CSCD 330
Network Programming
Winter 2020

Lecture 5
Application Layer

Reading: Chapter 2 – Still

Some Material in these slides from J.F Kurose and K.W. Ross
All material copyright 1996-2007
More Network Applications

• Looked at HTTP Web applications
  • HTTP – Very simple protocol
  • Text or Ascii based
• Today
  • SMTP – POP3 and IMAP
    • Simple Mail Transport Protocol
    • Pop3 and Imap
  • DNS – A very important part of Internet
    • Implemented as an application layer protocol
Chapter 2: Application layer

• Principles of network applications
• Web and HTTP
• FTP (Skip)
• Electronic Mail
  • SMTP, POP3, IMAP
• DNS

Stop Here

• Socket programming with TCP
• Socket programming with UDP
Internet Mail

- E-mail has been around since the beginning of the Internet
  - One of the first and most popular applications
  - Gotten more complex over time
    - Began as text ... Now
    - Send different media through the Internet, pictures, video, text and HTML formatted text
- View email protocols and how they work
SMTP Underlying Mail Protocol

- Defined in RFC 2821
  - Dates back to 1982!
  - Legacy protocol
  - Restricted to ASCII 7 bit values
    - No binary data, no hyphens
  - Way around this, encode multimedia and other binary data to ASCII and decode it back

http://www.ietf.org/rfc/rfc2821.txt
Past Email Statistics

Past Email Statistics (2004):

- 31 billion emails sent daily, expected to double by 2006
- Email generates about one billion Gigabytes of new “information” per year
- Spam accounts for about 40% of all email traffic (maybe)
  - Stats on spam varies

Current Email Statistics

• In 2015, the number of emails sent and received per day total over 205 billion.
  • This figure is expected to grow at an average annual rate of 3% over the next four years, reaching over 246 billion by the end of 2019

• Email use continues to see strong use in the business world, as well as among consumers
• Consumer email continues to grow mainly due to its use for notifications (e.g. for online sales) rather than simply as an interpersonal communication tool

Statistics from marketing business group, Radicati
http://www.radicati.com/
Projected Email Statistics - 2019

Email Statistics Report, 2015-2019 – Executive Summary

<table>
<thead>
<tr>
<th>Daily Email Traffic</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Worldwide Emails Sent/Received Per Day (B)</td>
<td>205.6</td>
<td>215.3</td>
<td>225.3</td>
<td>235.6</td>
<td>246.5</td>
</tr>
<tr>
<td>% Growth</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Business Emails Sent/Received Per Day (B)</td>
<td>112.5</td>
<td>116.4</td>
<td>120.4</td>
<td>124.5</td>
<td>128.8</td>
</tr>
<tr>
<td>% Growth</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Consumer Emails Sent/Received Per Day (B)</td>
<td>93.1</td>
<td>98.9</td>
<td>104.9</td>
<td>111.1</td>
<td>117.7</td>
</tr>
<tr>
<td>% Growth</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 2: Worldwide Daily Email Traffic (B), 2015-2019

Statistics from marketing business group, Radicati
http://www.radicati.com/

Other Statistics on email from Lifewire
https://www.lifewire.com/how-many-emails-are-sent-every-day-1171210
Webmail Popularity

In spite of the social network buzz, email still remains the core of business communication on the Internet

Below are the top ranked free email sites
How email works
Electronic Mail

Three major components

1. User agents
2. Mail servers
3. Simple mail transfer protocol: **SMTP**

User Agent

- Or, “mail reader”
- Composing, editing, reading mail messages
- **Example:** Eudora, Outlook, Mozilla Thunderbird, Opera
- Outgoing, incoming messages stored on server
Electronic Mail: Mail servers

Mail Servers

- **Mailbox** contains incoming messages for user
- **Message queue** of outgoing (to be sent) mail messages
- **SMTP protocol** between mail servers to send email messages
  - **Client**: Sends mail to server
  - **Server**: Receives mail from client
Electronic Mail: SMTP

• Which transport protocol is used for SMTP?
• Uses **TCP** to transfer email message from client to server,
  • **Port 25** or alternate **port 587**
• **Direct transfer:** Sending server to Receiving server

• **Three phases of transfer**
  • Handshaking (greeting)
  • Transfer of messages
  • Closure

• **Command/response Interaction**
  • **Commands:** ASCII text
  • **Response:** Status code and phrase
Scenario: Alice sends message to Bob

