

CSE/EE 461 – Lecture 8

David Wetherall
djw@cs.washington.edu

Last Time

- Focus
 - How do we calculate routes for packets?
- Routing Algorithms
 - Distance Vector routing (RIP)
 - Link State routing (OSPF)
 - Cost Metrics

Application
Presentation
Session
Transport
Network
Data Link
Physical

djw // CSE/EE 461, Winter 2000

L8.2

This Time

- Focus
 - How do we make routing scale?
- Topics
 - IP Addressing (ARP, DHCP, CIDR, subnets)
 - Inter-domain routing (EGP, BGP)

Application
Presentation
Session
Transport
Network
Data Link
Physical

djw // CSE/EE 461, Winter 2000

L8.3

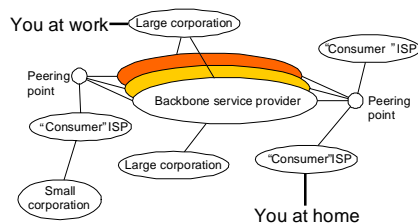
Scalability Concerns

- Routing burden grows with size of an internetwork
 - Size of routing tables
 - Volume of routing messages
 - Amount of routing computation
- RIP/OSPF do not scale to the size of the Internet
- We must apply further techniques:
 - Careful address allocation
 - Use of hierarchy
 - Route aggregation

djw // CSE/EE 461, Winter 2000

L8.4

Structure of the Internet



djw // CSE/EE 461, Winter 2000

L8.5

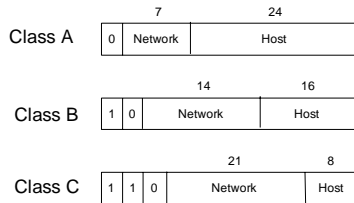
IP Addresses

- Reflect location in topology; used for scalable routing
 - Unlike "flat" Ethernet addresses
- Interfaces on same network share prefix
 - Prefix administratively assigned (IANA or ISP)
 - Addresses globally unique
- Routing only advertises entire networks
 - Local delivery within a single "network" doesn't involve router
 - (will make "network" precise later on)

djw // CSE/EE 461, Winter 2000

L8.6

IPv4 Address Formats



- 32 bits written in "dotted quad" notation, e.g., 18.31.0.135

djw // CSE/EE 461, Winter 2000

L8.7

IPv6 Address Format



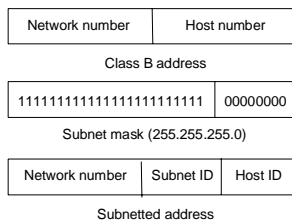
- 128 bits written in 16 bit hexadecimal chunks
- Still hierarchical, just more levels

djw // CSE/EE 461, Winter 2000

L8.8

Subnetting

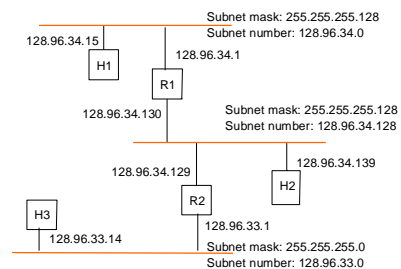
- Split up one network number into multiple physical networks
- Internal structure isn't propagated
- Helps allocation efficiency



djw // CSE/EE 461, Winter 2000

L8.9

Subnet Example



djw // CSE/EE 461, Winter 2000

L8.10

Getting an IP address

- Old fashioned way: sysadmin configured each machine
- Dynamic Host Configuration Protocol (DHCP)
 - One DHCP server with the bootstrap info
 - Host address, gateway address, subnet mask, ...
 - Find it using broadcast
 - Addresses may be leased; renew periodically
- "Stateless" Autoconfiguration (in IPv6)
 - Get rid of server - reuse Ethernet addresses for lower portion of address (uniqueness) and learn higher portion from routers

djw // CSE/EE 461, Winter 2000

L8.11

Address Resolution Protocol (ARP)

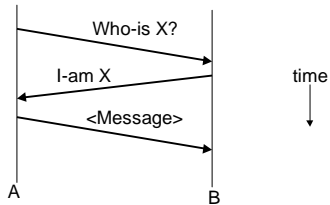
- On a single link, need Ethernet addresses to send a frame ... source is a given, but what about destination?
 - Requires mapping from IP to MAC addresses
- ARP is a dynamic approach to learn mapping
 - Node A sends broadcast query for IP address X
 - Node B with IP address X replies with its MAC address M
 - A caches (X, M); old information is timed out (~15 mins)
 - Also: B caches A's MAC and IP addresses, other nodes refresh

djw // CSE/EE 461, Winter 2000

L8.12

ARP Example

- To send a message ... common case doesn't require ARP

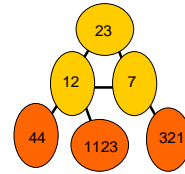


djw // CSE/EE 461, Winter 2000

L8.13

Inter-Domain Routing

- Network comprised of many Autonomous Systems (ASes) or domains
- To scale, use hierarchy: separate inter-domain and intra-domain routing
- Also called interior vs exterior gateway protocols (IGP/EGP)
 - IGP = RIP, OSPF
 - EGP = EGP, BGP

