CSC458 – Lecture 6

Inter-domain Routing
IP Addressing

Administrivia

- Homework:
  - #2 due today
  - #3 out today, due in two weeks (same date as midterm)
- No lecture next week
  - Reading Week
- Midterm in two weeks
  - 60 minute, close-book
  - No tutorial that Monday
  - Homework #4 will be out that day

Midterm will Cover Following Topics

- Protocols and layering
- Encoding bits with signals: NRZ, NRZI, Manchester, 4B/5B
- Error detection and correction: parity, CRC, Hamming distance
- Latency and RTT calculation
- MAC Protocols:
  - Wired: Aloha, CSMA/CD, Ethernet
  - Wireless: CSMA/CA, RTS/CTS
  - Contention-free: token ring, FDDI, DQDB

Midterm will Cover Following Topics (2)

- Bridging LANs, spanning tree algorithm
- IPv4
  - Header fields
  - Fragmentation
  - Path MTU
- ICMP
- Forwarding and routing
- Distance vector protocols: RIP, count-to-infinity, split horizon, split horizon with poison reverse
- Link state routing: Dijkstra algorithm, OSPF, cost metrics

Midterm will Cover Following Topics (3)

- Inter-domain routing: BGP, AS, path vectors, multi-homing
- IP addressing: ARP, CIDR, hierarchical addressing

- That was the long version
- The short version is…
  - Everything so far including today's lecture

Midterm will Cover following Readings

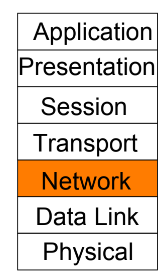
- Chapters 1-4
- Fishnet projects 1 and 2 is fair game

Any Questions?

This Lecture

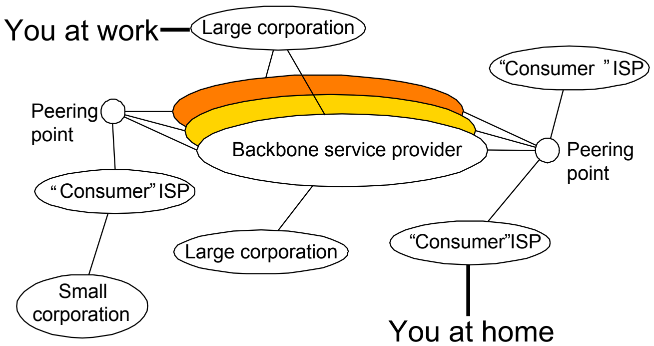
- Focus
  - How do we make routing scale?

- Inter-domain routing
  - ASes and BGP



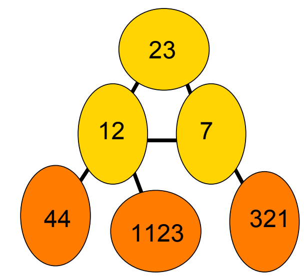
Structure of the Internet

- Inter-domain versus intra-domain routing



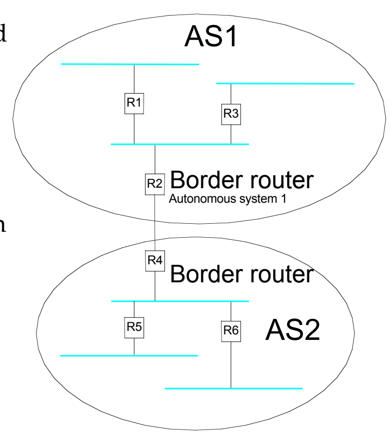
Inter-Domain Routing

- 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



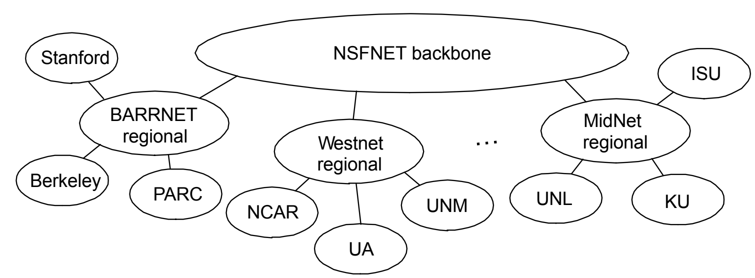
Inter-Domain Routing

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



Exterior Gateway Protocol (EGP)

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

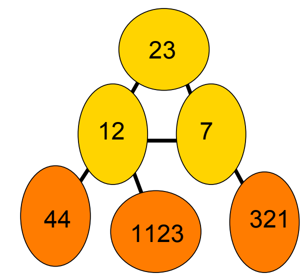


Border Gateway Protocol (BGP-4)

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

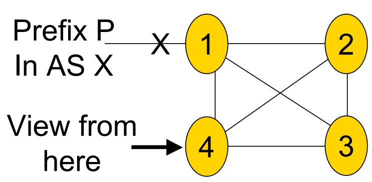
Path Vectors

- 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



An Ironic Twist on Convergence

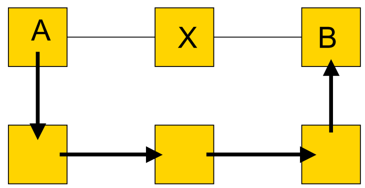
- Recently, it was realized that BGP convergence can undergo a process analogous to count-to-infinity!



