

Lecture 8: Internetworking, Naming

CSE 123: Computer Networks
Chris Kanich



Project 1 due 11:59pm tonight



Lecture 8 Overview



- Finish up IP Fragmentation, Route Aggregation
 - CIDR
 - Packet forwarding example
- User-friendly names (DNS)
- Discovering addresses (DHCP/ARP)
- End-to-end lookup and forwarding example

CSE 123 – Lecture 8: Internetworking & Naming

2

Costs of Fragmentation



- Interplay between fragmentation and retransmission
 - A single lost fragment may trigger retransmission
 - Any retransmission will be of entire packet (why?)
- Packet must be completely reassembled before it can be consumed on the receiving host
 - Takes up buffer space in the mean time
 - When can it be garbage collected?
- Why not reassemble at each router?

CSE 123 – Lecture 8: Internetworking & Naming

3

Path MTU Discovery



- Path MTU is the smallest MTU along path
 - Packets less than this size don't get fragmented
- Fragmentation is a burden for routers
 - We already avoid reassembling at routers
 - Avoid fragmentation too by having hosts learn path MTUs
- Hosts send packets, routers return error if too large
 - Hosts can set "don't fragment" flag
 - Hosts discover limits, can size packets at source
 - Reassembly at destination as before

CSE 123 – Lecture 8: Internetworking & Naming

4

IP Addresses



- 32-bits in an IPv4 address
 - Dotted decimal format a.b.c.d
 - Each represent 8 bits of address
- Hierarchical: Network part and host part
 - E.g. IP address 128.54.70.238
 - 128.54 refers to the UCSD campus network
 - 70.238 refers to the host ieng6.ucsd.edu
- Which part is network vs. host?

CSE 123 – Lecture 8: Internetworking & Naming

5

Class-based Addressing



- Most significant bits determines "class" of address

Class A

0	Network	Host
---	---------	------

 127 nets, 16M hosts

Class B

14		16	
1	0	Network	Host

 16K nets, 64K hosts

Class C

21			8	
1	1	0	Network	Host

 2M nets, 254 hosts

- Special addresses
 - Class D (1110) for multicast, Class E (1111) experimental
 - 127.0.0.1: local host (a.k.a. the loopback address)
 - Host bits all set to 0: network address
 - Host bits all set to 1: broadcast address

CSE 123 – Lecture 8: Internetworking & Naming

6

IP Forwarding Tables



- Router needs to know where to forward a packet
- Forwarding table contains:
 - List of network names and [next hop](#) routers
 - Local networks have entries specifying which interface
 - » Link-local hosts can be delivered with Layer-2 forwarding
- E.g. [www.ucsd.edu](#) address is 132.239.180.101
 - Class B address – class + network is 132.239
 - Lookup 132.239 in forwarding table
 - Prefix – part of address that really matters for routing

CSE 123 – Lecture 8: Internetworking & Naming

7

Subnetting



- Individual networks may be composed of several LANs
 - Only want traffic destined to local hosts on physical network
 - Routers need a way to know which hosts on which LAN
- Networks can be arbitrarily decomposed into [subnets](#)
 - Each subnet is simply a prefix of the host address portion
 - Subnet prefix can be of any length, specified with [netmask](#)



CSE 123 – Lecture 8: Internetworking & Naming

8

Subnet Addresses



- Every (sub)network has an address and a [netmask](#)
 - Netmask tells which bits of the network address is important
 - Convention suggests it be a proper prefix
- Netmask written as an all-ones IP address
 - E.g., Class B netmask is 255.255.0.0
 - Sometimes expressed in terms of number of 1s, e.g., /16
- Need to size subnet appropriately for each LAN
 - Only have remaining bits to specify host addresses

CSE 123 – Lecture 8: Internetworking & Naming

9

IP Address Problem (1991)



- Address space depletion
 - In danger of running out of classes A and B
- Why?
 - Class C too small for most organizations (only ~250 addresses)
 - Very few class A – very careful about giving them out (who has 16M hosts anyway?)
 - Class B – greatest problem

