Jan-16 Lecture

Objectives

- Super quick intra-domain routing summary
- Super quick switching question
- Super quick Internet architecture
- Router architectures
- More extensive inter-domain routing summary
- Inter-domain routing attacks
- Tie to monitoring research

Reading Summary

- Notes from when I taught cs176a—pretty similar to what is in Kurose & Ross (and we'll review today)
- J. Aweya, "IP Router Architectures: An Overview," International Journal of Communication Systems, vol. 14, num. 5, June 2001.
- YouTube AS hijacking article and video
  - Mostly just used as an example with a decent animiation
- Washington Post article
  - Short article about security concerns
- IBM article
  - Also short with some background—useful to re-see the basics of BGP
- Lots of details about peering
  - I'll go over some of it, but not all of it and likely not today

Routing Summary

- From a router perspective
  - Need to maintain a forwarding table and a routing table
- Forwarding table
  - Need the simplest, fastest, most efficient way of taking a packet with a given destination and determining an outgoing interface
- Routing table (or database)
  - Collect as much information as available and choose among the different options

Routing Summary

- From a router perspective
  - Need to maintain a forwarding table and a routing table
- Forwarding table
  - Need the simplest, fastest, most efficient way of taking a packet with a given destination and determining an outgoing interface
- Routing table (or database)
  - Collect as much information as available and choose among the different options
- Different protocols offer different amounts/types of information
    - They have different properties with respect to overhead, convergence, security, etc.
    - As a result, different protocols are used in different ways in different parts of the network
    - For example: different protocols in mobile networks

Couldn't You Just Use Switches?

- Switches are simpler, faster, and cheaper
  - Switches don't do routing, they do switching (duh)
- Recall: they act on L2 headers (e.g., Ethernet switch)
- Switches already:
  - monitor which destinations are attached to which ports
  - use buffering to avoid collisions as much as possible
- So, what's the problem?

Cascade of Switches

- *Problem:* broadcast traffic gets sent everywhere
- *Problem*: switches can be connected together *and* connected back to each other to form loops
  - Protocols to solve this (and other) problems
- Solutions
  - Use Virtual LANs (VLANs) to control traffic flow
  - Avoid loops and use designated port to point “upstream”
  - Run a “spanning tree” algorithm: goal is to identify ultimate upstream connection and build tree, thereby avoiding loops

Cascade of Switches

- *Problem:* broadcast traffic gets sent everywhere
- *Problem*: switches can be connected together *and* connected back to each other to form loops
  - Protocols to solve this (and other) problems
- Solutions
  - Use Virtual LANs (VLANs) to control traffic flow
  - Avoid loops and use designated port to point "upstream"
  - Run a "spanning tree" algorithm: goal is to identify ultimate upstream connection and build tree, thereby avoiding loops
- …but now switches aren't so simple/cheap anymore
- Ultimately, it becomes a question of what is needed
  - Features, CPU, backplane, line cards/connects

Intra-Domain Routing Summary

- Intra-domain routing protocols collect various amounts of information from various locations
  - From neighbors or from routers throughout the network
- Border routers advertise non-AS routes (inbound)
  - Could just be a default route
  - Or could be sets of prefixes if default routing not applicable
- Outbound routes/prefixes are usually pretty stable
- Routers run different shortest-path algorithms based on the kinds of information available
- Result is a forwarding table
- Repeat…

A Quick Peek at ISP Connections

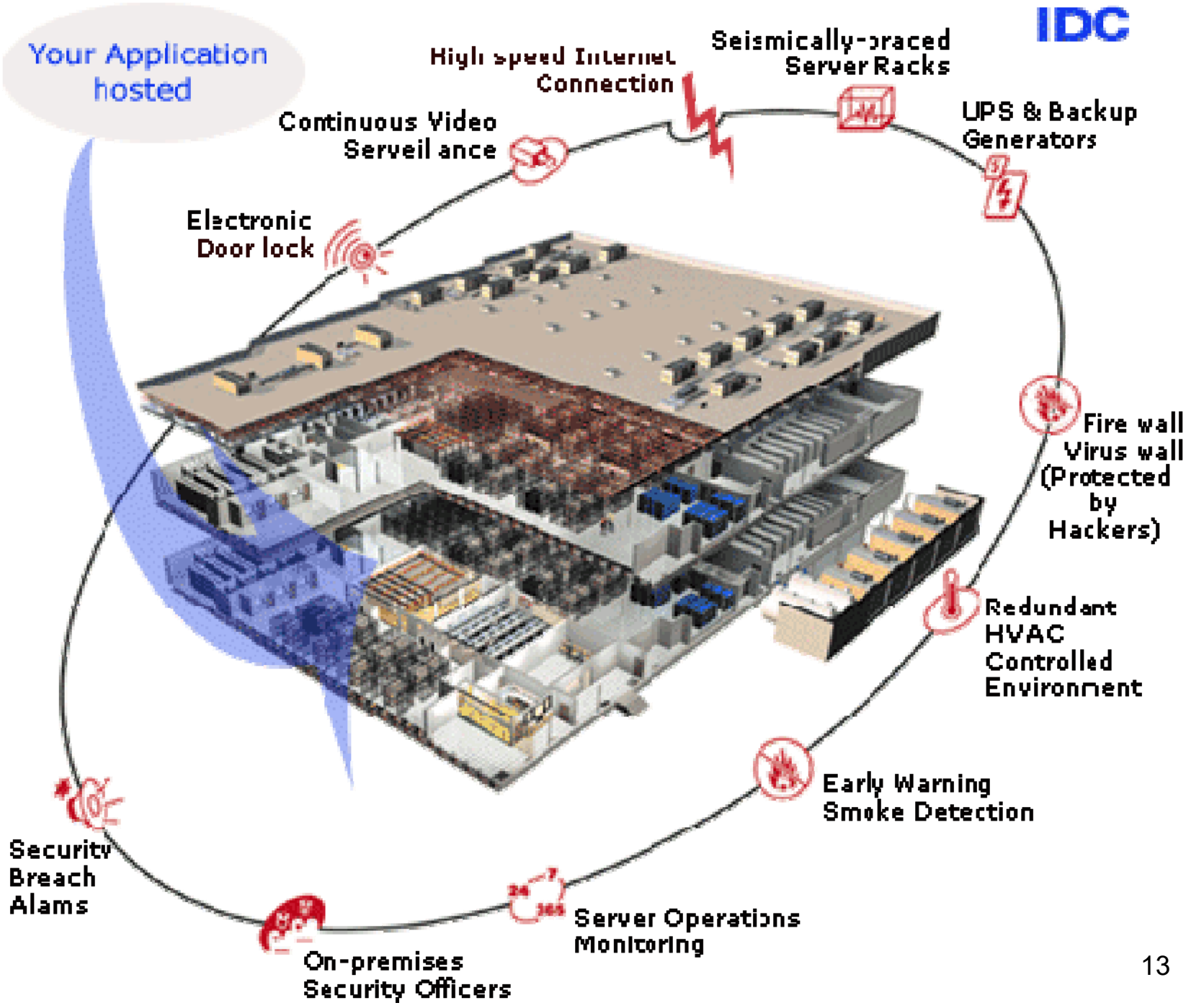
- Direct peering
  - Connect ISP A's router to ISP B's router
  - Connection is likely (these days) through direct fiber (or satellite or whatever)

