Cisco live!

What You Make Possible

TOMORROW starts here.

CISCO

Routing Operations in Cisco IOS Routers

BRKARC-2350

TOMORROW starts here.

Routing Operations in Cisco IOS Routers

Agenda

- Router Components
- Moving Packets
- CEF, CPU and Memory
- Outbound Load Sharing
- Routing Convergence Improvements

Router Components

Data and Control Planes

Router Components

Data and Control Planes

- Control Plane ← **Brains**
  - – Control Traffic
    - Routing Updates (BGP, OSPF etc)
    - SSH
    - SNMP
- Data Plane ← **Brawn**
  - – Through traffic (transit not for us!)
    - Routed packets (known via BGP etc)

Router Components

Software Based Routers

Router Components

Software Based Routers

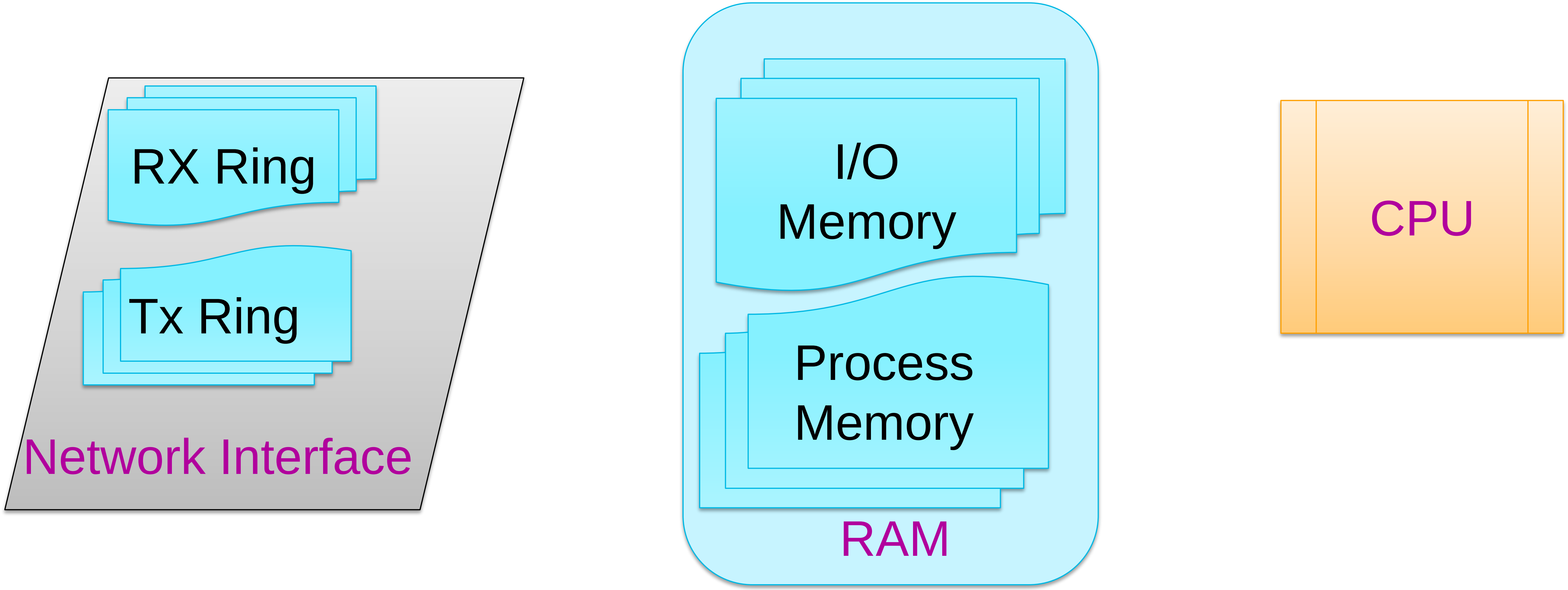
- Software Based
  - – **Shared** control and data plane
  - – General Purpose CPU (slow, but smart – feature rich)
  - – CPU responsible for all operations (features and forwarding)

