

Lecture 19:

Border Gateway Protocol (again)

CSE 123: Computer Networks

Aaron Schulman

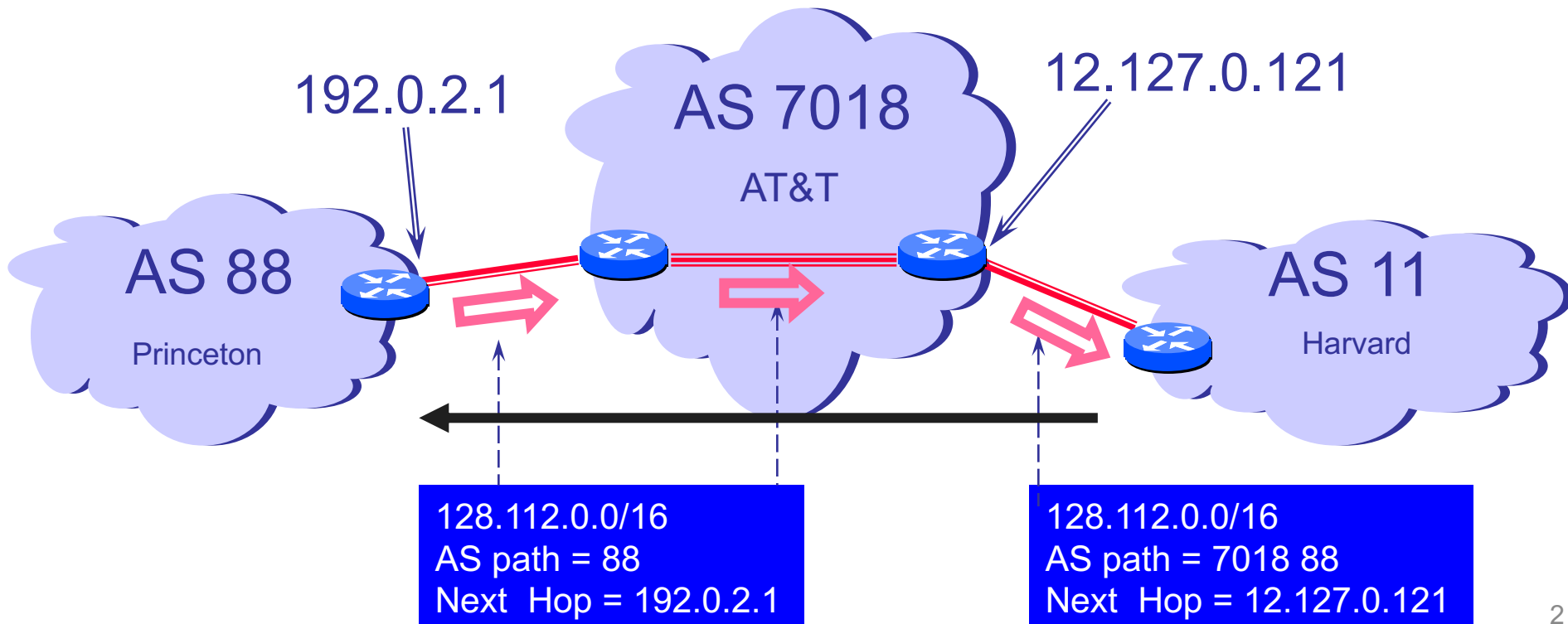


Some figures courtesy Mike Freedman



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)





