



# The Global Internet (cont'd)

\* Some slides are adapted from *Computer Networking: A Top-Down Approach*. J.F Kurose and K.W. Ross. All Rights Reserved.

Spring 2026

© CS 438 Staff, University of Illinois

1

# [ Learning Objectives ]

- Autonomous Systems
- Intradomain Routing
- Interdomain Routing
  - *BGP*

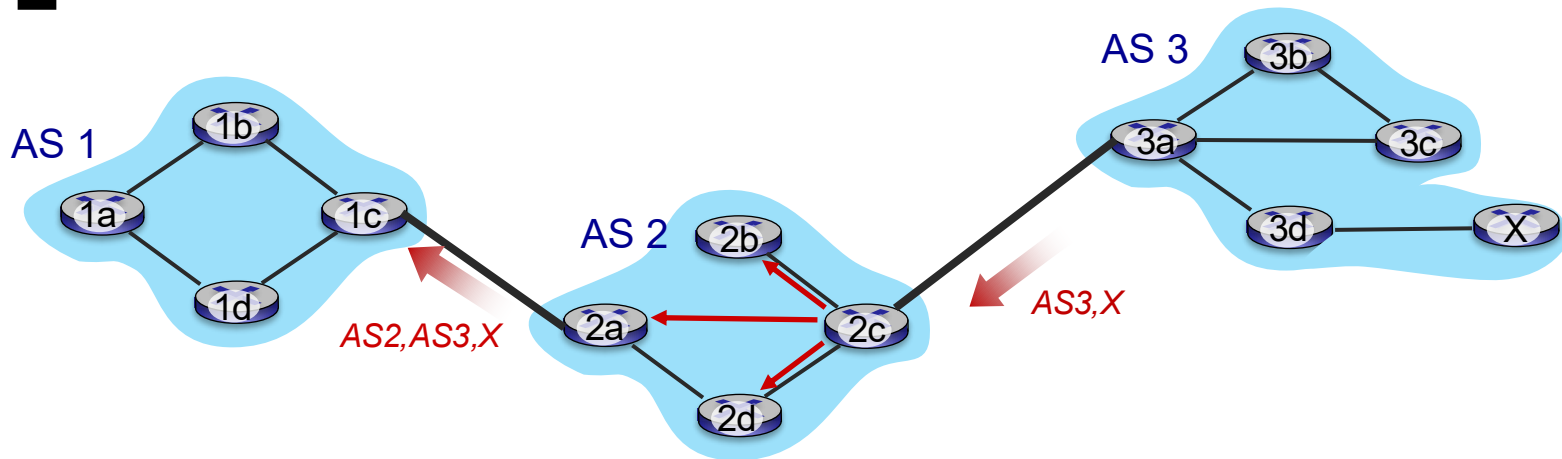


# [ BGP Recap (and History) ]

- Interdomain routing protocol for the Internet
  - Path-vector: AS path towards a destination IP prefix
  - Policy-based: path selection and export
  - Evolved over the years
- 1989 : BGP-1 [RFC 1105]
  - Replacement for EGP (1984, RFC 904)
- 1990 : BGP-2 [RFC 1163]
- 1991 : BGP-3 [RFC 1267]
- 1995 : BGP-4 [RFC 1771]
  - Support for Classless Interdomain Routing (CIDR)



# eBGP and iBGP



- AS2 router 2c receives path advertisement **AS3,X** (via **eBGP**) from AS3 router 3a
- based on AS2 policy, AS2 router 2c accepts path **AS3,X** and propagates (via **iBGP**) to all AS2 routers
- based on AS2 policy, AS2 router 2a advertises (via **eBGP**) path **AS2, AS3,X** to AS1 router 1c