2800/2900/3900/7200 Series Routers are software based



Router Components

Software Based Routers



Router Components

Hardware Based Routers

Router Components

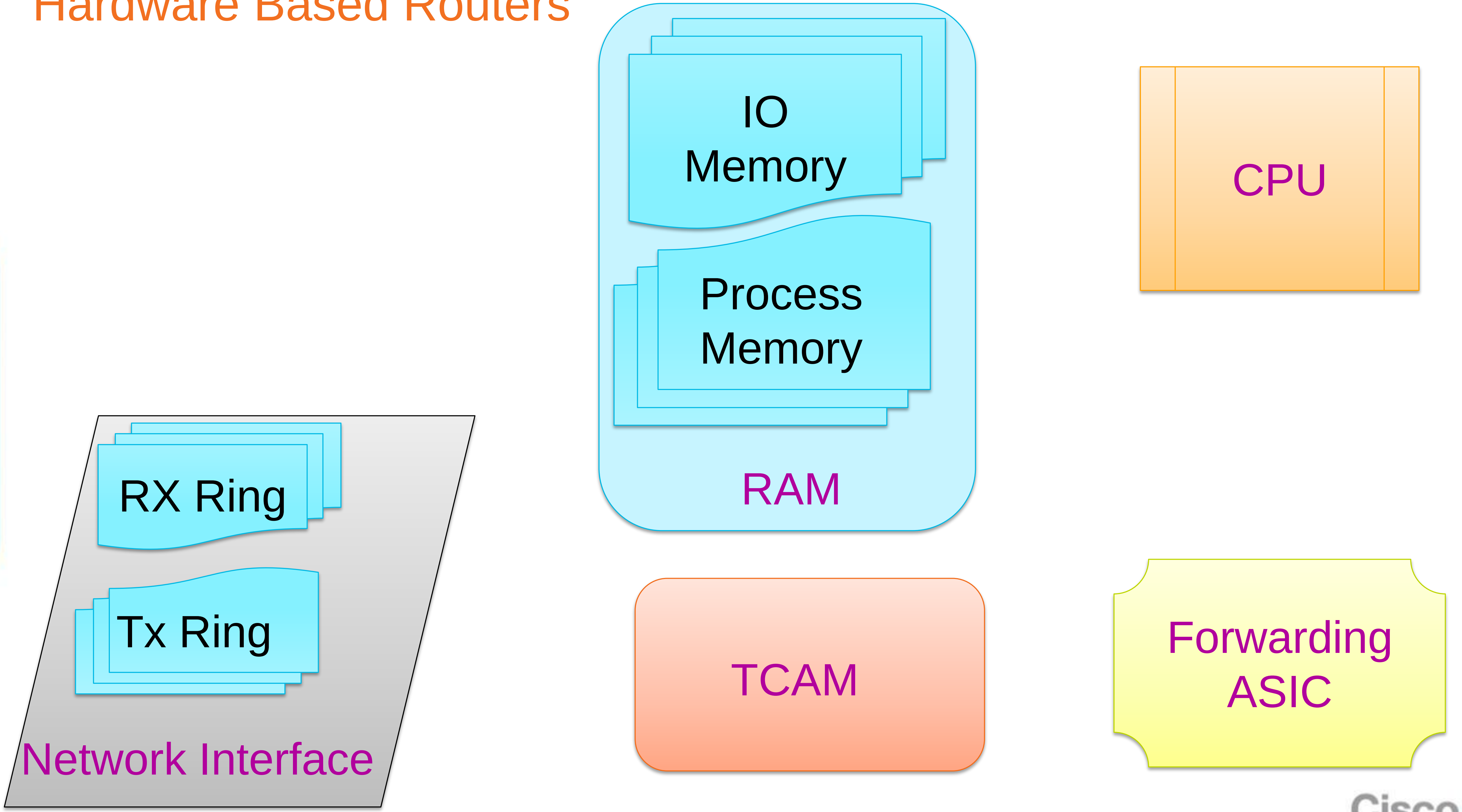
Hardware Based Routers

- Hardware based
  - **Separated** control and data plane
  - CPU + ASIC (Application Specific Integrated Circuit)
  - ASIC designed specifically to move packets (fast and dumb)
  - CPU manages control plane
  - CPU only moves packets the ASIC can't (options, fragmentation etc)
  - Data Plane packets sent to the CPU are "punted"

