BGP routing policies in ISP networks

Author:

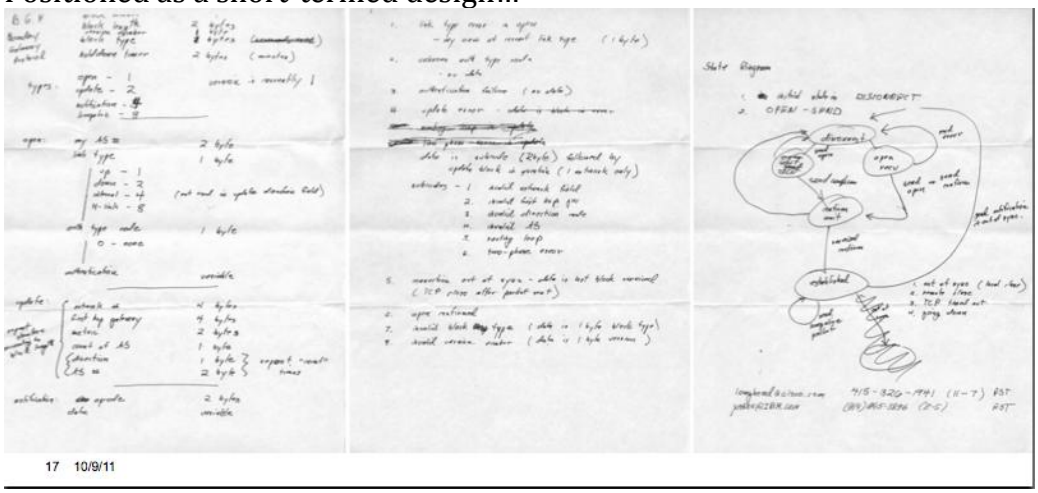
Matthew Caesar (PhD in UC Berkeley, post-doc in Princeton, now a professor in UIUC)
Jennifer Rexford (was a professor in Princeton, winner of Grace Hopper)

BGP Overview

Optional: Link State Routing Algorithm (Dijkstra), Distance-Vector Routing Algorithm (Bellman-Ford)

1. History: hierarchical routing: AS and ISP
2. prior to 1989 (when BGP was born), backbone NSFNET phase II and its regional networks used EGP, the architecture was "core routers – non-core routers"…
   GGP/EGP
   EGP: not a routing protocol, reachability exchange protocol, distance-vector
   - exchange the complete routing table every 3min no matter whether there is a update
   - no path info
   - difficult to scale or detect routing loops
   - constrain the topology to be a spanning tree
   - directly on IP

   You name the disadvantages…
3. BGP inventor: Kirk Loughee (Cisco), Yakov Rekhter, with help of Len Bosack (Cisco).
   BGP: three-napkin protocol, designed in 1989, IETF 12
   Positioned as a short-termed design…



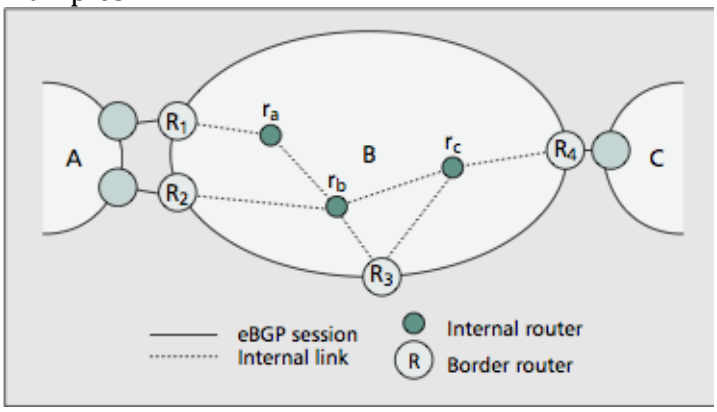
4. BGP overcame the problems of EGP described above. It evolved through the years

RFC Number	Date	Name	BGP Version
1105	June 1989	A Border Gateway Protocol (BGP)	BGP-1
1163	June 1990	A Border Gateway Protocol (BGP)	BGP-2
1267	October 1991	Border Gateway Protocol 3 (BGP-3)	BGP-3
1654	July 1994	A Border Gateway Protocol 4 (BGP-4)	BGP-4
1771	March 1995	A Border Gateway Protocol 4 (BGP-4)	BGP-4

   One big change of BGP4 is its support for **routing policies**.
5. By the time of publishing BGP4 RFC, it has already been deployed widely across the Internet!

