More Mysterious Root Query Traffic From a Large Cloud/DNS Operator

Christian Huitema, ICANN
Duane Wessels, Verisign

Discovery

- Analysis of unique names in ICANN Managed Root Serer data taking a long time
- Many weird (random) names
- Second-level label length 12 or 13

Instance	# Queries	# Names	Log Size (bytes)
aa01-fr-bfc	124,411,748	1,860,891	49,409,103
aa01-fr-lio	16,528,533	499,185	12,476,508
aa01-fr-mbv	63,649,855	691,235	16,951,212
aa01-fr-par	318,962,809	19,521,598	543,938,433

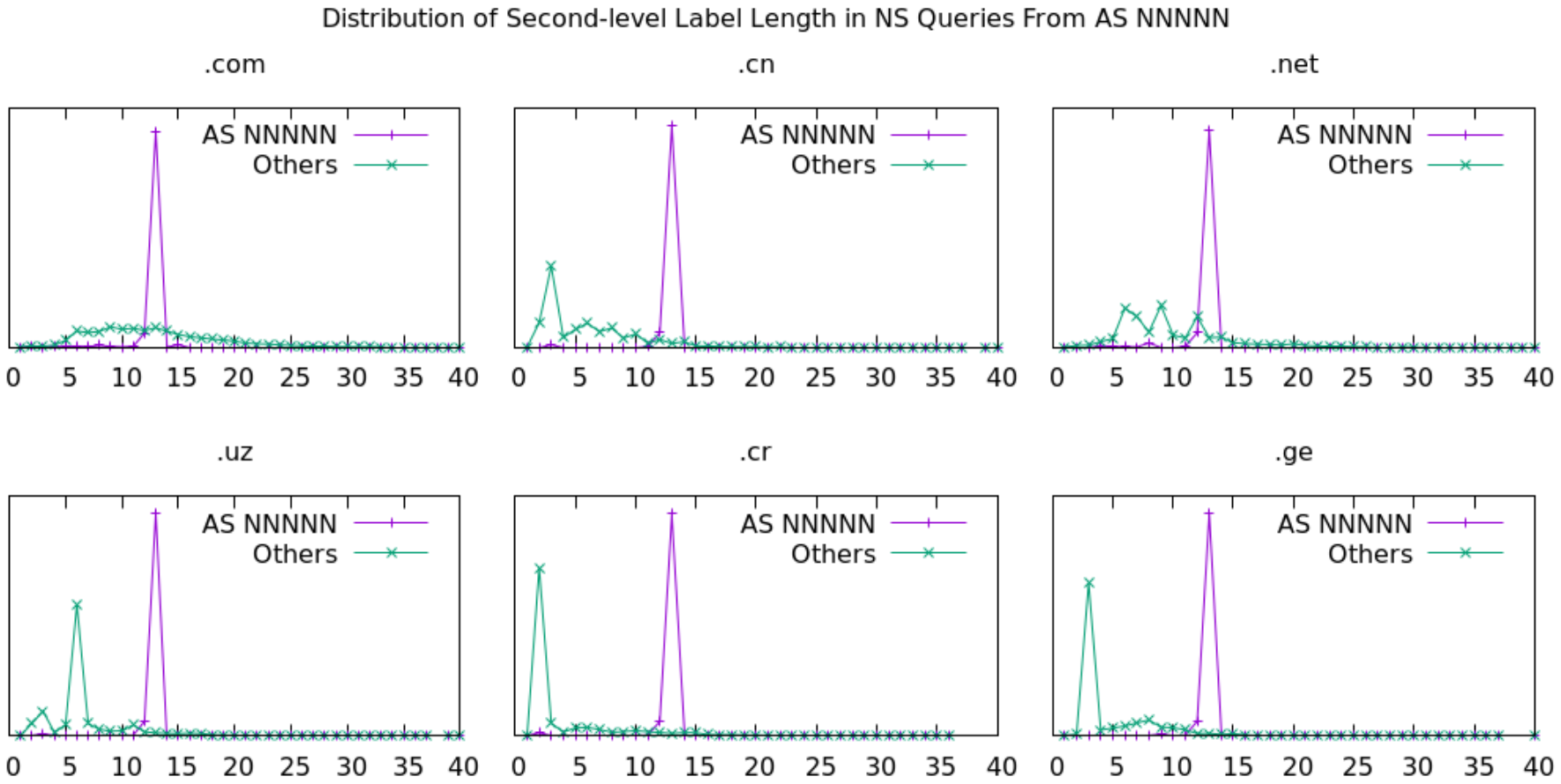
```
F863ZVV1XY2QF.SURGERY
BP639I-3NIRF.HIPHOP
VSQN7JSMF0IVL.LOAN
QO35JJK419GFM.NET
YYIF0AIJR21GN.COM
YIJOBSSQ-F0N.DE
BC9R7RSB9X0OH.HOUSE
TA2KN2ISP45QO.COM
XVSFZKPKVM4ID.COM
ZMF2QIG0ANJMM.CU
QYYX31MPJ0Z9H.SANOFI
DQ0G3MVB9NR-F.IT
U5VMT7XANB6RF.COM
B4E6RSNQ37KJK.LAW
```

