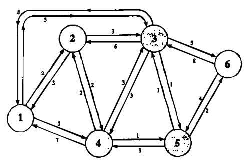
Exterior Gateway Protocols: EGP, BGP-4, CIDR



Shivkumar Kalyanaraman
Rensselaer Polytechnic Institute
shivkuma@ecse.rpi.edu
http://www.ecse.rpi.edu/Homepages/shivkuma

- **Cores, Peers, and the limit of default routes**
- **Autonomous systems & EGP**
- **BGP**
- **CIDR: reducing router table sizes**
- **Refs: Chap 10. Books: "Routing in Internet" by Huitema, "Interconnections" by Perlman, "Internetworking with TCP/IP" by Comer**

Default Routing

- **Default routes => partial information**
- **Routers/hosts w/ default routes rely on other routers to complete the picture.**
- **In general routing "signposts" should be:**
  - **Consistent, I.e., if packet is sent off in one direction then another direction should not be more optimal**
  - **Complete, I.e., should be able to reach all destinations**

Core

- A small set of routers that have consistent & complete information about all destinations.
- Outlying routers can have partial information provided they point default routes to the core
  - Partial info allows site administrators to make local routing changes independently.
- Initially, core routers were under a central authority and were synchronized for consistency => single backbone.
- Internet quickly outgrew single backbone (ARPANET + NSFNET). Core architecture does not scale well.

Peers

- Initially NSFNET had only one connection to ARPANET (router in Pittsburg) => only one route between the two.
- Addition of multiple interconnections => multiple possible routes => need for dynamic routing decision
- Single *core* replaced by a network of *peer* backbones => more scalable
  - Today there are over 30 backbones!
- The routing protocol used by cores & peers was called Gateway-Gateway Protocol (GGP). Replaced by EGP and now by BGP-4.

Autonomous Systems

- The “core” + edges were still considered “one network” => administrative problems like rebooting a router required coordination.
- Replace this n/w with “autonomous systems”(AS). “Stub” AS connect via “cores”
- AS = set of routers and networks under the same administration
  - No theoretical limit to the size of the AS
  - All parts within an AS remain connected.
  - If two networks rely on core-AS to connect, they don’t belong to a single AS

Autonomous Systems (contd)

- One router represents the AS to the external world (the core and other AS). This router also collects reachability info ("external routes") from other AS and diffuses it into its domain.
- AS is identified by a 16-bit AS number
- Traffic types: *Local* = traffic originating or terminating at AS. *Transit* = non-local traffic
- AS types:
  - Stub AS => only single connection to one other AS => it carries only local traffic.
  - Multihomed AS: Connected to multiple AS, but does not allow transit traffic
  - Transit AS: carries transit traffic under policy restrictions

Exterior Gateway Protocol (EGP)

- A mechanism that allows non-core routers to learn routes from core routers so that they can choose optimal backbone routes
- A mechanism for non-core routers to inform core routers about hidden networks
- Autonomous System (AS) has the responsibility of advertising reachability info to other ASs.
  - One or more routers may be designated per AS.
  - Important that info propagates to core routers

EGP weaknesses

- EGP does not interpret the distance metrics in routing update messages => cannot be compute shorter of two routes
- As a result it restricts the topology to a (possibly non-optimal) tree structure, with the core as the root
  - Rapid growth => many networks may be temporarily unreachable
  - Only one path to destination => no load sharing

Border Gateway Protocol (BGP)

- Uses a path-vector concept which enables loop prevention in complex topologies
- In AS-level, shortest path may not be preferred for policy, security, cost reasons.
  - Different routers have different preferences (policy) => as packet goes thru network it will encounter different policies
  - Same problem for link-state. Link state also has a more serious scaling problem. Aggregation needed.
  - Solution: use source-based routing and specify entire path

BGP (contd)

- BGP sets up TCP connection between peers
  - Exchange entire BGP table first
  - Later exchanges only incremental updates
  - Application (BGP)-level keepalive messages
- # of paths proportional to number of AS
- But, memory requirement: proportional to number of networks (one entry per network)
- Path attributes: list of traversed AS and list of reachable networks
- Interior and exterior peers: need to exchange reachability information among interior peers before updating intra-AS routing tables

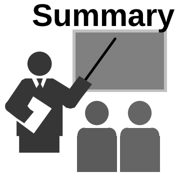
CIDR

- Supported by BGP-4
- Shortage of class Bs => give out many class Cs instead of one class B address
  - New problem: every class C network needs a routing entry !
  - Solution: Classless Inter-domain Routing (CIDR). Also called “supernetting”
- Key: allocate addresses such that they can be summarized.
  - Share same higher order bits (I.e. prefix)
- Routing tables and protocols must be capable of carrying a subnet mask.

CIDR

- Eg: allocate class Cs from 194.0.0.0 thru 195.255.255.255 for hosts in Europe (higher order 7 bits the same).
  - Allows one routing entry for Europe
- Allow other routing entries too. Eg: 194.0.160 + mask of 255.255.240.0
  - When an address matches multiple entries (eg 194.0.22.1), choose the one which had the longest mask (“longest-prefix match”)
- Routing decisions independent of class
- Slows down router table growth.
  - If hosts renumbered, router sizes would drastically reduce.

Summary



- Cores, peers, autonomous systems
- Early protocols: GGP, EGP
- BGP avoids EGP-induced tree structure and allows policy-based routing
- CIDR allows reduction of routing table sizes