PART XXII

BOOTSTRAP AND AUTOCONFIGURATION (DHCP)

System Startup

- To keep protocol software general
  - IP stack designed with many parameters
  - Values filled in when system starts
- Two possible sources of information
  - Local storage device (e.g., disk)
  - Server on the network

Bootstrapping

- BOOTstrap Protocol (BOOTP)
  - Early alternative to RARP
  - Provided more than just an IP address
  - Obtained configuration parameters from a server
  - Used UDP
- Dynamic Host Configuration Protocol (DHCP)
  - Replaces and extends BOOTP
  - Provides dynamic address assignment

Apparent Contradiction

- DHCP used to obtain parameters for an IP stack
- DHCP uses IP and UDP to obtain the parameters
- Stack must be initialized before being initialized

Solving The Apparent Contradiction

- DHCP runs as application
- Only needs basic facilities
- In particular:

An application program can use the limited broadcast IP address to force IP to broadcast a datagram on the local network before IP has discovered the IP address of the local network or the machine's IP address.

- Note: server cannot use ARP when replying to client because client does not know its own IP address

Two-Step Bootstrap

- DHCP provides information, not data
- Client receives
  - Name of file that contains boot image
  - Address of server
- Client must use another means to obtain the image to run (typically TFTP)

Dynamic Address Assignment

- Needed by ISPs
  - Client obtains an IP address and uses temporarily
  - When client finishes, address is available for another client
- Also used on many corporate networks

DHCP Address Assignment

- Backward compatible with BOOTP
- Can assign addresses in three ways
  - Manual (manager specifies binding as in BOOTP)
  - Automatic (address assigned by server, and machine retains same address)
  - Dynamic (address assigned by server, but machine may obtain new address for successive request)
- Manager chooses type of assignment for each address

Summary

- Two protocols available for bootstrapping
  - BOOTP (static binding of IP address to computer)
  - DHCP (extension of BOOTP that adds dynamic binding of IP addresses)
- DHCP
  - Server grants lease for an address
  - Lease specifies length of time
  - Host must renew lease or stop using address when lease expires
  - Actions controlled by finite state machine


Questions?

PART XXIII

DOMAIN NAME SYSTEM (DNS)

Names For Computers

- Humans prefer pronounceable names rather than numeric addresses
- Two possibilities
  - Flat namespace
  - Hierarchical namespace

Naming Hierarchy

- Two possibilities
  - According to network topology
  - By organizational structure (independent of physical networks)
- Internet uses the latter

Internet Hierarchy

In a TCP/IP internet, hierarchical machine names are assigned according to the structure of organizations that obtain authority for parts of the namespace, not necessarily according to the structure of the physical network interconnections.

Internet Domain Names

- Flexible hierarchy
  - Universal naming scheme (same everywhere)
  - Each organization determines internal naming structure
- Mechanism known as *Domain Name System* (*DNS*)
- Name assigned to a computer known as *domain name*

Domain Name Syntax

- Set of *labels* separated by delimiter character (period)
- Example

cs.*purdue*.*edu*

- Three labels: *cs*, *purdue*, and *edu*
- String *purdue*.*edu* is also a domain
- Top-level domain is *edu*

Original Top-Level Domains

Domain Name	Assigned To
com	Commercial organizations
edu	Educational institutions (4-year)
gov	Government institutions
mil	Military groups
net	Major network support centers
org	Organizations other than those above
arpa	Temporary ARPANET domain (obsolete)
int	International organizations
country code	Each country (geographic scheme)

