SOFTWARE ENGINEERING RADIO

Editor: **Robert Blumen**
SalesForce
robert@robertblumen.com

Bert Hubert on the Domain Name System

Gavin Henry

From the Editor
Bert Hubert, author of the open source PowerDNS nameserver, discusses Domain Name System (DNS) security and all aspects of DNS. Host Gavin Henry spoke with Hubert about what DNS is as well as its history, attacks, flaws, privacy, encryption, integrity, trust, defending against query failure, DNS over HTTPS, DNS over Transport Layer Security (TLS), DNS Security Extension (DNSSEC), HTTP/2 with DNS, Quick User Datagram Protocol (UDP) Internet Connections, monitoring DNS traffic, cache poisoning, amplification attacks, UDP attacks, URL fetch attacks via Twitter, browser DNS lookups, social-engineering attacks, and what you need to worry about DNS as a software engineer. We provide summary excerpts below; to hear the full interview, visit http://www.se-radio.net or access our archives via RSS at http://feeds.feedburner.com/se-radio.—*Robert Blumen*

Gavin Henry: What is a DNS?

Bert Hubert: When you type a name in the Internet, the computer cannot directly connect. It must look up the Internet Protocol (IP) address of that website—IPV4 or IPV6—which DNS provides. A lot happens to make DNS work well. Browsers ensure that their queries get answers quickly, because speed affects the user experience.

It's an old protocol that we rely on, and it's tricky. There is a difference between "this name doesn't exist" and "the name exists, but it doesn't have an IPV6 address." The latter says the name does exist, but the thing you asked for doesn't. But often, you get back the answer, "the name doesn't exist," or "we tried to resolve this name, but we it didn't work." This has led to websites disappearing from the web because a load balancer messed up this nuance.

What else is stored in a DNS?

DNS also denotes where the mail server for a domain is located, as an example. That's the mail exchanger (MX) record. If you have a network with many subscribers and you're worrying about bad behavior, such as servers that have been hacked, you can check in DNS how many MX domain queries the server is doing because hacked servers will attempt to spam widely. Then, you see tons of MX records.

More frequent than MX records are text records that store arbitrary bits of text in DNS. These are used by many spam lists to identify IP addresses known to be spamming or domain names known to be phishing. Many mail servers perform DNS text-record lookups to figure out if a sending mail server is known to spam. With all its faults and its age, DNS is still the one technology that

Digital Object Identifier 10.1109/MS.2020.3000883
Date of current version: 20 August 2020

functions as a low-investment worldwide distributed database; you can ask it many questions per second and get good answers.

What types of DNS servers are there?

There are authoritative name servers—those that say, “I know all about some website, and this is the IP address of that website.” Queries on the Internet start with root servers: more than 1,000 servers with 26 IP addresses. You can query these 26 IP addresses and get an answer. Root servers don’t know a lot, but they can give you a list of IP addresses for the dot-com servers. You can query the dot-com servers for an IP address, and a dot-com server will say, “I don’t know, but I do know where the name servers are for that website. Ask over there, and here are their IP addresses.” And then you end up with this authoritative server that I first mentioned, and that server can actually tell you the right IP address.

This process was originally thought to be too much work for simple computers. So a separate server was invented called a *resolver* (or a *cache*), which would traverse these authoritative servers, starting from the root server, to find the answer to your question. Thus, in your browser, you type the domain name you want to visit. The browser sends the query to its configured resolver that is part of your local network. And the resolver trolls the whole Internet and consults authoritative servers until it finds the right answer.

The resolvers come with a hints file that contains a list of about 26 IP addresses, which might be outdated if you haven’t turned the computer on for a few years. At startup, a resolving name server tries these hints records to find a name server that answers. Chances are that the resolver will find one IP address that works, and from there, it can learn the latest list of 26 IP addresses.

SOFTWARE ENGINEERING RADIO

Visit www.se-radio.net to listen to these and other insightful hour-long podcasts.

