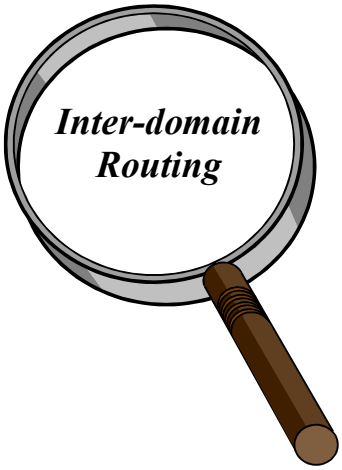
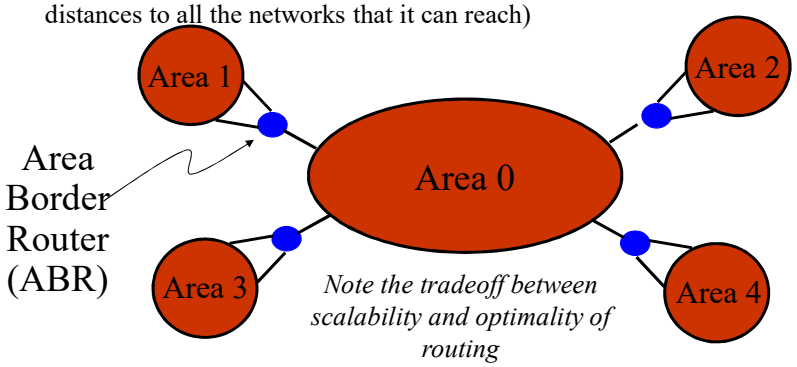


Growing the Internet: It's all the question of scalability

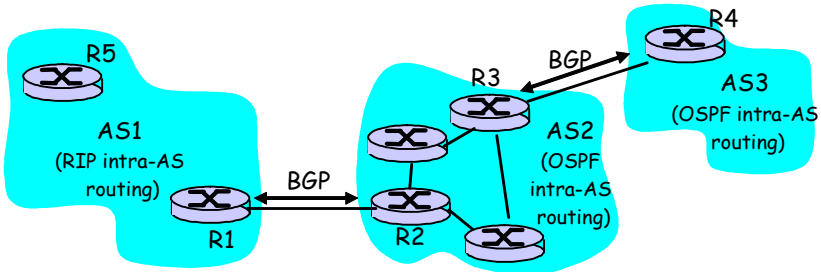
- The link-state routing protocols are more efficient, but still cannot support the "flat-routing" approach for very large networks.
- The key question is: how to grow the network from hundred of nodes to billion of nodes?
- The main approach to Internet scalability is hierarchical routing.
- The address space was another problem hindering growth; solution: move from IPv4 to IPv6.
- Another aspect of growth: the demand for new features, such as support for multicast and support for mobility.

Scalability: Introducing Hierarchy Through Areas

- Divide network into regions
  - Backbone (area 0) and non-backbone areas
  - Each area has its own link-state database
  - Routing within an area is based on intra-domain routing
  - Advertise only *path* distances at area boundaries (an ABR advertises the distances to all the networks that it can reach)



Inter-AS routing in the Internet: (BGP)



Autonomous System (AS) = Routing Domain

Inter-domain routing

- *BGP* provides
  - Inter-domain reachability in a scalable way
  - Enables ISPs to implement various policies
  - Helps ISPs to cooperate under competitive circumstances
- Inter-domain routing is hard because:
  - Large scale
    - More than 300,000 prefixes and ~20,000 AS-s
  - Autonomous nature of AS-s
    - Cost metrics may be different in different AS-s; interdomain routers advertise *reachability* only
  - Trust issues (do I trust the advertised routes?); "selective trust"
  - Need to support various types of policies

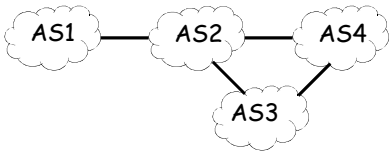
Reachability vs. Optimality

- Each AS can choose its own intra-AS routing protocol
  - Can use any cost metrics to its interior paths
- No consistency between AS-s
  - A cost of 1000 may be good for one, may be unacceptable for other
  - Impossible to find least cost path to a destination AS
    - Any path is okay!
- Each AS originates "reachability advertisements"
  - I'm AS1 and I own networks 129.110/16, 129.111/16
- AS-s advertise/exchange reachability to other AS-s
  - I'm AS2 and can reach AS1 that includes 129.110/16 and 129.111/16 (no cost is mentioned!)
    - So the advertised path is (AS2 AS1)

Path selection

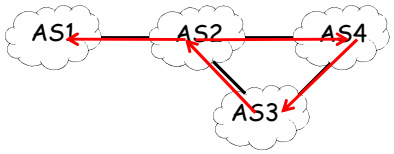
- Suppose AS3 sends its-path-to-AS1 (AS3 AS2 AS1) to its peer AS4
  - AS4 may or may not select the path offered by AS3
    - Cost, policy (don't route via competitors AS), loop prevention reasons
  - If AS4 select the path, then