CSE 123 – Lecture 8: Internetworking & Naming

10

CIDR



- Classless Inter-Domain Routing (1993)
 - Networks described by variable-length prefix and length
 - Allows arbitrary allocation between network and host address

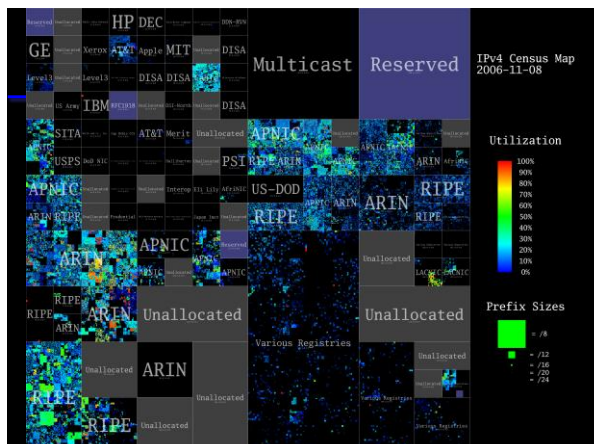


Prefix Mask=# significant bits representing prefix

- e.g. 10.95.1.2/8: 10 is network and remainder (95.1.2) is host
- Pro: Finer grained allocation; aggregation
- Con: More expensive lookup: [longest prefix match](#)

CSE 123 – Lecture 8: Internetworking & Naming

11



MAP OF THE INTERNET
THE IPV4 SPACE, 2006



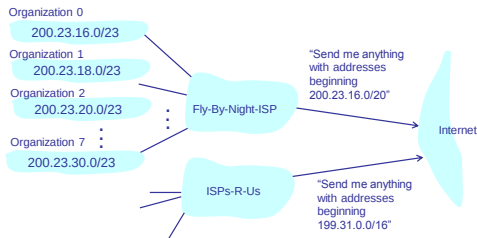
CSE 123 -

13

Route Aggregation



- Combine adjacent networks in forwarding tables
 - Helps keep forwarding table size down



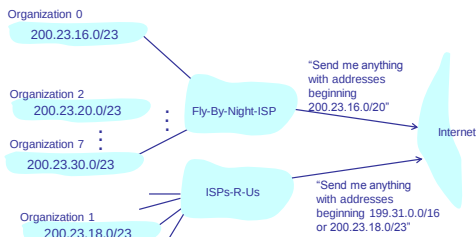
CSE 123 - Lecture 8: Internetworking & Naming

14

Most Specific Route



- But what if address range is not contiguous?



CSE 123 - Lecture 8: Internetworking & Naming

15

Longest Matching Prefix

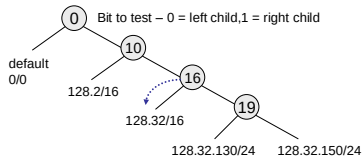


- Forwarding table contains many prefix/length tuples
 - They *need not* be disjoint!
 - E.g. 200.23.16.0/20 and 200.23.18.0/23
 - What to do if a packet arrives for destination 200.23.18.1?
 - Need to find the longest prefix in the table which matches it (200.23.18.0/23)
- Not a simple table, requires multiple memory lookups
 - Lots and lots of research done on this problem
 - Our own George Varghese is the master of this domain

PATRICIA Trie



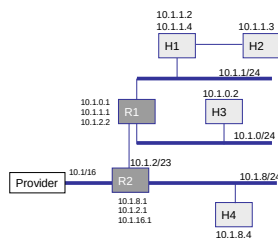
- Straightforward way to look up LMP
 - Arrange route entries into a series of bit tests
 - Worst case = 32 bit tests
 - Problem: memory speed is a bottleneck



Forwarding example



- Packet to 10.1.1.3 arrives
- Path is R2 – R1 – H1 – H2



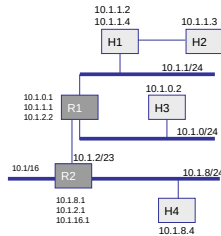
Forwarding example (2)