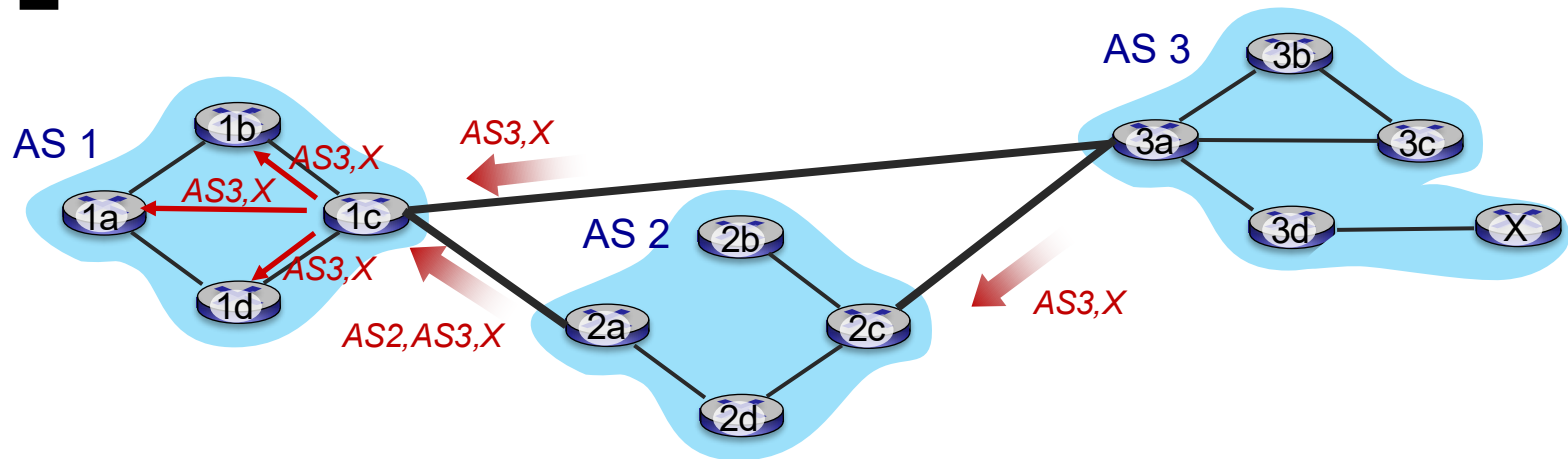


# [ BGP: Multiple Paths ]

- When a node learns multiple paths to destination
  - Stores all of the routes in a routing table
  - Applies policy to select a single active route
  - ... and may advertise the route to its neighbors



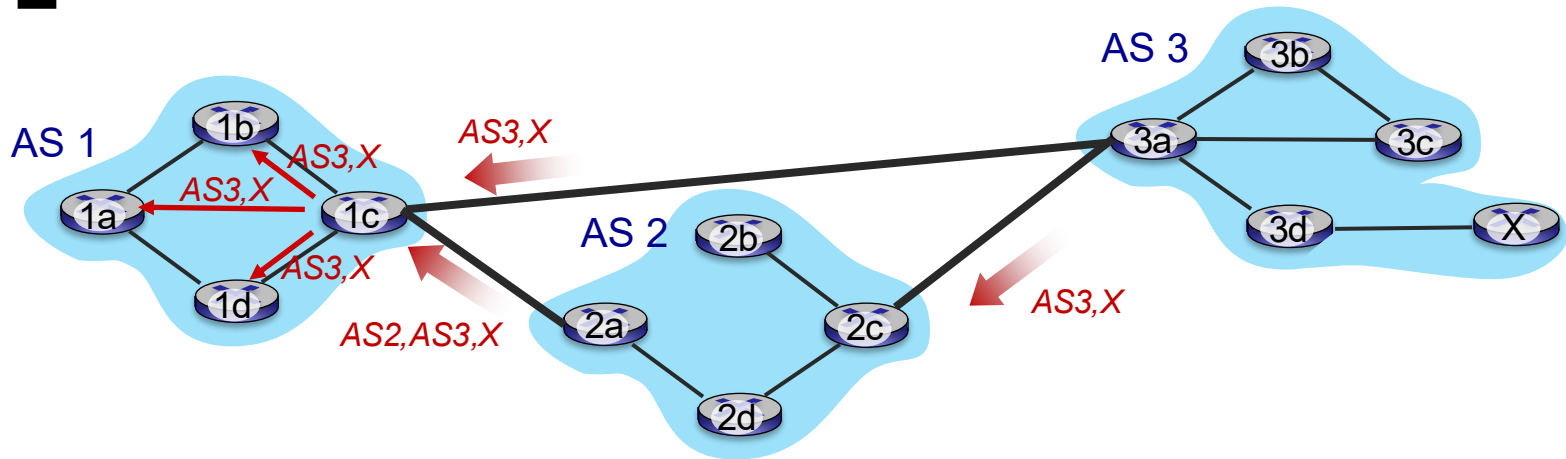
# BGP: Multiple Paths



Gateway router may learn about **multiple** paths to destination:

- AS1 gateway router 1c learns path **AS2,AS3,X** from 2a
- AS1 gateway router 1c learns path **AS3,X** from 3a
- based on *policy*, AS1 gateway router 1c chooses path **AS3,X** and advertises path within AS1

# Question Arises...



How does router 1c reach X?

- They are not even in the same AS
- Solution: tracking “BGP next hop”

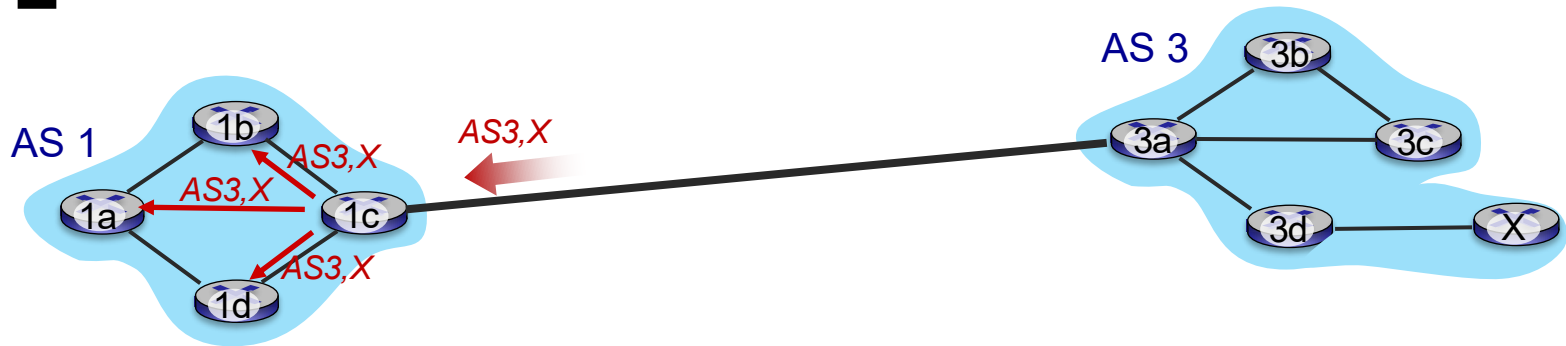


# [ BGP Next Hop ]

- BGP advertised route: IP prefix + attributes
- IP prefix: destination being advertised
- two important attributes:
  - AS-PATH: list of ASes through which prefix advertisement has passed
  - **NEXT-HOP**: IP of the next-hop router to the destination

