

Lecture 18:

Router Design

CSE 123: Computer Networks
Alex C. Snoeren



Homework 3 due FRIDAY

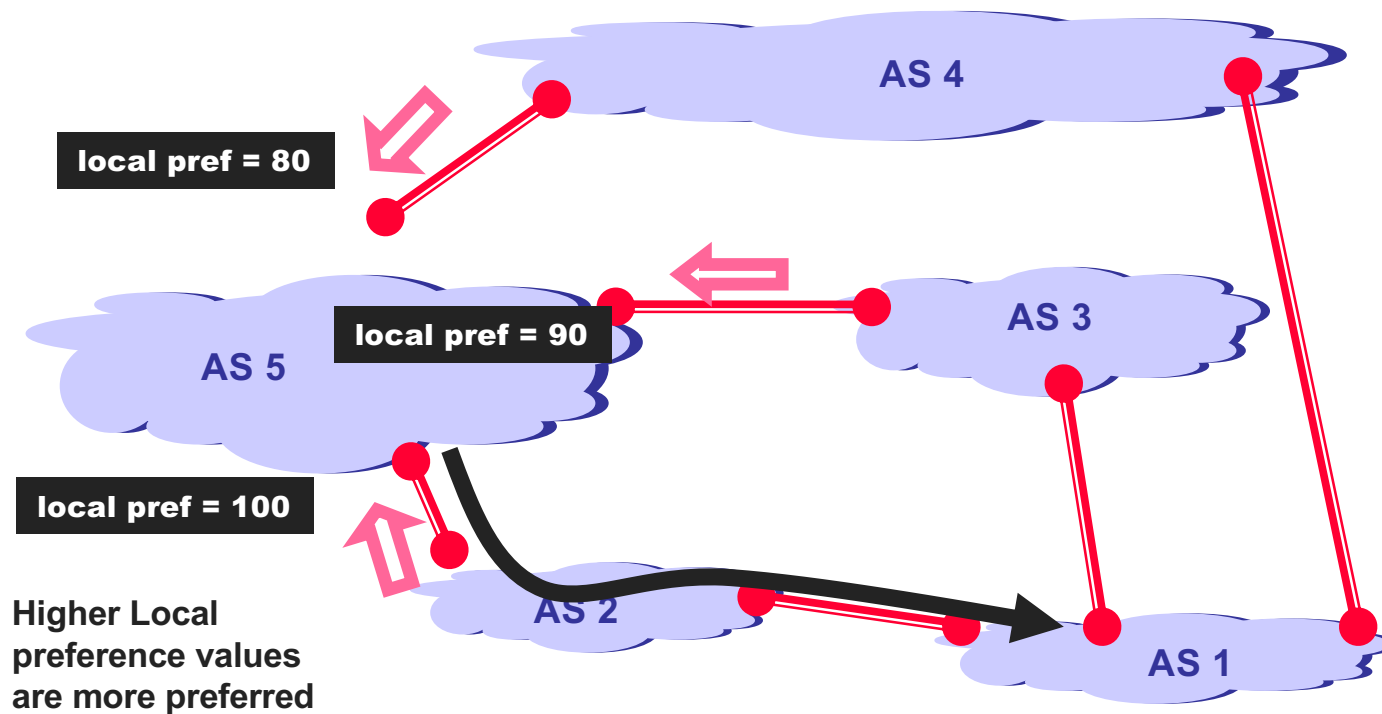


Lecture 18 Overview



- BGP examples
