ICANN NCAP Meeting #37 (2021-01-13) Report Recap

Introduction. The meeting opened with no reported changes to participant Statement of Interest (SOI) and a reminder that a SSAC review of Study 2 was coming due. This was the first in a series of “deep dives” into A+J root queries for select collision domains. This week began with a look at *.mail* which would ultimately foreshadow the next four weeks or so of analysis in other names under study (e.g. *.internal*, *.home*, *.corp*). The analysis showed how .mail query volume had seen a stable period of up until 2018 when there was a big drop. Then in late 2020 there was another big drop hypothesized as a result of a change in Chromium behavior. The discussion centered around what if anything this analysis could say about the problem or solution with collisions generally.

.MAIL Analysis. In 2017 and 2018, query volume for .mail was at a high, but stable rate year over year, then it dropped precipitously without explanation. Interestingly, the drop is primarily associated with A-root, but why? Matt guessed there may have been some systematic process, code, or other system artifact that suddenly changed behavior. This drop and the reasons behind the queries under study of particular interest to really understand what is going on in order to best suggest the potential harm or solution. There was a noticeable increase in query volume around March 2020, which would be seen with other names under study in the analysis in the weeks to come. Overall diversity for .mail queries is high, stemming from a long tail of countries around the world. A & J root each saw some shared sources, but there was a large degree of uniqueness between the sets.

Later in the analysis the vast majority of the query names were also one-time unique, which was attributed to changes in Chromium random domain name test behavior that was removed in December 2020. These appeared as two-label length domain names. Many of the other names were associated with infrastructure (e.g. large mail providers, well-known prefix labels such as *wpad*, namespaces intended for internal-only usage, and old sendmail configuration examples that were widely copied from an O'Reilly book). A noticeable number of queries were also attributed to some default Windows Exchange setups in a set of hosting providers.

> *[Editor's note: The topic of harm came up multiple times, but there was no discussion on the possibility of what the harm or even harm would be if those names were delegated for those examples. This would be a useful sidebar to have for those purported examples. It would not only help answer the harm question, but may lead to help come up with a "classification" of harm.]*

An interesting diversion explored the different in top second-level domain *.mail* queries seen in 2017 with those in 2021. Most notably missing from 2021 were a number of DoD-related names. It turns out DoD systems were dropping some .mil queries, but then they would apparently leak out using the *.mail* TLD. After this was reported to them, they "fixed" the problem.

> *[Editor's note: Not explicitly discussed, but this was a "mitigation" of a sort and considering it further may have been a fruitful discussion on the harm and mitigation of "some" query traffic that can be identified as potentially problematic.]*

The meeting concluded with a discussion on the idea of an experiment where .mail would be delegated, but the authoritative name servers would simply not respond to any queries. It was not clear what this would accomplish and it was suggested that this would do more harm than good since resolvers would only retry and timeout, exacerbating unnecessary query traffic and forcing resolvers to perform more work and hold open resources.