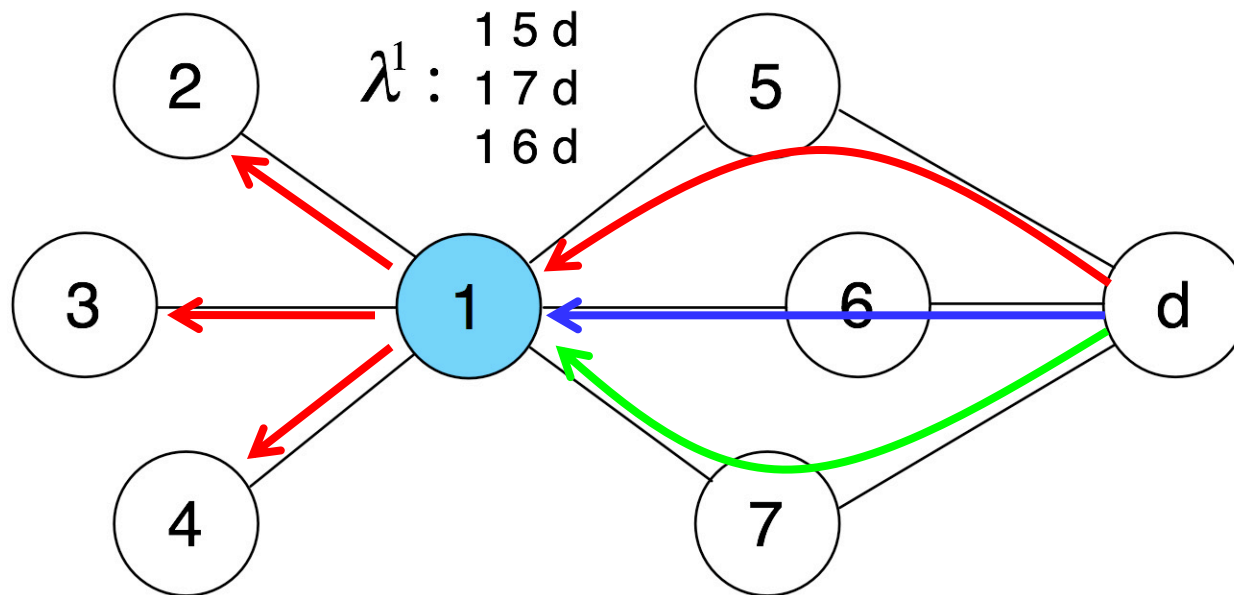
(some) BGP Attributes

- **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:** preference for where to *exit* network
- **Community:** opaque data used for inter-ISP policy



