

Lecture 17:

Border Gateway Protocol

CSE 123: Computer Networks
Alex C. Snoeren



Some figures courtesy Mike Freedman

Lecture 17 Overview



- Border Gateway Protocol (BGP)
 - The canonical path vector protocol
 - How routing gets done on the Internet today



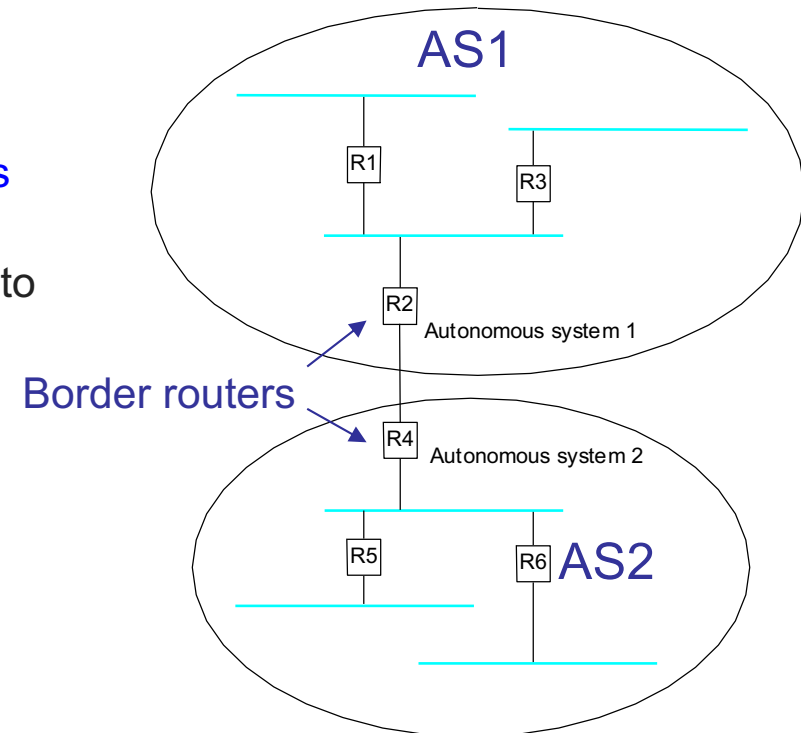
Autonomous Systems

- Internet is divided into **Autonomous Systems**
 - ◆ Distinct regions of administrative control
 - ◆ Routers/links managed by a single “institution”
 - ◆ Service provider, company, university, ...
- Hierarchy of Autonomous Systems
 - ◆ Large, “tier-1” provider with a nationwide backbone
 - ◆ Medium-sized regional provider with smaller backbone
 - ◆ Small network run by a single company or university
- Interaction between Autonomous Systems
 - ◆ Internal topology is not shared between ASes
 - ◆ ... but, neighboring ASes interact to coordinate routing

Inter-domain Routing



- **Border routers** summarize and advertise their 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
- But what routing protocol?





Issues with Link-state

- Topology information is flooded
 - ◆ High bandwidth and storage overhead
 - ◆ Forces nodes to divulge sensitive information
- Entire path computed locally per node
 - ◆ High processing overhead in a large network
- Minimizes some notion of total distance
 - ◆ Works only if policy is shared and uniform
- Typically used only inside an AS
 - ◆ E.g., OSPF and IS-IS



Distance Vector *almost there*

- Advantages
 - ◆ Hides details of the network topology
 - ◆ Nodes determine only “next hop” toward the destination
- Disadvantages
 - ◆ Minimizes some notion of total distance, which is difficult in an interdomain setting
 - ◆ Slow convergence due to the counting-to-infinity problem (“bad news travels slowly”)
- Idea: extend the notion of a distance vector
 - ◆ To make it easier to detect loops