- AS relationships
- Router internals

Example: Local Pref

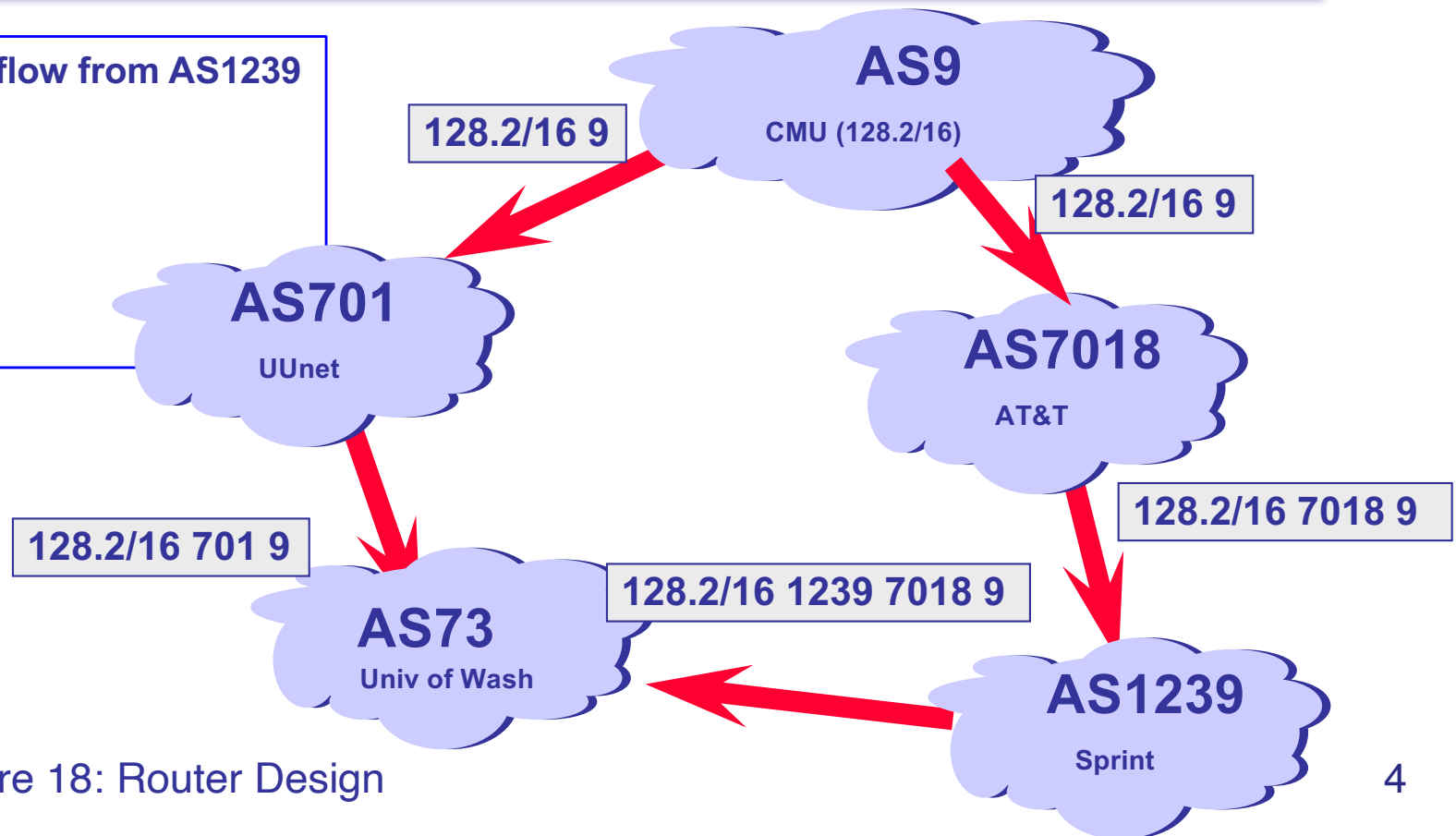




Example: Shortest AS Path

How does traffic flow from AS1239 to AS9?