- Meaning assigned to each
- Three domains considered generic

  .com

  .net

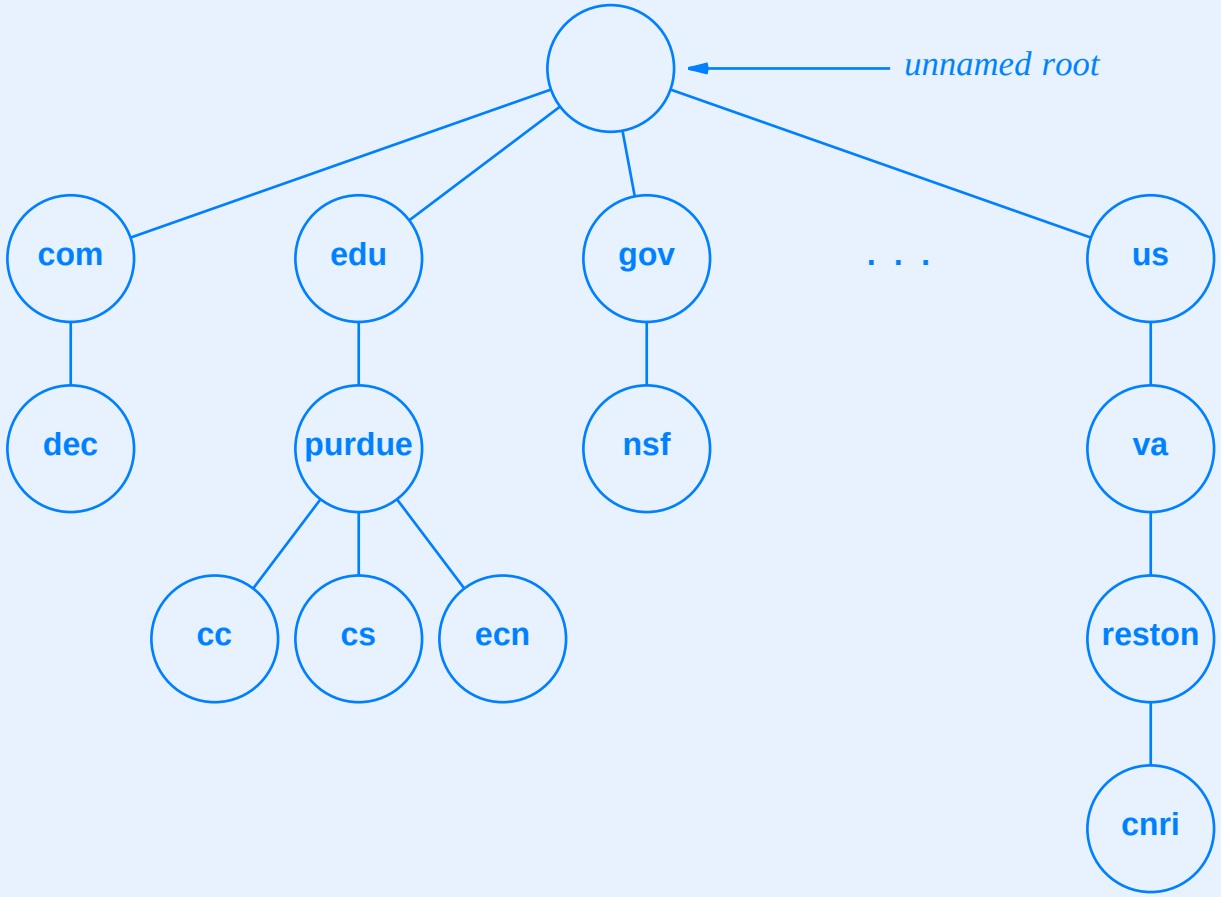
  .org

New Top-Level Domains

Domain Name	Assigned To
aero	Air-Transport Industry
biz	Businesses
coop	Non-Profit Cooperatives
info	Unrestricted
museum	Museums
name	Individuals
pro	Professionals (accountants, lawyers, physicians)

- Proponents argued (incorrectly) that DNS would collapse without additional TLDs
- New TLDs created legal nightmare

Illustration Of Part Of The DNS Tree



Authority For Names

- Authority delegated down the tree
- Example
  - Purdue University registers under top level domain *.edu* and receives authority for domain *purdue . edu*
  - Computer Science Department at Purdue registers with the Purdue authority, and becomes the authority for *cs . purdue . edu*
  - Owner of a lab in the CS Department registers with the departmental authority, and becomes the authority for *xinu . cs . purdue . edu*

DNS Database

- Record has (name, class)
- Class specifies type of object (e.g., computer, email exchanger)
- Consequence:

A given name may map to more than one item in the domain system. The client specifies the type of object desired when resolving a name, and the server returns objects of that type.