RECENT EPISODES

- 413—Spencer Kimball, cofounder and CEO of Cockroach Labs, talks with host Akshay Manchale about CockroachDB, an open source distributed database system.
- 412—Host Felienne speaks with Gavis-Hughson about how to prepare for the dreaded “whiteboard interview,” and how to get better at doing interviews.
- 409—Joe Kutner, software architect for Heroku at salesforce.com, discusses the 12-factor app with host Kanchan Shringi.

UPCOMING EPISODES

- Adam Shostack discusses threat modeling with host Justin Beyer.
- Host Adam Bell talks about incident response with Berkay Mollamustafaoglu.
- Vladimir Khorikov discusses functional programming in enterprise applications with host Jeremy Jung.

What about integrity and encryption with a DNS?

DNSSEC adds integrity to a DNS; if you have an answer, it can validate the answer. But applications that use a DNS typically don’t validate DNS answers. Even browsers simply query a name server and trust the answer. It’s technically possible for a browser to validate DNS answers using DNSSEC, and many domain names are now DNSSEC signed. But browsers have decided to put trust in TLS certificates instead.

TLS provides transport security. It validates that the data you received are what you should have received. DNSSEC works more like Pretty Good Privacy or secure Multipurpose Internet Mail Extension programs—it signs the final answer to your question, which is the actual name/IP-address combination. That has made DNSSEC difficult to implement.

The idea behind TLS is that you trust the certificate’s final verification of HTTPS. If you trust HTTPS for message integrity or transport integrity, you no longer care about the IP address because the transport has integrity. The common view is that you don’t need to worry about DNSSEC if you’re encrypting. But there have been brief hijacks of the DNS name in which people have been able to get new TLS certificates during the hijack. That is where DNSSEC validation would have helped.

What about privacy?

DNS traffic tells you everything about a person. If you have access only to someone's DNS records, you will be able to tell where they live, what phone they have, or what brand of TV they have, and so forth. Per bit, DNS may be the most privacy-sensitive material on the Internet. It is worthwhile to protect DNS records because we don't want too many people to have access to those.

If someone has a Tesla and I know that they work at a certain company and which sports team they like, that narrows down a search for a person. All this information radiates from that person's Internet connection through a DNS.

A common current view is that we should encrypt DNS wherever we can and then send it to a new third party, which then gets access to all your browsing data. I'm not convinced that that is actually progress. Previously, we had a highly regulated telecommunications company, at least here in Europe, that knew what sites you are visiting. But it's not progress if now some American company knows what sites you are visiting. However, the largest Internet service providers are now either offering or developing fully encrypted DNS. So there is not much benefit anymore in sending all encrypted data to a third party that you did not select.

What one thing should a software engineer remember?

Monitor DNS. When it breaks, everything will break, so it is worth adding diagnostics. If you have long-lived applications, the DNS answer that you got at the start-up of the application that you cached from weeks ago may no longer be the right IP address. DNS is more dynamic than it used to be.

ABOUT THE AUTHOR

GAVIN HENRY is the founder and managing director of SureVoIP, an Internet telephony service. Contact him at ghenry@surevoip.co.uk.

IEEE Software (ISSN 0740-7459) is published bimonthly by the IEEE Computer Society. IEEE headquarters: Three Park Ave., 17th Floor, New York, NY 10016-5997. IEEE Computer Society Publications Office: 10662 Los Vaqueros Cir., Los Alamitos, CA 90720; +1 714 821 8380; fax +1 714 821 4010. IEEE Computer Society headquarters: 2001 L St., Ste. 700, Washington, DC 20036. Subscribe to *IEEE Software* by visiting www.computer.org/software.

Postmaster: Send undelivered copies and address changes to *IEEE Software*, Membership Processing Dept., IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854-4141. Periodicals Postage Paid at New York, NY, and at additional mailing offices. Canadian GST #125634188. Canada Post Publications Mail Agreement Number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8, Canada. Printed in the USA.