1) Alice uses User Agent (UA) to compose message
   The “to” is bob@someschool.edu

2) Alice’s UA sends message to her mail server; message placed in message queue

3) Client side of SMTP opens TCP connection with Bob’s mail server

4) SMTP client sends Alice’s message over TCP connection

5) Bob’s mail server places message in Bob’s mailbox

6) Bob invokes his (UA) user agent to read message
Sample SMTP Interaction

telnet cmu.edu 587
S – Mail server C – Client
cmu.edu ewu.edu

S: 220 cmu.edu
C: HELO ewu.edu
S: 250 Hello ewu.edu, pleased to meet you
C: MAIL FROM: alice@ewu.edu
S: 250 alice@ewu.edu... Sender ok
C: RCPT TO: bob@cmu.edu
S: 250 bob@cmu.edu ... Recipient ok
C: DATA
S: 354 Enter mail, end with "." on a line by itself
C: How are you?
C: Want to get together this Saturday?
C: Call me.
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 cmu.edu closing connection
Mail Message Format

Another RFC defines the Actual message format
RFC 822: standard for text Message format:

• Header lines, e.g.,
  Required
  • To:
  • From:
  Optional
  • Subject:
  • Body
  • the “message”, ASCII characters only
Sending multimedia through email

• E-mail designed for text.
• How do you send multimedia through the email system?
  • Pictures, video etc.
Message Format: Multimedia Extensions

- **MIME: Multipurpose Internet Mail Extension**, MIME is specified in six linked RFC documents RFC 2045, RFC 2046, RFC 2047, RFC 4288, RFC 4289 and RFC 2049, which together define the specifications.
- Additional lines in message header declare MIME content type

```
From: alice@ewu.edu
To: bob@cmu.edu
Subject: Picture of yummy crepe.
MIME-Version: 1.0
Content-Transfer-Encoding: base64
Content-Type: image/jpeg

base64 encoded data ....
........................
......base64 encoded data
```
More MIME Encoding

• Why do we need MIME at all?
• Binary data – takes up 8 bits in a byte
• Ascii data no extensions – takes up 7 bits
• Email servers used to be 7 bit ascii – set high bit to 0
• Could only send 7 bits of data reliably
• MIME extensions recodes the 8 bit binary into 7 bit ascii values so it makes it to the other end without being altered
• Other end, MIME gets decoded back to binary
Base 64 bit Encoding

- **Steps**
  1. a-z – 26 characters
  2. A-Z – 26 characters
  3. 0 – 9 – 10 characters
  4. +,/- – 2 more chars
  5. = padd

Read stream, divide into 3 8-bit groups

3 groups become 24 bits, divide into 4 groups of 6 bits

Map these to characters after shifting and padding

ORACLE - > T1JBQ0xF (64 bit encoded), fit in 7 bits

**Nice explanation below**

https://blogs.oracle.com/rammenon/entry/base64_explained
Email Headers

• Can tell a lot about email by the headers
• If it is legitimate, where it came from, its path ...
• Spam is another story
• Link below is a tutorial on reading email headers
  https://pepipost.com/blog/how-to-read-email-headers-to-identify-spam/
• Try it out on some of your email
• First, need to see entire header
• All web email allows you to see entire header
  – Gmail, Yahoo, Comcast etc.
Mail Access Protocols

- **SMTP**: Simple delivery/storage to receiver’s server
- **Mail access protocols**: Allow retrieval from server
  - **POP**: Post Office Protocol [RFC 1939]
    - Performs authorization (agent <-> server) and download
  - **IMAP**: Internet Mail Access Protocol [RFC 1730]
    - More features (more complex)
    - Manipulation of stored messages on server
  - **HTTP (Web)**: gmail, Hotmail, Yahoo! Mail, etc.
(a) Sending and reading mail when the receiver has a permanent Internet connection and the user agent runs on the same machine as the message transfer agent  
(b) Reading e-mail when the receiver has a dial-up connection to an ISP.
Email Examples

Can look at some examples … with full headers

Another reference for Mime encoded emails:

http://webmarketingtoday.com/articles/html-email-multi/
DNS – Domain Name System
DNS Overview

• **Domain Name System (DNS)** associates information with domain names
• Serves as "phone book" for Internet
• Translates human-readable hostnames into IP addresses
• Also stores other information
• Such as the list of mail exchange servers that accept email for a given domain
• **DNS** is an essential component of the modern Internet
DNS: Domain Name System

Internet hosts, routers:
- IP address
  - Addresses of datagrams
  - Network level

Computer’s prefer this
IP = 87.248.113.14

But, humans prefer this
Domain name - www.yahoo.com

Runs over Port 53
Uses UDP to request/receive messages

Why UDP?

