Network Layer, Part 1
Internet Architecture

These slides are created by Dr. Huang of George Mason University. Students registered in Dr. Huang's courses at GMU can make a single machine readable copy and print a single copy of each slide for their own reference as long as the slide contains the copyright statement, and the GMU facilities are not used to produce the paper copies. Permission for any other use, either in machine-readable or printed form, must be obtained form the author in writing.

History

- U.S. DoD Advanced Research Project Agency (**DARPA**) created ARPANET in 1968.
- ARPANET is the first wide-area general purpose packet network.
- In early 70's, the concept of internetworking was advanced.
  - Most networks are established independently.
  - Bridged networks have scalability problems.
  - Demand for universal connection soon arose.
  - The ARPANET became the core of the Internet experiment.

- By 1985, the ARPANET was heavily used and congested; the National Science Foundation (NSF) initiated **NSFNET** as the new backbone.
  - original NSFNET links were 56K, updated to T1 (1.544 Mbps) in 1988 and later to T3 (43 Mbps) in 1991
  - the NSFNET was decommissioned in April 1995

- The contemporary Internet dose not have an official backbone; instead Network Access Points (**NAP**s) were established for multiple backbone networks to exchange traffic:
  - Spring NAP - Pennsauken, NJ
  - PacBell NAP - San Francisco, CA
  - Ameritech NAP - Chicago, IL
  - MFS Datanet - Washington, D.C.
- Each NAP is essentially a high speed LAN, using FDDI or ATM technologies.

Basic Architecture

- Packet-switching network
  - Packets are routed independently
- Unreliable delivery
  - No guarantee by the routing infrastructure to deliver your packets
  - Possibly out-of-order delivery
  - Possibly corrupted delivery
  - Delivery model similar to postal systems
- Interconnecting all kinds of networks/DLLs.

- An addressing scheme independent of the underlying DLL addresses
- No congestion control in the routing infrastructure
  - Users assumed cooperative
  - Congestion control is performed by TCP at endpoints
  - No way to punish network abusers

IP Addresses

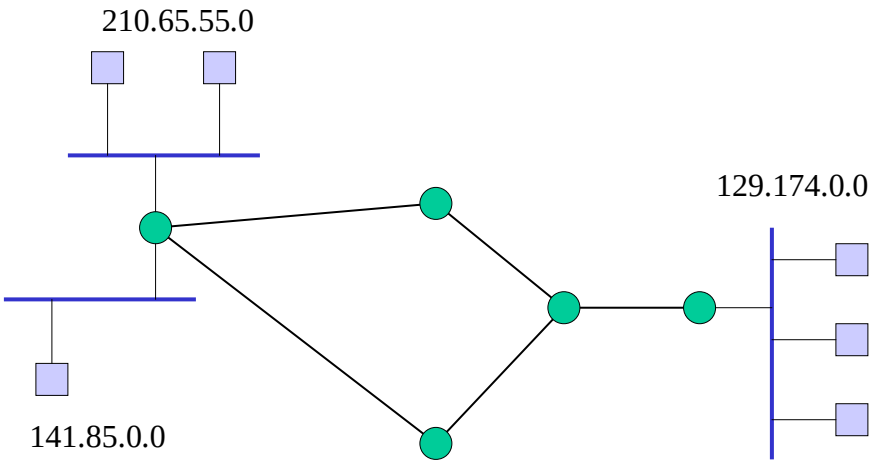
- Fixed-length, 32-bit address
- **Class A**: 128 nets with 16M hosts each
  `0nnnnnnn hhhhhhhh hhhhhhhh hhhhhhhh`
- **Class B**: 16K nets with 64K hosts each
  `10nnnnnn nnnnnnnn hhhhhhhh hhhhhhhh`
- **Class C**: 4M nets with 256 hosts each
  `110nnnnn nnnnnnnn nnnnnnnn hhhhhhhh`
- **Class D**: multicast addresses, 256M groups
  `1110gggg gggggggg gggggggg gggggggg`

IP Number Notation

- A **dotted decimal** notation is used by humans, regardless the class of the address.