- A. 1239 7018 9
- B. 1239 73 701 9
- C. 1239 9
- D. You can't tell

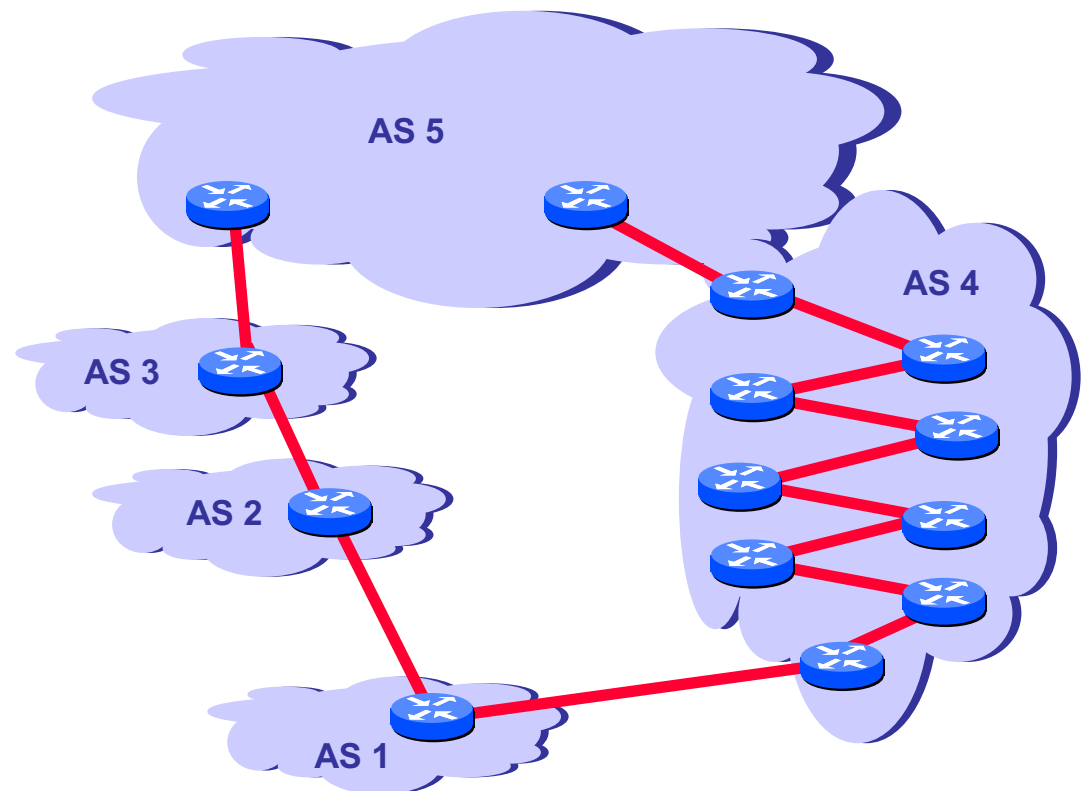


AS Paths vs. Router Paths



Assuming identical local pref, how does traffic flow from AS5 to AS1?

- A. 5 3 2 1
- B. 5 4 1
- C. 3 2 1
- D. You can't tell



Business Relationships



- Neighboring ASes have business contracts
 - ◆ How much traffic to carry
 - ◆ Which destinations to reach
 - ◆ How much money to pay
- Common business relationships
 - ◆ Customer-provider
 - » E.g., UCSD is a customer of Sprint
 - » E.g., MIT is a customer of Level3
 - ◆ Peer-peer
 - » E.g., UUNET is a peer of Sprint
 - » E.g., Harvard is a peer of Harvard Business School

A History of Settlement



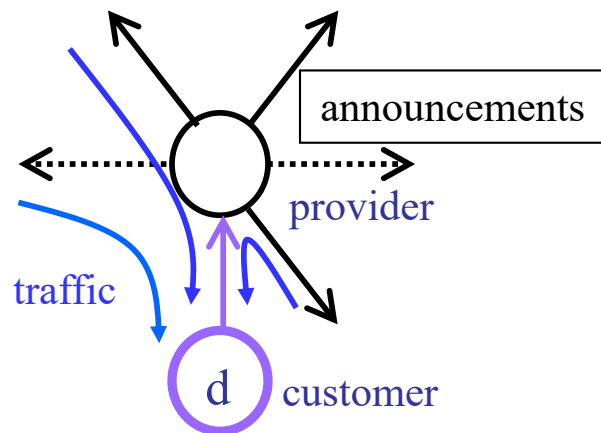
- The telephone world
 - ◆ LECs (local exchange carriers) (e.g., PacBell, NYNEX)
 - ◆ IXC (inter-exchange carriers) (e.g., Sprint, AT&T)
- LECs MUST provide IXCs access to customers
 - ◆ This is enforced by laws and regulation
- When a call goes from one phone company to another:
 - ◆ Call billed to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

Customer/Provider

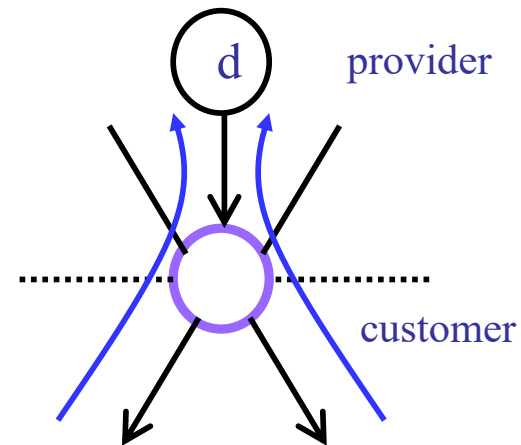


- Customer needs to be reachable from everyone
 - ♦ Provider tells all neighbors how to reach the customer
- Customer does not want to provide transit service
 - ♦ Customer does not let its providers route through it