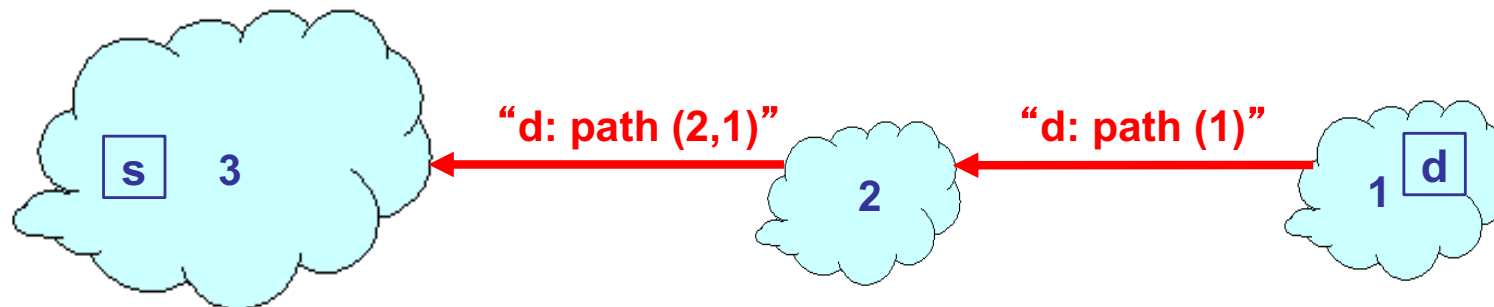


Path-vector Routing

- Extension of distance-vector routing
 - ◆ Support flexible routing policies
 - ◆ Avoid count-to-infinity problem
- Key idea: advertise the entire path
 - ◆ Distance vector: send *distance metric* per destination
 - ◆ Path vector: send the *entire path* for each destination

How does traffic flow from d to s?

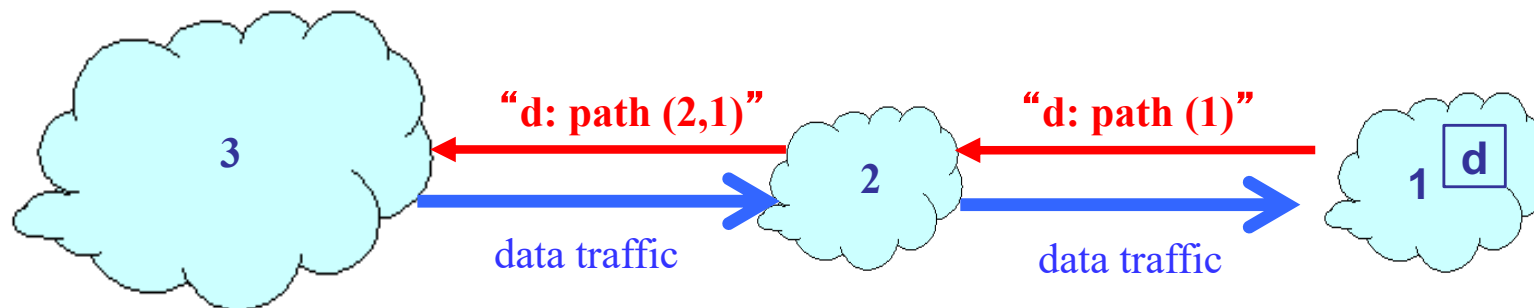
- A. 3, 2, 1
- B. 1, 2, 3
- C. 3, 1
- D. You can't tell





Path-vector Routing

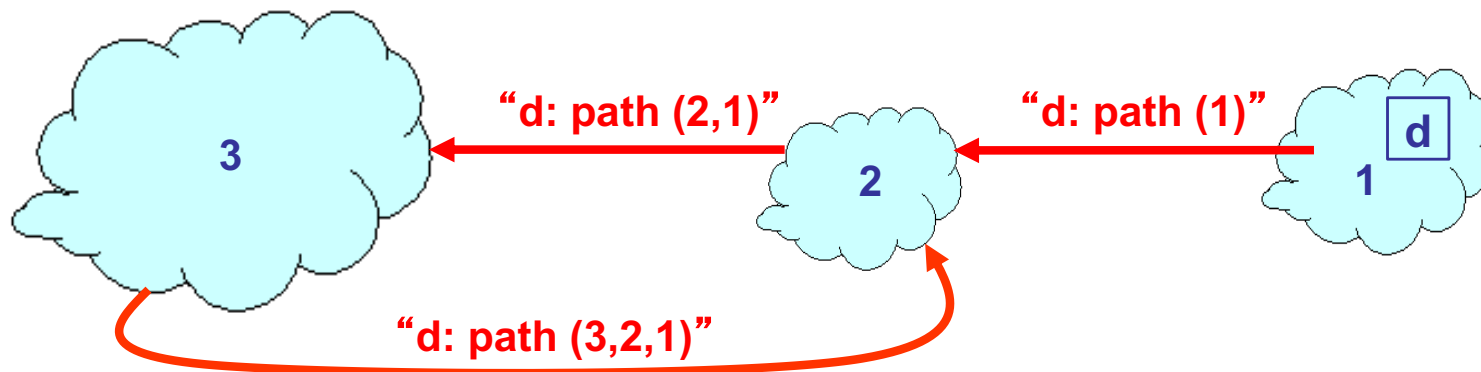
- Extension of distance-vector routing
 - ◆ Support flexible routing policies
 - ◆ Avoid count-to-infinity problem
- Key idea: advertise the entire path
 - ◆ Distance vector: send *distance metric* per destination
 - ◆ Path vector: send the *entire path* for each destination