From 24 hours of IMRS data on June 15, 2021, provided by ICANN.

The Signal

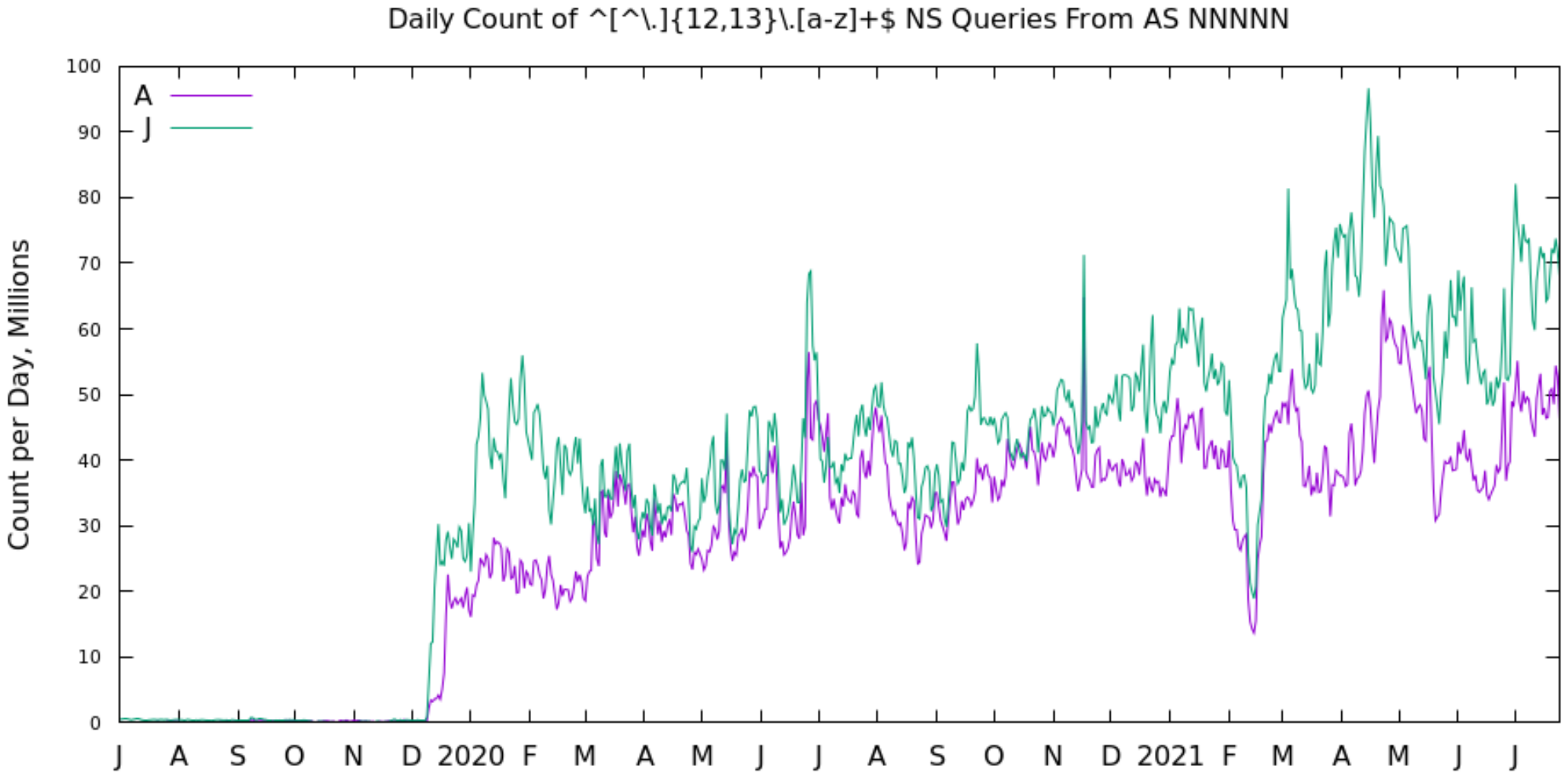
- Query type: NS
- Second-level label length: 12 or 13 characters