Traffic **to** the customer



Traffic **from** the customer



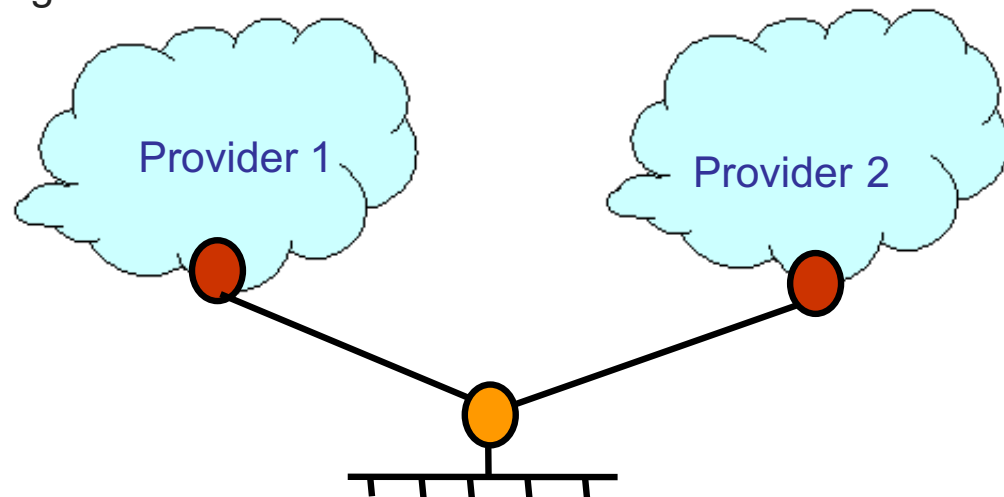
Multi-Homing



- 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

Does the customer forward route announcements from Provider 1 to Provider 2?

- A. Yes**
- B. No**
- C. It depends on the customer's policy**



Peer-to-Peer Relationship

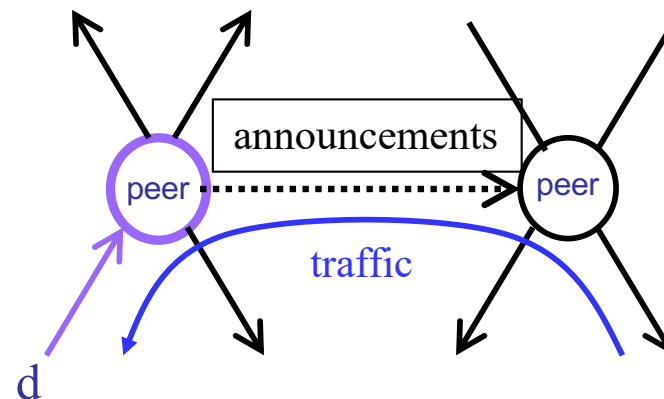


- Peers exchange traffic between customers
 - ◆ AS exports *only* customer routes to a peer
 - ◆ AS exports peer's routes *only* to its customers
 - ◆ Often the relationship is settlement-free (i.e., no \$\$\$)

Can a peer's providers send traffic to it over the peering link?

- A. Yes
- B. No
- C. It depends on the peer's policy

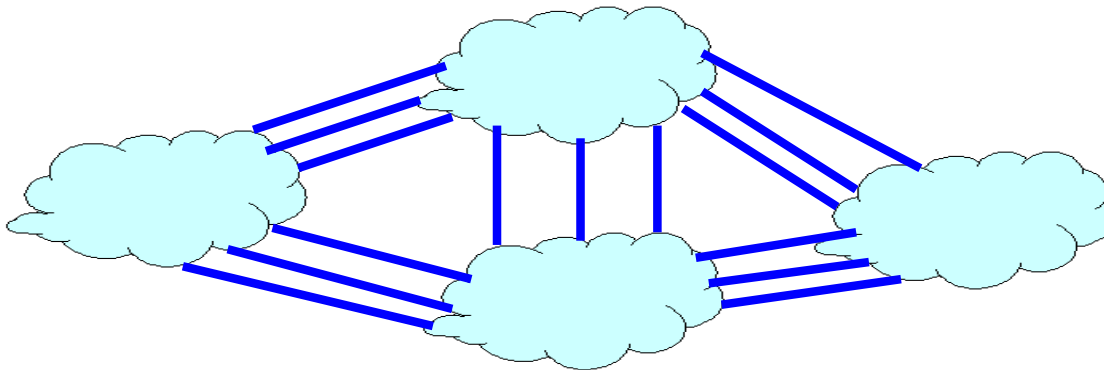
Traffic to/from the peer and its customers