You type: www.yahoo.com
DNS resolves this to: 87.248.113.14
History

- Paul Mockapetris, a program manager at ARPA, invented Domain Name System in 1983 and wrote first implementation.
  
  Original specifications appear in RFC 882 and 883.

- Mockapetris recognized problems with early Internet system of holding name to address translations in a single table on a single host (HOSTS.TXT).

- Several more-recent RFCs have proposed various extensions to the core DNS protocols.

  https://en.wikipedia.org/wiki/InterNIC
DNS

Domain Name System

- Distributed database
  Hierarchy of name servers

- Application-layer protocol
  Hosts, routers, name servers communicate to resolve names known as resolvers
  (address/name translation)

- Core Internet function, implemented as application-layer protocol
DNS: Domain Name System

Why not centralize DNS?

• Single point of failure
• Traffic volume
• Distance of a centralized database from clients
• Maintenance

Not possible ... doesn’t scale!
How DNS Works - Theory

Have **Domain Names**

Arranged in a tree

- Cut into **zones**
  - Each served by a **nameserver**
  - Not all nameservers are equal – different types
DNS is Distributed and Hierarchical

• DNS has hierarchy of Nameservers

• 3 Classes of servers
  1. Root DNS servers
     • Global servers, most are in North America
  2. Top-level Domains
     • com, net, org, edu, gov, and uk, fr, jp ... other countries
  3. Authoritative servers
     • Every group with public hosts on Internet must provide accessible DNS records for IP-to-hostname mapping
     • Groups can do it themselves or pay an ISP to maintain their IP records
DNS is Distributed and Hierarchical

• How many domains and country codes?
  As of June 2019, the root domain contains 1530 top-level domains,
  • http://www.iana.org/domains/root/db
  For just the country code domains in the root domain
  http://www.iana.org/cctld/cctld.htm
  Full List in one place is at Wikipedia
  https://en.wikipedia.org/wiki/List_of_Internet_top-level_domains
Client wants IP for www.amazon.com
1. Client queries a root server to find com DNS server (root)
2. Client queries com DNS server to get amazon.com DNS server (top-level)
3. Client queries amazon.com DNS server to get IP address for www.amazon.com (authoritative)
The DNS Name Space

A portion of the Internet domain name space.
DNS: Root name servers

• Contacted by local name server can't resolve name
• Root name server
  • Contacts authoritative name server if name mapping not known
  • Gets mapping
  • Returns mapping to local name server

http://root-servers.org/  Actual map
Top Level Domains Explained

The following are the original 7 common top-level domains:

• COM -- commercial Web sites, though open to everyone
• NET -- network Web sites, though open to everyone
• ORG -- non-profit organization Web sites, though open to everyone
• EDU -- restricted to schools and educational organizations
• MIL -- restricted to the U.S. military
• INT – International organizations
• GOV -- restricted to the U.S. government

Later, country codes and other domains were added:

• US, UK, RU and other two-letter country codes -- each is assigned to a domain name authority in the respective country
SubDomains

A subdomain combines a unique identifier with domain name to become a "domain within a domain."

- Unique identifier simply replaces the www in the web address.
- Yahoo!, for example, uses subdomains
  - mail.yahoo.com and music.yahoo.com to reference its mail and music services, under the umbrella of www.yahoo.com

NOTE: Subdomains can be created "at will"
- For any domain, you can create as many subdomains as you like ... do NOT need registrar and do NOT pay an annual fee for each subdomain
Registrars Enforce Uniqueness

• Names in given domain need to be unique, there has to be some way to control the list ...