    Path (AS4,AS1) = AS4, Path(AS3,AS1)
  - AS3 can control incoming traffic by controlling its route advertisements
    - E.g., don't want to route traffic to AS1 -> don't advertise any routes to AS1 to neighbor AS4



Routing Loops

- Loops are bad!
  - Traffic interruption
  - Congestion in the network



- Example:
  - AS2 announces to AS3 it can reach AS1
  - AS3 announces to AS4 it can reach AS1
  - AS4 announces to AS2 it can reach AS1
  - AS2 chooses AS4 as its next hop to AS1

Three Types of Routing Domains

- The three AS types:
  - Stub AS – connected to only one other AS (e.g., small cooperation)
  - *Multihomed AS* – connected to more than one AS, but carries no *transit* traffic (e.g., large cooperation)
  - *Transit AS* - connected to more than one AS and carries *transit* traffic (e.g., backbone provider)
- Goals of inter-domain routing:
  - Loop-free paths to destinations
  - Compliant with AS policies along the paths

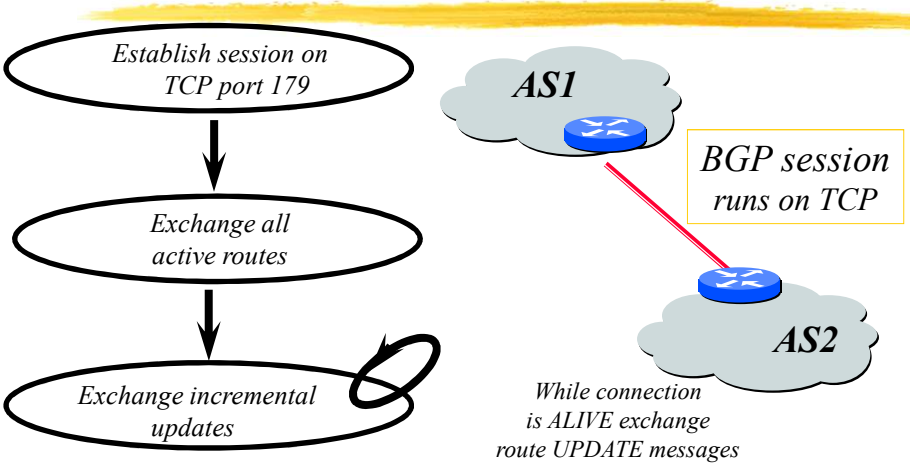
BGP-4

- BGP (Border Gateway Protocol)
  - Current standard protocol for inter-domain routing
- Each AS has
  - One or more border routers
    - Connects an AS to the Internet
    - Used for external routes
  - One or more BGP routers (BGP speakers)
    - Routers that participate in the inter-domain routing
- BGP routers
  - Receive, filter route advertisements from neighbor BGP routers
  - Perform route selection
    - To route to destination X, which path (of several advertised) will be taken
  - Send route advertisements to neighbors

BGP-4 (con't)

- BGP is a *Path Vector* protocol
  - Similar to distance vector protocol
  - BGP routes to networks, not individual hosts
  - BGP advertisements include *complete path* (i.e., uses source routing)
    - *Complete path: list of AS-s to reach a particular network*
    - *Why? What is the benefit of including the complete path?*
  - Requires unique AS identities, which are assigned by a central authority. Speakers advertise routes based on their local policy
- No obligation to advertise any particular route (thus enforcing no-transit policies)
- *Withdrawn Route* advertisement cancels previously advertised route
- BGP updates include both *withdrawn routes* and *reachability information*
- BGP routes are sent over TCP, thus ensuring reliability
- Only changes are advertised, while *keepalive* messages are sent in lieu of updates
- When no *keepalive* packets are received from a router, its routes are assumed to be invalid.

BGP Operations (Simplified)



A Sample BGP table

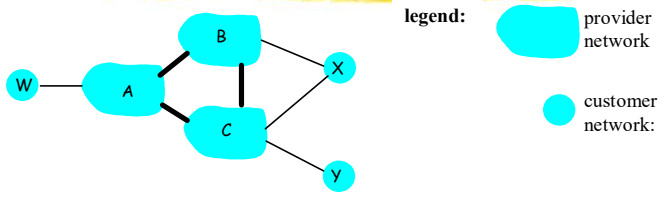
router> show ip bgp

		next hop	AS path
*	12.104.140.0/22	213.200.87.254	0 3257 1239 701 6985 i
*		209.123.12.51	0 8001 7911 701 6985 i
*		203.62.252.26	0 1221 4637 701 6985 i
		64.50.230.1	0 4181 3356 701 6985 i
*	12.104.147.0/24	213.200.87.254	0 3257 7132 21744 i
*		209.123.12.51	0 8001 1784 7132 21744 i
*		64.50.230.1	0 4181 7132 21744 i
*		216.191.65.118	0 15290 7132 21744 i
*		213.140.32.146	0 12956 7132 21744 i
*	12.104.244.0/24	213.200.87.254	0 3257 1239 701 15343 i
*		64.50.230.1	0 4181 3356 701 15343 i
*		209.123.12.51	0 8001 7911 701 15343 i
*		157.130.182.254	0 19092 701 15343 i
*		203.62.252.26	0 1221 4637 701 15343 i