Loop Detection

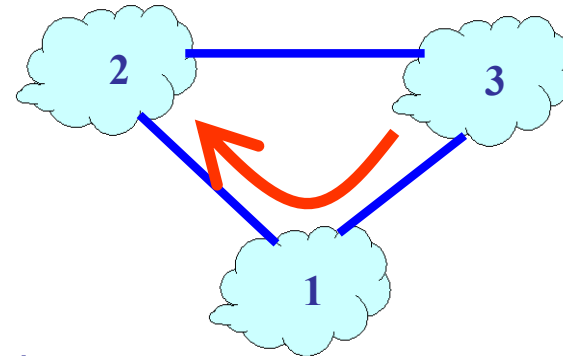
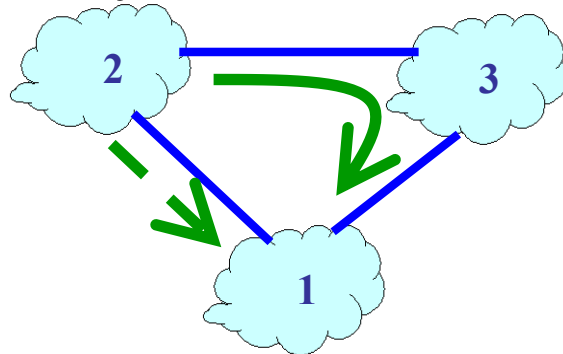
- Routers can easily detect a loop
 - ♦ Look for its own node identifier in the path
 - ♦ E.g., AS 2 sees itself in the path “3, 2, 1”
- Routers discard paths with loops
 - ♦ E.g., AS 2 simply discards the advertisement



Policy Support



- Each AS can apply local policies
 - ◆ Path **selection**: Which path to use?
 - ◆ Path **export**: Which paths to advertise?
- Examples
 - ◆ AS 2 may prefer the path “2, 3, 1” over “2, 1”
 - ◆ AS 1 may not let node 3 hear the path “1, 2”