A Quick Peek at ISP Connections

- Direct peering
  - Connect ISP A's router to ISP B's router
  - Connection is likely (these days) through direct fiber (or satellite or whatever)

- Peering at a Network Access Point (NAP)
  - Special purpose location that allows ISPs to (co-)locate a router in their network with routers from other ISPs


CITYNAP


IDC
Your Application hosted
High speed Internet Connection
Seismically-braced Server Racks
UPS & Backup Generators
Continuous Video Serveil ance
Electronic Door lock
Fire wall
Virus wall
(Protected by Hackers)
Redundant HVAC Controlled Environment
Early Warning Smoke Detection
Server Operations Monitoring
On-premises Security Officers
Security Breach Alams

Router Architecture

- So I decided last minute to throw in some stuff about router architectures

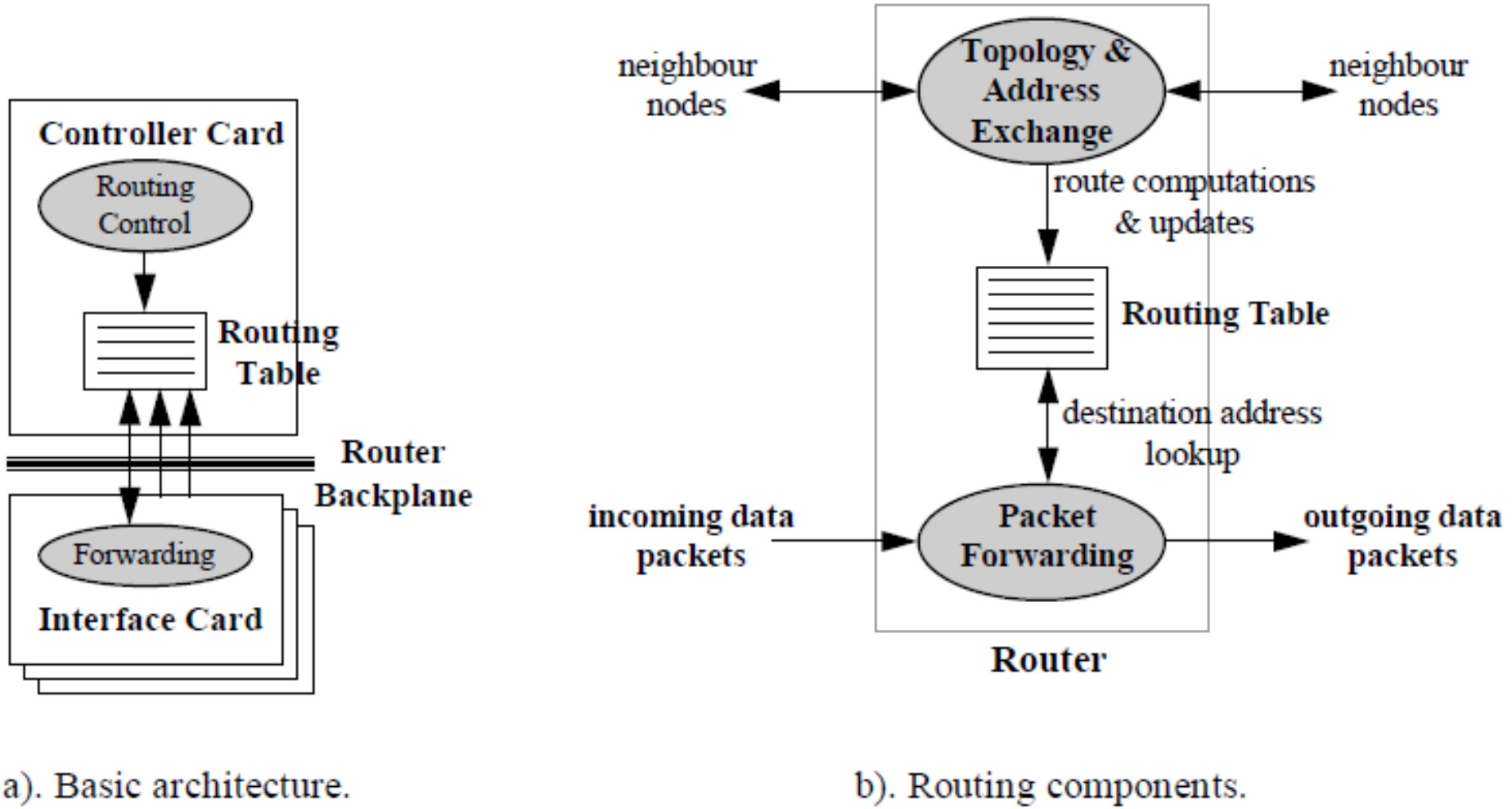
- See: J. Aweya, "IP Router Architectures: An Overview," International Journal of Communication Systems, vol. 14, num. 5, June 2001.


Juniper
M10

OC12=622 Mbps
OC48=2.4 Gbps
OC192=10 Gbps


STM-4/OC12
ATM2 MMF
OFFLINE
STATUS
TX
RX
PORT 0
PORT 1
STM-16/OC48
SONET SFP
STATUS
OFFLINE
STM-16/OC48
SONET SFP
STATUS
OFFLINE
STM-64/OC-192
SONET SMF-SR-2
STATUS
TX
RX
ONLINE/
OFFLINE
STM-64/OC-192
SONET SMF-SR-2
STATUS
TX
RX
ONLINE/
OFFLINE
STM-4/OC-12
SONET SMF-IR
ONLINE/OFFLINE
STATUS
TX
RX
PORT 2
PORT 3
STM/OC-12
SONET SMF-IR
ONLINE/OFFLINE
STATUS
TX
RX
PORT 1
PORT 2
STM/OC-12
SONET SMF-IR
ONLINE/OFFLINE
STATUS
TX
RX
PORT 1
PORT 2
STM-64/OC-192
SONET SMF-SR-2
STATUS
TX
STM-64/OC-192
SONET SMF-SR-2
STATUS
TX

Basic Router Architecture



a). Basic architecture.

b). Routing components.

Figure 1. Generic architecture of a router.