Export Active Routes

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



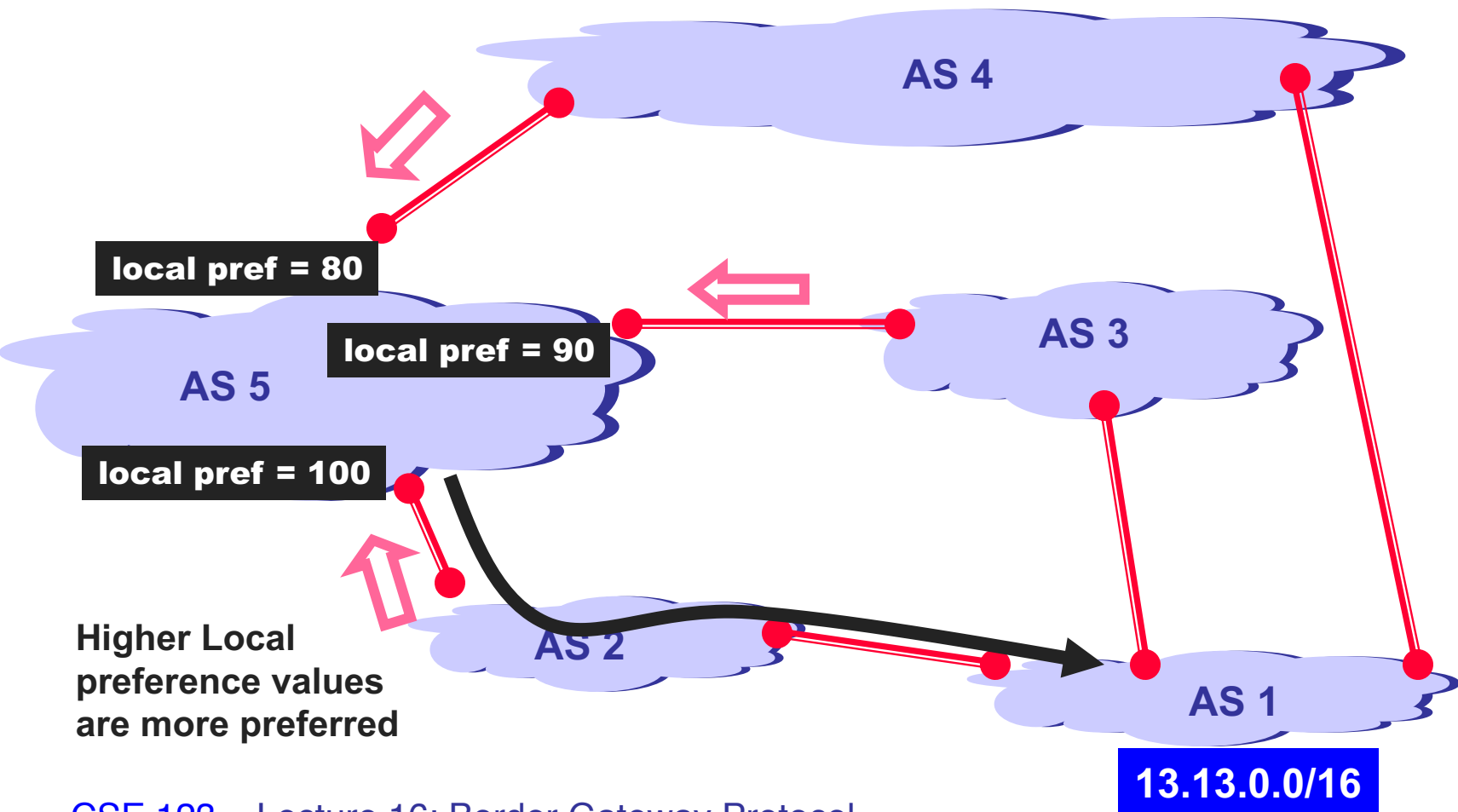


BGP Decision Process

- 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 principal
 - » 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