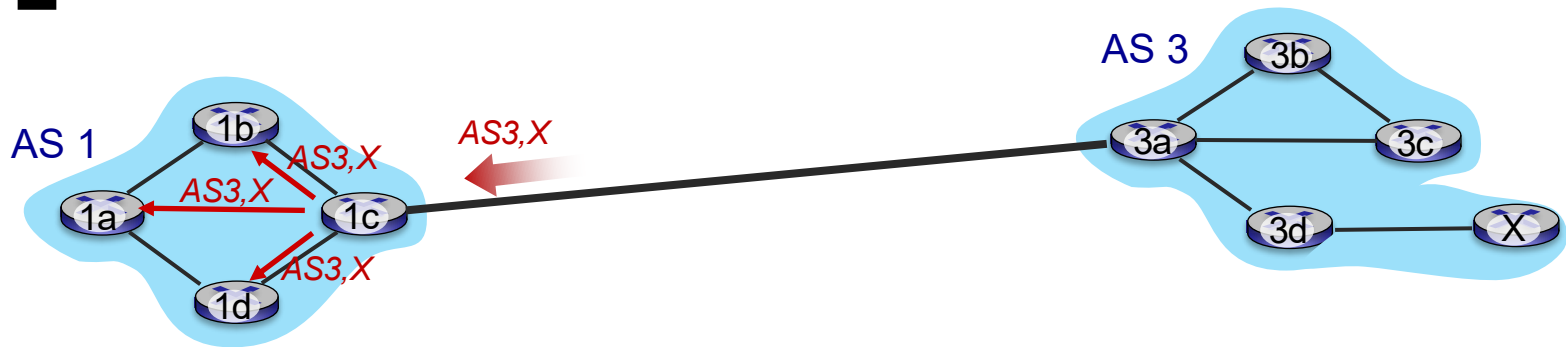


# BGP Next Hop



- 1c learns from 3a via eBGP:
  - Destination: **X**; AS\_PATH: **AS3**; NEXT\_HOP: **3a**
- At 1c:
  - To reach X, next hop is 3a (according to BGP)

# BGP Next Hop

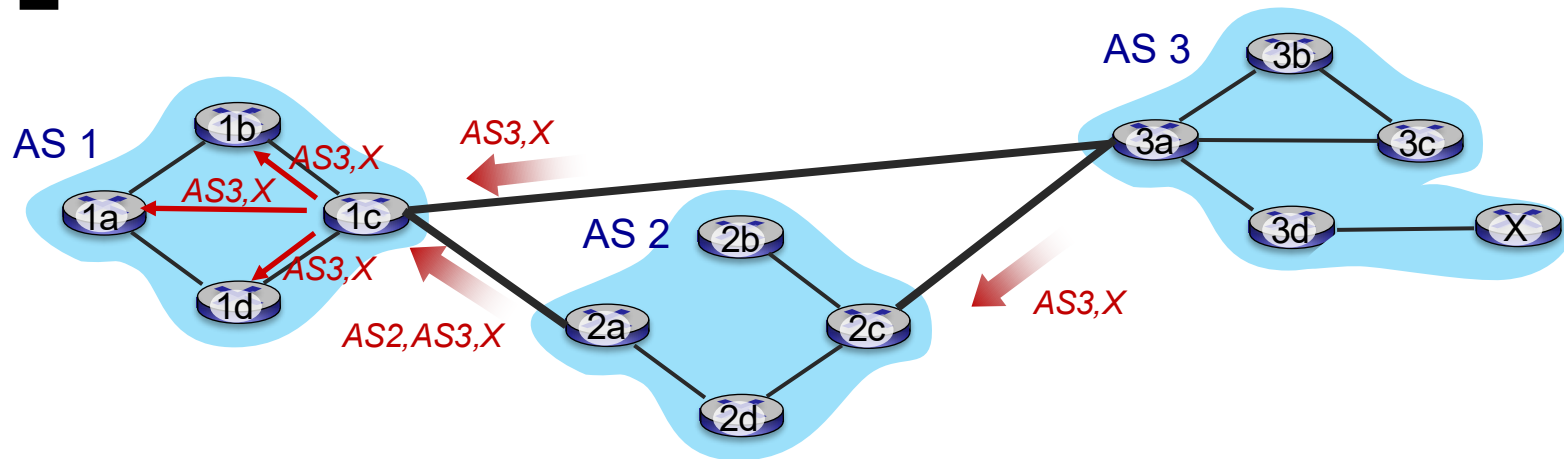


- 1d learns from 1c via iBGP:
  - Destination: **X**; AS\_PATH: **AS3**; NEXT\_HOP: **3a**
  - NEXT\_HOP is *not* updated by iBGP by default
- At 1d:
  - To reach X, next hop is 3a (by BGP) – how to reach 3a?
  - Solution: IGP (e.g., OSPF)