Use of a Common Bus

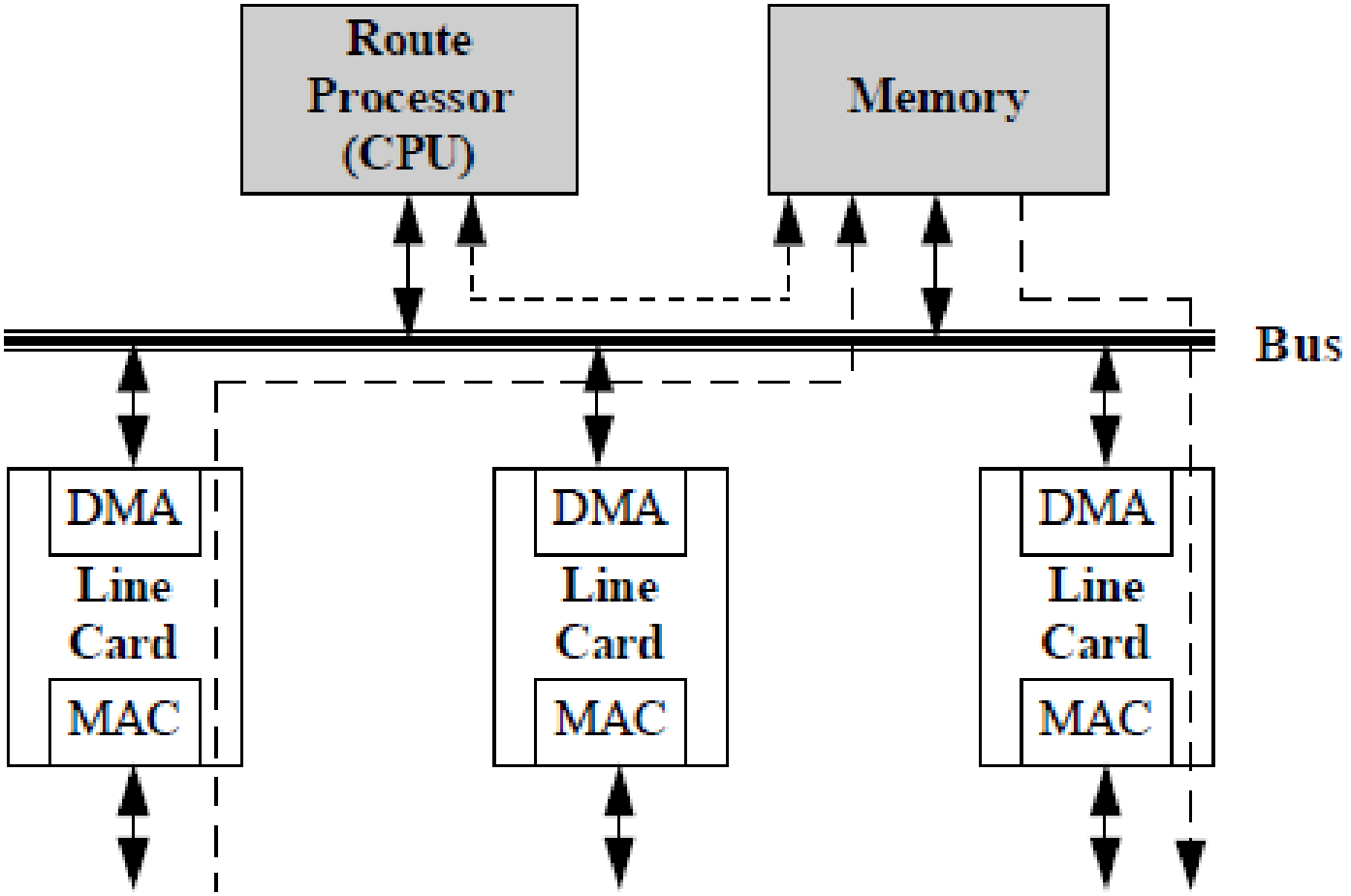


Figure 2. Traditional bus-based router architecture.

Use of a Route Cache

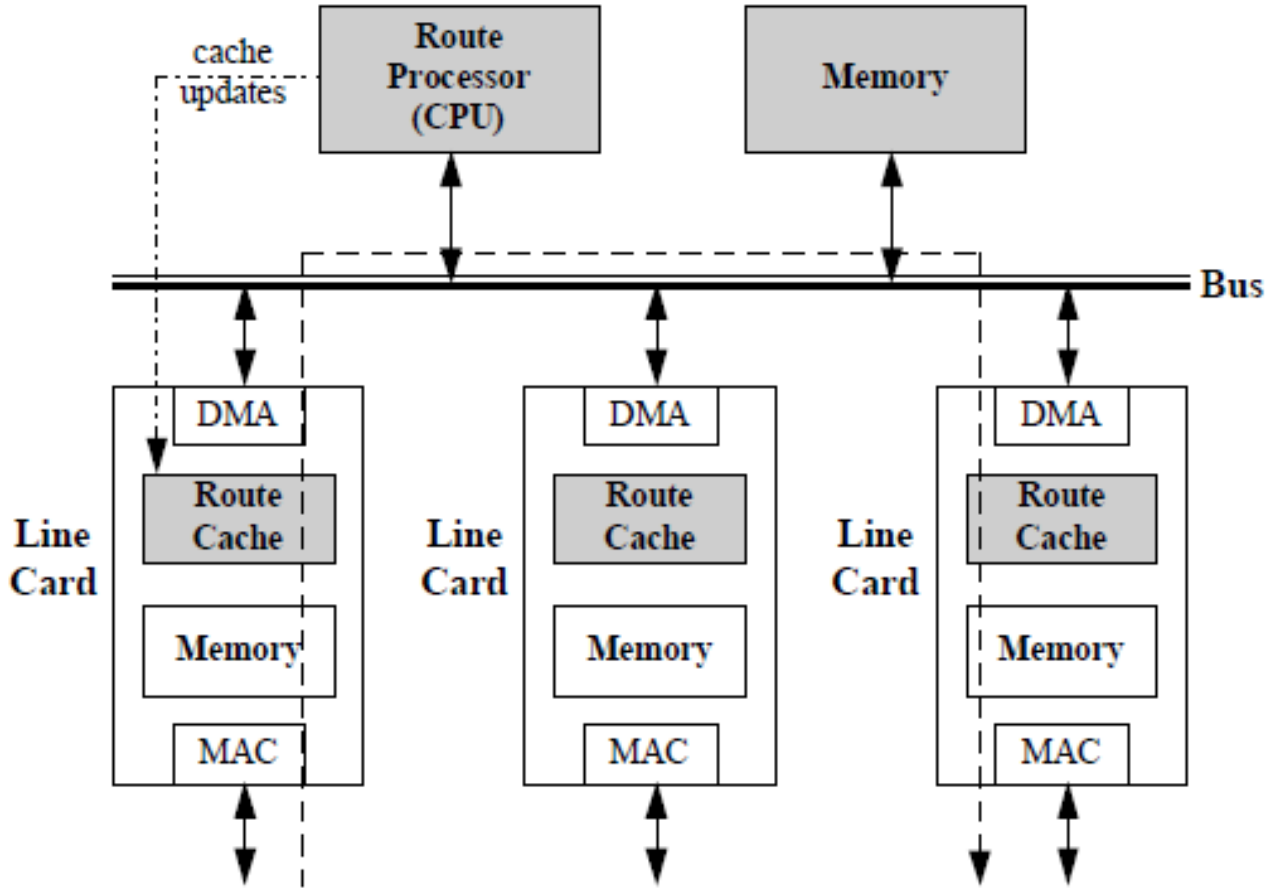
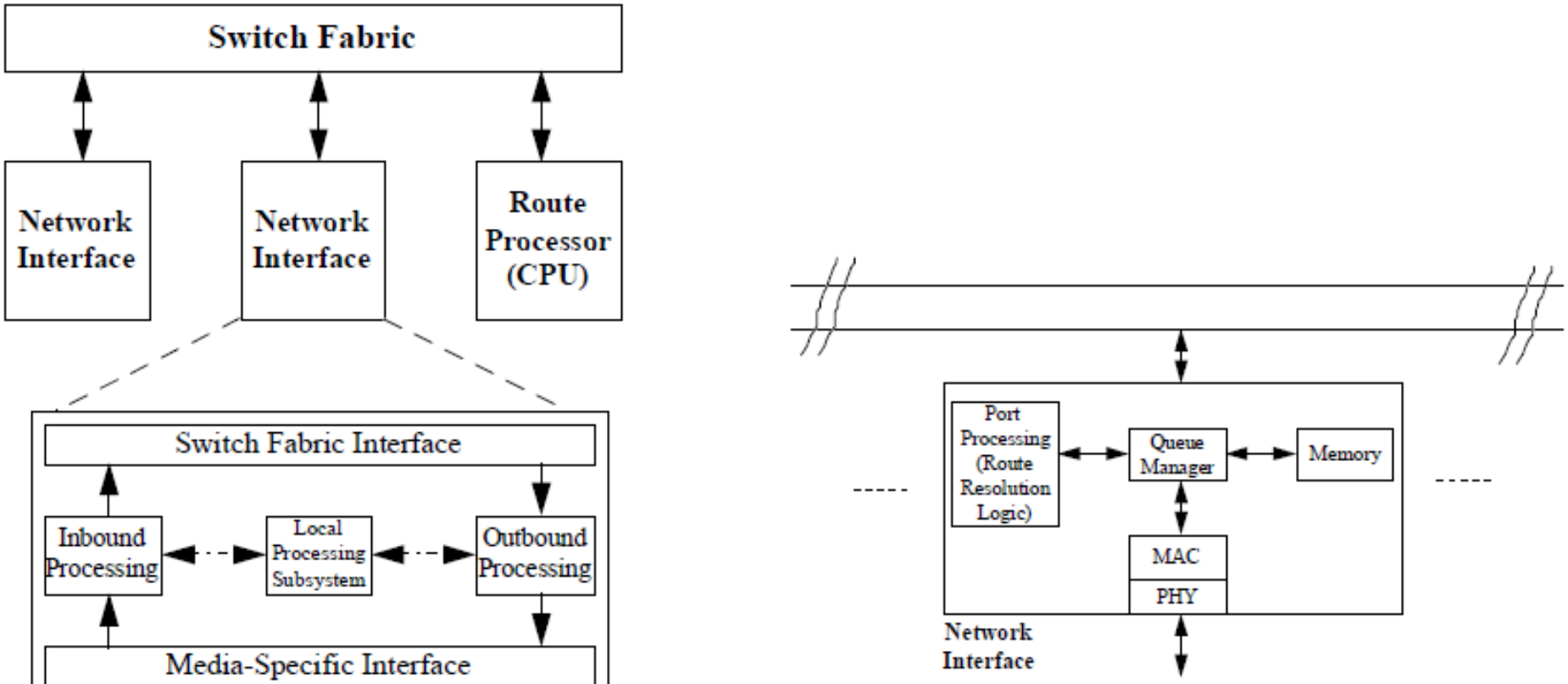