Confirmed in A-Root and J-Root Traffic



Data from July 6, 2021

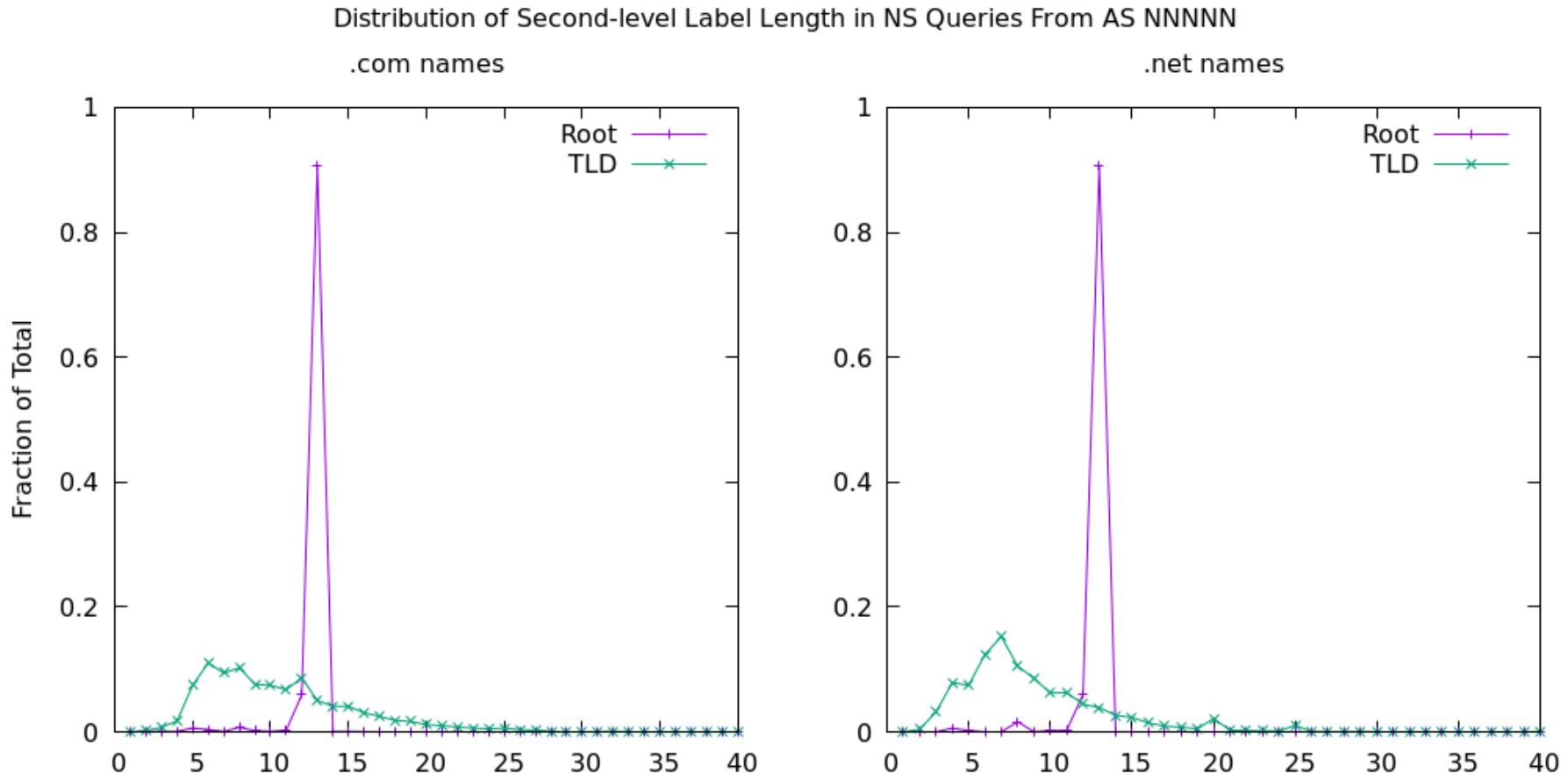
Further Research: Occurring Since Late 2019



DNS Resolver or Cloud Customer Traffic?