  Binary: 11000000 01111111 11111101 00000001
  Decimal: 192. 127. 253. 1
- Examples:
  - cs.gmu.edu = 129.174.40.13
  - site.gmu.edu = 129.174.40.83
  - bacon.gmu.edu = 129.174.65.1

Examples



What Does an IP Address Mean ?

- Each network (subnet) has its network ID.
  - Each point-to-point link considered a network
- Every interface in that network has an IP address comprising the network ID and a host ID.
- *IP addresses identify interfaces*.
  - There is no 1-to-1 correspondence between IP addresses and nodes (hosts or routers)

Domain Name System (DNS)

- Of course, we humans don't even want to memorize decimal numbers; we use names.
- The DNS is like a directory hierarchy: you start with a top-level domain and specify sub-domain name and sub-sub-domain name, and so on, in a right-to-left manner.
  - `bach.cs.gmu.edu.`

Top-level Domains

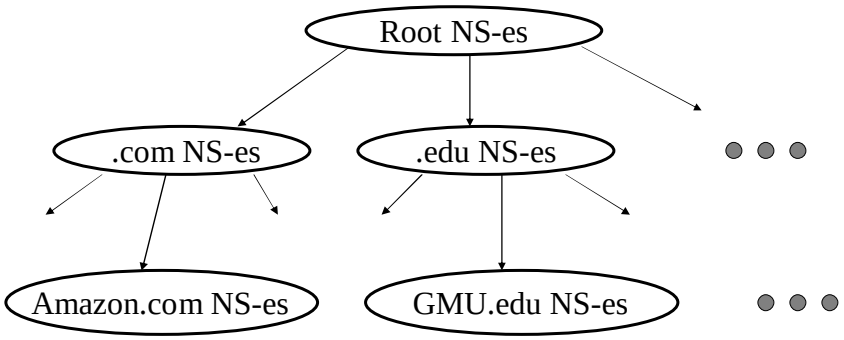
- seven traditional top-level domains
  - com, edu, gov, mil, net, org, int
- new domains established recently
  - .biz, .name, .museum, etc.
- 2-character ISO 3166 country codes
  - Notice that `.tv` is the country code of Tuvalu, a small country in South Pacific Ocean with a population less than 12K.
  - We lease the code for $50M over 12 years.

- Top-level domain names are maintained by the Internet Cooperation for Assigned Names and Numbers (ICANN, `www.icann.org`).
- Second level domains (*i.e.*, `www.IhateCS455.com`) can be obtained from registrars accredited by ICANN.
  - NetworkSolutions.com, DynDNS.org, GoDaddy.com, Joker.com, and many many more

- Management of lower-level names is delegated.
  - For example, GMU is responsible for managing the domain `gmu.edu`.
  - GMU further delegates the responsibility of `cs.gmu.edu` to the CS department.
- Some companies/organizations give 3rd-level domains for free.
  - `DynDNS.org`, `www.hn.org`
  - My free domain: `huangyih.homeip.net`
  - Why not get one for yourself !

DNS Servers

- A hierarchy of name servers (NS) are used to translate domain names to IP addresses.