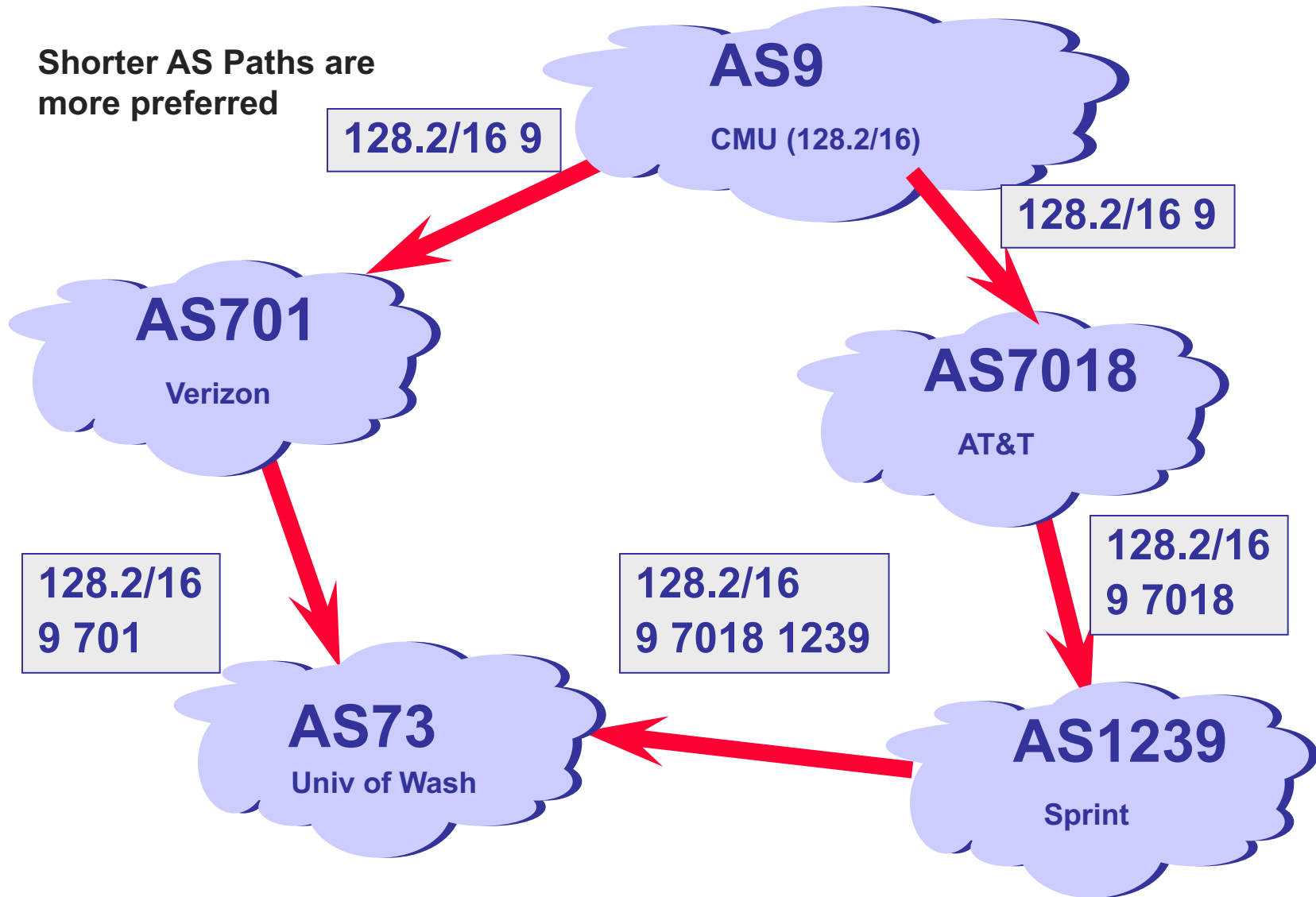
Example: Local Pref



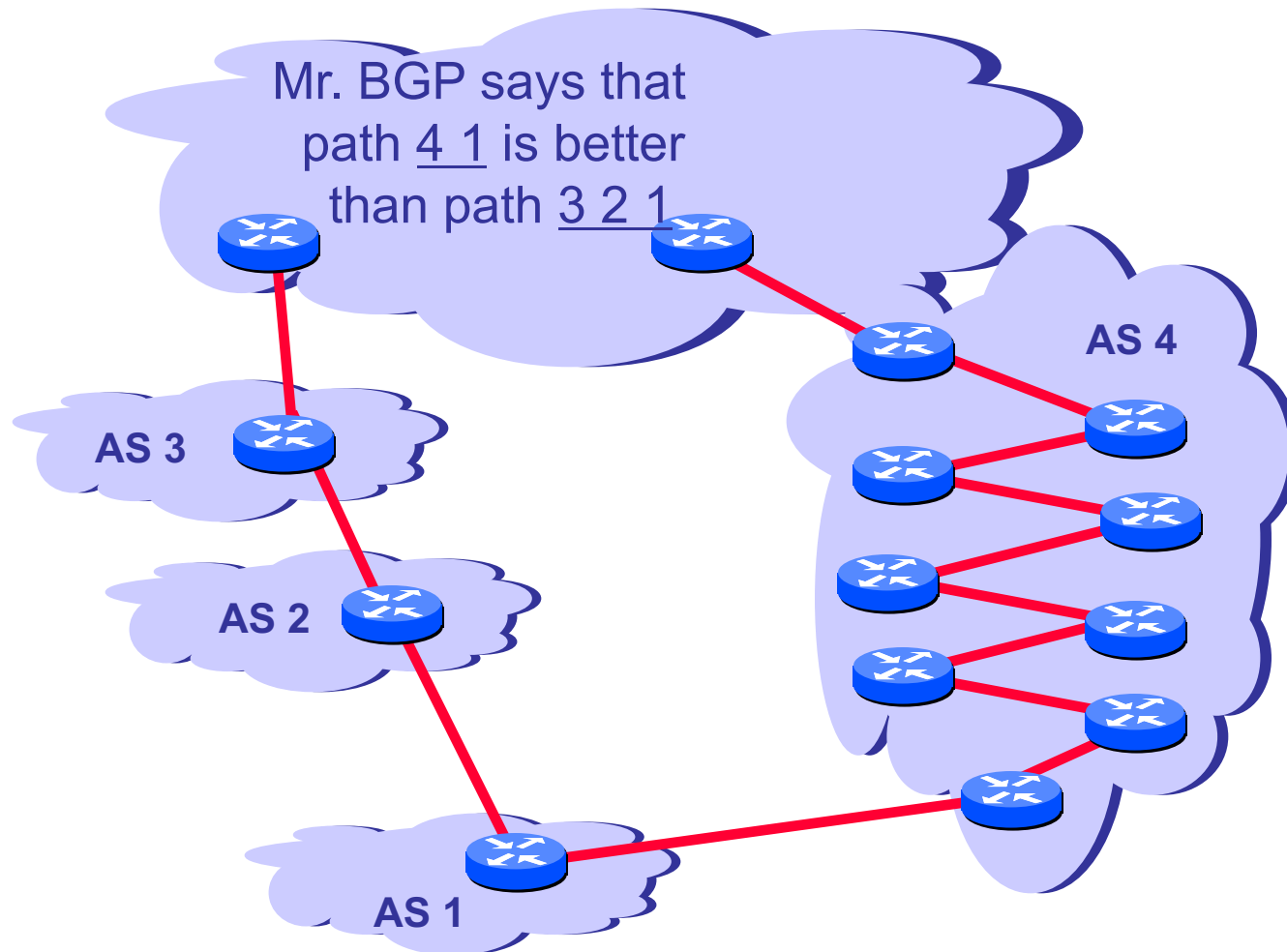


Example: Short AS Path

Shorter AS Paths are more preferred



AS Paths vs. Router Paths



BGP Has Lots of Problems



- Instability
 - ◆ Route flapping (network x.y/z goes down... tell everyone)
 - ◆ Long AS-path decision criteria defaults to DV-like behavior (bouncing)
 - ◆ Not guaranteed to converge, NP-hard to tell if it does
- Scalability still a problem
 - ◆ >500,000 network prefixes in default-free table today
 - ◆ Tension: Want to manage traffic to very specific networks (eg. multihomed content providers) but also want to aggregate information.