Figure 3. Reducing the number of bus copies using a route cache in the network interface.

Switching and Routing Together



a). Functional diagram [34][35][36].

b). Generic architecture.

Figure 6. A generic switch-based distributed router architecture.

Fast Path v. Slow Path

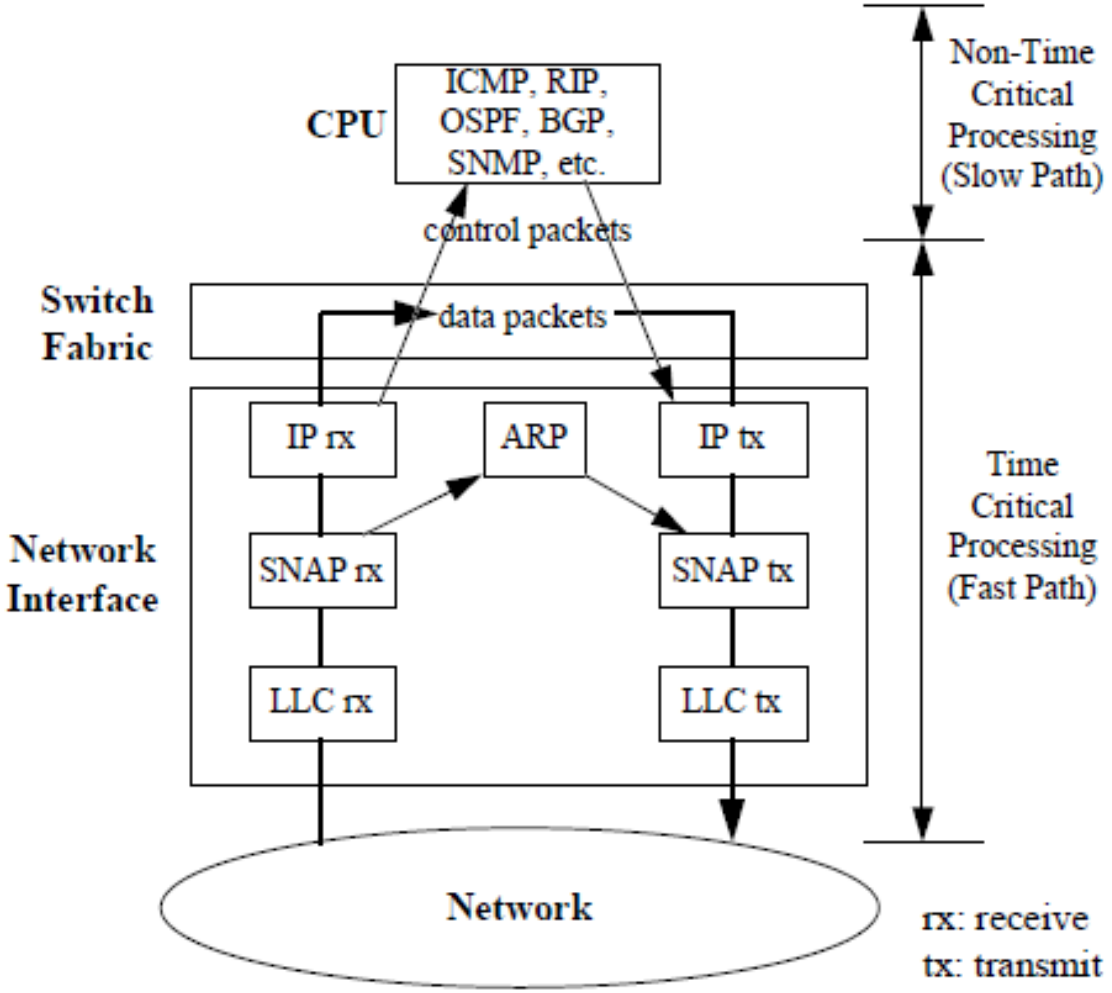


Figure 7. Example IEEE 802 protocol entities in an IP router [Adapted from 34].