BGP Messages and Policies

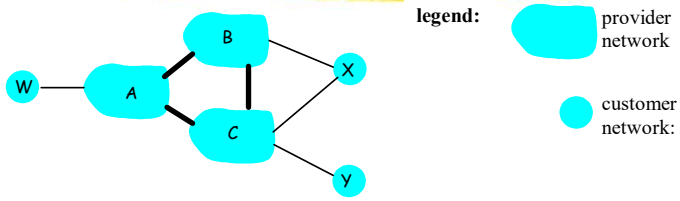
- BGP messages:
  - *OPEN*: opens TCP connection to peer and authenticates sender
  - *UPDATE*: advertises new path (or withdraws old)
    - IP Prefix: Attributes
  - *KEEPALIVE*: keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - *NOTIFICATION*: reports errors in previous message; also used to close connection
- AS Policies:
  - *Provider-Customer:* advertise all routes to/from customers
  - *Customer-Provider:* advertise all routes to/from providers, but not between providers
  - *Peer-Peer:* advertise routes to/from customers to peers, but not between peers and providers

Examples of BGP advertising routes



- A,B,C are provider networks
- X,W,Y are customer (of provider networks)
- X is dual-homed: attached to two networks
  - X does not want to route from B via X to C
  - .. so X will not advertise to B a route to C

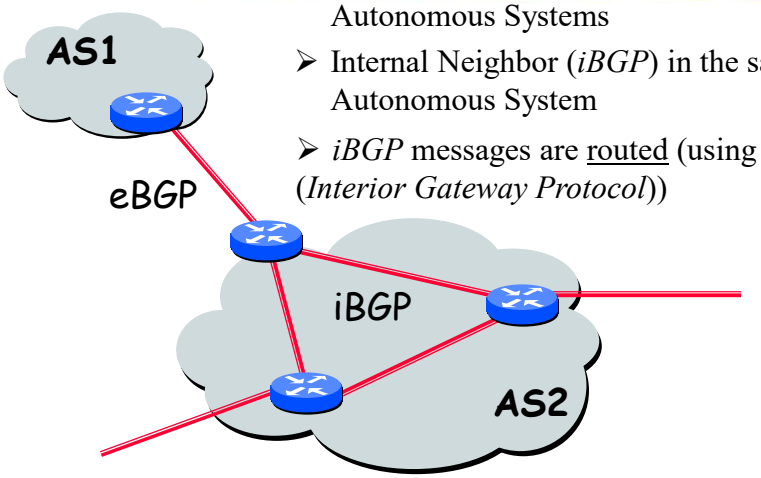
Examples of BGP advertising routes



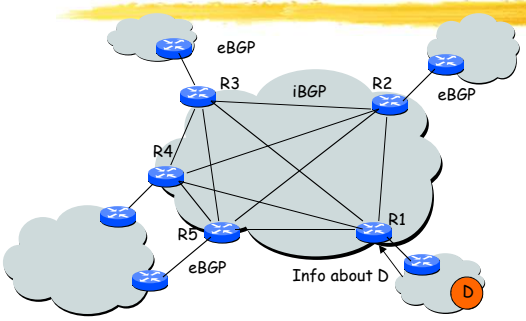
- A advertises to B the path AW
- B advertises to X the path BAW
- Should B advertise to C the path BAW?
  - No way! B gets no "revenue" for routing CBAW since neither W nor C are B's customers
  - B wants to force C to route to W via A
  - B wants to route only to/from its customers!

Two Types of BGP Neighbor Relationships

- External Neighbor (*eBGP*) in different Autonomous Systems
- Internal Neighbor (*iBGP*) in the same Autonomous System
- *iBGP* messages are routed (using *IGP* (*Interior Gateway Protocol*))



Why do we need iBGP?



Tables at R5

Prefix	BGP Next Hop
18.1/16	R2
18.7.7/24	R3
24.34/16	R1
24.34/16	R1

Router	IGP Path
R1	R2
R2	R2
R3	R4
R4	R4

Prefix	BGP Next Hop
18.1/16	R2
18.7.7/24	R4
24.34/16	R2
24.34/16	R2

- *eBGP* (*exterior BGP*) runs between AS-s
- *iBGP* (*interior BGP*) redistribute information from speaker BGP to other routers in the AS

Why do we need iBGP?

- We want all BGP routers to have the exact same information for all remote networks
  - Note that BGP routers may learn different routes to the same network thru their eBGP pairs
- Why not use IGP to exchange routes learned thru eBGP?
  - There are some differences between BGP and IGPs
    - BGP does not use periodic updates but IGPs use it
    - BGP messages carry various types of attributes but IGPs do not have them which causes loss of important route info

How to integrate intra-domain and inter-domain routing

- Stub-AS-s
  - Default routes to provider network
- Regional AS-s (small ISPs)
  - Inject selected routes into IGP, for others use default routes
- Backbone AS-s
  - No default routes
  - Use iBGP to effectively redistributed info about external routes
    - Enables IGP routers to learn the best border router for an external route
    - Each IGP router also keeps track of how to reach each border router