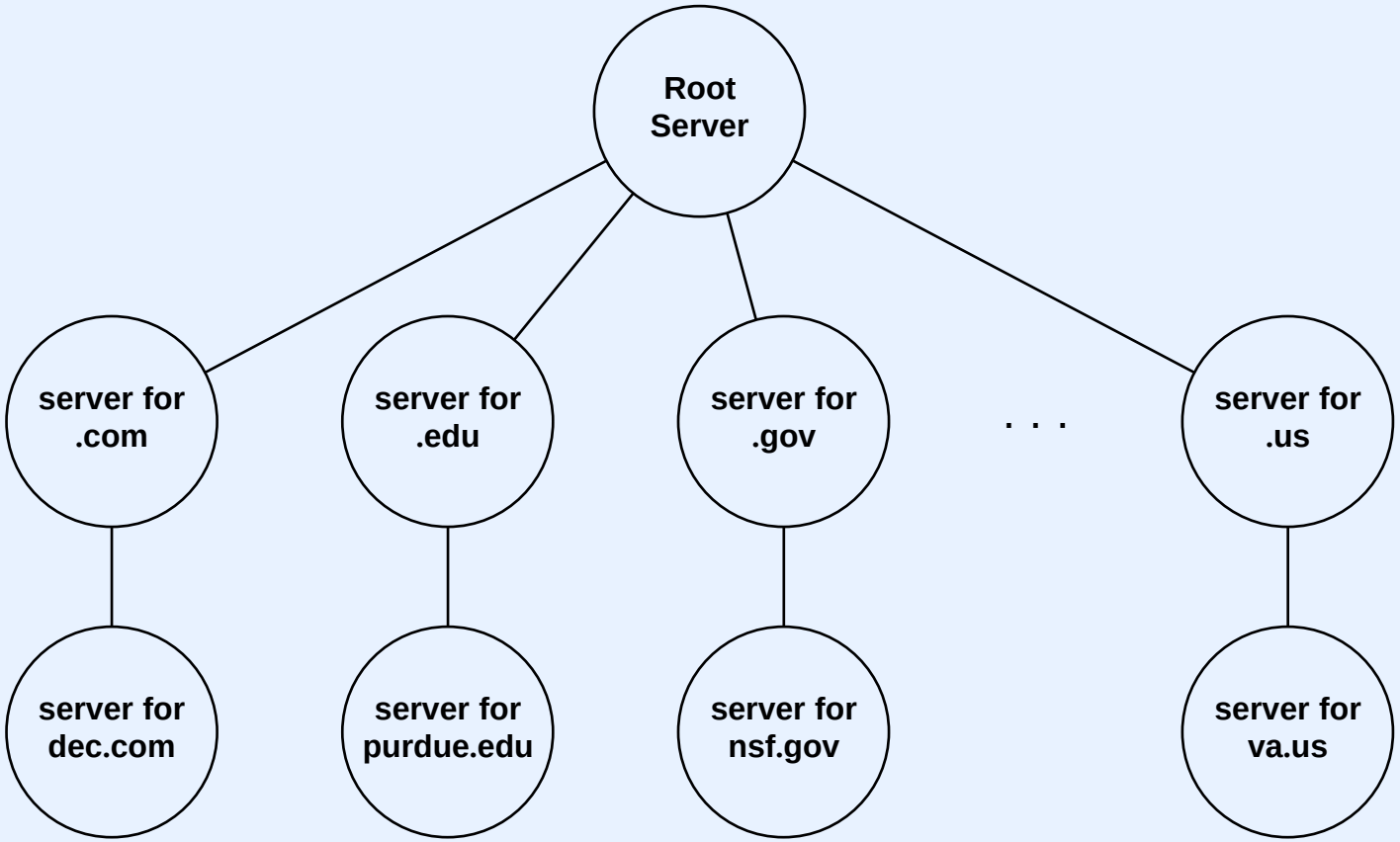
Mapping Domain Names To Addresses

- DNS uses a set of on-line servers
- Servers arranged in tree
- Given server can handle entire subtree
  - Example: ISP manages domain names for its clients (including corporations)

Terminology

- DNS server known as *name server*
- DNS client software known as *resolver*

Illustration Of Topology Among DNS Servers



In Practice

- Single server can handle multiple levels of the naming tree
- Example: root server handles all top-level domains

