CSE561 – Interdomain routing

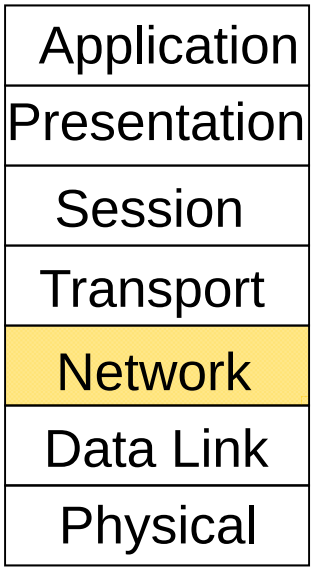
David Wetherall
djw@cs.washington.edu

Interdomain routing

- Focus:
  - Routing across internetworks made up of different parties

- Route scaling
- Route policy

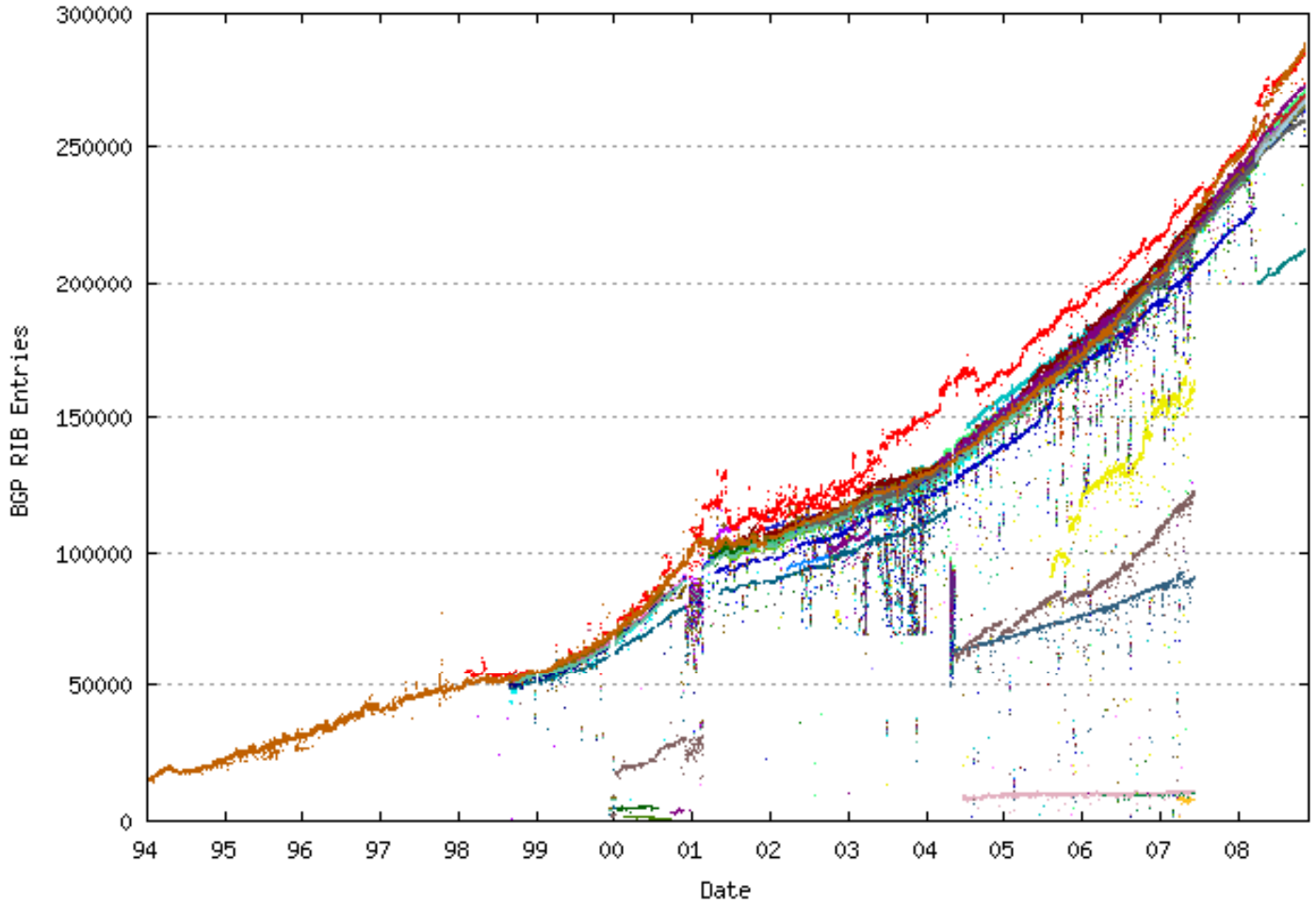
- The protocol: BGP



Two key problems beyond intradomain

- Scale
  - Size of routing tables, computation, messages
  - All grow with the size of the network
- Policy
  - Different parties with different goals make different decisions