More Examples of Slow Path Functions

Typical Router Slow Path Functions	
Packet-by-Packet Operations	**Background Tasks**
- Fragmentation and reassembly - Source routing option - Route recording option - Timestamp option - ICMP message generation	- Routing protocols (RIP, OSPF, BGP, etc.) - Network management (SNMP) - Router configuration (BOOTP, DHCP, etc.)

Figure 8. IP router slow-path functions.

Slow/Fast Path Functions in the Architecture

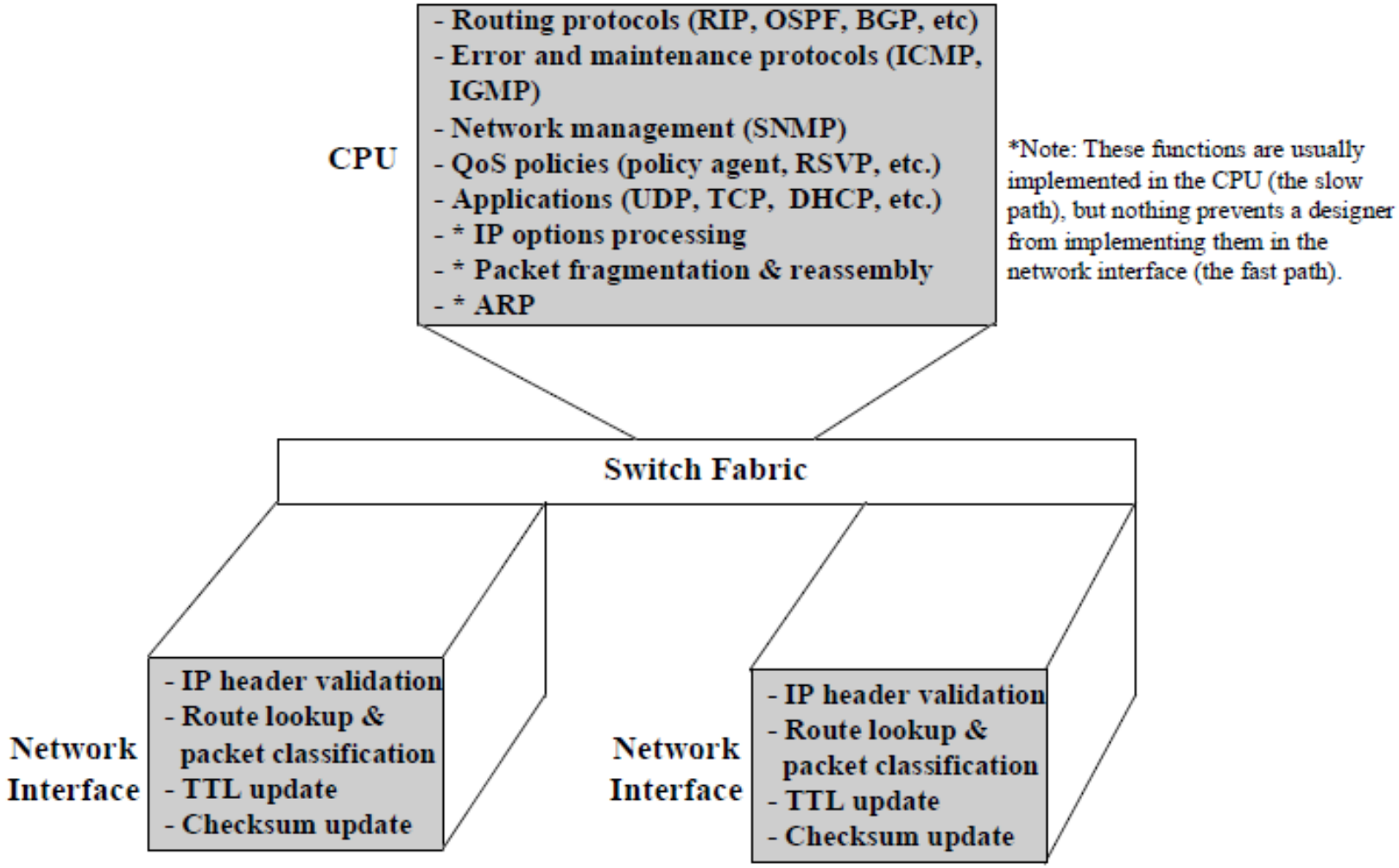


Figure 9. An example functional partitioning in the distributed router architecture.