- AS 4 uses path 4 1 X. A link fails and 1 withdraws 4 1 X.
- So 4 uses 4 2 1 X, which is soon withdrawn, then 4 3 2 1 X, …
- Result is many invalid paths can be explored before convergence

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

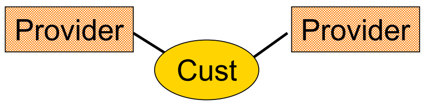


Simplified Policy Roles

- Providers sell Transit to their customers
  – Customer announces path to their prefixes to providers in order for the rest of the Internet to reach their prefixes
  – Providers announces path to all other Internet prefixes to customer C in order for C to reach the rest of the Internet
- Additionally, parties Peer for mutual benefit
  – Peers A and B announce path to their customer's prefixes to each other but do not propagate announcements further
  – Peering relationships aren't transitive
  – Tier 1s peer to provide global reachability

Multi-Homing

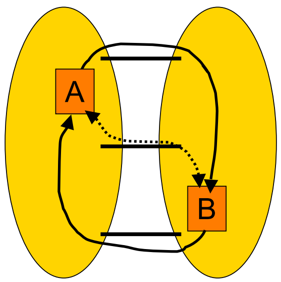
- Connect to multiple providers for reliability, load sharing



- Customer can choose the best outgoing path from any of the announcements heard from its providers
  – Easy to control outgoing traffic, e.g, for load balancing
- Less control over what paths other parties will use to reach us
  – Both providers will announce that they can reach to the customer
  – Rest of Internet can choose which path to take to customer
    - Hard for the the customer to influence this

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



Operation over TCP

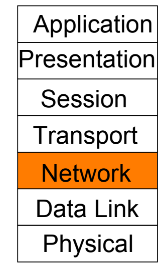
- Most routing protocols operate over UDP/IP
- BGP uses TCP
  – TCP handles error control; reacts to congestion
  – Allows for incremental updates
- Issue: Data vs. Control plane
  – Shouldn't routing messages be higher priority than data?

Key Concepts

- Internet is a collection of Autonomous Systems (ASes)
  - Policy dominates routing at the AS level
- Structural hierarchy helps make routing scalable
  - BGP routes between autonomous systems (ASes)

This Lecture

- Focus
  - How do we make routing scale?
- IP Addressing
  - Hierarchy (prefixes, class A, B, C, subnets)
  - Also allocation (DHCP, ARP)



Scalability Concerns

- Routing burden grows with size of an internetwork
  - Size of routing tables
  - Volume of routing messages
  - Amount of routing computation
- To scale to the size of the Internet, apply:
  - Hierarchical addressing
  - Use of structural hierarchy
  - Route aggregation

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 by prefix
  - Local delivery in a single “network” doesn’t involve router
  - (will make “network” precise later on)

Getting an IP address

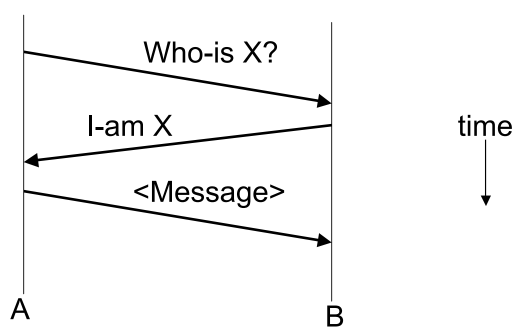
- Old fashioned way: sysadmin configured each machine
- Dynamic Host Configuration Protocol (DHCP)
  - One DHCP server with the bootstrap info
    - Host address, gateway address, subnet mask, …
    - Find it using broadcast
  - Addresses may be leased; renew periodically
- "Stateless" Autoconfiguration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

Address Resolution Protocol (ARP)

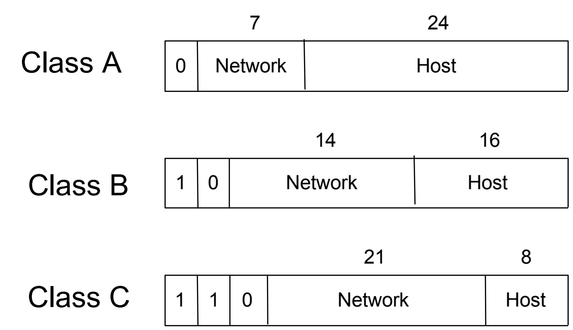
- On a single link, need Ethernet addresses to send a frame … source is a given, but what about destination?
  - Requires mapping from IP to MAC addresses
- ARP is a dynamic approach to learn mapping
  - Node A sends broadcast query for IP address X
  - Node B with IP address X replies with its MAC address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