- Packet to 10.1.1.3
- Matches 10.1.0.0/23

Forwarding table at R2

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.0.1
10.1.8.0/24	interface1
10.1.2.0/23	interface2
10.1.0.0/23	10.1.2.2



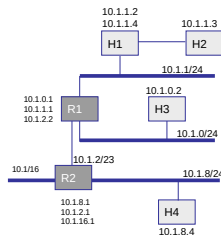
Forwarding example (3)



- Packet to 10.1.1.3
- Matches 10.1.1.2/31
 - Longest prefix match

Routing table at R1

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.2.1
10.1.0.0/24	interface1
10.1.1.0/24	interface2
10.1.2.0/23	interface3
10.1.1.2/31	10.1.1.2



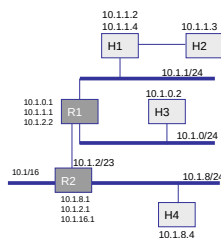
Forwarding example (4)



- Packet to 10.1.1.3
- Direct route
 - Longest prefix match

Routing table at H1

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.1.1
10.1.1.0/24	interface1
10.1.1.3/31	interface2



Layers of Identifiers



- **Host name** (e.g., `www.ucsd.edu`)
 - Used by *humans* to specify host of interest
 - Unique, selected by host administrator
 - Hierarchical, variable-length string of alphanumeric characters
- **IP address** (e.g., `128.54.70.238`)
 - Used by *routers* to forward packets
 - Unique, topologically meaningful locator
 - Hierarchical namespace of 32 bits
- **MAC address** (e.g., `58:B0:35:F2:3C:D9`)
 - Used by *network adaptors* to identify interesting frames
 - Unique, hard-coded identifier burned into network adaptor
 - Flat name space (of 48 bits in Ethernet)

CSE 123 – Lecture 8: Internetworking & Naming

22

Naming Hierarchy for Scale



- Host name: `www.ucsd.edu`
 - **Domain**: registrar for each top-level domain (e.g., `.edu`)
 - **Host name**: local administrator assigns to each host
- IP addresses: `128.54.70.238`
 - **Prefixes**: ICANN, regional Internet registries, and ISPs
 - **Hosts**: static configuration, or dynamic using DHCP
- MAC addresses: `58:B0:35:F2:3C:D9`
 - **OIDs**: assigned to vendors by the IEEE
 - **Adapters**: assigned by the vendor from its block

CSE 123 – Lecture 8: Internetworking & Naming

23

Mapping Between Identifiers



- Domain Name System (**DNS**)
 - Given a host name, provide the IP address
 - Given an IP address, provide the host name
- Address Resolution Protocol (**ARP**)
 - Given an IP address, provide the MAC address
 - To enable communication within the Local Area Network
- Dynamic Host Configuration Protocol (**DHCP**)
 - Automates host boot-up process
 - Given a MAC address, assign a unique IP address
 - ... and tell host other stuff about the Local Area Network

CSE 123 – Lecture 8: Internetworking & Naming

24

Domain Name System (DNS)

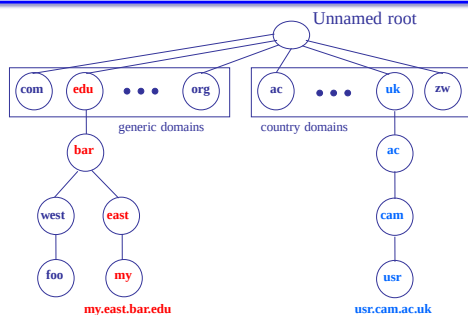


- Distributed administrative control
 - Hierarchical name space divided into zones
 - Distributed over a collection of DNS servers
- Hierarchy of DNS servers
 - Root servers
 - Top-level domain (TLD) servers
 - Authoritative DNS servers
- Performing the translations
 - Local DNS servers
 - Resolver software