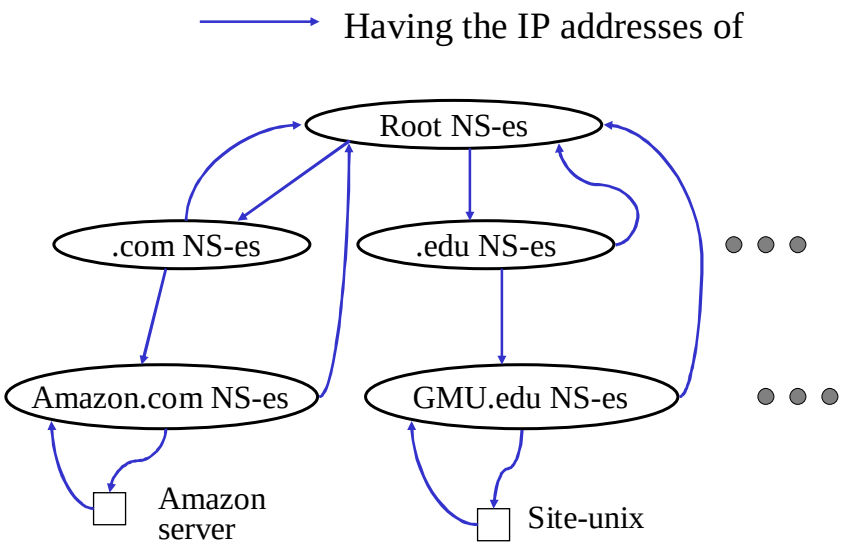


- There are 13 root NS around the world, maintaining 13 identical databases of top-level domain NS.
  - 3 of them in Herndon VA, 1 in Vienna !
- Every root NS *knows* all .com NS, .edu NS, .net NS, .org NS, …
- Each .com NS is also "complete;" it *knows* the NS of all 2nd-level .com domains.
  - It *knows* the NS of `amazon.com`, `yahoo.com`, etc
- The same applies to every .net NS, .edu NS, .jp NS, and so on

"Knows" means having the IP addresses of

- It is the lower-level NS that actually maintain machine addresses.
  - An amazon NS knows the exact IP address of www.amazon.com
  - A GMU NS knows the exact IP address of site-unix.gmu.edu
- Each low-level NS *knows* all machines in its domain.
- Every NS in the world has the list of root NS.
- Each host is configured with the IP addresses of one or two local NS.

NS Hierarchy Again



Local DNS Lookup

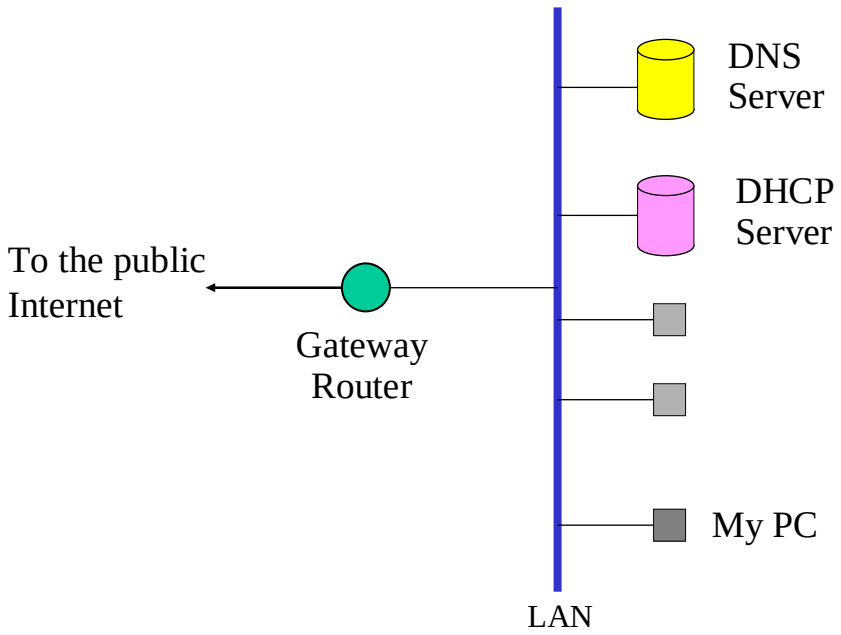
- You are logging into site-unix from a campus workstation.
- You workstation sends a DNS query to a GMU NS.
- The GMU NS responds with the IP address of site-unix.

Looking Up Non-local Domains

- You are on site-unix browsing amazon.com.
- Site-unix queries a GMU NS.
- The GMU NS queries a root NS.
- The root NS responds with a .com NS.
- The GMU NS queries the .com NS.
- The .com NS responds with an amazon NS.
- The GMU NS queries the amazon NS.
- The amazon NS responds with the IP address of www.amazon.com.
- The GMU NS sends site-unix the IP address.