  - ISPs are out to make money (locally good paths), not save the world (global shortest path)

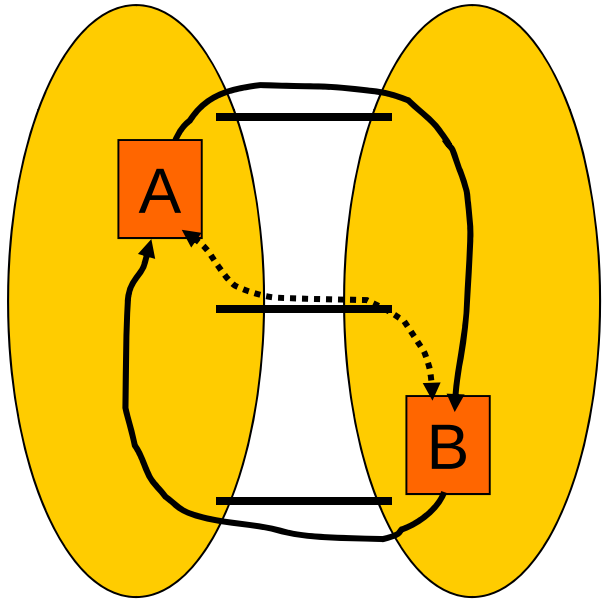
Core BGP Table Growth 1994 - 2008



www.cidr-report.org November 2008

Impact of independent decisions

- Early Exit / Hot Potato policy
  - "if it's not for you, get rid of it"
- Combination of best local policies not globally best
  - Shorter paths exist
- Side-effect: route asymmetry

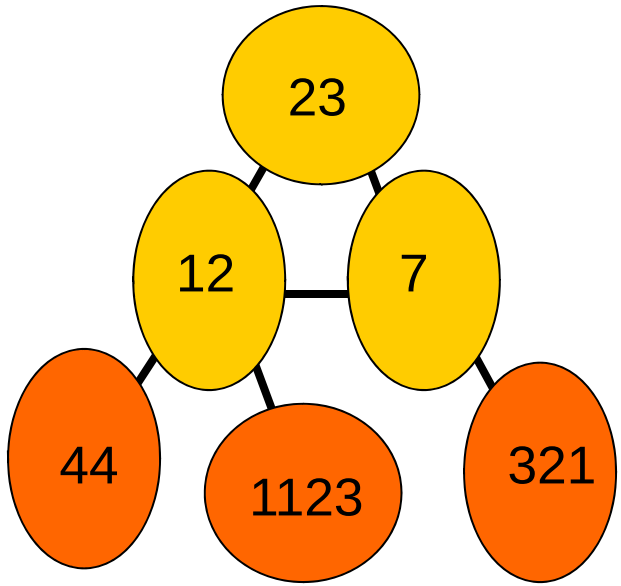


Solutions?

