query name

.INTERNAL Analysis :: Label Analysis



WHY RANCHER? PRODUCTS CUSTOM

Why Rancher?

Rancher is a complete software stack for teams adopting containers. It addresses the operational and security challenges of managing multiple Kubernetes clusters across any infrastructure, while providing DevOps teams with integrated tools for running containerized workloads.

The DNS entry is added to the `resolv.conf` file but it's overridden by rancher.

```
## cat /etc/resolv.conf
search 10.42.173.171 dns.rancher.internal busy.dns.rancher.internal rancher.internal
# nameserver 10.42.173.171
nameserver 169.254.169.250
```

<https://github.com/rancher/rancher/issues/15304>

web page for Cooper Nettie technology that is Called rancher and it suggests putting your DNS under a suffix search list for rancher .internal ee're seeing a change that the rancher software is probably very likely the the root cause of many of these others. So I have made it a item to submit a ticket to GitHub and disclose this

Warren: believe this issue was with an older version of .rancher; ended in 2018 with new version. The fact that this is still seen gives us a metric for expectations at how quickly a software related issue ends even after software is changed.

.INTERNAL Analysis :: SVC Segway and Kubernetes

A and J Root Traffic for SVC



Table 6: Kubernetes References to URLs containing SVC

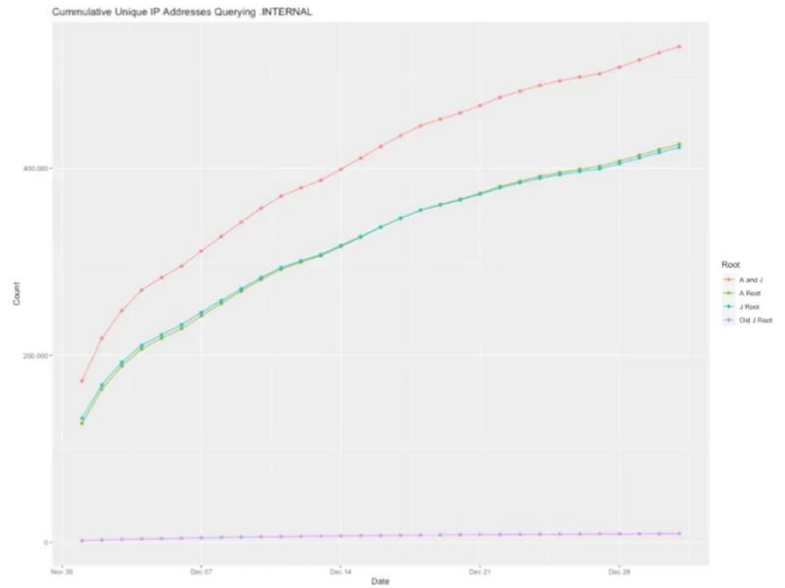
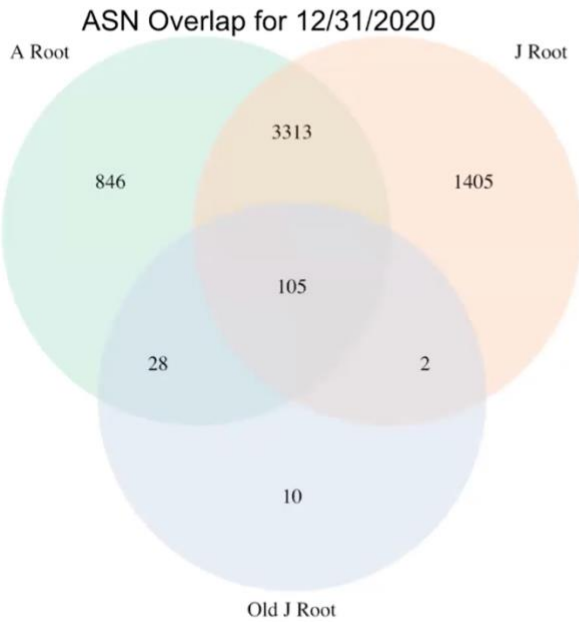
File	Line	Text
kubernetes/cmd/kubeadm/app/phases/controlplane/manifests.go	141	"service-account-issuer": fmt.Sprintf("https://kubernetes.default.svc.%s", cfg.Networking.DNSDomain),
kubernetes/staging/src/k8s.io/kube-aggregator/artifacts/self-contained/etcd-pod.yaml	22	-"--advertise-client- url=https://etcd.kube-public.svc:4001"
kubernetes/staging/src/k8s.io/kube-aggregator/artifacts/self-contained/etcd-pod.yaml	28	-"--initial-advertise-peer- url=https://etcd.kube-public.svc:7001"
kubernetes/staging/src/k8s.io/kube-aggregator/artifacts/self-contained/etcd-pod.yaml	33	-"--initial-cluster=default=https://etcd.kube-public.svc:7001"
kubernetes/staging/src/k8s.io/kube-aggregator/artifacts/self-contained/kubernetes-discover-pod.yaml	43	-"--etcd-servers=https://etcd.kube-public.svc:4001"

This is for .svc, which started getting 127million queries per day. Looks similar to .internal....85% came from a company whose cloud infrastructure was using a Kubernetes configuration leaking out svc queries.