Even More Functionality Added

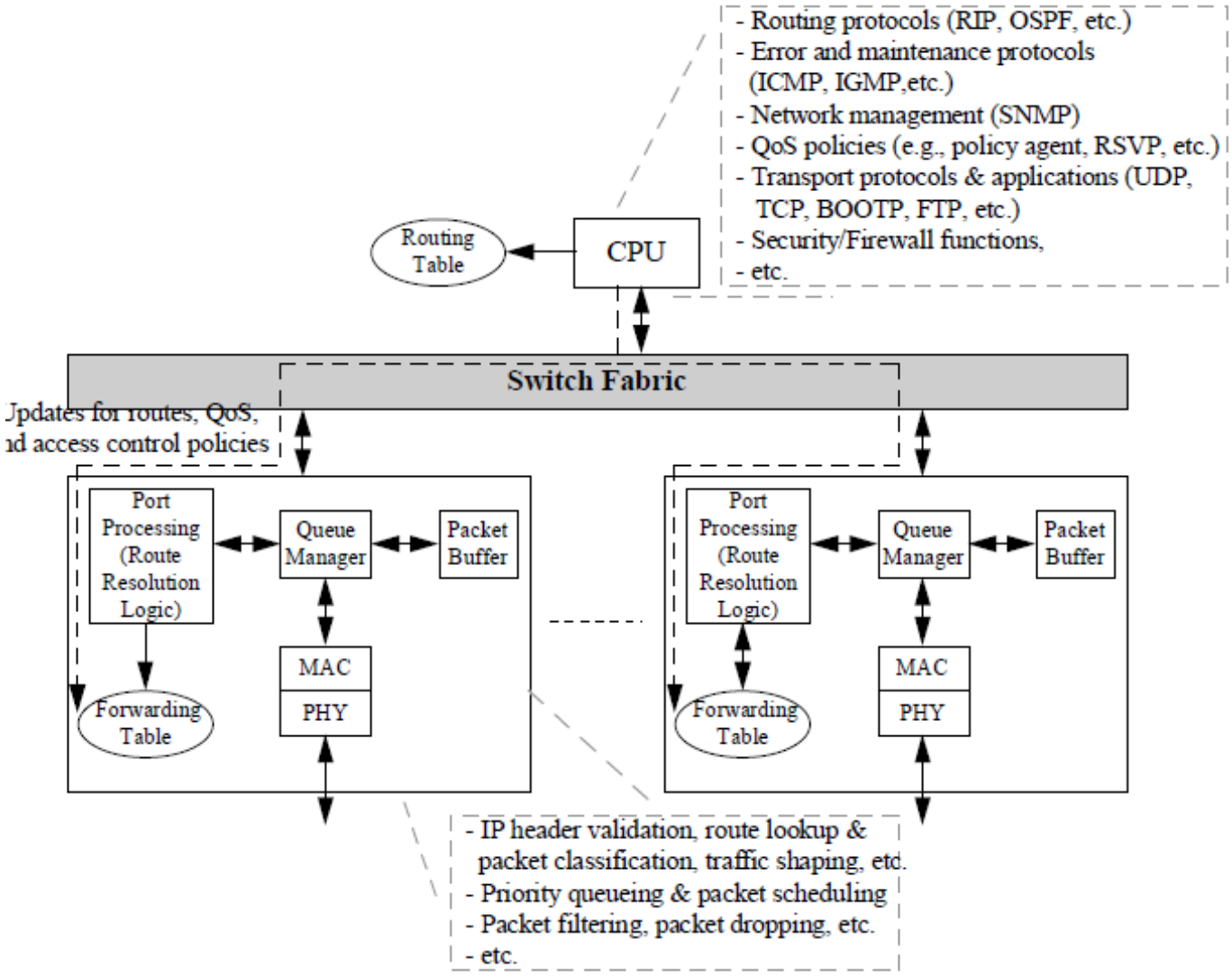


Figure 10. A high level functional diagram of a distributed router architecture.

Pretty Close to Modern Functionality

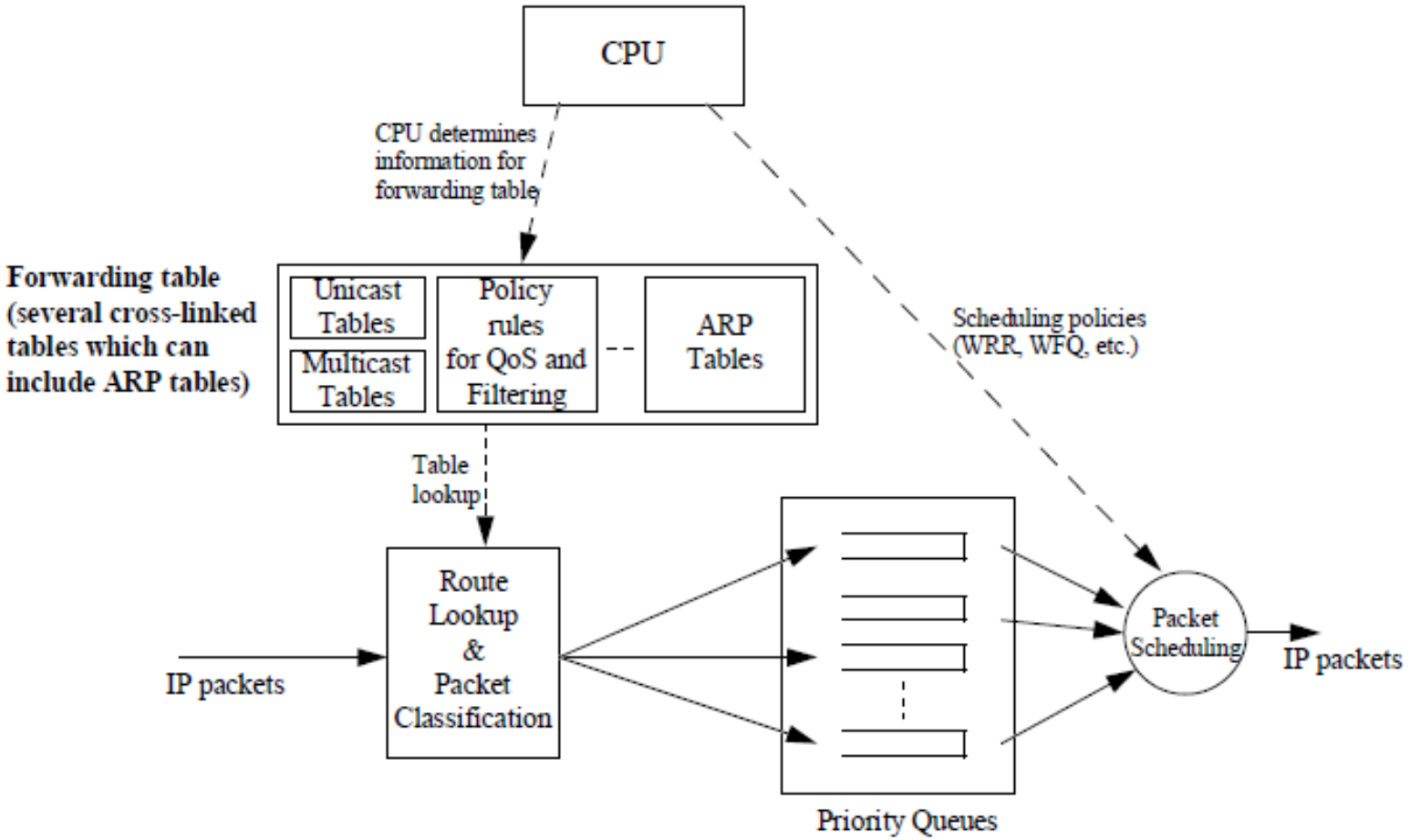


Figure 11. Forwarding database consisting of several cross-linked tables.

Hierarchical Routing

- Now, we've talked about (1) the basics of switching and routing, (2) what routers are, (3) how they connect to each other, (4) how networks connect to each other, now…

Hierarchical Routing

- Now, we've talked about (1) the basics of switching and routing, (2) what routers are, (3) how they connect to each other, (4) how networks connect to each other, now…

- *Problem*: Internet is too big to just have a single domain and a single routing protocol
  - If there were just a flat topology, every router would have to hear from every other router in order to populate its routing table and build a forwarding table

Hierarchical Routing

- Ways of dealing with the massive scale of the Internet
  - *Default routing*: networks (and hosts) on the edges of the Internet don't need full routing tables, they just need a route into the core

Hierarchical Routing

- Ways of dealing with the massive scale of the Internet
  - *Default routing*: networks (and hosts) on the edges of the Internet don't need full routing tables, they just need a route into the core
  - *Address aggregation*: when packets are far from their destinations, they only need to know the general direction to go (still, as specified by a next hop)
    - By aggregating addresses, the number of entries in the routing and forwarding tables can be greatly reduced.

Hierarchical Routing

- Ways of dealing with the massive scale of the Internet
  - *Default routing*: networks (and hosts) on the edges of the Internet don't need full routing tables, they just need a route into the core
  - *Address aggregation*: when packets are far from their destinations, they only need to know the general direction to go (still, as specified by a next hop)
    - By aggregating addresses, the number of entries in the routing and forwarding tables can be greatly reduced.
  - *Border/Gateway routers*: like default routing, a border router can act as a proxy for everything that isn't local (*e.g.*, all traffic *not* for 128.111/16 goes to the border router)
    - The benefit of border/gateway routers also works in the reverse direction: the rest of the Internet does not need to know about routing within networks at the edges—core routers just get one prefix for a stub network (like UCSB).