- Performance
 - ◆ Non-optimal, doesn't balance load across paths



Business Relationships

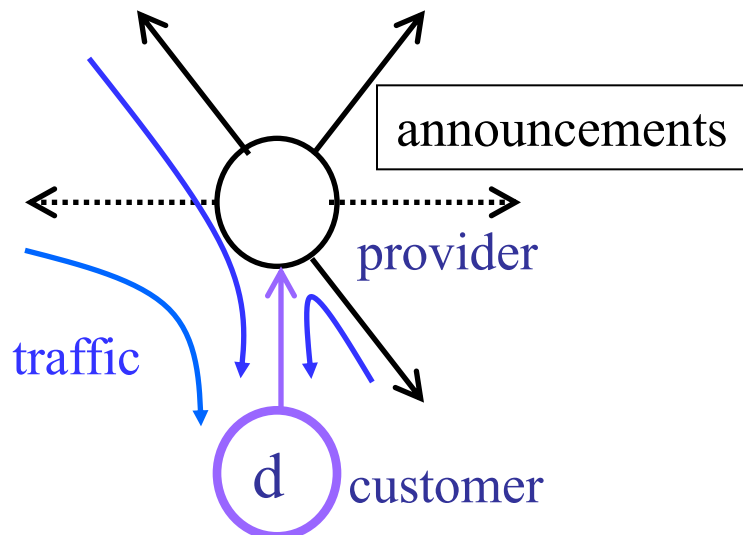
- Neighboring ASes have business contracts
 - ◆ How much traffic to carry
 - ◆ Which destinations to reach
 - ◆ How much money to pay
- Common business relationships
 - ◆ Customers pay providers (95% billing model)
 - » E.g., Princeton is a customer of USLEC
 - » E.g., MIT is a customer of Level3
 - ◆ Peers don't pay peers (exchange equal traffic for free)
 - » E.g., UUNET is a peer of Sprint
 - » E.g., Harvard is a peer of Harvard Business School
- Routing follows the money (allow inexpensive paths)