Bootstrap Configurations

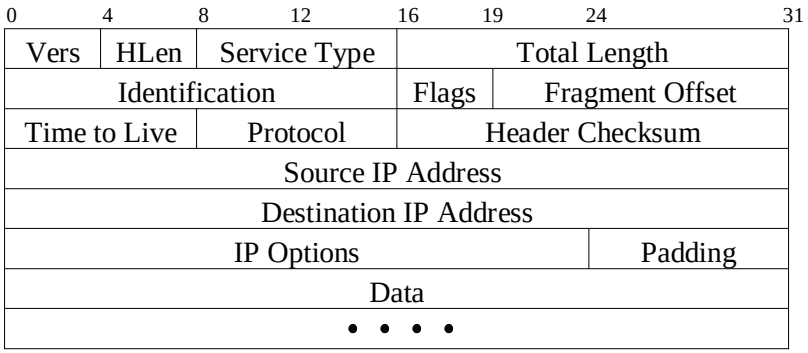
- Each host in the Internet needs the following
  - An IP address per interface
  - Gateway router address
  - DNS server address
  - Network mask (discussed later)
- Typically dynamically configured by DHCP
- Each router needs
  - An IP address per interface
  - (optional) DNS server address
- Manually, statically configured



DHCP: Dynamic Host Configuration Protocol

- A PC broadcasts a DHCP request on LA using MAC broadcast
- The server responds with a DHCP reply, comprising
  - Client IP address,
  - DHCP server IP address
  - DNS server IP address
  - Network mask.

IP Header



- Vers: protocol version (current=4, new=6)
- HLen: Header length in 32-bit words (max=60 bytes)

- Service Type: precedence, delay-sensitive, reliability-sensitive, …
- Total Length: packet length in bytes (max 65535)
- Time to Live (TTL): maximum number of hops the packet can survive; set by the sender and decremented by each intermediate router.
  - One use of TTL is to restrict the damage of routing loops.

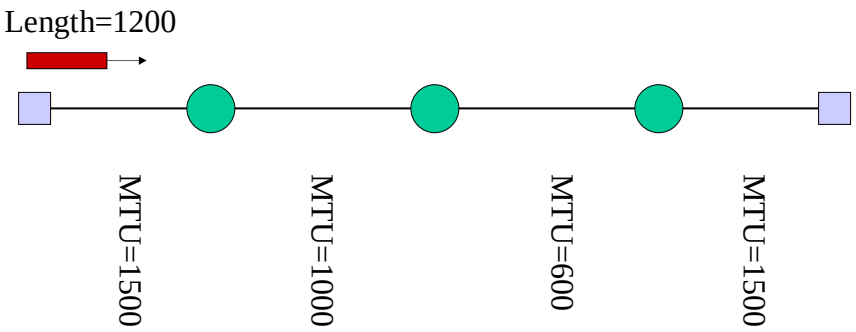
- Identification, Flags, Fragment Offset: used in fragmentation.
- Protocol: code for transport protocol
- Header Checksum: error check for the header
  - uses exclusive-OR, rather than CRC
  - easy for incremental updates
    - ☞fragment offset
    - ☞time to live
- Source/Destination IP Addresses

- Options
  - Record route: each intermediate router records its IP address in the options data area
  - Time stamp: each router records the time
  - Loose source route: the sender lists a series of routers that must be visited in the specified order; other routers may be visited when the packet moves from one listed router to the next
  - Strict source route: only listed routers can be visited.

Fragmentation

- Datagrams transmitted across the physical network/link in frames
- Each network imposes a maximum transmission unit (**MTU**); the MTUs along the delivery path of a packet may vary.
- When an IP datagram of *b* bytes is about to go across a network with an MTU less than *b*, the datagram must be divided into fragments.
- Re-assembly takes place only at the destination
  - this is true even if all subsequent links can transmit the entire datagram in one frame

Example



Fragmentation-Related Fields

- Each datagram from a sender to a recipient has a unique Identification.
- This Identification is copied in every fragments of the datagram.
- The Offset field contains the position of the first byte of the fragment in the entire datagram.
- A more-fragment bit in the Flags field is turned on for all fragments except the last.
- Fragments are treated like regular datagrams.
  - each fragment contains complete source and destination addresses, is routed independently, and could be further fragmented.