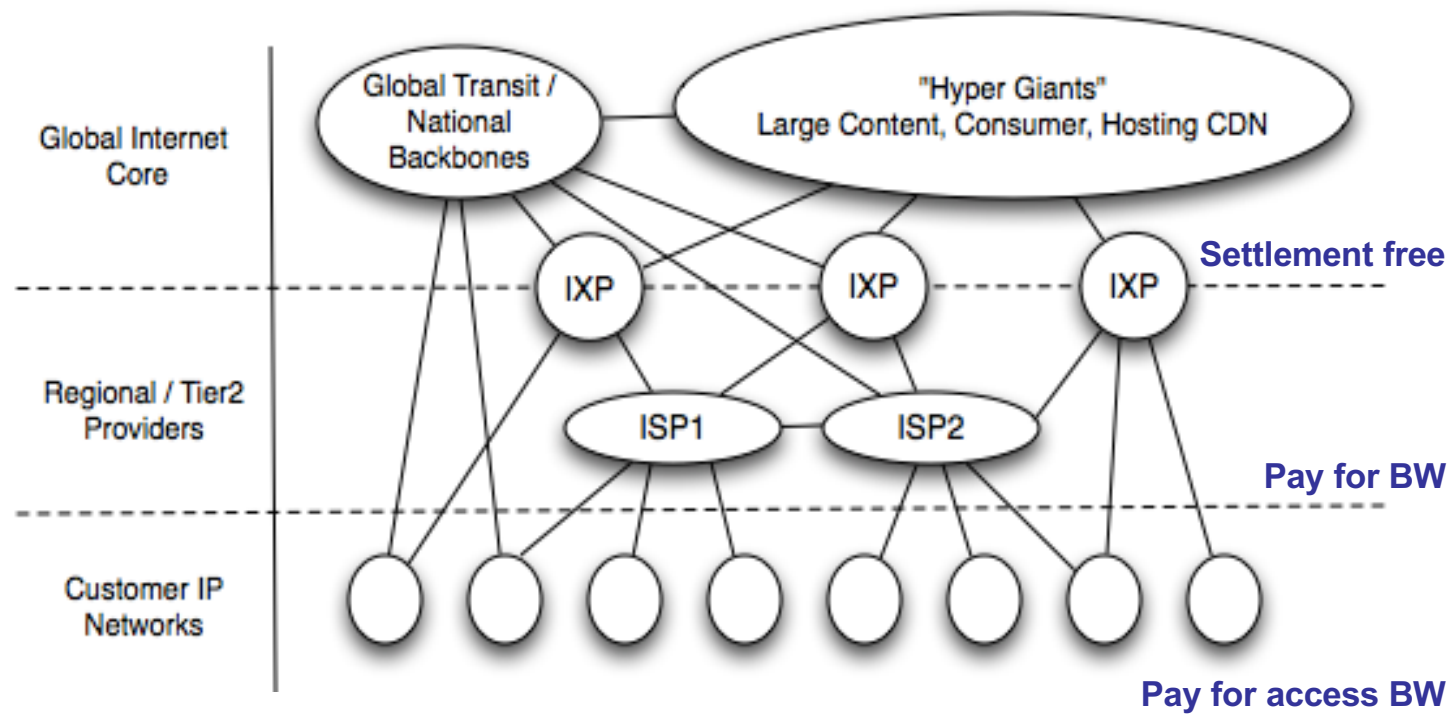


Tier-1 Providers

- 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 (Global Crossing), NTT/Verio, Century Link (formerly Qwest), Sprint, Verizon
 - ◆ Full peer-peer connections between tier-1 providers



The Internet Hierarchy



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

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



What's in a Router?