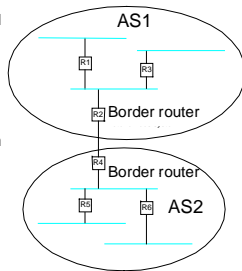


djw // CSE/EE 461, Winter 2000

L8.14

Inter-Domain Routing

- Border routers summarize and advertise internal routes to external neighbors and vice-versa
- Border routers apply policy
- Internal routers can use notion of default routes
- Core is "default-free"; routers must have a route to all networks in the world

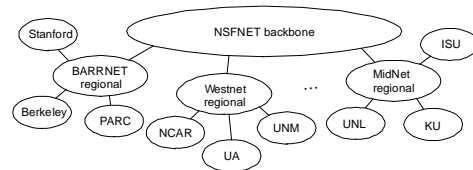


djw // CSE/EE 461, Winter 2000

L8.15

Exterior Gateway Protocol (EGP)

- First major inter-domain routing protocol
- Constrained Internet to tree structure; no longer in use



djw // CSE/EE 461, Winter 2000

L8.16

Border Gateway Protocol (BGP-4)

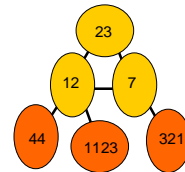
- EGP used in the Internet backbone today
- Features:
 - Path vector routing
 - Application of policy
 - Operates over reliable transport (TCP)
 - Works with CIDR

djw // CSE/EE 461, Winter 2000

L8.17

Path Vectors

- Similar to distance vector, except send entire paths
 - e.g. 321 hears [7,12,44]
 - stronger avoidance of loops
 - supports policies (later)
- Modulo policy, shorter paths are chosen in preference to longer ones
- Reachability only – no metrics

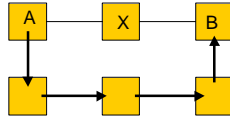


djw // CSE/EE 461, Winter 2000

L8.18

Policies

- Choice of routes may depend on owner, cost, AUP, ...
 - Business considerations
- Local policy dictates what route will be chosen and what routes will be advertised!
 - e.g., X doesn't provide transit for B, or A prefers not to use X

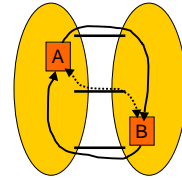


djw // CSE/EE 461, Winter 2000

L8.19

Impact of Policies – Example

- Early Exit / Hot Potato
 - “if it's not for you, bail”
- Combination of best local policies not globally best
- Side-effect: asymmetry



djw // CSE/EE 461, Winter 2000

L8.20

Operation over TCP

- Most routing protocols operate over UDP/IP
- BGP uses TCP
 - TCP handles error control; reacts to congestion
 - Allows for incremental updates
- Issue: Data vs. Control plane
 - Shouldn't routing messages be higher priority than data?

djw // CSE/EE 461, Winter 2000

L8.21

CIDR (Supernetting)

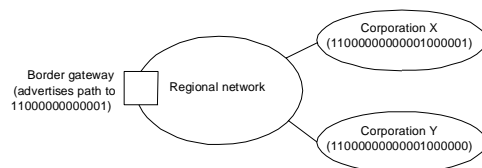
- CIDR = Classless Inter-Domain Routing
- Aggregate advertised network routes
 - e.g., ISP has class C addresses 192.4.16 through 192.4.31
 - Really like one larger 20 bit address class ...
 - Advertise as such (network number, prefix length)
 - Reduces size of routing tables
- But IP forwarding is more involved
 - Based on Longest Matching Prefix operation

djw // CSE/EE 461, Winter 2000

L8.22

CIDR Example

- X and Y routes can be aggregated



djw // CSE/EE 461, Winter 2000

L8.23

Key Concepts

- Scalable routing is requires that we minimize:
 - Size of routing tables
 - Exchange of routing messages
- Scalable routing is based on:
 - Careful address allocation
 - Use of hierarchy and route aggregation
- Internet is a collection of Autonomous Systems (ASes)
 - Policy dominates routing at the AS level

djw // CSE/EE 461, Winter 2000

L8.24