Domain Name Resolution

- Conceptually, must search from root of tree downward
- In practice
  - Every name server knows location of a root server
  - Only contacts root if no subdomain known
  - Lookup always starts with local server first (host can learn address of DNS server from DHCP)

Abbreviation Of Domain Names

- DNS only recognizes full domain names
- Client software allows abbreviation

Example Of Domain Name Abbreviation

- Client configured with suffix list
  - .cs.purdue.edu
  - .cc.purdue.edu
  - .purdue.edu
  - *null*
- User enters abbreviation *xinu*
- Client tries the following in order
  - xinu.cs.purdue.edu
  - xinu.cc.purdue.edu
  - xinu.purdue.edu
  - xinu

The Point About Abbreviation

The Domain Name System only maps full domain names into addresses; abbreviations are not part of the Domain Name System itself, but are introduced by client software to make local names convenient for users.

Object Types That DNS Supports

Type	Meaning	Contents
A	Host Address	32-bit IP address
CNAME	Canonical Name	Canonical domain name for an alias
HINFO	CPU & OS	Name of CPU and operating system
MINFO	Mailbox info	Information about a mailbox or mail list
MX	Mail Exchanger	16-bit preference and name of host that acts as mail exchanger for the domain
NS	Name Server	Name of authoritative server for domain
PTR	Pointer	Domain name (like a symbolic link)
SOA	Start of Authority	Multiple fields that specify which parts of the naming hierarchy a server implements
TXT	Arbitrary text	Uninterpreted string of ASCII text
AAAA	Host Address	128-bit IPv6 address

Summary

- Domain Name System provides mapping from pronounceable names to IP addresses
- Domain names are hierarchical; top-level domains are dictated by a central authority
- Organizations can choose how to structure their domain names
- DNS uses on-line servers to answer queries
- Lookup begins with local server, which caches entries


Questions?

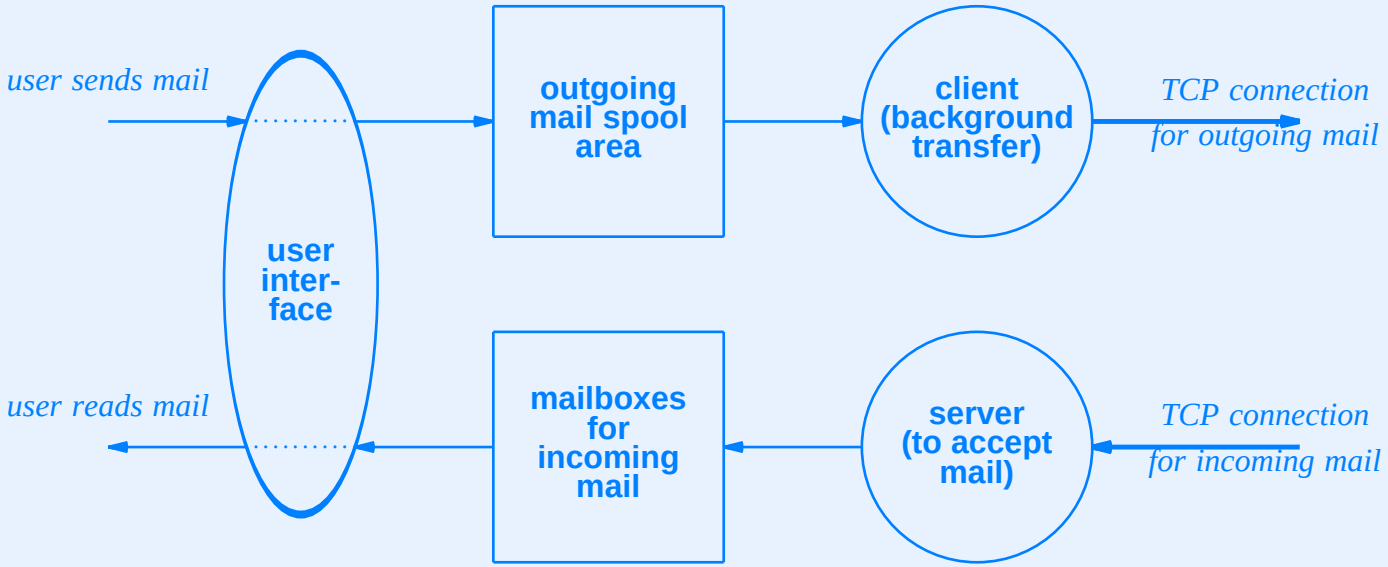
PART XXVI

APPLICATIONS: ELECTRONIC MAIL (SMTP, POP, IMAP, MIME)