ARP Example

- To send first message use ARP to learn MAC address
- For later messages (common case) don't need to ARP



IPv4 Address Formats



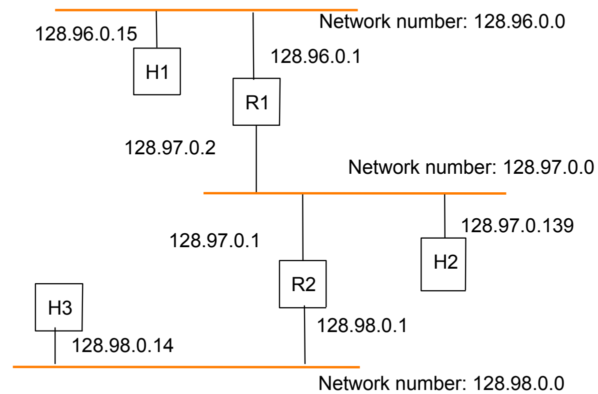
- 32 bits written in "dotted quad" notation, e.g., 18.31.0.135

IPv6 Address Format

001	RegistryID	ProviderID	SubscriberID	SubnetID	InterfaceID

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

Network Example



Updated Forwarding Routine

- Used to be “look up destination address for next hop”
- Now addresses have network and host portions:
  - If host: if destination network is the same as the host network, then deliver locally (without router). Otherwise send to the router
  - If router: look up destination network in routing table to find next hop and send to next router. If destination network is directly attached then deliver locally.
- (Note that it will get a little more complicated later)

Subnetting – More Hierarchy

- Split up one network number into multiple physical networks
- Helps allocation efficiency -- can hand out subnets
- Rest of internet does not see subnet structure
  - subnet is purely internal to network
  - aggregates routing info

Network number	Host number

Class B address

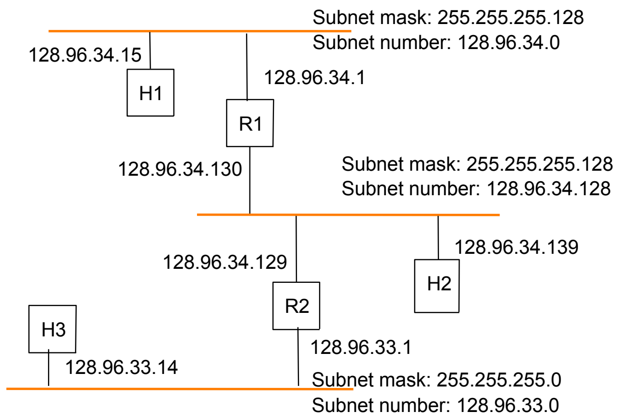
111111111111111111111111	00000000

Subnet mask (255.255.255.0)

Network number	Subnet ID	Host ID

Subnetted address

Subnet Example



Updated Forwarding Routine

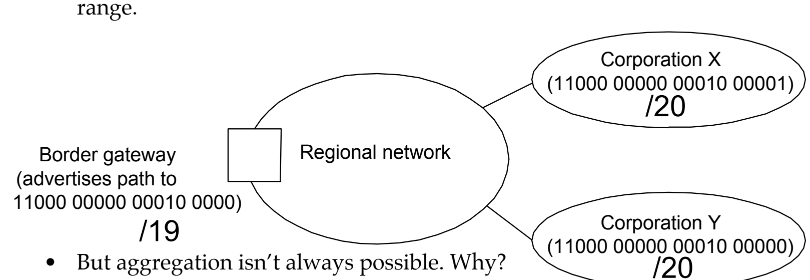
- Used to know network from address (class A, B, C)
- Now need to "search" routing table for right subnet
  - If host: easy, just substitute "subnet" for "network"
  - If router: search routing table for the subnet that the destination belongs to, and use that to forward as before
- (Note that it will get a little more complicated later :)

CIDR (Supernetting)

- CIDR = Classless Inter-Domain Routing
- Generalize class A, B, C into prefixes of arbitrary length; now must carry prefix length with address
- Aggregate adjacent 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

CIDR Example

- X and Y routes can be aggregated because they form a bigger contiguous range.



- But aggregation isn't always possible. Why?
  - can only aggregate power of 2

IP Forwarding Revisited

- Routing table now contains routes to “prefixes”
  - IP address and length indicating what bits are fixed
- Now need to “search” routing table for longest matching prefix, only at routers
  - Search routing table for the prefix that the destination belongs to, and use that to forward as before
  - There can be multiple matches; take the longest prefix
- This is the IP forwarding routine used at routers.

Key Concepts

- Hierarchical address allocation helps routing scale
  - Addresses are constrained by topology
  - Only need to advertise and compute routes for networks
  - Hide internal structure within a domain via subnets
  - Keep host simple and let routers worry about routing
- ARP learns the mapping from IP to MAC address