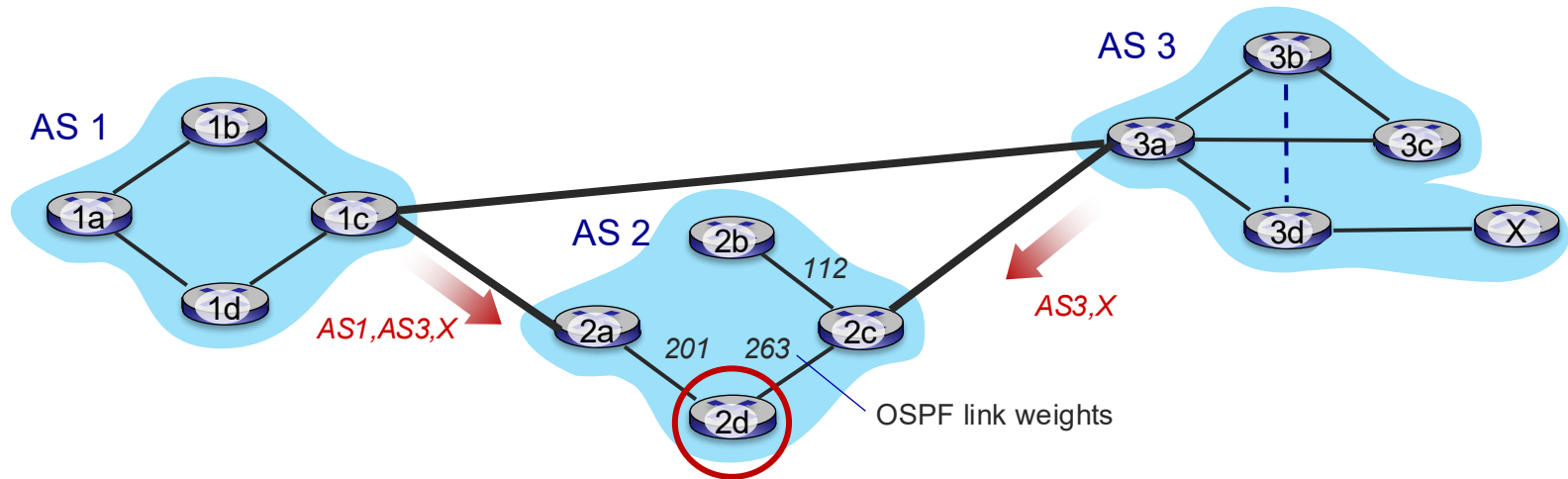
# Joining BGP and IGP



- BGP
  - Announces reachability to external destinations
  - Maps a destination prefix to an egress point (BGP next hop)
- IGP
  - Used to compute paths “within” the AS
  - Maps the egress point to an outgoing link/interface



# Simplest: Hot Potato Routing



- 2d learns (via iBGP) it can route to X via two routes
  - 2a → 1c ... X
  - 2c → 3a ... X
- **hot potato routing**: choose local gateway that has least *intra-domain* cost
  - e.g., 2d chooses 2a, even though more AS hops to X
  - don't worry about inter-domain cost! — not a good idea in practice



# [ BGP route selection in practice ]

- BGP router selects route based on:
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
  3. closest NEXT-HOP router (hot potato routing)
  4. additional criteria



# [ BGP router's routing table ]

```
ner-routes>show ip bgp
```

```
BGP table version is 6128791, local router ID is 4.2.34.165
```