Data Sensitivity Analysis

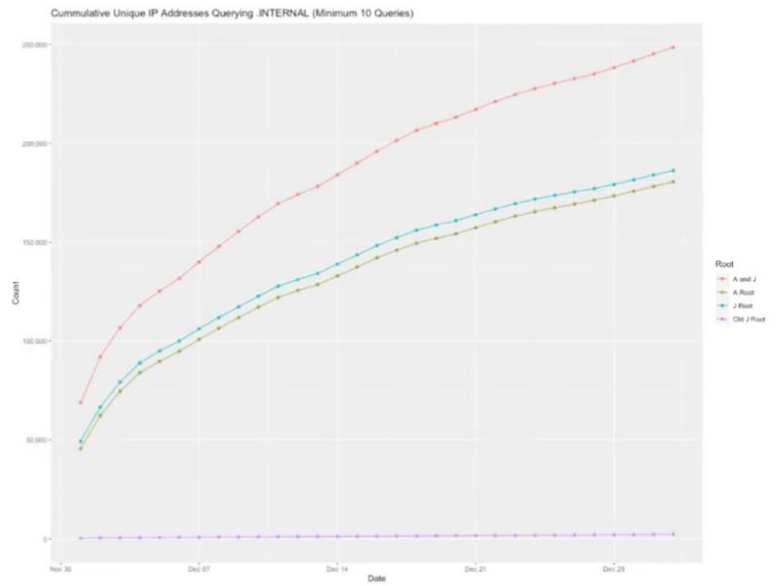
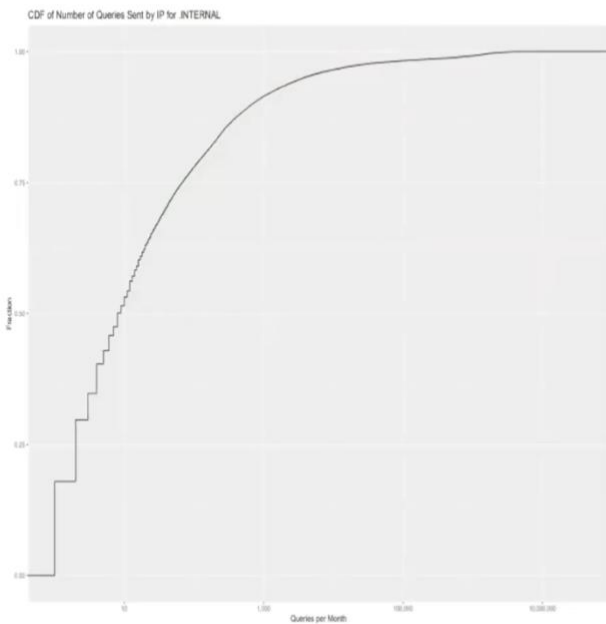
Slide 11:

.INTERNAL Analysis :: Root ASN Overlap and IP growth



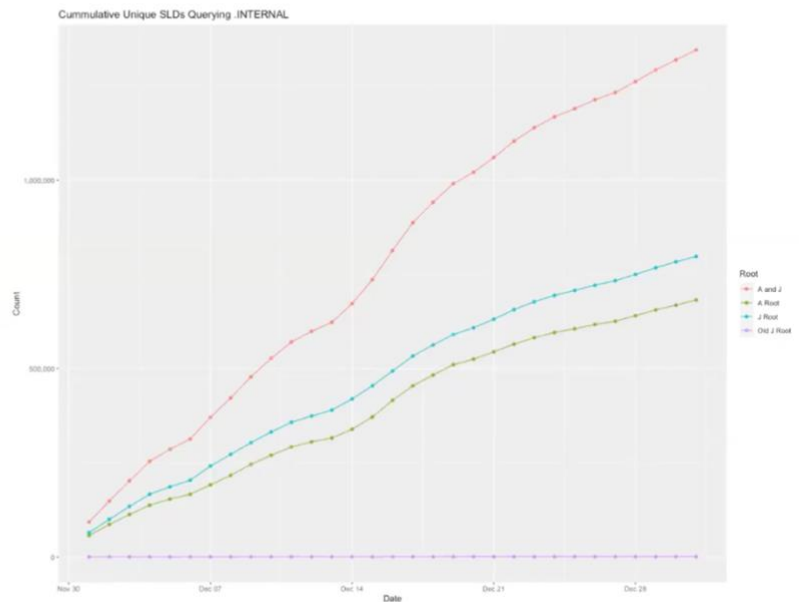
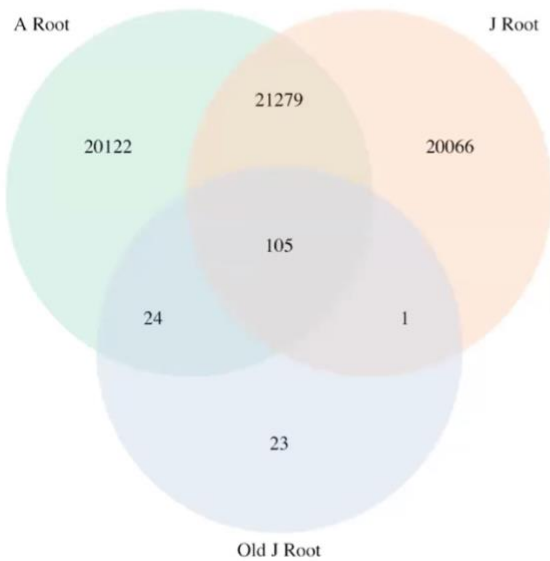
Different than .mail.

.INTERNAL Analysis :: Root ASN Overlap and IP growth



Slide 14: SLD Overlap Analysis

.INTERNAL Analysis :: SLD Overlap Analysis



The catchment of A and J differ. 80% of SLDs only seen once in the month.