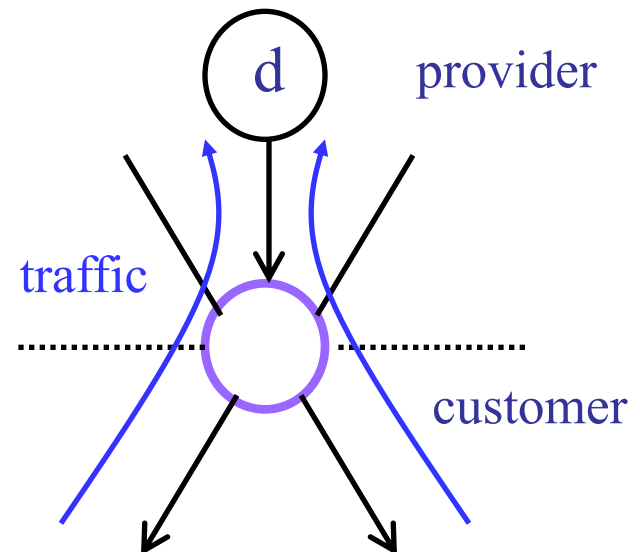
Customer/Provider

- Customer needs to be reachable from everyone
 - ◆ Provider tells all neighbors how to reach the customer
- Customer does not want to provide transit service
 - ◆ Customer does not let its providers route through it