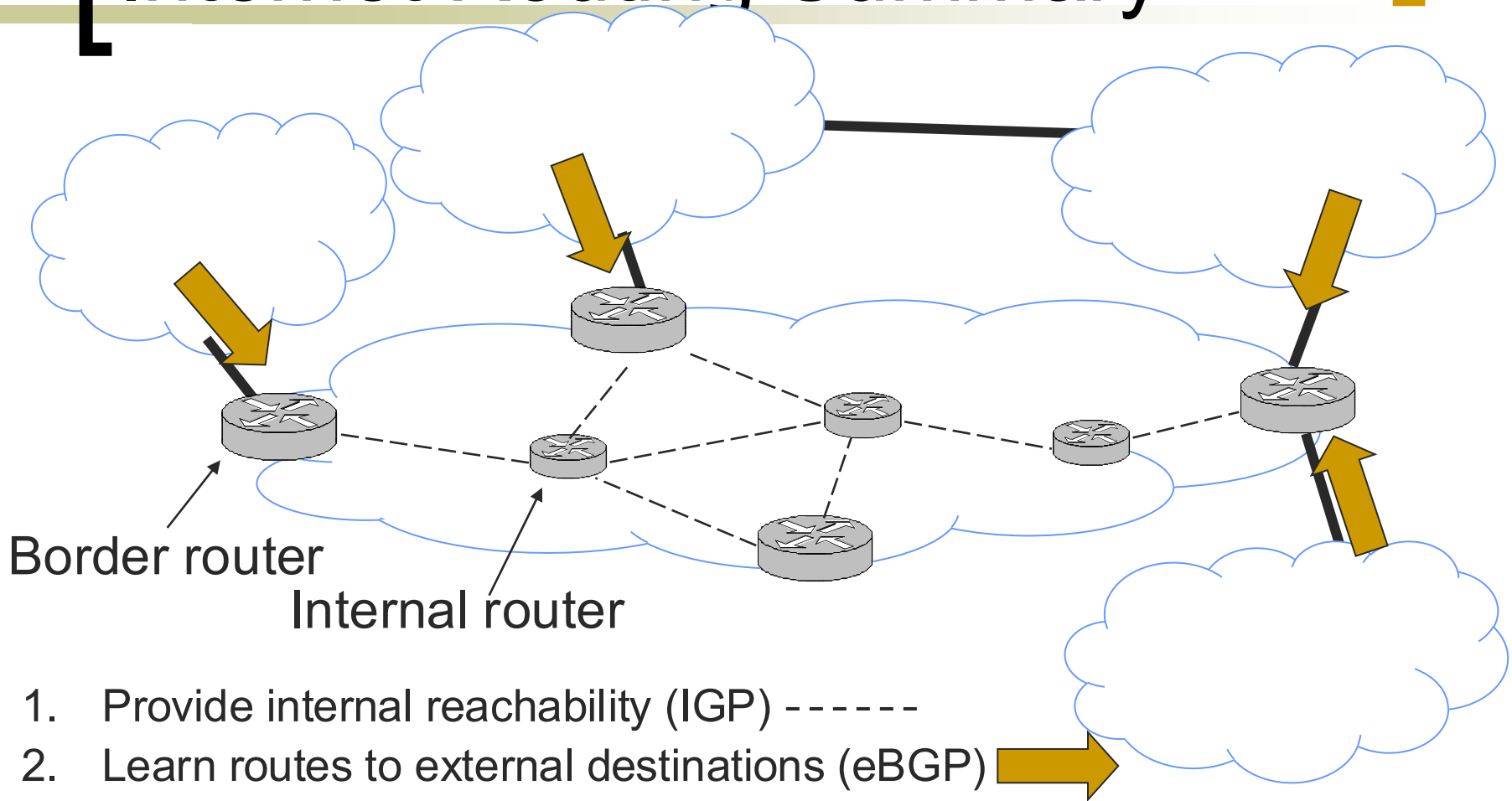
```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

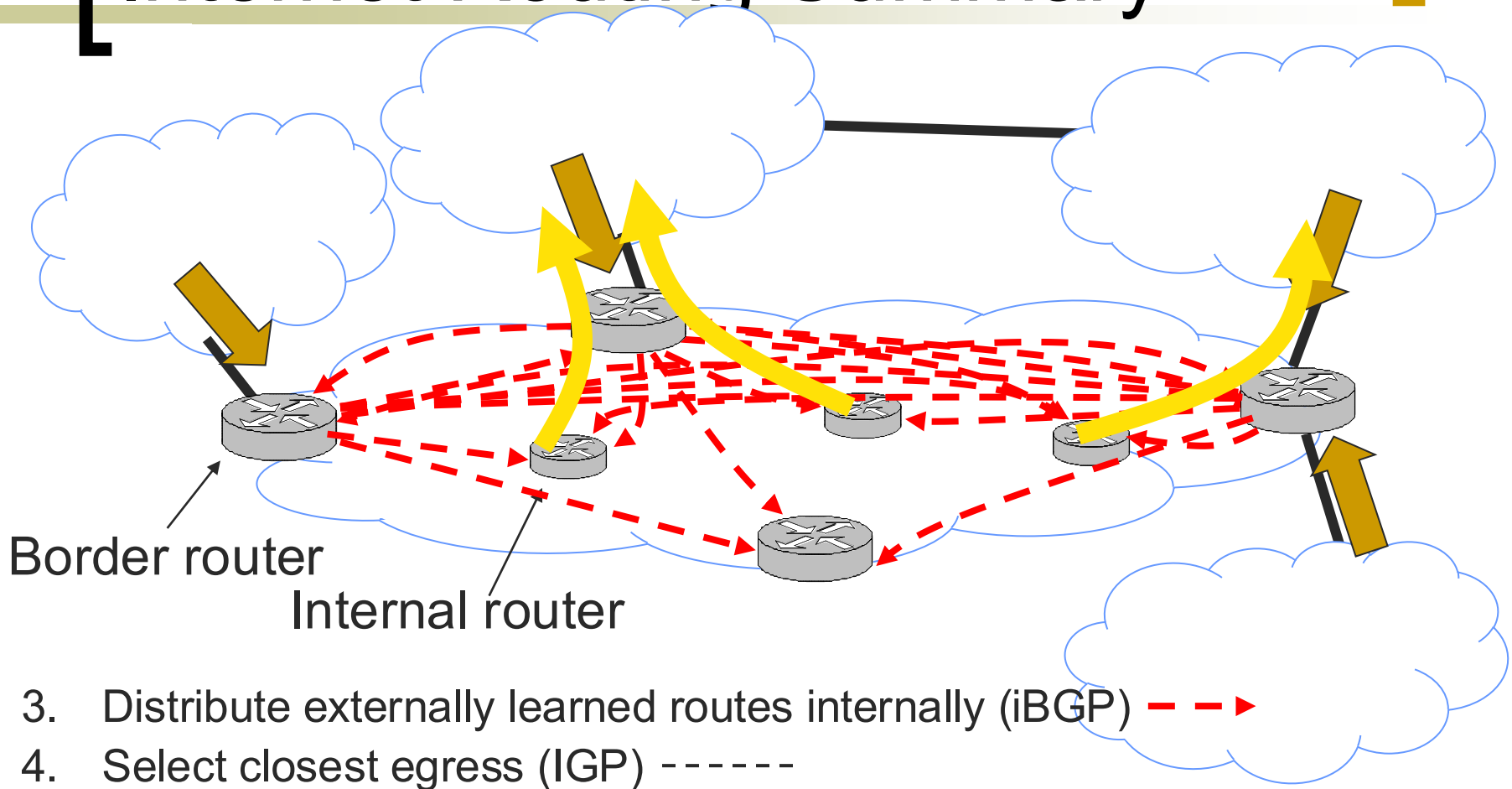
Network	Next Hop	Metric	LocPrf	Weight	Path
* i3.0.0.0	4.0.6.142	1000	50	0	701 80 i
* i4.0.0.0	4.24.1.35	0	100	0	i
* i12.3.21.0/23	192.205.32.153	0	50	0	7018 4264 6468 ?
* e128.32.0.0/16	192.205.32.153	0	50	0	7018 4264 6468 25 e



# [ Internet Routing Summary ]

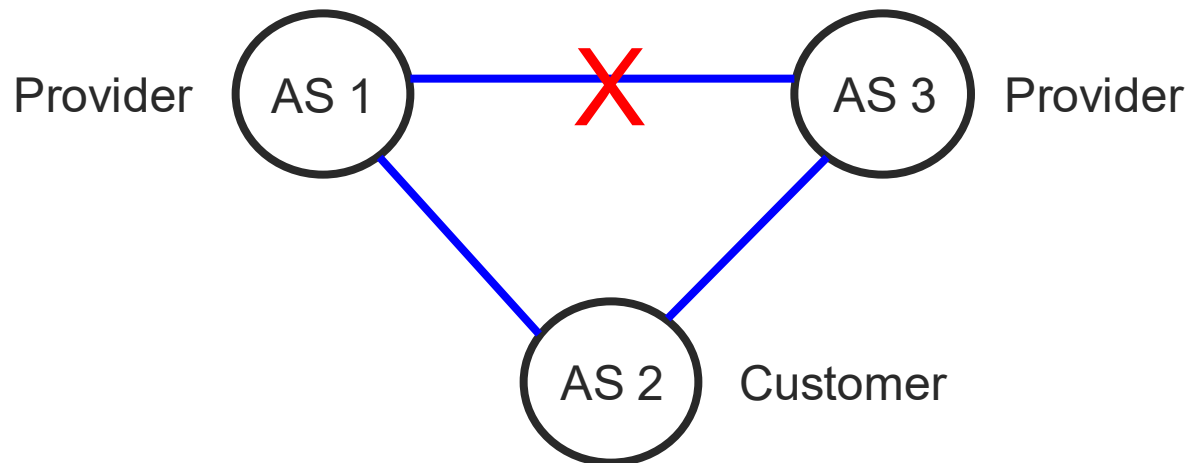


# [ Internet Routing Summary ]



# BGP Issues: Reachability

- Normal routing
  - If graph is connected, reachability is assured
- Policy routing
  - Does not always hold



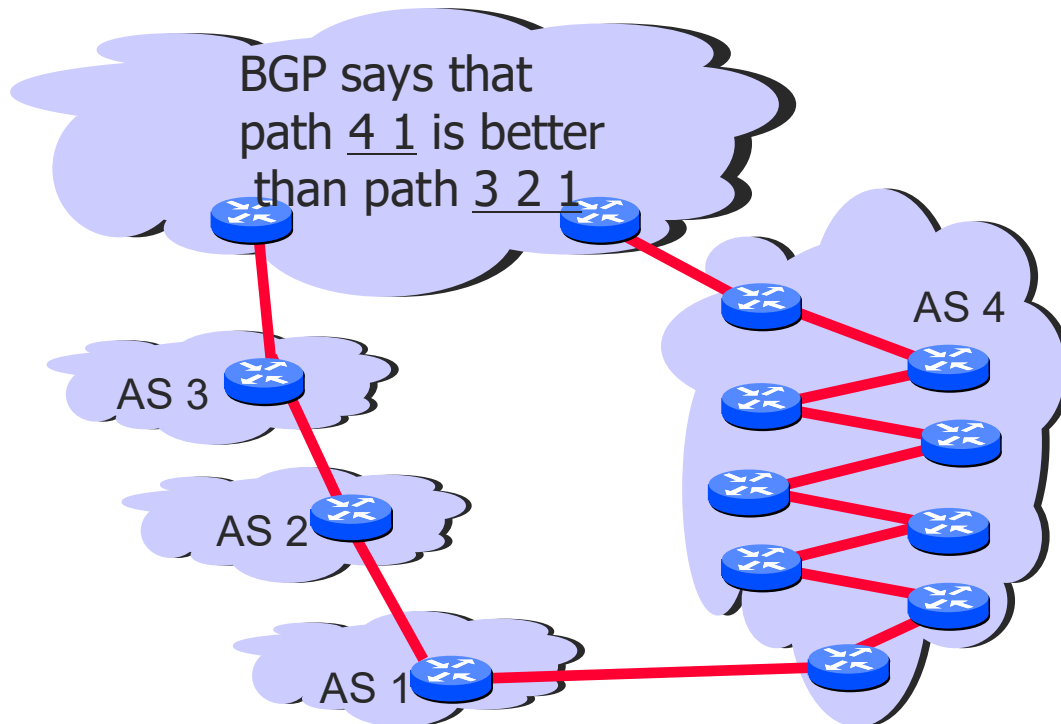
# [ BGP Issues: Performance ]

- BGP designed for policy not performance
- BGP “shortest paths” are not shortest
  - Fewest ASes != Fewest number of routers