- Scale solution
  - Standard approach of hierarchy / information hiding
  - In the forms of prefixes and ASes
- Policy solution
  - No great solutions here!
  - Let everyone make their own decisions to the extent possible
  - Economic model gives rise to common commercial policies, e.g, peering vs transit

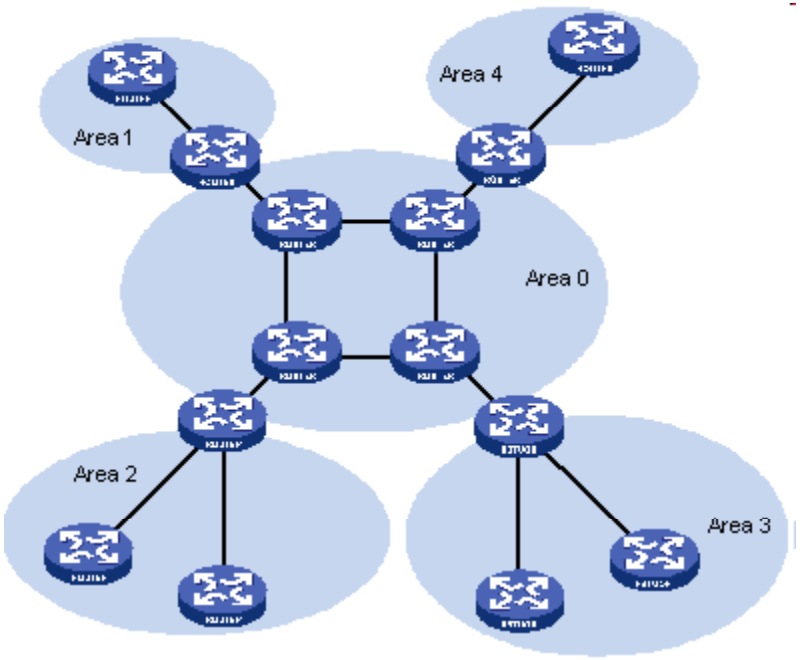
Inter-Domain Routing

- Network comprised of many Autonomous Systems (ASes) or domains
- To scale could use hierarchy to separate inter-domain (BGP) and intra-domain (OSPF) routing
- But not really how BGP works!



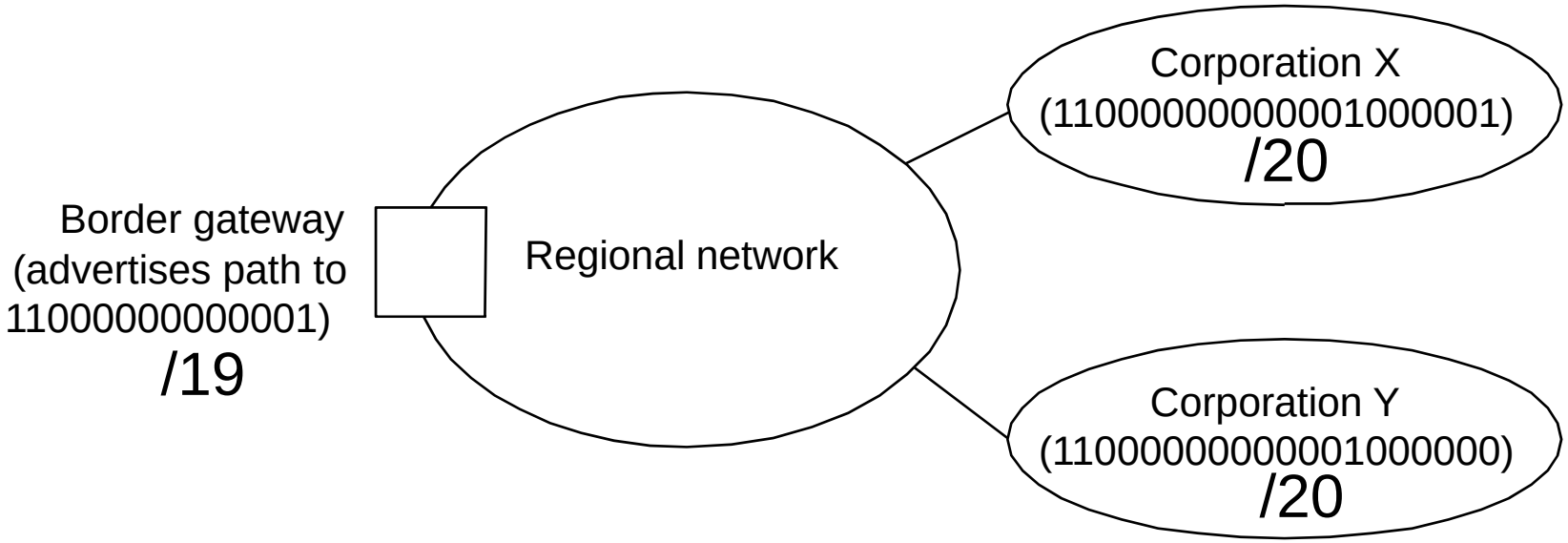
Scaling OSPF with Areas

- Split a large network into “areas”
  - Areas connect via border routers
  - Backbone area connects to all
- Border routers send a summary of the area routes to other areas
  - Hides internal area detail
- Example of applying hierarchy



