

CSE/EE 461 – Lecture 12

Inter-domain Routing

David Wetherall
djw@cs.washington.edu

Last Time

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

Application
Presentation
Session
Transport
Network
Data Link
Physical

This Lecture

- Focus
 - How do we make routing scale?
- Inter-domain routing
 - ASes and BGP
 - CIDR for route aggregation

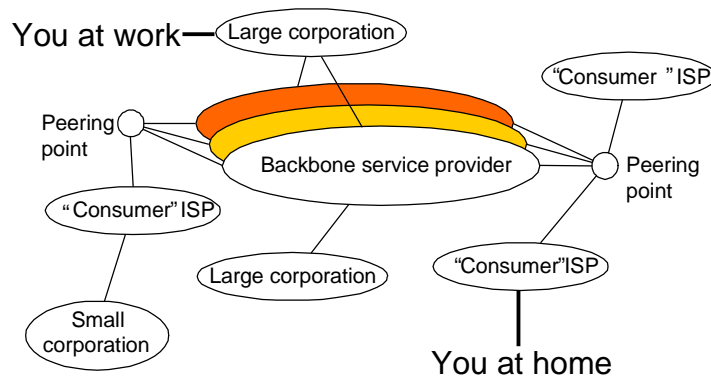
Application
Presentation
Session
Transport
Network
Data Link
Physical

djw // CSE/EE 461, Autumn 2001

L12.3

Structure of the Internet

- Inter-domain versus intra-domain routing

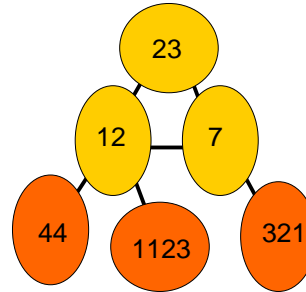


djw // CSE/EE 461, Autumn 2001

L12.4

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

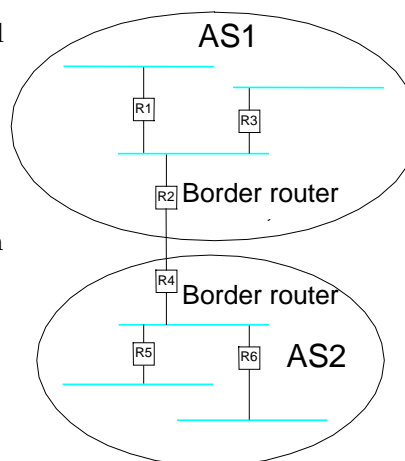


djw // CSE/EE 461, Autumn 2001

L12.5

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

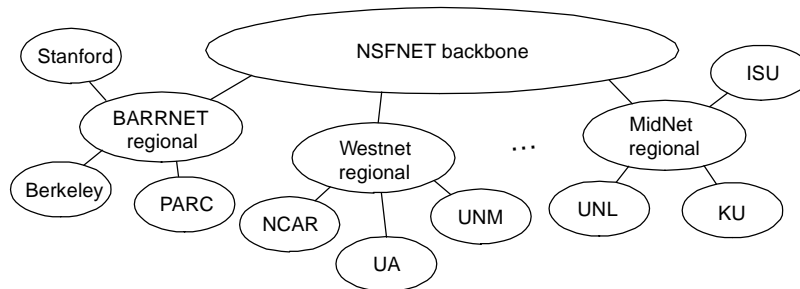


djw // CSE/EE 461, Autumn 2001

L12.6

Exterior Gateway Protocol (EGP)

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



djw // CSE/EE 461, Autumn 2001

L12.7

Border Gateway Protocol (BGP-4)

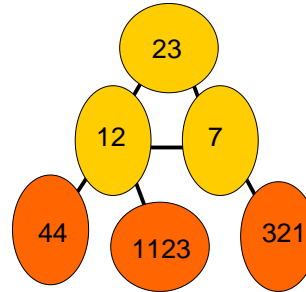
- EGP used in the Internet backbone today
- Features:
 - Path vector routing
 - Application of policy
 - Operates over reliable transport (TCP)
 - Uses route aggregation (CIDR)

djw // CSE/EE 461, Autumn 2001

L12.8

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

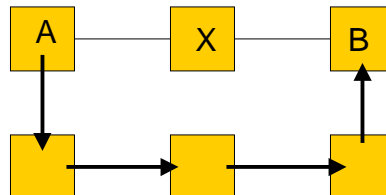


djw // CSE/EE 461, Autumn 2001

L12.9

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

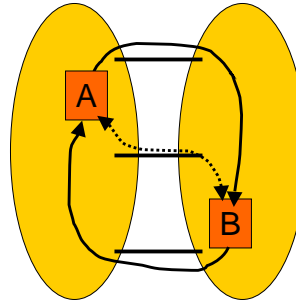


djw // CSE/EE 461, Autumn 2001

L12.10

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, Autumn 2001

L12.11

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, Autumn 2001

L12.12

CIDR (Supernetting)

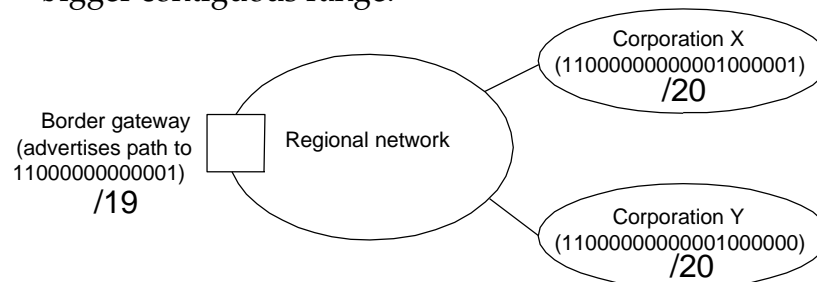
- CIDR = Classless Inter-Domain Routing
- Aggregate adjacent advertised network routes
 - e.g., ISP has class C addresses 192.4.16 through 192.4.31
 - Really like one larger 20 bit address class ...
 - Advertise as such (network number, prefix length)
 - Reduces size of routing tables
- But IP forwarding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

djw // CSE/EE 461, Autumn 2001

L12.13

CIDR Example

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



- But aggregation isn't always possible. Why?

djw // CSE/EE 461, Autumn 2001

L12.14

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

- 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)
- Route aggregation (CIDR) improves scalability