Electronic Mail

- Among most widely used Internet services
- Two major components
  - User interface
  - Mail transfer software
- Paradigm: transfer is separate background activity

Illustration Of Email System Components



Mailbox Names And Aliases

- Email destination identified by pair

(mailbox, computer)

- Aliases permitted (user enters alias that is expanded)

TCP/IP Standards For Email

- Syntax for email addresses
- Format of email message
- Protocols for email transfer and mailbox access

Email Address Syntax

- Mailbox identified by string

mailbox@computer

- String *computer* is domain name of computer on which a mailbox resides
- String *mailbox* is unique mailbox name on the destination computer

Format Of Email Message

- Message consists of
  - Header
  - Blank line
  - Body of message
- Headers have form

keyword : information

- Standard given in RFC 2822

Protocol For Email Transfer

- Specifies interaction between transfer components
  - Transfer client
  - Transfer server
- Standard protocol is *Simple Mail Transfer Protocol* (*SMTP*)

SMTP

- Application-level protocol
- Uses TCP
- Commands and responses encoded in ASCII

Example Of SMTP

```
S: 220 Beta.GOV Simple Mail Transfer Service Ready
C: HELO Alpha.EDU
S: 250 Beta.GOV

C: MAIL FROM:<Smith@Alpha.EDU>
S: 250 OK

C: RCPT TO:<Jones@Beta.GOV>
S: 250 OK

C: RCPT TO:<Green@Beta.GOV>
S: 550 No such user here

C: RCPT TO:<Brown@Beta.GOV>
S: 250 OK

C: DATA
S: 354 Start mail input; end with <CR><LF>.<CR><LF>
C: ...sends body of mail message...
C: ...continues for as many lines as message contains
C: <CR><LF>.<CR><LF>
S: 250 OK

C: QUIT
S: 221 Beta.GOV Service closing transmission channel
```

Protocol For Mailbox Access

- Used when user's mailbox resides on remote computer
- Especially helpful when user's local computer is not always on-line
- Two protocols exist
  - *Post Office Protocol* version 3 (*POP3*)
  - *Internet Message Access Protocol* (*IMAP*)
- Each provides same basic functionality
  - User authentication
  - Mailbox access commands

Multipurpose Internet Mail Extensions (MIME)

- Permits nontextual data to be sent in email
  - Graphics image
  - Voice or video clip
- Sender
  - Encodes binary item into printable characters
  - Places in email message for transfer
- Receiver
  - Receives email message containing encoded item
  - Decodes message to extract original binary value

MIME Header

- Header in email message describes encoding used
- Example

```
From: bill@acollege.edu
To: john@example.com
MIME-Version: 1.0
Content-Type: image/jpeg
Content-Transfer-Encoding: base64

...data for the image...
```

Seven Basic MIME Types

Content Type	Used When Data In the Message Is
text	Textual (e.g. a document).
image	A still photograph or computer-generated image
audio	A sound recording
video	A video recording that includes motion
application	Raw data for a program
multipart	Multiple messages that each have a separate content type and encoding
message	An entire e-mail message (e.g., a memo that has been forwarded) or an external reference to a message (e.g., an FTP server and file name)

Example Of Mixed / Multipart Message

```
From: bill@acollege.edu
To: john@example.com
MIME-Version: 1.0
Content-Type: Multipart/Mixed; Boundary=StartOfNextPart

--StartOfNextPart
Content-Type: text/plain
Content-Transfer-Encoding: 7bit
John,
   Here is the photo of our research lab I promised
to send you.  You can see the equipment you donated.

Thanks again, Bill

--StartOfNextPart
Content-Type: image/jpeg
Content-Transfer-Encoding: base64
     ...data for the image...
```

Summary

- Email operates at application layer
- Conceptual separation between
  - User interface
  - Mail transfer components
- Simple Mail Transfer Protocol (SMTP)
  - Standard for transfer
  - Uses ASCII encoding
- Post Office Protocol (POP) And Internet Mail Access Protocol (IMAP) allow access of remote mailbox.
- Multipurpose Internet Mail Extensions (MIME) permits transfer of nontextual information (e.g., images)


Questions?