Prefixes and Aggregation (CIDR)

- Route to blocks of IP addresses called prefixes, e.g., 18/8
- Combine (aggregate) routes to X and Y where they form a larger contiguous range.

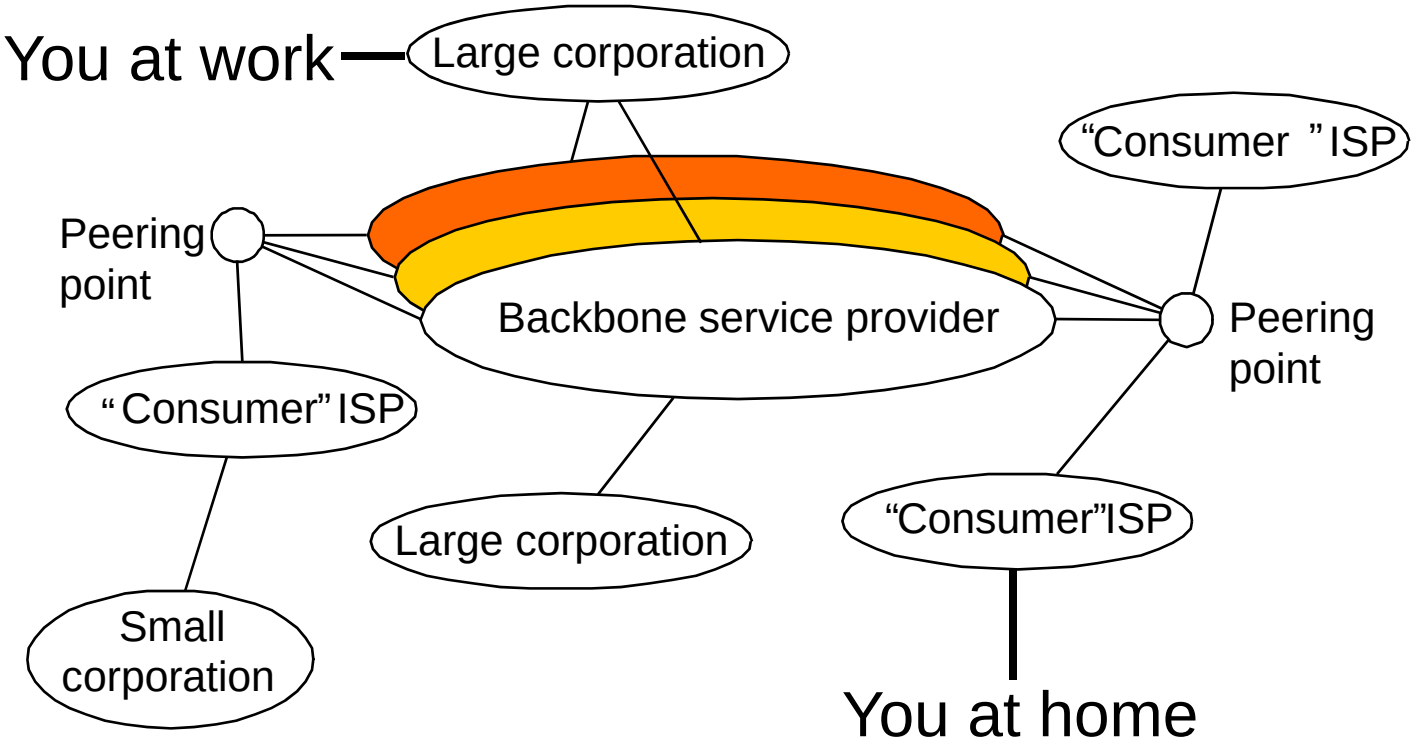


BGP

- Interdomain routing protocol of the Internet
- Each AS tells other ASes the paths it is offering
  - Paths are summaries to prefixes via the sequence of ASes
  - No detailed paths of cost metrics to particular IPs
  - This happens at each border router of the AS
- Each AS picks the paths it wants to use to send traffic
  - Default rule: prefer shortest AS path, then shortest internal path
  - But selection heavily customized by ISPs
  - This happens at each border router of the AS