CSE 123 – Lecture 8: Internetworking & Naming

25

DNS: Distributed Database



CSE 123 – Lecture 8: Internetworking & Naming

26

DNS Root Servers



- 13 root servers (see <http://www.root-servers.org/>)
 - Labeled A through M



CSE 123 – Lecture 8: Internetworking & Naming

27

Using DNS



- Local DNS server (“default name server”)
 - Usually near the end hosts who use it
 - Local hosts configured with local server (e.g., `/etc/resolv.conf`) or learn the server via DHCP
- Client application
 - Extract server name (e.g., from the URL)
 - Do `gethostbyname()` to trigger resolver code
- Server application
 - Extract client IP address from socket
 - Optional `gethostbyaddr()` to translate into name

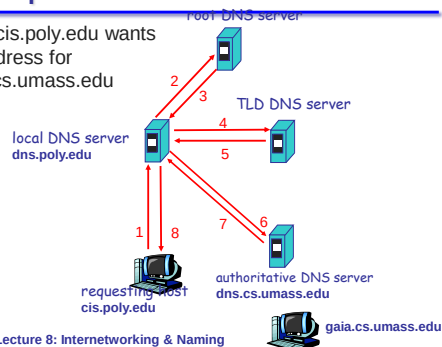
CSE 123 – Lecture 8: Internetworking & Naming

28

Example



Host at `cis.poly.edu` wants IP address for `gaia.cs.umass.edu`



CSE 123 – Lecture 8: Internetworking & Naming

29

Reliability



- DNS servers are replicated
 - Name service available if at least one replica is up
 - Queries can be load balanced between replicas
- UDP used for queries
 - Need reliability: must implement this on top of UDP
 - Try alternate servers on timeout
 - Exponential backoff when retrying same server
- Cache responses to decrease load
 - Both at end hosts and local servers

CSE 123 – Lecture 8: Internetworking & Naming

30

Whence come IP Addresses?



- You already have a bunch from the days when you called Jon Postel and asked for them (e.g. BBN)
- You get them from another provider
 - E.g. buy service from Sprint and get a /24 from one of their address blocks
- You get one directly from a routing registry
 - ARIN: North America, APNIC (Asia Pacific), RIPE (Europe), LACNIC (Latin America), etc.
 - Registries get address from IANA (Internet Assigned Numbers Authority)

CSE 123 – Lecture 8: Internetworking & Naming

31

How Do You And I Get One?



- Well from your provider!
- But how do you know what it is?
- Manual configuration
 - They tell you and you type that number into your computer (along with the default gateway, DNS server, etc.)
- Automated configuration
 - Dynamic Host Resolution Protocol (DHCP)

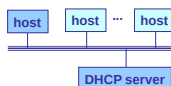
CSE 123 – Lecture 8: Internetworking & Naming

32

Bootstrapping Problem



- Host doesn't have an IP address yet
 - So, host doesn't know what source address to use
- Host doesn't know who to ask for an IP address
 - So, host doesn't know what destination address to use
- Solution: shout to discover a server who can help
 - Install a special server on the LAN to answer distress calls



CSE 123 – Lecture 8: Internetworking & Naming

33

DHCP



- Broadcast-based LAN protocol algorithm
 - Host broadcasts "DHCP discover" on LAN (e.g. Ethernet broadcast)
 - DHCP server responds with "DHCP offer" message
 - Host requests IP address: "DHCP request" message
 - DHCP server sends address: "DHCP ack" message w/ IP address
- Easy to have fewer addresses than hosts (e.g. UCSD wireless) and to *renumber* network (use new addresses)
- What if host goes away (how to get address back?)
 - Address is a "lease" not a "grant", has a timeout
 - Host may have different IP addresses at different times?