# BGP Issues: Performance

- AS path length can be misleading
  - An AS may have many router-level hops



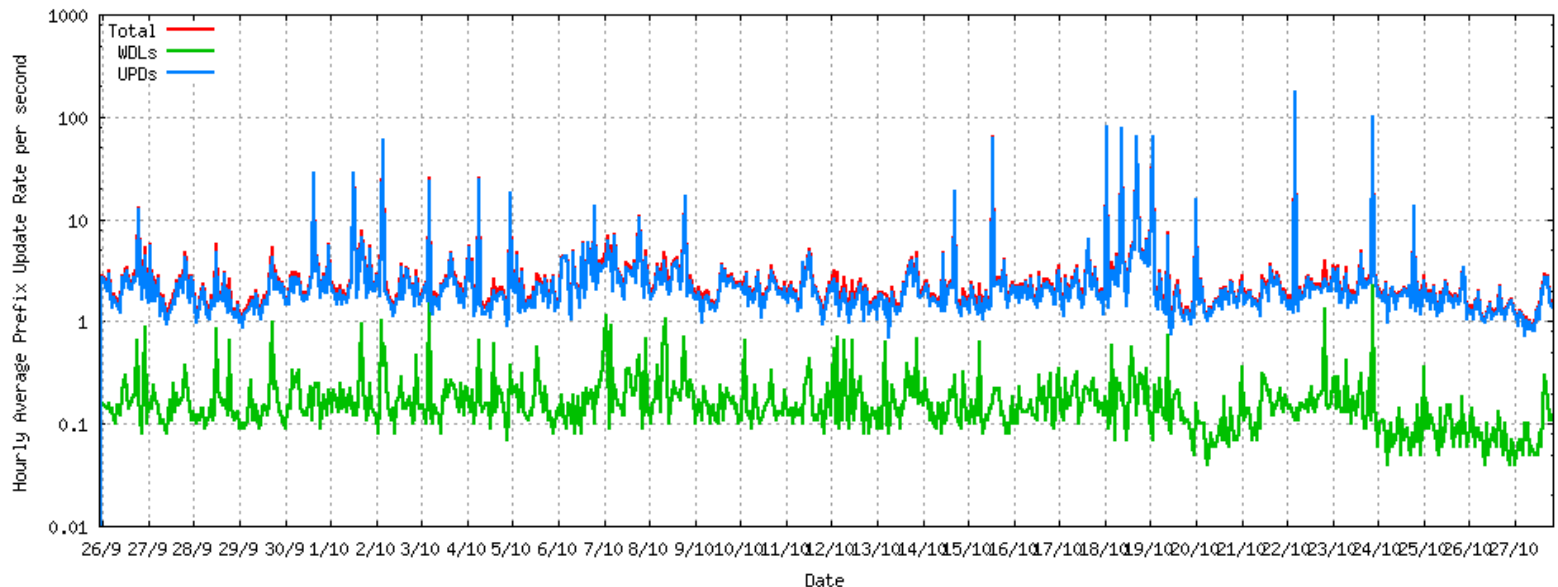
# [ Recall that BGP is distributed ]

- Incremental updates
  - Announcement
    - Upon selecting a new active route, add node id to path
    - ... and (optionally) advertise to neighbors
  - Withdrawal
    - If the active route is no longer available
    - ... send a withdrawal message to the neighbors



# BGP Issues: Dynamics

- Change in the path
  - Path must be re-advertised to every node using the path



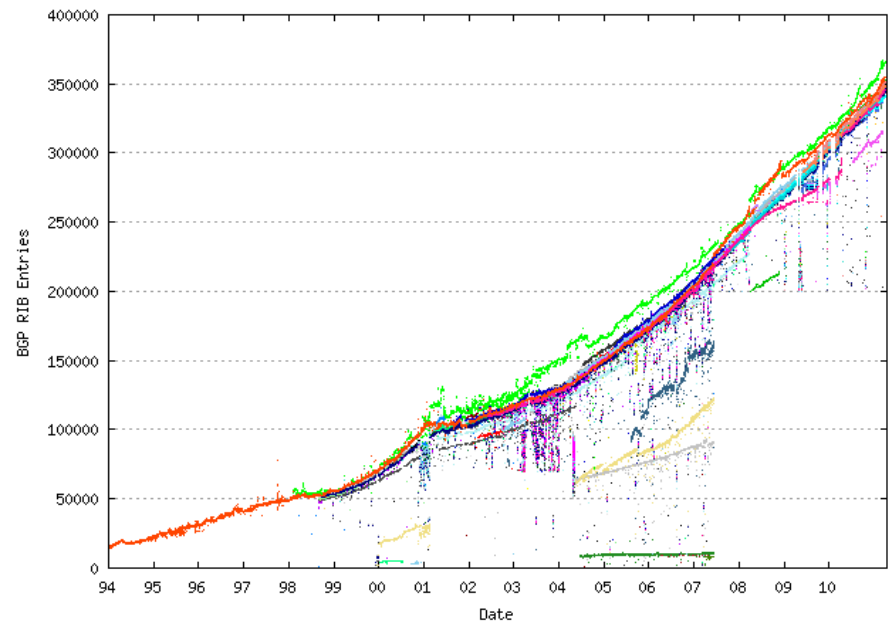
BGP updates per hour per prefix



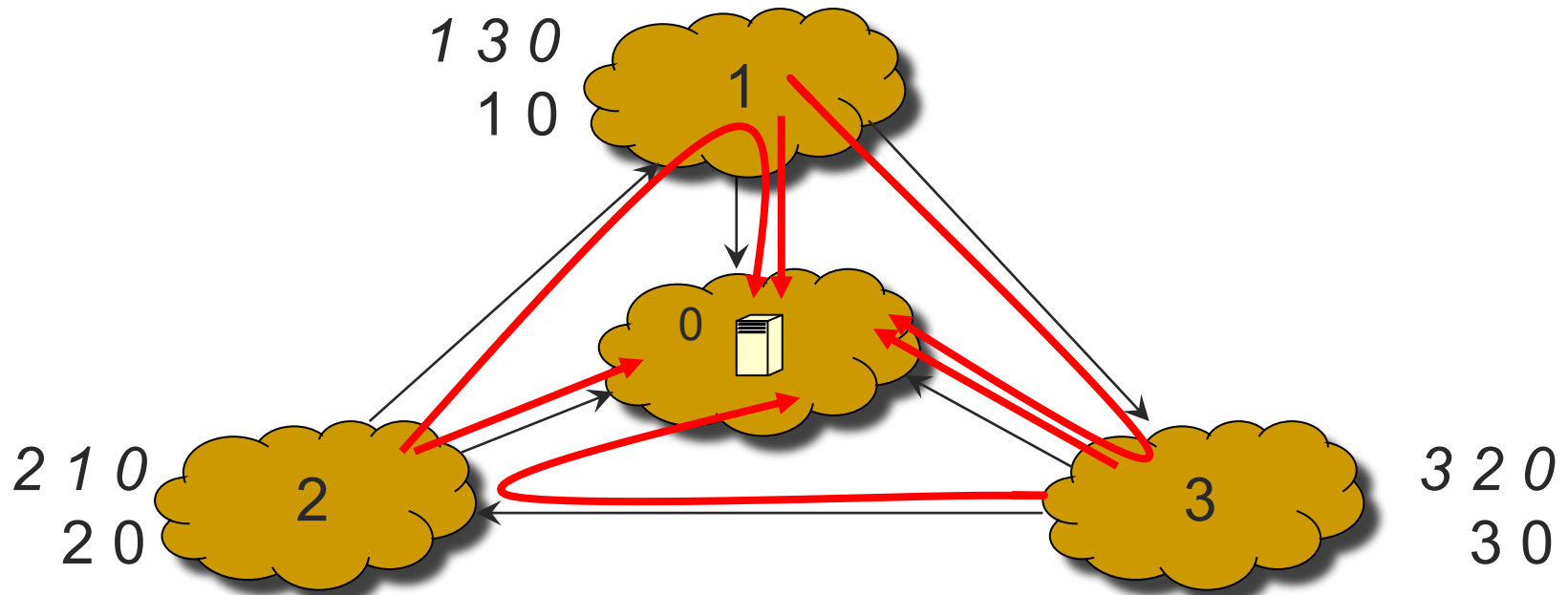
# BGP Issues: Routing Table Size

- Each BGP router must know path to every other IP prefix
  - But router memory is expensive and thus constrained