6500/7600/ASR9K/CRS and Nexus 7000 switches are hardware based

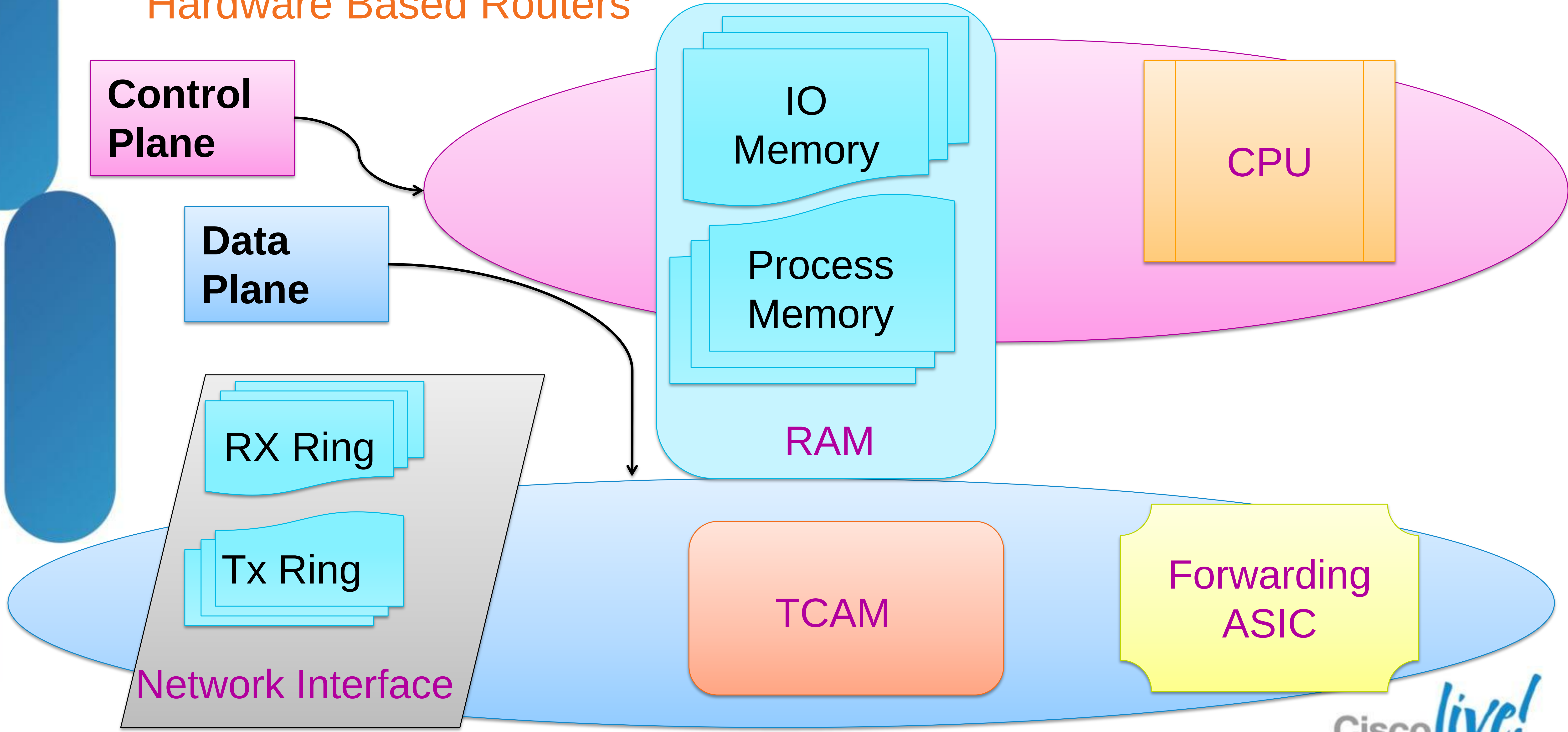
Router Components

Hardware Based Routers



Router Components

Hardware Based Routers