- Physical components
 - ◆ One or more **input interfaces** that receive packets
 - ◆ One or more **output interfaces** that transmit packets
 - ◆ A chassis (box + power) to hold it all
- Functions
 - ◆ **Forward** packets
 - ◆ **Drop** packets (congestion, security, QoS)
 - ◆ **Delay** packets (QoS)
 - ◆ **Transform** packets? (Encapsulation, Tunneling)

Router Functions



1. Receive incoming packet from link input interface
2. Lookup packet destination in forwarding table (destination, output port(s))
3. Validate checksum, decrement ttl, update checksum
4. Buffer packet in input queue
5. Send packet to output interface (interfaces?)
6. Buffer packet in output queue
7. Send packet to output interface link

Functional architecture

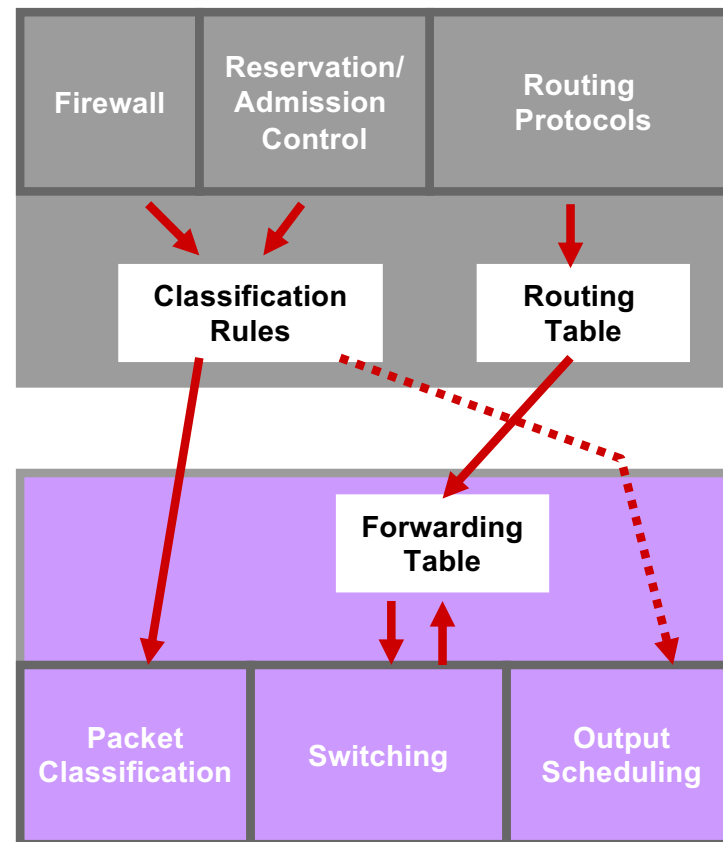


Control Plane

- Complex
- Per-control action
- May be slow

Data plane

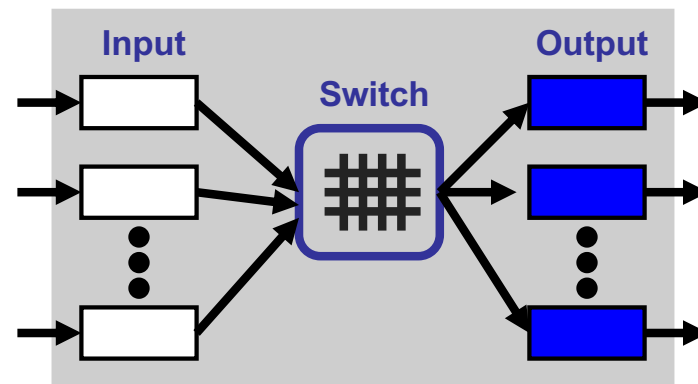
- Simple
- Per-packet
- Must be fast



Interconnect architecture



- Input & output connected via switch fabric
- Kinds of switch fabric
 - ♦ Shared Memory
 - ♦ Bus
 - ♦ Crossbar
- How to deal with transient contention?
 - ♦ Buffering