Hierarchical Routing

- Routing grows much more complicated the further away from the edges
- Examples
  - When (transit) domains have multiple egress points
  - When a domain has multiple equal provider connections
  - When money is involved
- Most difficult for transit providers
  - E.g., CENIC

Hierarchical Routing

- Border/Gateway routers serve another purpose as well
  - Provide a place to wall off an administrative domain from the rest of the Internet
  - Border routers provide a “demarcation point” (or border) between what is considered “inside” the network and what is considered “outside”
  - A place to implement “routing policy”

Hierarchical Routing Analogy

- One of the more significant challenges in running a border router is dealing with the different routing protocols that communicate with the border router

- Analogy: an information station at a “T” intersection
  - It receives updates on traffic conditions from all of the different roads
  - Yet travelers arrive and want to know how to get to a particular destination
  - And lots of different languages are being spoken
  - The information station translates everything and stores it away, probably processing everything it hears to come up with a best road to take to get to each of the different destinations.
  - Ultimately, the advice is either “take Road X” or “take Road Y”

Routing Protocols

- At the end of the day, how information about possible paths and the conditions on each are defined as part of the routing protocol

- The most commonly used protocols are must simpler than you might imagine
  - 99% of the time, there isn't a lot going on
  - The routing *algorithms* are very simple: most don't consider path conditions, the metric is simply hop count

- From CS theory: because the scale is so large, small increments in complexity and state require significant resources as the network grows larger

Border Gateway Protocol

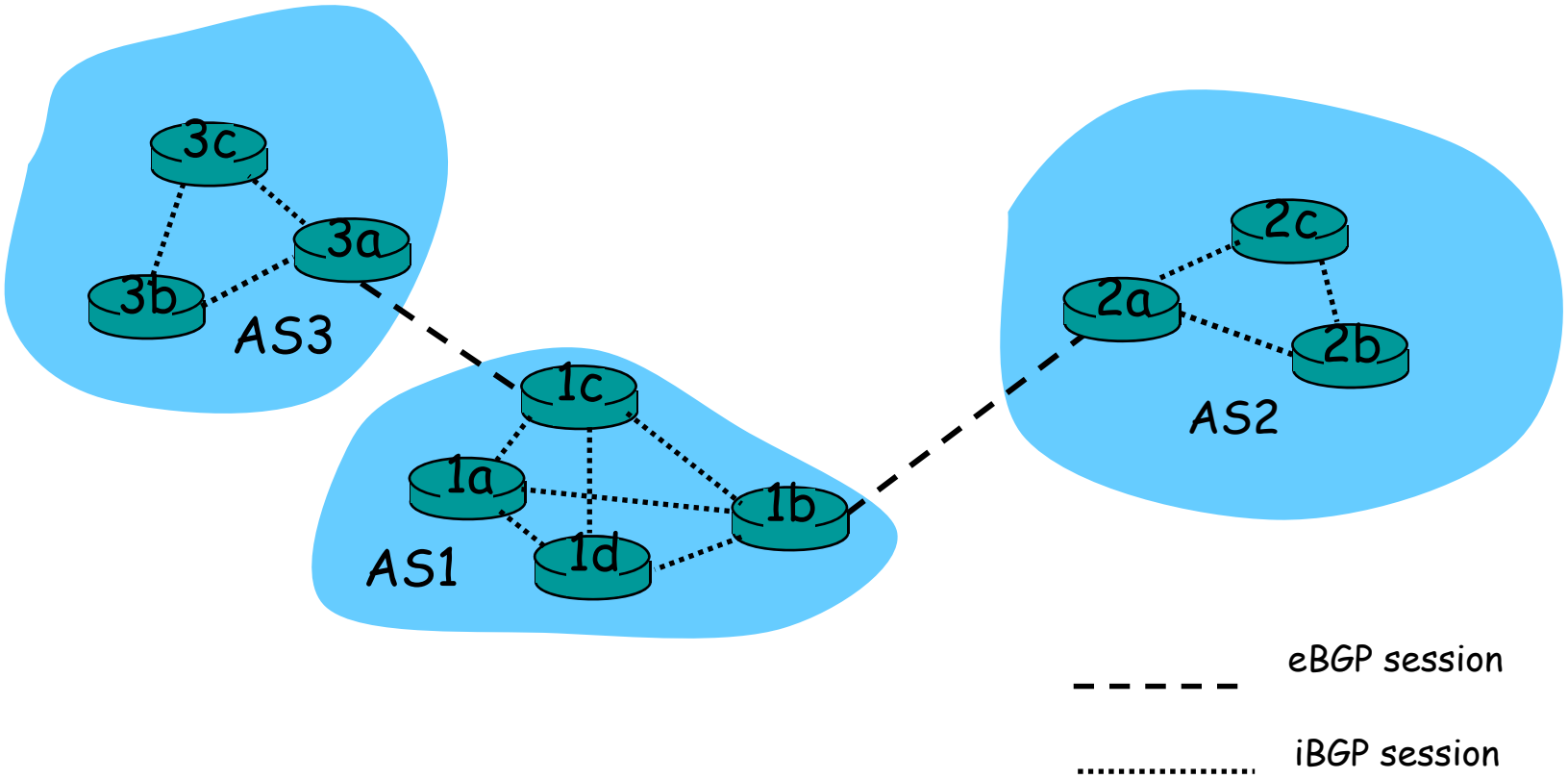
- *THE* inter-domain routing protocol
  - Designed to provide scale
  - Designed to provide security
  - Designed to provide policy mechanisms
- High level BGP routing steps are about the same as for an intra-domain routing protocol:
  - Receive routing updates from neighbors
  - Propagate that information to other neighbors
  - Compute best routes for a particular prefix
- The issues are:
  - What information is exchanged
  - How the route is determined
  - What exactly is determined

BGP Terminology

- *Stub networks*: a single border connection
- *Multi-Connected*: no third-party network reachability
- *Transit*: inter-connected ASes & provide traffic through the AS (not sourced or destined within the AS)

- *External BGP (eBGP):* receiving routes from routers outside the AS
- *Internal BGP (iBGP)*: receiving external routes from within the AS
  - Structured as a mesh