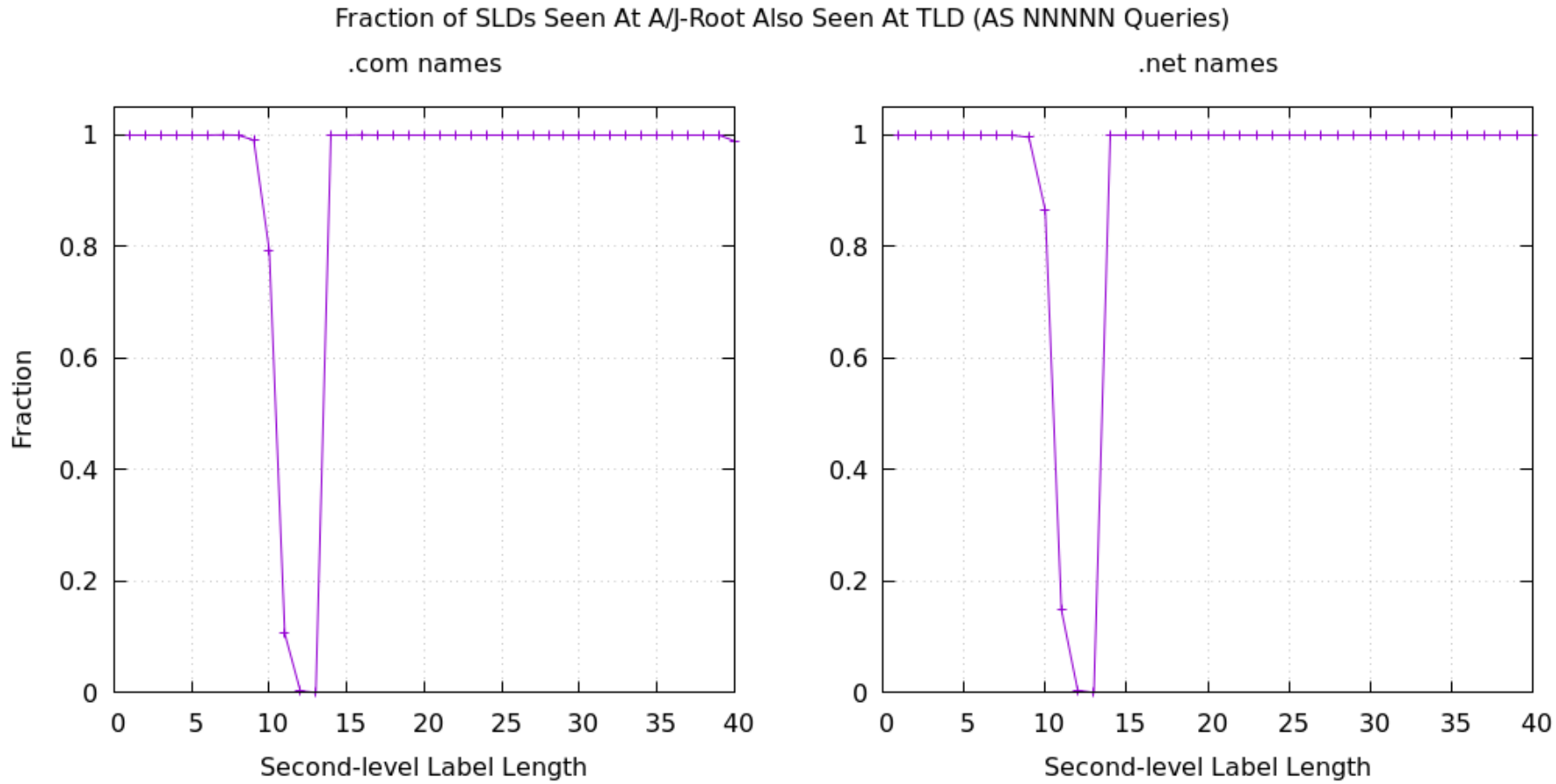
- For A-root & J-root, 99.9% of query sources fall within IPv4 and IPv6 prefixes listed as backend query sources for the DNS resolver service
- So, probably coming from the recursive DNS, and not other cloud customers

Why Observed at Root, but Not TLD?



Data from July 6, 2021

Why Observed at Root, but Not TLD?



Data from August 16, 2021

Could Be An Embedded Dot in the Label?

- Verified that qnames consist of separate labels
- Root server returns delegation referral and not RCODE 3 (NXDomain)

In Summary

- Excessive amount of unusual queries observed at root name servers
- Type NS
- Second-level label 12-13 random characters, across multiple TLDs
- From the autonomous system of a large DNS/cloud provider
- IPv4 and IPv6
- Since late 2019
- Source IP analysis implies queries come from recursive resolver
- But no observed follow-up queries at the TLD name servers

Answers?