Need for Registrars
  – A registrar is an authority that can assign domain names directly under one or more top-level domains and registers them with InterNIC, a service of ICANN
  – Enforces uniqueness of domain names across Internet

• Each registration becomes part of a central domain registration database Whois database

• Part of registering a domain you must identify one or more name servers as having authority for that name

• Resolves host names in that domain
Back to History of the Internet .....
Internet Corporation for Assigned Names and Numbers, ICANN

• In November of 1998,
• Department of Commerce sponsored a proposal for creation of a Domain Naming and Management System
• ICANN is registered in California because its where Jon Postal, manager of IANA name space was located
  – IANA was started in 1970's, part of its purpose, Domain Names
    ICANN, is a private, non-profit corporation  http://www.icann.org/

ICANN Responsibilities
1. Establish and implement a procedure for registrar accreditation
2. Manage Internet Names and Numbers under IANA
   IANA is then responsible for management ...
ICANN and InterNIC

- InterNIC, short for Internet Network Information Center
  - Governing body primarily responsible for domain name and IP address allocations **until 1998**

- But, their role was assumed by ICANN in 1998

- They still have web presence and offer whois database lookup services

http://www.internic.net/whois.html
Internet Assigned Numbers Authority (IANA)

- Internet Assigned Numbers Authority (IANA)
  - Entity that oversees global IP address allocation, autonomous system number allocation, root zone management in Domain Name System (DNS), and other Internet Protocol-related symbols and numbers, has been around since 1970's
  - IANA is operated by ICANN,
  - Basically is Under ICANN
    - http://www.iana.org/
  - Specifically ...
IANA Duties

• IANA responsible for allocation of globally unique names and numbers used in Internet protocols that are published as RFC documents
• IANA delegates allocations of IP address blocks to Regional Internet Registries (RIRs)
  
  Each RIR allocates addresses for a different area of world
    – Ex: 0.0.0.0 – 10.10.10.0 - Europe
• IANA administers data in root DNS nameservers
How DNS Actually Works ......
Local Name Server

- Does not strictly belong to hierarchy
- Each ISP
  - Residential ISP, company, university has one
  - Also called default name server
- When host makes DNS query, query is sent to its local DNS server
  - Acts as proxy, forwards query into hierarchy
Local DNS from CS Department
DNS name resolution example

- Host at **ewu.edu** wants IP address for **gaia.cs.umass.edu**

**Iterated query:**
- Contacted server replies with name of server to contact
- “I don’t know this name, but ask this server”
DNS: Caching and Updating records

- Once (any) name server learns mapping, it caches mapping
  - Cache entries timeout (disappear) after some time
    - Typical to have times of 24 to 48 hours
  - TLD servers typically cached in local name servers
    - Thus root name servers not often visited
Creating a Corporate Website Example
Creation of DNS Records

• Company wants a corporate Web site
• What do you do?
  1. Research domain name
     It must be unique and not already registered
     Access whois database to see if your name is already taken.
     One popular site that offers this service, Network Solutions
     http://www.networksolutions.com/whois/index.jsp

  2. Purchase domain name and Register it
     Need to register it annually, or pay for a number of years
     Use a recognized registrar
     GoDaddy – Large Registrar
     InterNIC has complete registrar list
     http://www.internic.net/regist.html
Creation of DNS Records

- **3. Need a Web hosting company**
  
  
  - Probably don't want to go with free unless it's a hobby website
  - Hosting company will provide besides space for website, DNS nameservers
  - Paper on pros and cons of International Hosting
    

  - DNS nameservers hold IP address records to your site
  - Next few slides show these records ...
DNS Records for a Domain

• This information is kept in a zone file on the DNS server
• What is a Zone file?

The DNS Zone file is the representation of the DNS Zone - it is the actual file, which contains all the records for a specific domain.

In a DNS Zone file, each line can hold only one record, and each DNS Zone file must start with:

- TTL (Time to Live), which specifies for how long the records should be kept in the DNS Server's cache.
DNS Records for a Domain

Following are the most common types of records you can configure for your DNS server:

- **Host (A)** -- This is the basic mapping of IP address to host name.
- **Canonical Name (CNAME)** -- indicates the true, or canonical, host name of a computer that its aliases are associated with.
- **Mail Exchanger (MX)** -- This maps e-mail traffic to a specific server.
- **Name Server (NS)** -- Contains name server information for zone.
- **Start of Authority (SOA)** -- This record at beginning of every zone file with primary name server for zone and some other information. Go over it with an example.

Good Article on SOA
http://www.peerwisdom.org/2013/05/15/dns-understanding-the-soa-record/
1. Research Available Domain Names

• Whois DB’s
  • To get your domain name registered as official, must go through an approved registrar
  • For the .net, .com and .org top level domains
  • List of Registrars here
    http://www.internic.net/regist.html
  • If you needed information on specific domain, enter: http://www.internic.net/whois.html which allows a user to enter a domain name
1. Research Available Domain Names

- Whois DB’s
  - For other countries, use
    http://www.allwhois.com/home.html
  - Military sites, use
    http://whois.nic.mil
  - Education, use
    http://www.networksolutions.com
Internic.net/whois.html
Example from Internic.net/whois

Whois Server Version 1.3

Domain names in the .com and .net domains can now be registered with many different competing registrars. Go to http://www.internic.net for detailed information.

