Doctoral Dissertation

Academic Year 2013

Client Based Naming

Graduate School of Media and Governance

Keio University

William Carl Manning

Abstract

The DNS has been described in the literature as the most successful distributed naming system that has ever been designed. However, the limitations of the fundamental design premises are increasingly difficult to work around, particularly the assumption that certain nodes are always reachable. Mobility and transient connectivity are becoming the standard for nodes in the Internet. In general, the DNS as implemented is constructed as a "soft-fail" service, with authoritative node replicas, caching, etc. However, changes to the DNS itself and the changes to the Internet infrastructure are degrading the robustness and reachability of parts of the system at the same time that client expectations about resolvability are rising.

This work documents how to eliminate the inherent reliance of DNS on fixed third-party servers by using a solution I call Client Based Naming (CBN) which operates optimally in fluid environments, including self-organizing networks. To eliminate reliance on reachability, the re-implemented DNS uses existing technologies in three new ways as follows:

- Using configuration changes, i.e., placing the node at the root of the DNS hierarchy for query purposes,
- Using multicast in queries to perform service discovery, and
- Using cryptographic techniques for name discrimination.

These techniques, plus additional identity credentials, form the basis of a globally persistent name that is usable inside and outside the DNS. Persistent names are critical for service delivery since the nodes' location (based on IP address)

changes when the node or network is moved in the Internet. Persistent names can be constructed using the DNS namespace without the need to migrate to Distributed Hash tables (DHT) or crypto-hash based names for resolving ambiguity or to retain a persistent name. Research has shown that multicast transport for DNS is a commercially viable tactic as long as scoping is restricted to local scope. Concerns still remain about ambiguous naming, resolution and robust crypto key distribution, but this discourse lays out a path for future work.

Challenges to DNS Scaling

The Domain Name System (DNS) is a distributed, coherent, reliable, autonomous, hierarchical database, designed for converting strings (hostnames and domain names) into Internet Protocol (IP) addresses on the Internet or on local networks that use the TCP/IP protocol. It is often considered to be the most successful distributed naming system that has ever been designed. Much of the commercial success of the Internet depends on a reliable and robust DNS service being available to the millions of nodes in the Internet. Despite all of its success in its present form, the existing implementation suffers from a potentially fatal flaw that is directly tied to original implementation assumptions and deployment choices. The flaw is the inherent dependency on an "always on, always connected" set of nodes that hold the apex of the DNS namespace known as the root of the DNS. This apex and each of the cut points in the ephemeral namespace are coded as "zone files". In an environment where both nodes and networks are only transiently connected and are mobile, it is increasingly more difficult to manage robust access to the authoritative servers for the root and other parts of the DNS namespace. Coupling these mobility and connectivity considerations with the fact that infrastructure operators are increasingly manipulating access to and through their networks suggests that it is not enough to focus on just implementation considerations or protocol modifications or underlying transport

constraints. A successful evolutionary path will focus on the needs of the client system in such an environment and will address each of these concerns.

A successful evolution of any naming system must take these three attributes into consideration. Client Based Naming (CBN) moves the origin of the name resolution process from someplace far removed from the Client node to the Client node itself. With the origin now anchored at the Client it is now possible for the client to have reliable name resolution when the client is mobile, or is in disconnected or administratively controlled topologies and cannot reach the traditional Internet root apex nodes. This shift in focus does take into consideration those three attributes, making it a good candidate path for the evolution of the Internet naming system.

This thesis will show that this flaw, i.e., a dependence on third parties, can be mitigated by 1) moving the origin of the root lookup to the local nodes, 2) using multicast for DNS service discovery instead of a list of "well-known" addresses, and by 3) using additional credentials to mitigate spoofing. Changes in the scope of lookup minimize the adverse affects of transport manipulation or interference. The resultant changes remove the modified node from the dependence on access to the traditional root context, thus bypassing the original implementation flaw.

1.1 Existing DNS implementation

To date recently the root zone of the Domain Name System (DNS) has enjoyed the following two important stabilizing properties:

- **It is relatively small**. Currently, the root zone holds delegation information for 280 generic, country-code, and special-purpose top-level domains (TLDs); the size of the root zone file is roughly 80,000 bytes
- **It changes slowly**. On an average, the root zone absorbs fewer than one change per TLD per year, and the changes tend to be minor

The root system has therefore evolved in an environment in which information about a small number of familiar TLDs remains stable for long periods of time. However, the type, amount, and volatility of the information that is contained in the root zone are expected to change as a result of the following four recent or pending policy decisions by U.S. Department of Commerce (DoC) and the Internet Corporation for Assigned Names and Numbers (ICANN):

- Support for DNS security (DNSSEC), or "signing the root"
- The addition of "internationalized" top-level domain names (IDN TLDs)
- Support for the additional larger addresses associated with Internet Protocol version 6 (IPv6)
- The addition of an unbounded number of new TLDs

These four changes—and the changes in the type, amount, and volatility of the information that they will incur—will unfold against the backdrop of an Internet infrastructure that is fundamentally changing from a tethered, "always on and connected" status to one where 1) nodes are mobile and only connected at random and at sporadic intervals, and where 2) infrastructure operators are manipulating access including providing locally tailored views of the DNS namespace. This change in the dynamics of the nodes and infrastructure interference removes a third attribute of a stable DNS that was the presumption of a common transport protocol with well-defined constraints on a limited set of physical media.

1.2 Core design principles

The first core DNS principle is maximum reachability. DNS has been designed so that queries and responses have the greatest chance of survival and broadest reachability by utilizing an IPv4 default UDP packet size of 512 bytes for the initial bootstrapping. This is also known as the priming query. (ICANN Security

and Stability Advisory Committee, 2008)[999]. Larger packet sizes are supported and TCP was defined and used as an alternate transport protocol, but traditionally only to be infrequently used, primarily for zone transfers. For queries and their replies, the best chance for successful delivery is to use the IPv4 UDP packet size.

With this core principle intact, the DNS has been able to evolve successfully into a highly decentralized dynamic system. The geographic and organizational decentralization of the root system arose from a deliberate design decision in favor of diversity and minimal fate-sharing coordination. This second core principle confers substantial stability and robust benefits on the global Internet but at some cost. The ability to decentralize—both geographically and organizationally—has lead to diffused spans of control for the various delegation points in the DNS namespace hierarchy. With this diaspora of control and segmentation of data elements creates incentives to have a well-connected mesh so that packets can reach across infrastructure to reach authoritative services.

1.3 Threats to Core Principles

Simple quantitative extrapolation from a baseline model of the current DNS does not predict realistic future states of the system beyond the very short term. This is because each part of the system adapts in different ways to changes in the quantity, type, and update frequency of information, while also responding to changes in the rest of the Internet. These adaptations are not and, cannot be effectively coordinated. And for some, if not all, of the actors (e.g., operators of authoritative and caching servers, telecoms operators and other transport operators), non-quantifiable considerations dominate their individual adaptation behavior both strategically (in a planning context) and tactically (in an operations context).