

Board Question #7

suggested criteria for determining whether an undelegated string should be considered a string that manifest name collisions, (i.e.) placed in the category of a Collision String

Draft Answer (workspace below)

By "Collision String", the NCAP Discussion Group assumes the Board is asking for a method of identifying a Top Level string that, due to the high risk of name collisions associated with the string and the potential for harm (as described in response to Board Question #3), that particular string should not be delegated and should be reserved. The DG notes that there is, in fact, a spectrum or "range" of risks of harm associated with delegation of each new string as was observed in connection with ICANN's Alternate Path to Delegation process followed in the 2012 round, which permitted more rapid delegation of certain strings provided certain names were "blocked" at the second level. In this regard, the subsequent Name Collision Framework which adopted a system of 90-day "controlled interruption" is a system which lends itself to identification of name collisions, but does not, in and of itself, mitigate those collisions.

Workspace: for data and documentation

Applicable notes from Study 1

- * most obviously queries in the root server are a consideration
- * what else could be considered?
- * how often does it occur? why does it occur? we should use frequency of occurrence to drive any recommendation?
- * IANA special use domain registry is a source of collision
- * Similar things like "http" are probably problematic
- * Distribution of queries - coming from a lot of places versus localized sources
- * how does changes in DNS eco-system affect our analysis, e.g., DOH, QNAME minimization?
- * RFC7706 and its "bis" are a consideration; hyper local root and aggressive NSEC
- * closely related to question (8); together they are "threshold questions"

Proposed Gap: * Need to consider sooner rather than later what questions we might to ask of the data to help understand this

* this is particularly true for resolver data

Questions from [Study 2 Proposal, Appendix 3](#)

The board created a term, "collision string" that is the list of strings that will not be delegated. Is there criteria as to whether a string should not be delegated?

Unclear whether its trying to get to a lexical answer (i.e., pattern matching) which is likely impossible because context around how its being used matters. A broad sweeping lexical analysis is probably too much, but we should get into a discussion of what elements we can consider.

Is this about identifying the strings that would undergo controlled interruption (that was intended to stop the collisions)? We haven't identified a mechanism that would stop these collisions from continuing. Study 3 was to examine mitigation strategies, and ther are question sin this list that get to whether mitigation is possible. There is a relationship here between this and mitigation. ("These could be collision strings but here are mitigation strategies that may apply.")

Is "collision string" a "yes" or "no" category? In Sub Pro we talked about "Do not delegate" versus other categories of risk for certain strings. In the Board Question 7, are we saying "collision string" = Do not delegate? Do we need to say this is "high risk"? Probably some things we can identify that make whether it is a collision string obvious, but we won't know all of them in advance. We may talk about reserved strings. So, yes or no up to a point, but judgement will need to be applied.

Do we have any actual mitigations that have been used? When we say anything about mitigations, what are we referring to? We need to wait for the root cause analysis study. We also need to allow for the possibility that someone will propose a mitigation strategy for a particular type of name collision.

- "controlled interruption" for 90 days is not great as a mitigation strategy. Majority in Sub Pro affirmed this as mitigation, however. It is a collision identification strategy and an opportunity to collect information about it

- Agree with Jim's characterization of "controlled interruption" as a way to discover name collisions. It does not prevent name collisions. Should we be saying that in response to this question to clarify? I think Jeff Neuman would say it's a way to mitigate "harm" as opposed to a way to mitigate "name collisions". So we may need to define further what we mean by "mitigate".
- CBA was mitigated if I recall correctly. Not sure controlled interruption was involved.
- there is a partial presumption that the source of the collision will stop the collision once made aware of the fact.
- Mitigations don't necessarily need to be technical; it may be targeted outreach instead that can be effective

Should we try to build a decision workflow model? Does that kind of effort tie to this in particular, or more generally to all the questions?