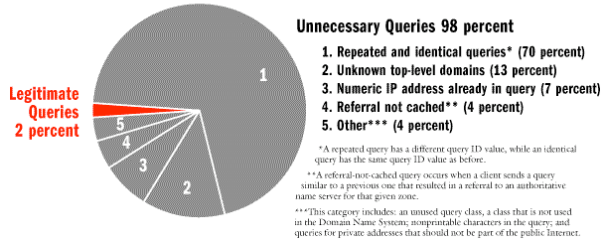


Macroscopic Internet Data Collection and Analysis



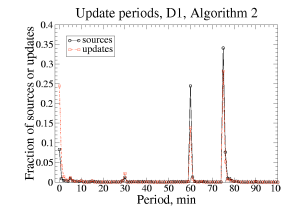
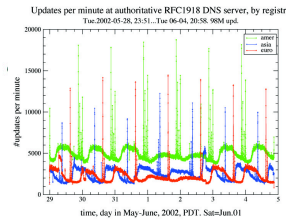
Summary of the types of queries received on Oct. 4, 2002 by a Domain Name System (DNS) root server in California



Analysis of DNS root name server queries reveals large amounts of unnecessary and illegitimate traffic.

Impact

- Discovered that only a small portion of queries to a DNS root nameserver were legitimate. Categorized the unnecessary traffic and discovered that much of it is the result of simple misconfigurations of firewalls and poorly chosen defaults for vendor-specific DNS implementations.
- Matched periodicity of undesirable, spurious private (RFC1918) DNS updates to documented behavior of Windows2000 and WindowsXP DNS.
- Analyzed DNS root and gTLD performance during the October 2002 DNS DoS attack.
- Provided real data for the NMS Integrated Demo: Supplied packet level traces to ISI for use in their RAMP tool. Also provided flow summary data to GeorgiaTech to model background traffic not modeled by RAMP.
- Customized CoralReef passive monitoring tools to provide network performance data for potential use by DoD in the ForceNET Limited Objective Experiment (LOE).



New Ideas

- Devised a methodology for simulating the Domain Name System in order to test different caching DNS nameserver implementations.
- Used binary auto-correlation algorithm to discover periodic behavior within undesirable private RFC1918 DNS updates to the root nameservers.
- Refined Internet routing geopolitical analysis and visualization, especially for exchange (IX) and peering points.
- Enhanced automated analysis of DNS root, gTLD, and ccTLD nameserver performance using NeTraMet passive monitors.

Deploy additional skitter, CoralReef, and NeTraMet monitors	Establish macroscopic topology archive for community access to data	Develop prototype hyperbolic viewer that can handle a million nodes or more	Refine methodology for identifying 'core' Internet nodes, prefixes, Autonomous Systems, or geographic regions	Make data analysis publicly available (passive, workload, routing) on website	Publish macroscopic topology data (Internet Topology Data Kit)	Develop methodology for identifying critical infrastructure hot-spots, exchange points, and other central resource locations	Expand/refine formatting and size/type of data in response to community feedback	Develop analysis and visualization tools to depict differences in routing table and real-time routing dynamics.	Build prototype model of peering points	Complete analysis and visualization of Internet core using continuously gathered skitter data	Perform additional DNS simulations with new experimental conditions of interest. For example: a) Allow multiple IP addresses per hostname; b) Use some GNAMEs, perhaps across domains; c) Force some 'lame delegations'.	Test Windows2003's DNS implementation to see how it performs.	Repeat study of private RFC1918 traffic to see whether conditions have changed since 2002.
Jul 2001	Dec 2001	Jul 2002	Dec 2002	Jul 2003	Dec 2003	Jul 2003	Dec 2003	Jul 2003	Dec 2003	Jul 2003	Dec 2003	Jul 2003	Dec 2003