BGP as an Inter-Domain Protocol

1. BGP basics: iBGP and eBGP
2. Functionalities:
   For each AS: obtain reachability information, reachability propagation within AS, make decisions
   For the Internet: Reachability advertisement
3. Route: prefix and its path attributes (e.g. AS-PATH, NEXT-HOP)
   Examples:



4. Difference between inter-AS routing protocol and intra-AS routing protocol
   - Policy
   - Scale
   - Performance

Routing Policies in Internet

1. Decision process of BGP

Step	Attribute	Controlled by local or neighbor AS?
1.	Highest LocalPref	Local
2.	Lowest AS path length	Neighbor
3.	Lowest origin type	Neither
4.	Lowest MED	Neighbor
5.	eBGP-learned over iBGP-learned	Neither
6.	Lowest IGP cost to border router	Local
7.	Lowest router ID (to break ties)	Neither

2. 3 steps to process route advertisement:
    - import policies
    - decision process
    - export policies

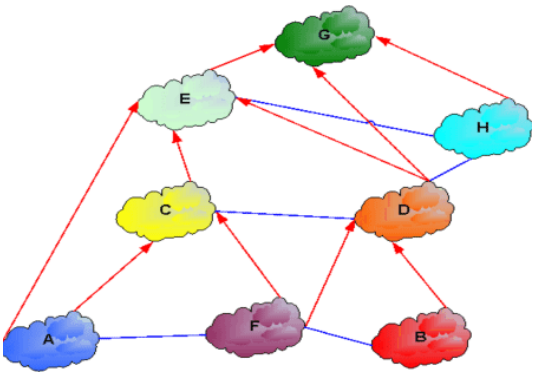
    ISPs tweak the above 3 steps in order to implement the routing policies

Policy Taxonomy

1. Business Relationships
    - Customer-provider

      Tier-1:AT&T
        - They only provide transit traffic (need payment) or peer with other ISPs.
        - Disconnection between two peering Tier-1 ISPs will partition the Internet.
        - Example: 2005, cogent and level3. Pop question: 3 scenarios to determine if you may have access problem
        - Sidebar: Being a pure Tier-1 is not an easy job…

      Tier-2: Comcast

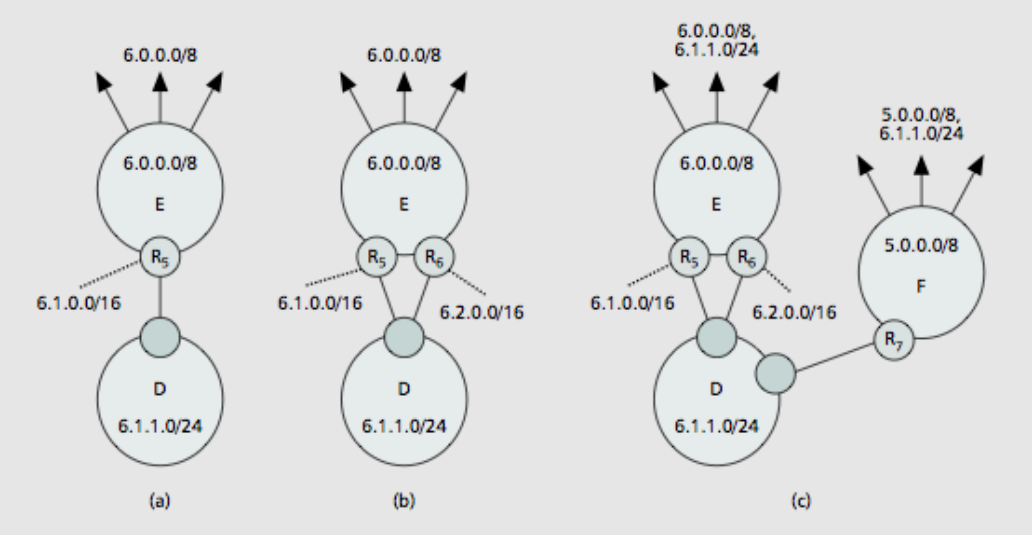
      Example: A->B, C->H

      Tier-3: resell transit

    - Peer-peer

- Backup
- Paper: change the local-perf to influence decision process
- Paper: controlling exports, using community attributes

2. Traffic Engineering
   - Outbound Traffic Engineering
     Change local-perf
     Achieve load balancing---goal of Traffic Engineering
   - Inbound Traffic Engineering
     Change MED
     Inflate the AS-PATH
   - Remote Control: community attributes
3. Scalability
   - Limit the routing table size
     Protect local ISP from other ISPs
     Protect other ISPs: route aggregate
     3 scenarios:



   - Limit the number of routing changes:
     Flap damping: improving routing instability
4. Security
   - Discard Invalid routes: import policies
   - Rewrite path attributes to prevent violations of peering agreements
   - Export policies: filter out sensitive information
   - Prevention of DoS attack
     Paper: discard offensive routes/updates
     IP Hijack Example: Feb 2008, Pakistan telecom hijacked Youtube by creating a black hole, propagated by BGP through the whole Internet, result: Youtube unavailable for 2 hours. (by using longest-prefix match in routing table) This can also be used to eavesdrop….

However, BGP remains as an in-secure protocol: BGP was designed when everyone was trusted…