PART XXVII

APPLICATIONS: WORLD WIDE WEB (HTTP)

World Wide Web

- Distributed hypermedia paradigm
- Major service on the Internet
- Use surpassed file transfer in 1995

Web Page Identifier

- Known as *Uniform Resource Locator* (*URL*)
- Encodes
  - Access protocol to use
  - Domain name of server
  - Protocol port number (optional)
  - Path through server's file system (optional)
  - Parameters (optional)
  - Query (optional)
- Format

http: // *hostname [: port] / path [; parameters] [? query]*

Web Standards

- Separate standards for
  - Representation
  - Transfer

Representation

- *HyperText Markup Language* (*HTML)*
- Document contains text plus embedded links
- HTML gives guidelines for display, not details
- Consequence: two browsers may choose to display same document differently

Transfer

- Used between browser and web server
- Protocol is *HyperText Transfer Protocol* (*HTTP*)
- Runs over TCP

HTTP Characteristics

- Application level
- Request / response paradigm
- Stateless
- Permits bi-directional transfer
- Offers capability negotiation
- Support for caching
- Support for intermediaries

HTTP Operation

- Browser sends requests to which server replies
- Typical request: *GET* used to fetch document
- Example

GET http://www.cs.purdue.edu/people/comer/ HTTP/1.1

- Relative URL also permitted

GET /people/comer/ HTTP/1.1

Error Messages

- HTTP includes set of error responses
- Server can format error as HTML message for user or use internal form and allow browser to format message

Persistent Connections

- HTTP version 1.0 uses one TCP connection per transfer
  - Browser forms TCP connection to server
  - Browser sends GET request
  - Server returns header describing item
  - Server returns item
  - Server closes connection
- HTTP version 1.1 permits connection to persist across multiple requests

HTTP Headers

HTTP uses MIME-like headers to carry meta information. Both browsers and servers send headers that allow them to negotiate agreement on the document representation and encoding to be used.

Handing Persistence

To allow a TCP connection to persist through multiple requests and responses, HTTP sends a length before each response. If it does not know the length, a server informs the client, sends the response, and then closes the connection.

Headers And Length Encoding

- HTTP headers use same syntax as email headers
  - Lines of text followed by blank line
  - Lines of text have form *keyword:information*
- For persistent connection header specifies length (in octets) of data item that follows

Items That Can Appear In An HTTP Header

Header	Meaning
Content-Length	Size of item in octets
Content-Type	Type of item
Content-Encoding	Encoding used for item
Content-Language	Language(s) used in item

Example Of Header

```
Content-Length:   34
Content-Language: english
Content-Encoding: ascii

<HTML> A trivial example. </HTML>
```

- Note: if length is not known in advance, server can inform browser that connection will close following transfer

Connection: close

Negotiation

- Either server or browser can initiate
- Items sent in headers
- Can specify representations that are acceptable with preference value assigned to each
- Example

Accept: text/html, text/plain; q=0.5, text/x-dvi; q=0.8

Items For Negotiation

```
Accept-Encoding:
Accept-Charset:
Accept-Language:
```

Conditional Request

- Allows browser to check cached copy for freshness
- Eliminates useless latency
- Sends *If-Modified-Since* in header of GET request
- Example

If-Modified-Since: Wed, 31 Dec 2003 05:00:01 GMT

Proxy Servers

- Browser can be configured to contact proxy
- Permits caching for entire organization
- Server can specify maximum number of proxies along path (including none)

Caching Of Web Pages

- Caching essential to efficiency
- Server specifies
  - Whether page can be cached
  - Maximum time page can be kept
- Intermediate caches and browser cache web pages
- Browser can specify maximum age of page (forces intermediate caches to revalidate)

Summary

- Web is major application in the Internet
- Standard for representation is HTML
- Standard for transfer is HTTP
  - Request-response protocol
  - Header precedes item
  - Version 1.1 permits persistent connections
  - Server specifies length of time item can be cached
  - Browser can issue conditional request to validate cached item


Questions?