Domain Name: PHPTR.COM
Registrar: REGISTER.COM, INC.
Whois Server: whois.register.com
Referral URL: http://www.register.com
Name Server: USRXDNS1.PEARSONTC.COM
Name Server: OLDTXDNS2.PEARSONTC.COM
Status: REGISTRAR-LOCK
Updated Date: 21-Jan-2005
Creation Date: 04-Jul-1997
Expiration Date: 03-Jul-2006

>>> Last update of whois database: Thu, 3 Feb 2005 08:01:23 EST <<<

NOTICE: The expiration date displayed in this record is the date the registrar's sponsorship of the domain name registration in the registry is scheduled to end. This is the date the domain is available for redisc...
Other info Sites

http://www.dnsstuff.com
  • Whois services and other services for testing your dns paths

http://www.zoneedit.com/lookup.html
  • Lookup services – enter a specific host
2. Purchase Domain Name and Register it Use Web Hosting

• At this stage, you have identified a unique name and bought it
  • networkutopia.com

• In the process you have registered it with ISP service with DNS Servers

3. Web Hosting Service to host web site
Inserting Records into DNS

- Example: new startup **Network Utopia**
- Register name **networkutopia.com**
  - **DNS registrar** (e.g., Network Solutions)
  - Provide names, IP addresses of authoritative name server(s)
- The Registrar inserts two Resource Records into .com TLD server:
  - (networkutopia.com, dns1.networkutopia.com, NS)
  - (dns1.networkutopia.com, 212.212.212.1, A)

- **Create authoritative server records**
  - Type A record for **www.networkutopia.com**
  - Type MX record for **networkutopia.com**
Other DNS Information Lookup

- In addition to Web sites where you can lookup information, another way is to run commands from your computer.
- Built in to Linux and Mac OS, Windows - it varies.
- Command below allow lookup of IP Addresses of sites.

Two commands

- nslookup – from most hosts, Windows and Linux.
- Dig, host – Linux and Mac.

Two tutorials on using Dig

http://kb.mediatemple.net/questions/909/Understanding+the+dig+command
http://www.madboa.com/geek/dig/
What is Dig?

`dig` is a command-line tool for querying DNS name servers for information about host addresses, mail exchanges, name servers, and related information.

You can do this from the comfort of your own computer ... if you have Linux or Mac.

However, you can install it in Windows 7

https://www.danesparza.net/2011/05/using-the-dig-dns-tool-on-windows-7/

Installing `dig` in Windows 10

http://nil.uniza.sk/linux-howto/how-install-dig-dns-tool-windows-10
NSLookup

• Examples

ctaylor@ctaylor-lt:~$ nslookup penguin.ewu.edu
Server: 146.187.134.22
Address: 146.187.134.22#53

Non-authoritative answer:  
Name: penguin.ewu.edu
Address: 146.187.134.7

To look up just the mail servers:
ctaylor@ctaylor-lt:~$ nslookup
> set q=mx
> ewu.edu

What does this mean?
• Dig Example

c台账@ctaylor-lt:~$ dig ewu.edu any
; <<>> DiG 9.4.2-P2.1 <<>> ewu.edu any
;; global options: printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 10555
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 9, AUTHORITY: 5, ADDITIONAL: 4

;; QUESTION SECTION:
;ewu.edu.           IN   ANY

;; ANSWER SECTION:
ewu.edu. 86400 IN TXT "v=spf1 "ip4:146.187.218.101" "ip4:146.187.218.102" "ip4:146.187.218.103" "include:hotmail.com"
    "include:spf.messaging.microsoft.com" "~all"
ewu.edu. 1800 IN MX 100 mail.messaging.microsoft.com.
ewu.edu. 86400 IN SOA dns1.ewu.edu. networks.ewu.edu. 17384 900 900 604800 7200
ewu.edu. 86400 IN NS ruler.wa-k20.net.
ewu.edu. 86400 IN NS apple.wa-k20.net.
ewu.edu. 86400 IN NS dns3.ewu.edu.
ewu.edu. 86400 IN NS dns2.ewu.edu.
ewu.edu. 86400 IN NS dns1.ewu.edu.
The End