DNS Cause and Effect

Geoff Huston
APNIC

Abstract

When we look at DNS query logs we have no clear idea of why we see these queries. In theory we could surmise that a caching recursive resolver would only emit queries when the recursive resolver's client has submitted a query that is not answerable from the local cache, or when the TTL of the locally held response has expired. This does not appear to be the case in a number of dimensions. We have query storms, persistent re-querying of supposed once labels, and similar. Much of the research work in this area has involved examination of queries and proposing plausible models of both end system and DNS resolution behaviour that could explain the observed query patterns. Is there a way to examine these pathologies of the DNS by encoding the label with information concerning the use of a label? In other words, can we relate cause and effect in such studies of DNS behaviour? We explore the technique of active measurement using seeded unique labels to expose various DNS pathologies.

In this work we generate unique DNS labels in a known manner and placing information concerning the end system and time and location of the ad placement into the DNS label. By understanding why this particular label has been used in a DNS name resolution query we can relate observed DNS behaviours relating to these labels with knowledge the original DNS query event. The technique used here relies on online advertisements to perform the 'seeding' part of the measurement. Advertisement systems distribute scripted objects to a wide diversity of end user devices. The script is executed on impression of the ad and the script uses an initial exchange with the measurement director to generate a collection of unique query labels that include in each label various attributes of the ad placement. This allows us to understand the context of subsequent queries seen at the measurement program's authoritative servers.

Some findings concerning DNS behaviours will be presented, illustrating the application of this technique to various DNS measurement questions