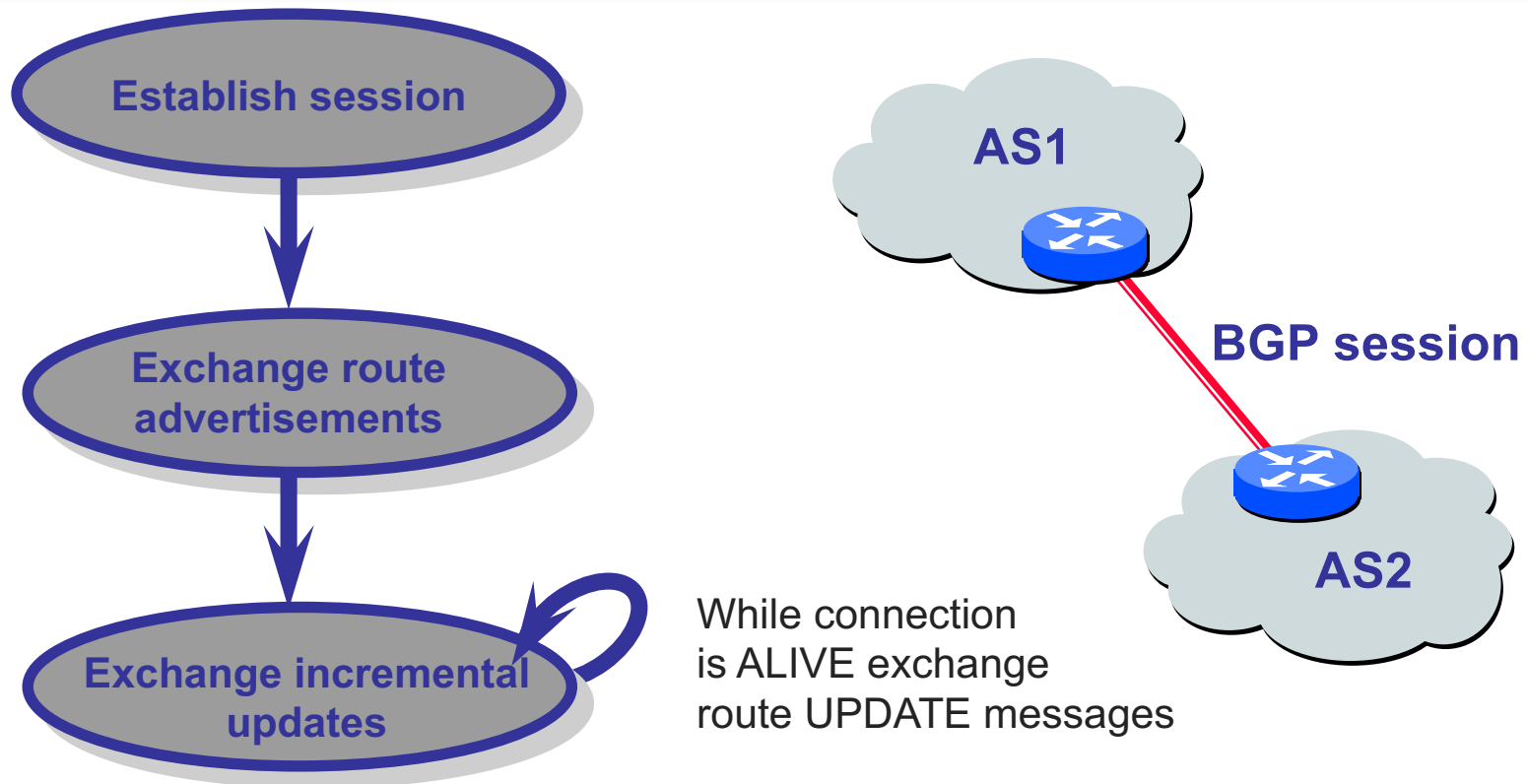
Border Gateway Protocol



- Interdomain routing protocol for the Internet
 - ◆ Prefix-based path-vector protocol
 - ◆ Policy-based routing based on AS Paths
 - ◆ Evolved during the past 30 years

- **1989 : BGP-1 [RFC 1105], replacement for EGP**
- **1990 : BGP-2 [RFC 1163]**
- **1991 : BGP-3 [RFC 1267]**
- **1995 : BGP-4 [RFC 1771], support for CIDR**
- **2006 : BGP-4 [RFC 4271], update**

Basic BGP Operation



Step-by-Step

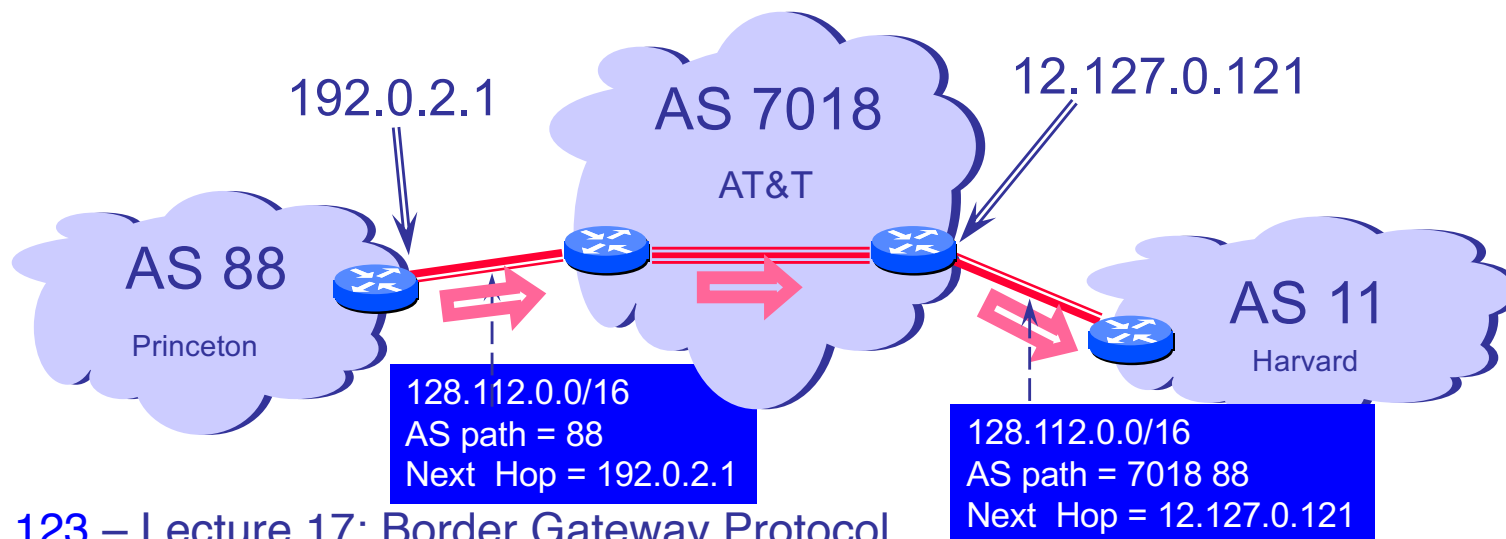


- A node learns multiple paths to destination
 - ◆ At most one from each neighbor
 - ◆ Stores all of the learned routes in a routing table
 - ◆ Applies policy to select a single active route
 - ◆ ... and may advertise the route to its neighbors
- Two types of incremental updates
 - ◆ Announcement
 - » Upon selecting a new active route, add own node id to path
 - » ... and (optionally) advertise to each neighbor
 - ◆ Withdrawal
 - » If the active route is no longer available
 - » ... send a withdrawal message to the neighbors