Routing Operations in Cisco IOS Routers

Agenda

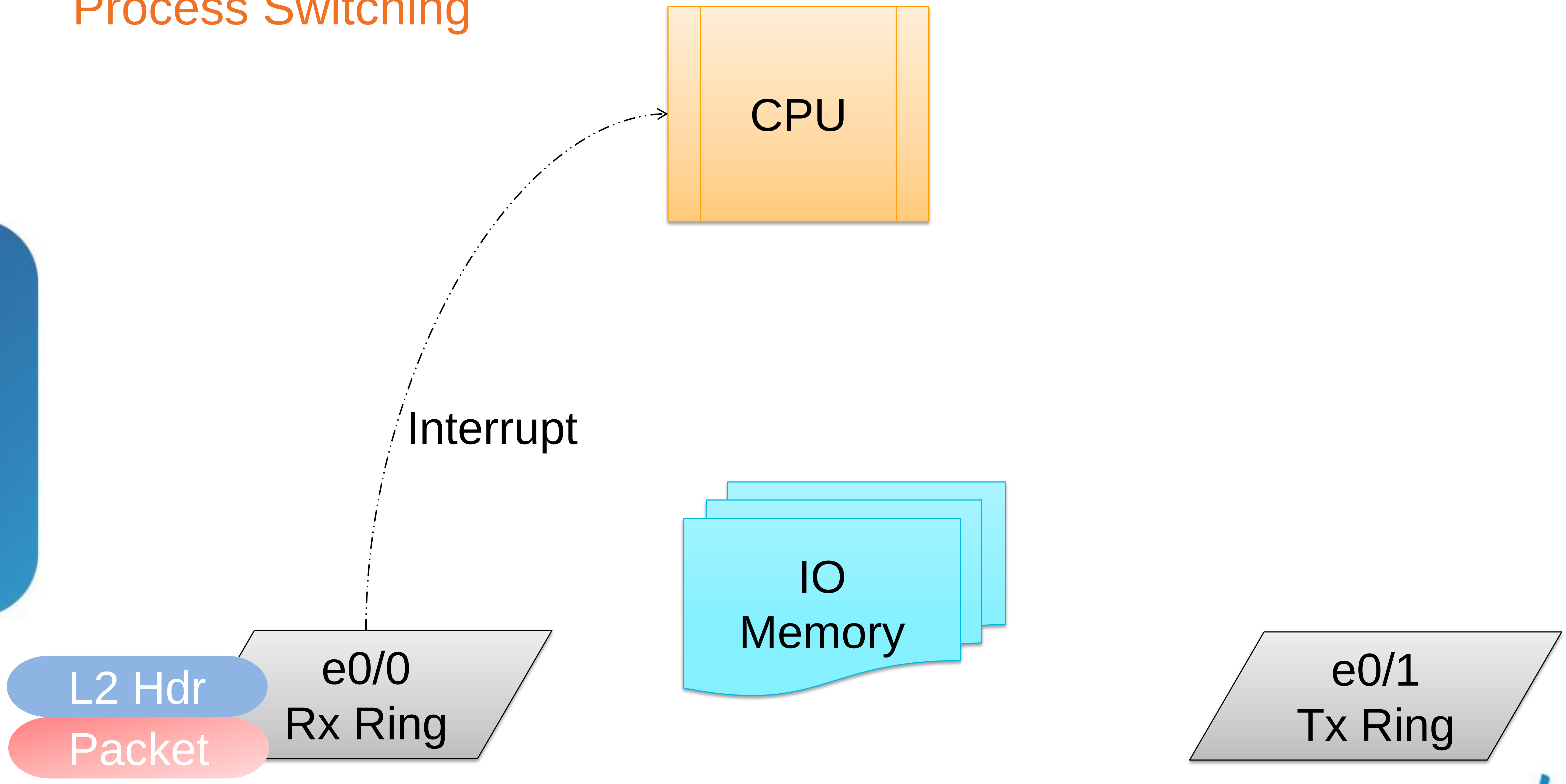
- Router Components
- Moving Packets
- CEF, CPU and Memory
- Outbound Load Sharing
- Routing Convergence Improvements