Traffic **to** the customer



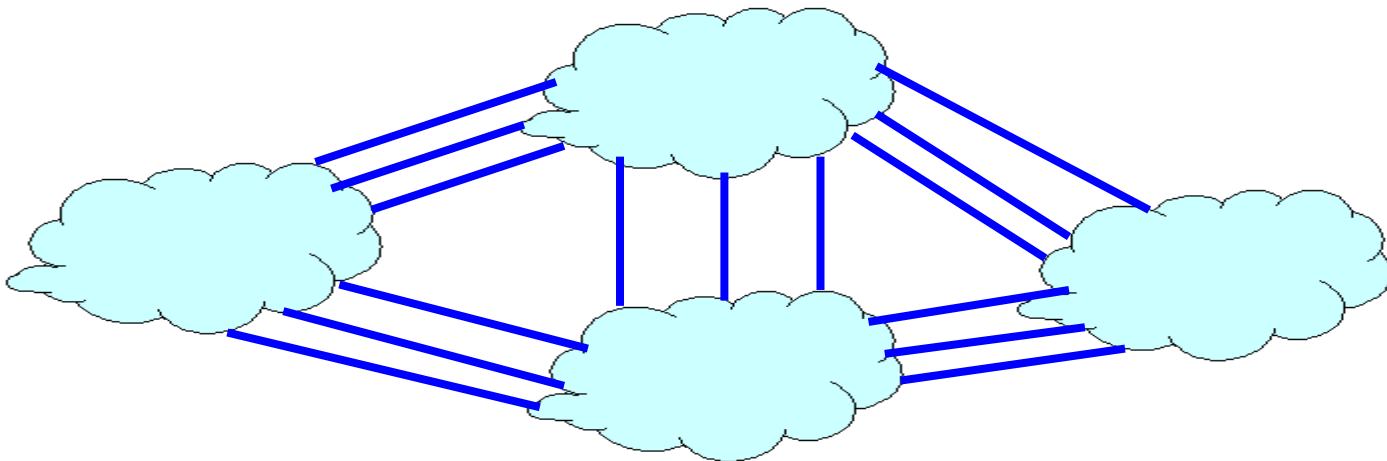
Traffic **from** the customer





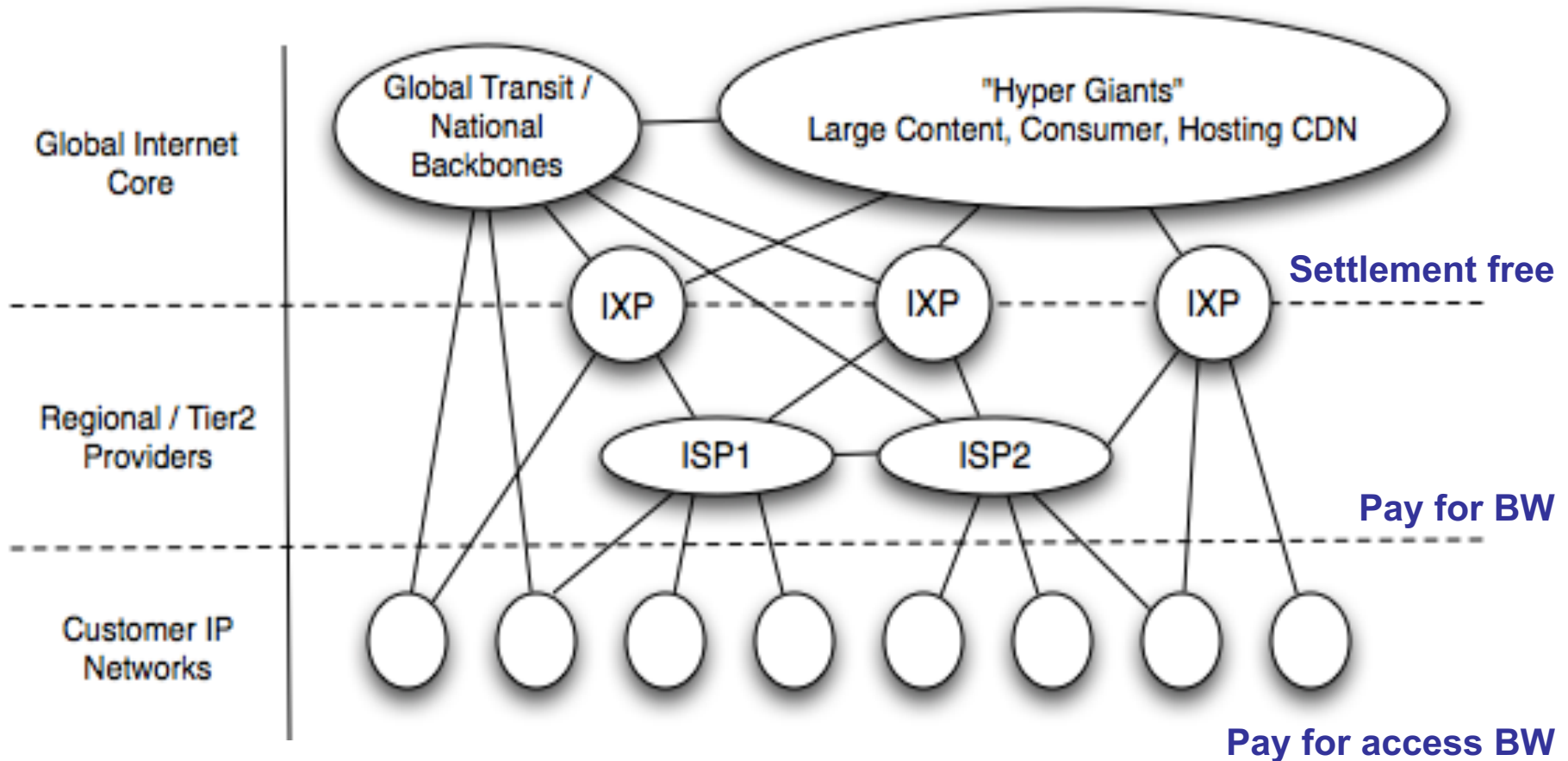
Tier-1 Providers

- Make up the “core” of the Internet
 - ◆ Has no upstream provider of its own
 - ◆ Typically has a national or international backbone
- Top of the Internet hierarchy of ~10-20 ASes
 - ◆ E.g., AT&T, Level 3 (CenturyLink), NTT/Verio, Century Link (formerly Qwest), Sprint, Verizon
 - ◆ Full peer-peer connections between tier-1 providers





The Internet Hierarchy





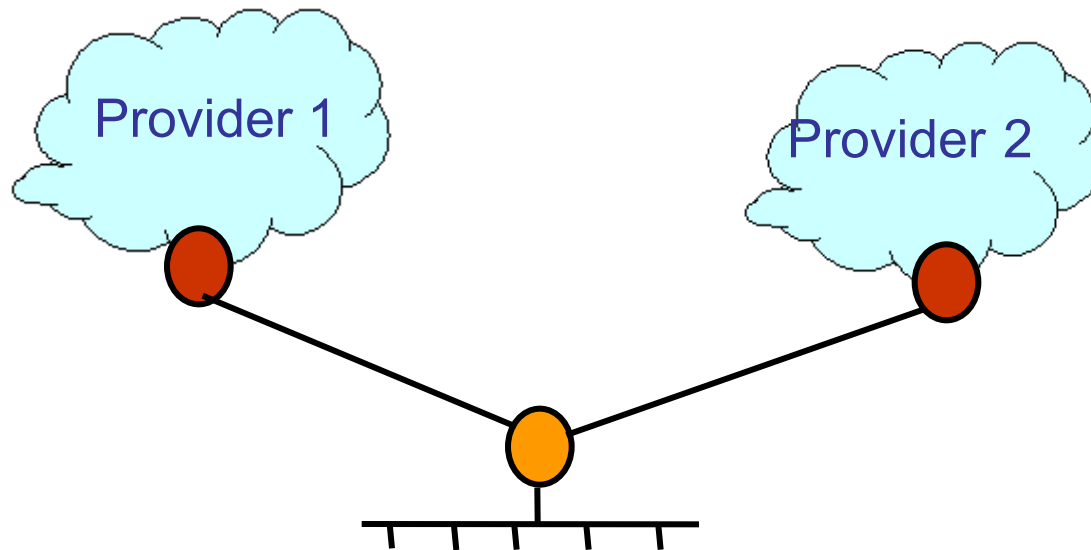
A History of Settlement

- The telephone world
 - ◆ LECs (local exchange carriers) (e.g., PacBell, NYNEX)
 - ◆ IXC (inter-exchange carriers) (e.g., Sprint, AT&T)
- LECs MUST provide IXCs access to customers
 - ◆ This is enforced by laws and regulation
- When a call goes from one phone company to another:
 - ◆ Call billed to the caller
 - ◆ The money is split up among the phone systems – this is called “settlement”



Multi-Homing

- Customers may have more than one provider
 - ◆ Extra reliability, survive single ISP failure
 - ◆ Financial leverage through competition
 - ◆ Better performance by selecting better path
 - ◆ Gaming the 95th-percentile billing model





BGP Summary

- Interdomain-routing
 - ◆ Exchange reachability information (plus hints)
 - ◆ BGP is based on path vector routing
 - ◆ Local policy to decide which path to follow
- Traffic exchange policies are a big issue \$\$\$
 - ◆ Complicated by lack of compelling economic model (who creates value?)
 - ◆ Can have significant impact on performance



For Next Time

- Read P&D 3.5 (Router Implementation)
- Homework 3 due Wed and Project 3 due