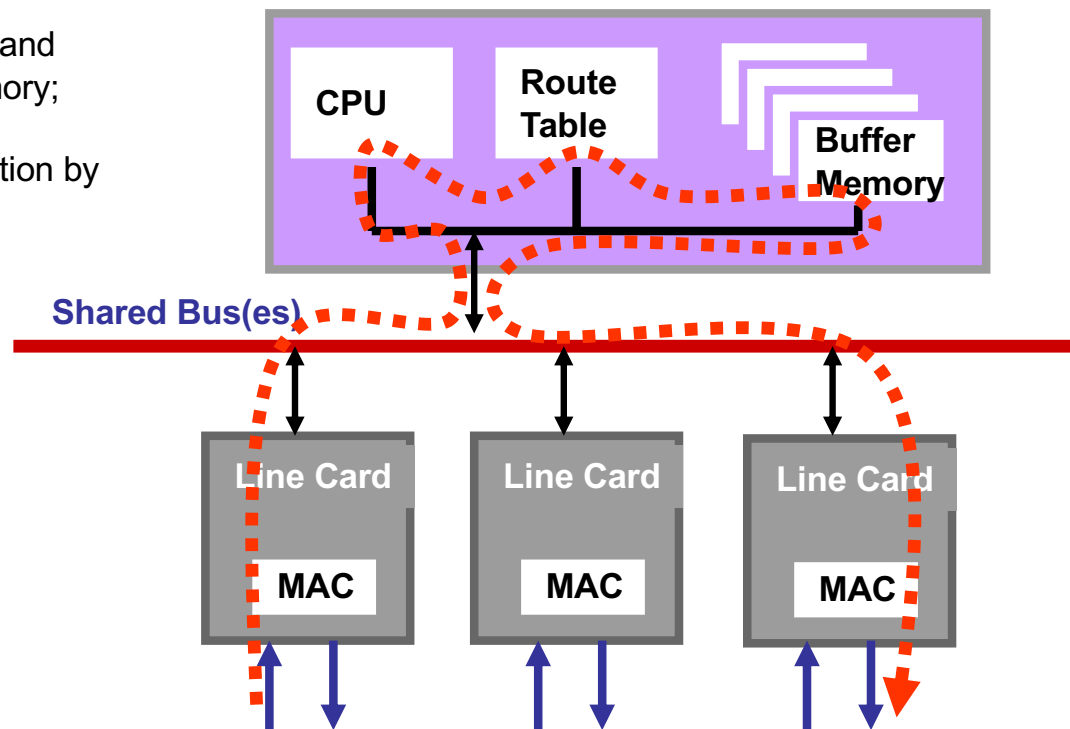




First Generation Routers

Single CPU and
shared memory;