CSE 123 – Lecture 8: Internetworking & Naming

34

Address Resolution Protocol

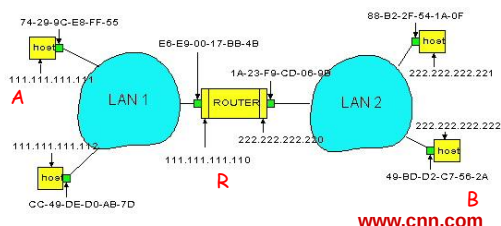


- Every node maintains an ARP table
 - (IP address, MAC address) pair
- Consult the table when sending a packet
 - Map destination IP address to MAC address
 - Encapsulate and transmit the data packet
- What if the IP address is not in the table?
 - Broadcast: "Who has IP address x.x.x.x?"
 - Response: "MAC address yy:yy:yy:yy:yy:yy"
 - Sender caches the result in its ARP table

CSE 123 – Lecture 8: Internetworking & Naming

35

Example: Sending to CNN



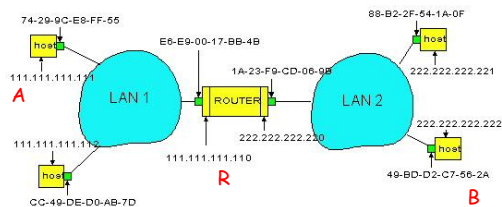
CSE 123 – Lecture 8: Internetworking & Naming

36

Basic Steps



1. Host A must learn the IP address of B via DNS
2. Host A uses gateway R to reach external hosts
3. Router R forwards IP packet to outgoing interface
4. Router R learns B's MAC address and forwards frame



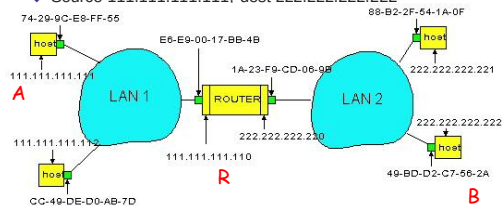
CSE 123 – Lecture 8: Internetworking & Naming

37

Host A Learns B's IP Address



- Host A does a DNS query to learn B's address
 - Suppose gethostbyname() returns 222.222.222.222
- Host A constructs an IP packet to send to B
 - Source 111.111.111.111, dest 222.222.222.222



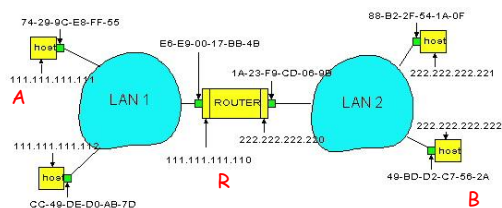
CSE 123 – Lecture 8: Internetworking & Naming

38

Host A Learns B's IP Address



- IP packet
 - From A: 111.111.111.111
 - To B: 222.222.222.222
- Ethernet frame
 - From A: 74-29-9C-E8-FF-55
 - To gateway: ????

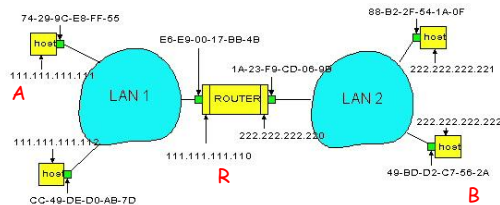


CSE 123 – Lecture 8: Internetworking & Naming

39

A Decides to Send Through R

- Host A has a gateway router R
 - Used to reach dests outside of 111.111.111.0/24
 - Address 111.111.111.110 for R learned via DHCP
- But, what is the MAC address of the gateway?

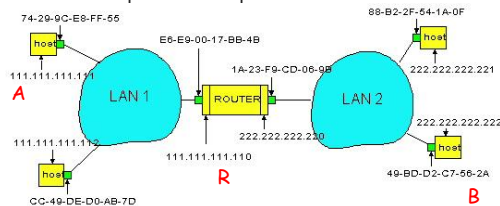


CSE 123 – Lecture 8: Internetworking & Naming

40

A Sends Packet Through R

- Host A learns the MAC address of R's interface
 - ARP request: broadcast request for 111.111.111.110
 - ARP response: R responds with E6-E9-00-17-BB-4B
- Host A encapsulates the packet and sends to R