BGP Terminology

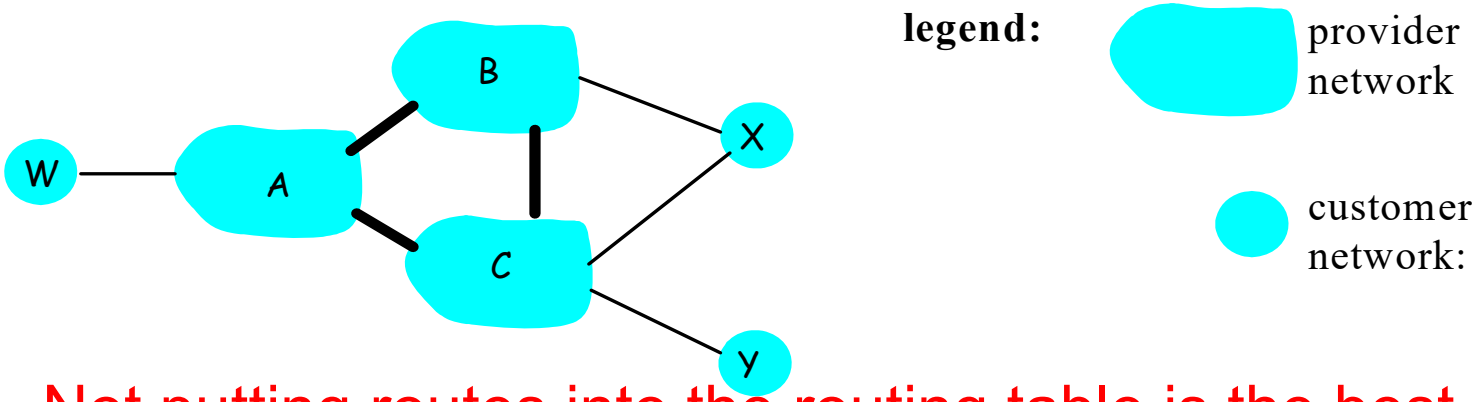


BGP Route Information

- *Origin*: AS that originated that update
- *AS-Path*: set of ASes to reach the destination prefix
- *Next-Hop*: IP address of the router that begins the AS-Path
- *Multi-Exit Description (MED):* tie-breaker to determine what entry/exit point in an AS should be
- *Local Pref*: Used in iBGP to determine best exit point (if multiple exit)
- *Communities*: set of BGP speakers that are trusted

BGP Policy Information

- Two examples where received routes are often not put into the routing table:
  1. Multi-homed network that does not want to act as a transit network
  2. Transit networks that do not want to do the majority of the work for its neighbors/peers/competitors



- Not putting routes into the routing table is the best policy enforcement mechanism
  - Preferring one route over another to avoid a particular destination is not guaranteed to work
  - Remember! Forwarding is always hop-by-hop

BGP Policy Information

- Another input that affects which routing paths are used is based on economics
  - This decision is usually made by a network operator
  - Use MED (multi-exit descriptor) attribute or LocalPref (used in iBGP to identify preferred exit point)
- Most common current billing model is “95/5”: 95% percentile of traffic, sampled at 5 minute intervals
  - https://www.semaphore.com/95th-percentile-bandwidth-metering-explained-and-analyzed/
- Past billing models
  - Based on access capacity (Ex: buy a T1 or OC3)
  - Average usage: skewed higher based on bursty nature of Internet traffic
- Multi-homed networks or transit providers monitor usage and adjust preferences accordingly

Bandwidth Pricing

Internet Transit Pricing (1998-2015)			
Source: http://DrPeering.net			
Year	Internet Transit Price		% decline
1998	$1,200.00	per Mbps	
1999	$800.00	per Mbps	33%
2000	$675.00	per Mbps	16%
2001	$400.00	per Mbps	41%
2002	$200.00	per Mbps	50%
2003	$120.00	per Mbps	40%
2004	$90.00	per Mbps	25%
2005	$75.00	per Mbps	17%
2006	$50.00	per Mbps	33%
2007	$25.00	per Mbps	50%
2008	$12.00	per Mbps	52%
2009	$9.00	per Mbps	25%
2010	$5.00	per Mbps	44%
2011	$3.25	per Mbps	35%
2012	$2.34	per Mbps	28%
2013	$1.57	per Mbps	33%
2014	$0.94	per Mbps	40%
2015	$0.63	per Mbps	33%

BGP Protocol Details

- Uses TCP, Port 179
- *Open*: used to exchange configuration and parameters for the session
- *Update*: distribute routing information including add, update, withdraw
- *KeepAlive*: sent at 1/3 of the "hold time" to keep the TCP session alive
  - Usually 90s or 120s
- *Notification*: Error detected, terminate TCP session
- *Route Refresh*: complete re-transmission of routing information

BGP Attacks and Weaknesses

- Interject a session reset (DoS attachk)
  - Through TCP or BGP
- Malicious advertisements
  - Create route flaps: send updates frequently
  - Inject black holes: send updates for routes that don't exist
- Human mistakes
  - Allow unusual routes to be advertised (misconfiguration)
  - Allow unusual routes to be accepted (too much trust)

YouTube Hickjacking by Pakistan Telecom, 2008

- **Before, during and after Sunday, 24 February 2008:** AS36561 (YouTube) announces 208.65.152.0/22
- **Sunday, 24 February 2008, 18:47 (UTC):** AS17557 (Pakistan Telecom) starts announcing 208.65.153.0/24.
- **Sunday, 24 February 2008, 20:07 (UTC):** AS36561 (YouTube) starts announcing 208.65.153.0/24.
- **Sunday, 24 February 2008, 20:18 (UTC):** AS36561 (YouTube) starts announcing 208.65.153.128/25 and 208.65.153.0/25.
- **Sunday, 24 February 2008, 21:01 (UTC):** AS3491 (PCCW Global) withdraws all prefixes originated by AS17557 (Pakistan Telecom), thus stopping the hijack of 208.65.153.0/24.

http://www.youtube.com/watch?v=IzLPKuAOe50#t=62

Washington Post Article

- "In 2010 a Chinese ISP momentarily hijacked the Internet. Due to a misconfiguration, some traffic that should have gone to Dell, CNN, Starbucks and Apple was sent to China instead. The incident lasted for only a few minutes and the responsible party claimed it was an accident. But it highlights a dangerous security weakness in one of the Internet's fundamental protocols."

- "Research released this week has revealed two more cases in which misconfigurations re-routed traffic far from their intended destination. For example, in one of the attacks, traffic traveling from Mexico to the United States took a circuitous and illogical route to Belarus."

Washington Post Article

- “According to Renesys, on July 31, 2013, [Opin Kerfi] it began announcing origination routes for 597 IP networks, despite the fact that it normally only originates three IP networks and has no downstream autonomous system customers. The faulty routes appear to have exclusively come through one of Opin Kerfi's ISPs, Síminn.”

- “Renesys believes this kind of attack is a serious threat to Internet security, but may have a very limited shelf life. "This is not a very subtle attack -- you can't carry it out without publishing your false routes all over the planet," said Cowie. ‘If everyone would take care to watch how their networks are being advertised around the world it would disappear overnight.’”

BGP Attacks and Weaknesses

- Interject a session reset
  - Through TCP or BGP
- Malicious advertisements
  - Create route flaps: send updates frequently
  - Inject black holes: send updates for routes that don't exist
- Human mistakes
  - Allow unusual routes to be advertised (misconfiguration)
  - Allow unusual routes to be accepted (too much trust)

http://www.cs.ucsb.edu/~almeroth/classes/W19.176B/papers/ibm-17.pdf

Reading Summary for Next Time

- Paxson paper
  - 1997: defines route stability metrics
- Global Internet paper
  - 2010: revisited some of the same metrics
- Will probably then talk about peering…