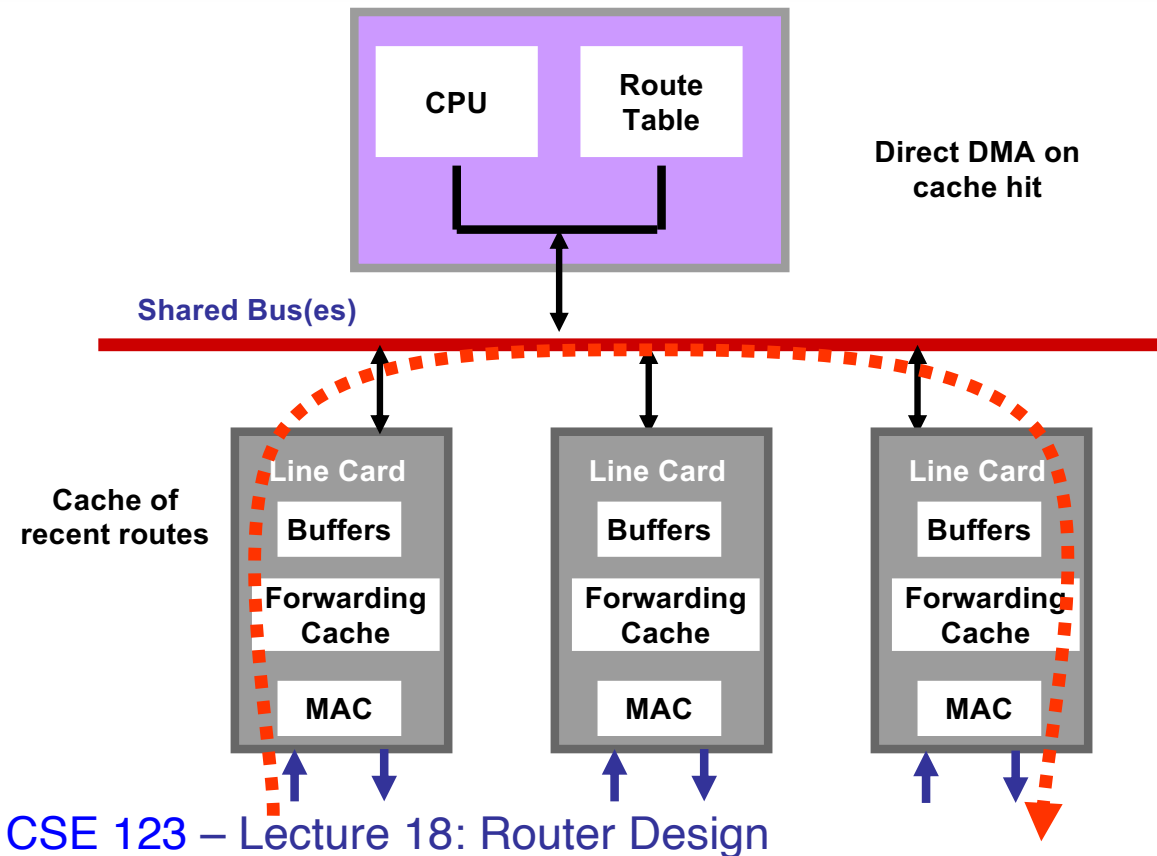
All classification by
main CPU



Which resource most
limits router capacity in
first-generation routers?

- A. Line cards
- B. Main memory
- C. CPU
- D. Shared bus

Second Generation Routers

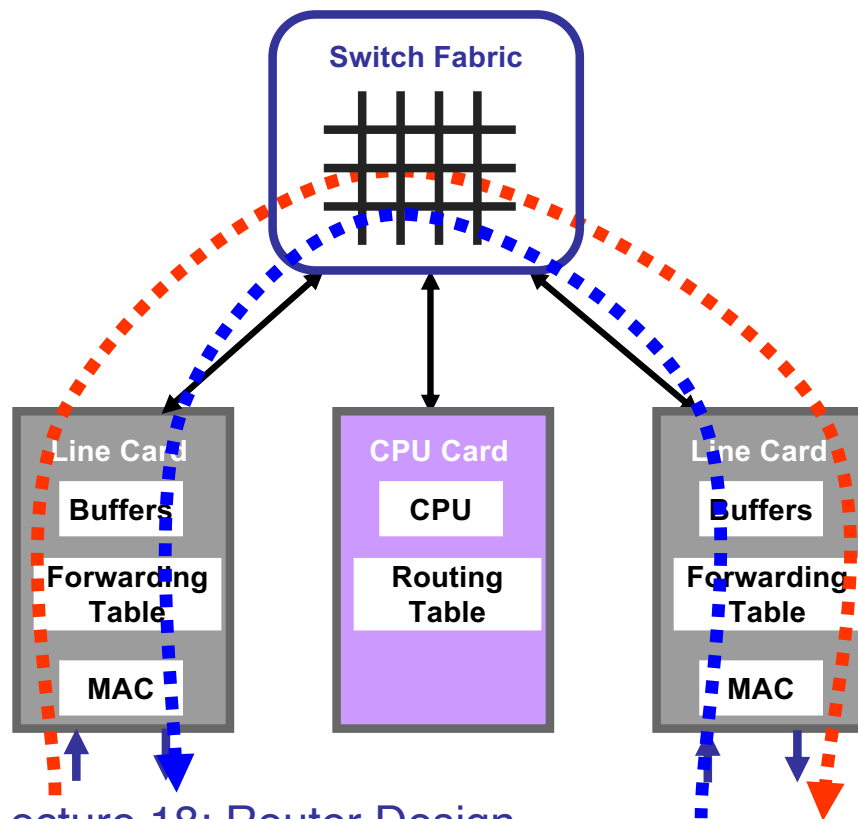


Which resource most limits router capacity in second-generation routers?

- A. Line cards
- B. Main memory
- C. CPU
- D. Shared bus



Third Generation Routers



Shared interconnect
(frequently crossbar)

Centralized scheduler

Full forwarding table in
line card

Fixed cells

For Next Time



- Read P&D 6.1
- Homework 3 due Friday
- Keep chugging on P2a