- Number of prefixes growing more than linearly
- Subject of current research

Number of prefixes in BGP table  
(often half a million routes)



# [ BGP Issues: Routing Stability ]



# BGP Issues: Routing Stability

- Policy autonomy vs. network stability
- Not an easy problem
  - Difficult to decide whether given policies will eventually converge!
- However, if policies follow normal business practices, stability is guaranteed

IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 9, NO. 6, DECEMBER 2001

681

## Stable Internet Routing Without Global Coordination

Lixin Gao, *Member, IEEE*, and Jennifer Rexford, *Senior Member, IEEE*



# [ BGP Issues: Security ]

- An AS can claim to serve a prefix that they actually don't have a route to
  - Blackholing traffic
  - This is a trust issue
- Even worse: snoop on all traffic to almost any destination
  - Stealthy
- Fixable: make ASes “prove” that they own a prefix (RPKI) and an AS path (BGPsec)
  - RPKI increasingly adopted, BGPsec very limited adoption



# [ BGP Summary ]

- BGP is essential to the Internet
  - ties different organizations together
- Poses fundamental challenges....
  - leads to use of path vector approach
- ...and myriad details



# Why different intradomain and interdomain routing?

- Policy:
  - inter-AS: admin wants control over how its traffic routed, who routes through its network
  - intra-AS: single admin, so policy less of an issue
- Scale:
  - hierarchical routing saves table size, reduced update traffic
- Performance:
  - intra-AS: can focus on performance
  - inter-AS: policy dominates over performance