A Simple BGP Route

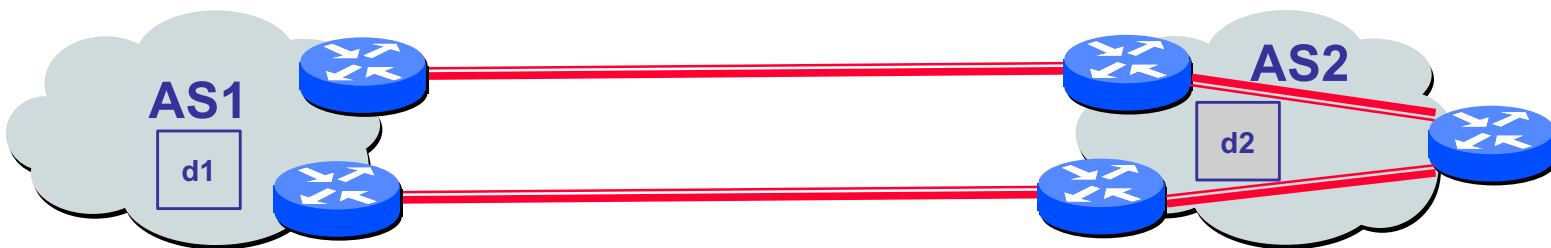
- Destination prefix (e.g., 128.112.0.0/16)
- Route attributes, including
 - ♦ AS path (e.g., “7018 88”)
 - ♦ Next-hop IP address (e.g., 12.127.0.121)



(some) BGP Attributes



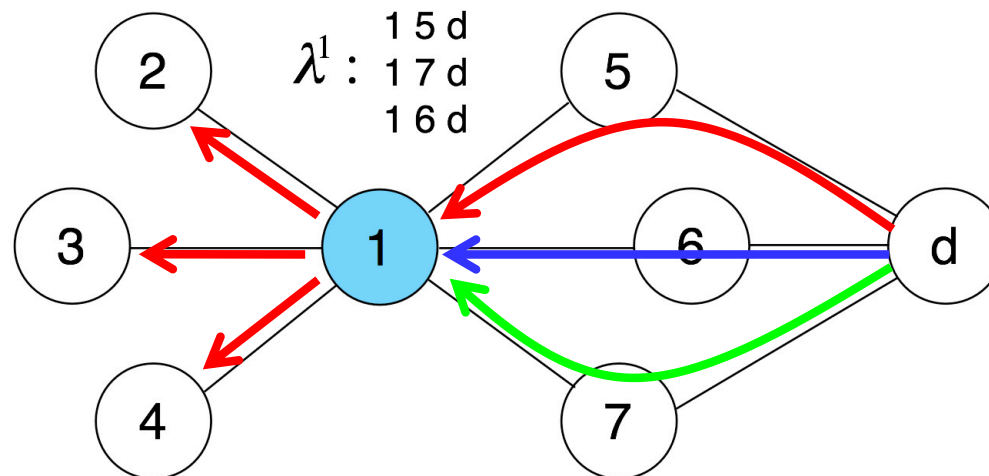
- **AS path:** ASs the announcement traversed
- **Next-hop:** where the route was heard from
- **Origin:** Route came from IGP or EGP
- **Local pref:** Statically configured ranking of routes within AS
- **Multi Exit Discriminator (MED):** preference for where to *exit* network
- **Community:** opaque data used for inter-ISP policy



Export Active Routes



- In conventional path vector routing, a router has one **ranking function**, which reflects its routing policy



BGP Decision Process



- Default decision for route selection
 - ♦ Highest local pref, shortest AS path, lowest MED, prefer eBGP over iBGP, lowest IGP cost, router id
- Many policies built on default decision process, but...
 - ♦ Possible to create arbitrary policies in principle
 - » Any criteria: BGP attributes, source address, prime number of bytes in message, ...
 - » Can have separate policy for inbound routes, installed routes and outbound routes
 - ♦ Limited only by power of vendor-specific routing language

For Next Time



- Read P&D 3.5
- Homework 3 due Friday