CIS 457 - Data Communications

Nathan Bowman

Images taken from Kurose and Ross book

DNS

Every host on internet needs some unique way to refer to it. Otherwise, how could we specify where to send messages?

Humans use **hostnames** to refer to hosts

For example, `www.gvsu.edu`

Easy for us to remember, but has drawbacks for computers

Hostnames can be of variable length

`www.amazon.com`, `mail.exchange.gvsu.edu`, etc.

Also, hostname gives no information about where in network host resides

Computers refer to hosts by **IP address**

Solves both issues: fixed length and gives information about location

IP address is 4 bytes, represented as numbers from 0-255

XXX . XXX . XXX . XXX

Address is hierarchical from left to right -- each byte gives more specific information about where host resides

When you go to `www.google.com`, how does web broswer know IP address for TCP connection?

Domain Name System (DNS) is distributed database containing translations from hostnames to IP addresses

DNS refers to both database and application-layer protocol used to access information in database

Other application-layer protocols rely on DNS. For example:

Visiting `www.gvsu.edu`

Emailing `fakestudent@mail.gvsu.edu`

In either case, browser must first make DNS query

DNS provides a few other services in addition to hostname -> IP mapping

- host aliasing
- mail server aliasing
- load distribution

Host aliasing -- allows single host to have more than one name

Typically used to give more mnemonic alias to canonical hostname

Canonical name:

```
relay1.west-coast.enterprise.com
```

Alias:

```
www.enterprise.com
```

Mail server aliasing -- allow mail server to share hostname with one other server

Visiting `yahoo.com` and emailing `bob@yahoo.com` require two different addresses for `yahoo.com` -- a web server and a mail server

Mail server aliasing is common enough and convenient enough that this exception is built right in to DNS

Load distribution -- spread traffic among various duplicate servers

Most large websites have more than one duplicate web server to handle traffic

DNS returns IP addresses of all servers, but shuffles order

Client will generally choose first IP address returned, so traffic is spread among servers