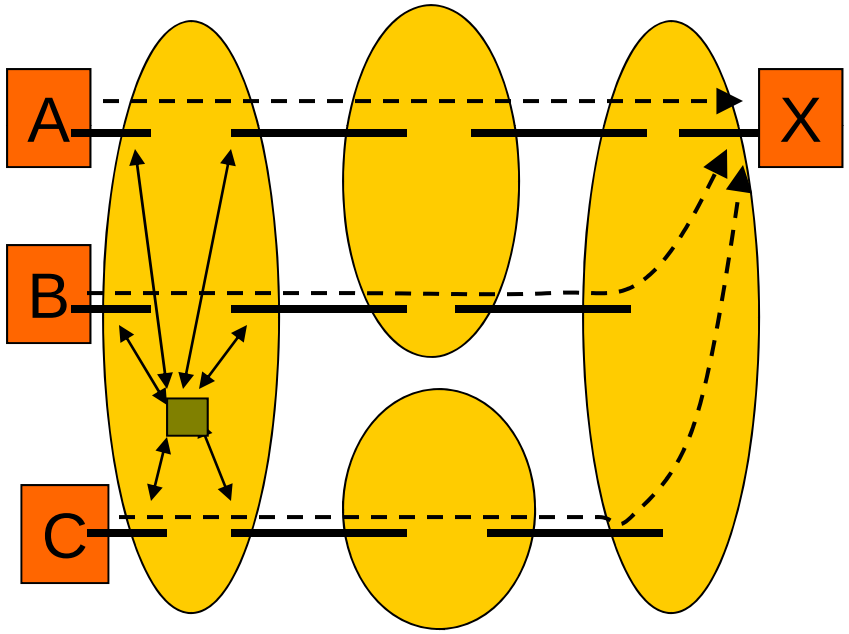
Structure of the Internet

- Consider each different entity to be an "AS"



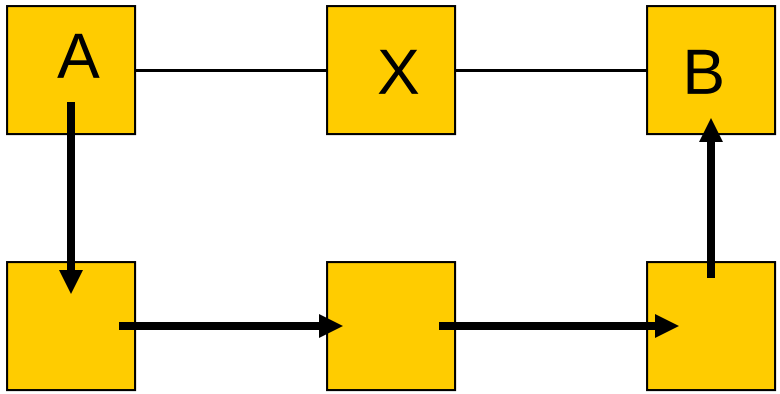
Integration of Intra- and Inter-domain

- Each location makes its own decision based on the routes it sees; not one decision per ISP.
  - e.g, paths A→X, B→X, C→X
- Externally facing routers at exchanges hear routes from other ISPs
- Internally, routers share the routes they know about (green box).



Policies

- Choice of routes may depend on owner, cost, AUP, …
  - Business considerations
- ISPs decide which routes to advertise, and which to use
  - e.g., X doesn't provide transit for B, or A prefers not to use X



- Q: why will this procedure find working routes?

Example 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

WISER routing paper discussion

- How circuitous are Internet paths?
- Why are they circuitous at all?
- What is Pareto-optimal?
- Is early-exit routing Pareto-optimal?
- What is the key idea for improving the routes?
- What key problems need to be tackled for better routes?
- What are the costs of finding these better routes?