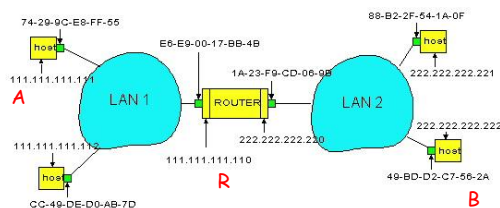


CSE 123 – Lecture 8: Internetworking & Naming

41

A Sends Packet Through R

- IP packet
 - From A: 111.111.111.111
 - To B: 222.222.222.222
- Ethernet frame
 - From A: 74-29-9C-E8-FF-55
 - To R: E6-E9-00-17-BB-4B



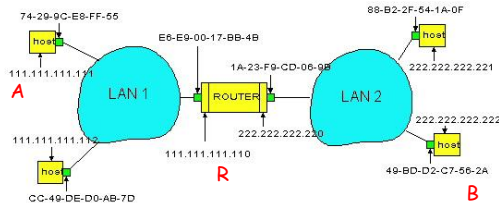
CSE 123 – Lecture 8: Internetworking & Naming

42

R Looks up Next Hop



- Router *R*'s adapter receives the packet
 - R* extracts the IP packet destined to 222.222.222.222
- Router *R* consults its forwarding table
 - Packet matches 222.222.222.0/24 via other interface



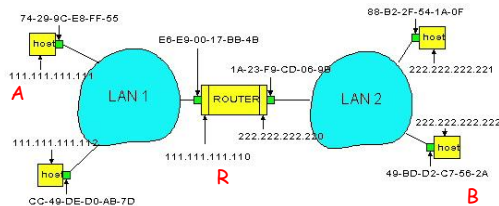
CSE 123 – Lecture 8: Internetworking & Naming

43

R Wants to Forward Packet



- IP packet
 - From A: 111.111.111.111
 - To B: 222.222.222.222
- Ethernet frame
 - From R: 1A-23-F9-CD-06-9B
 - To B: ???



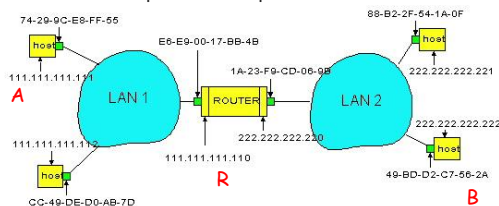
CSE 123 – Lecture 8: Internetworking & Naming

44

R Sends Packet to B



- Router *R*'s learns the MAC address of host *B*
 - ARP request: broadcast request for 222.222.222.222
 - ARP response: *B* responds with 49-BD-D2-C7-56-2A
- Router *R* encapsulates the packet and sends to *B*



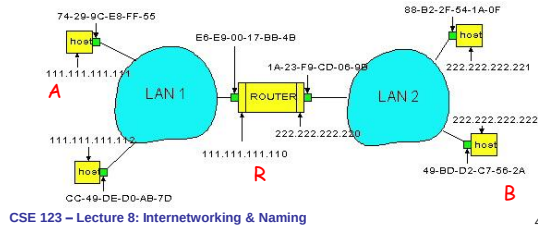
CSE 123 – Lecture 8: Internetworking & Naming

45

R Wants to Forward Packet



- IP packet
 - From A: 111.111.111.111
 - To B: 222.222.222.222
- Ethernet frame
 - From R: 1A-23-F9-CD-06-9B
 - To B: 49-BD-D2-C7-56-2A



46

Summary



- Domain Name System
 - Distributed, hierarchical database
 - Distributed collection of servers
 - Caching to improve performance
- IP to MAC Address mapping
 - Dynamic Host Configuration Protocol (DHCP)
 - Address Resolution Protocol (ARP)

CSE 123 – Lecture 8: Internetworking & Naming

47

For Next Time



- Read 5.1-2 in P&D
- Finish up Project 1!
 - Lose a letter grade for each day late.

CSE 123 – Lecture 8: Internetworking & Naming

48