Moving Packets
Overview

- CEF Switching and Process Switching
  – Fast Switching is deprecated as of 12.4(20)T
  – Not covered today
- CEF Switching is the default
- Process Switching is the fallback
  – Anything CEF can't handle

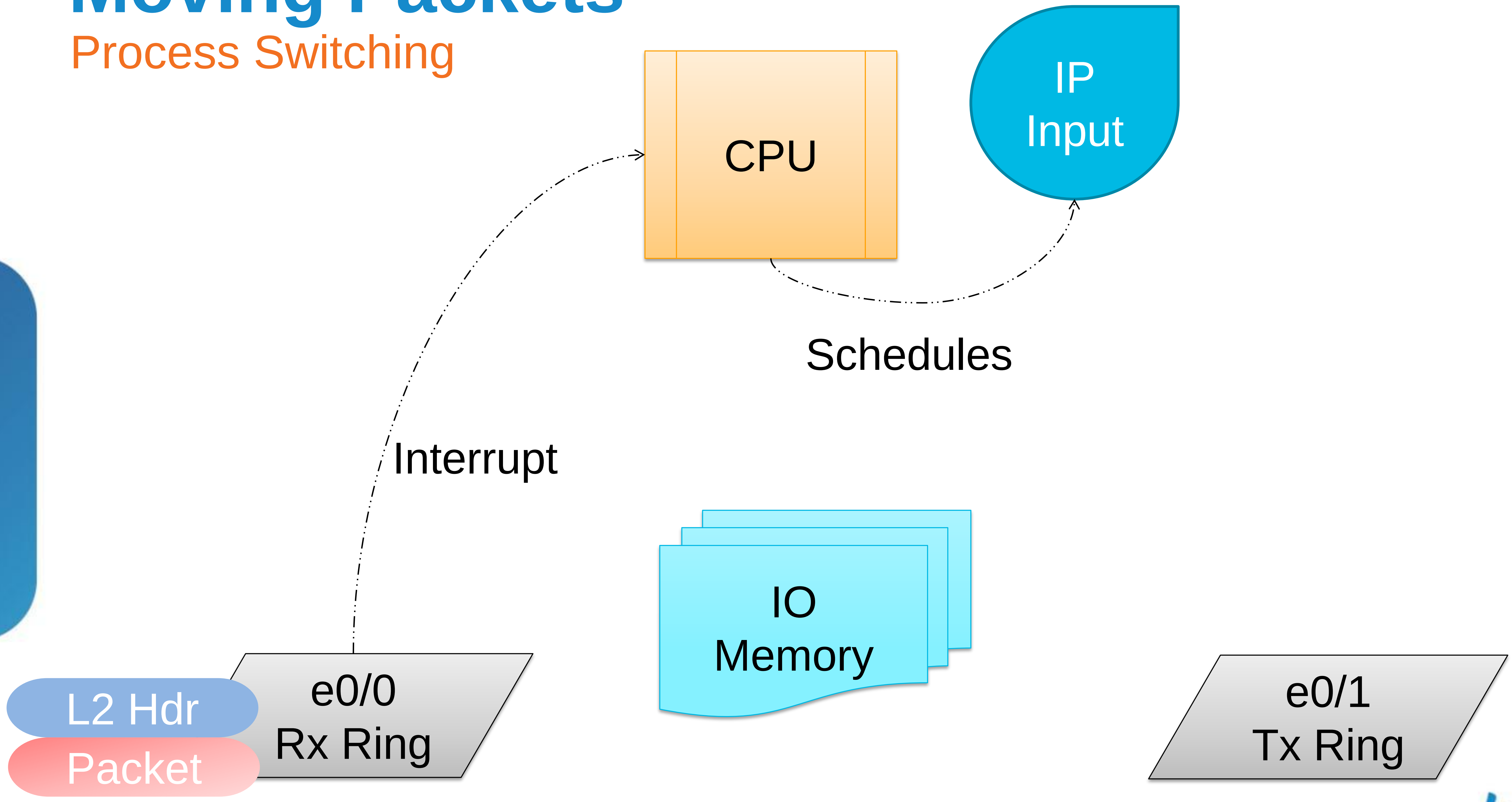
Moving Packets

Process Switching

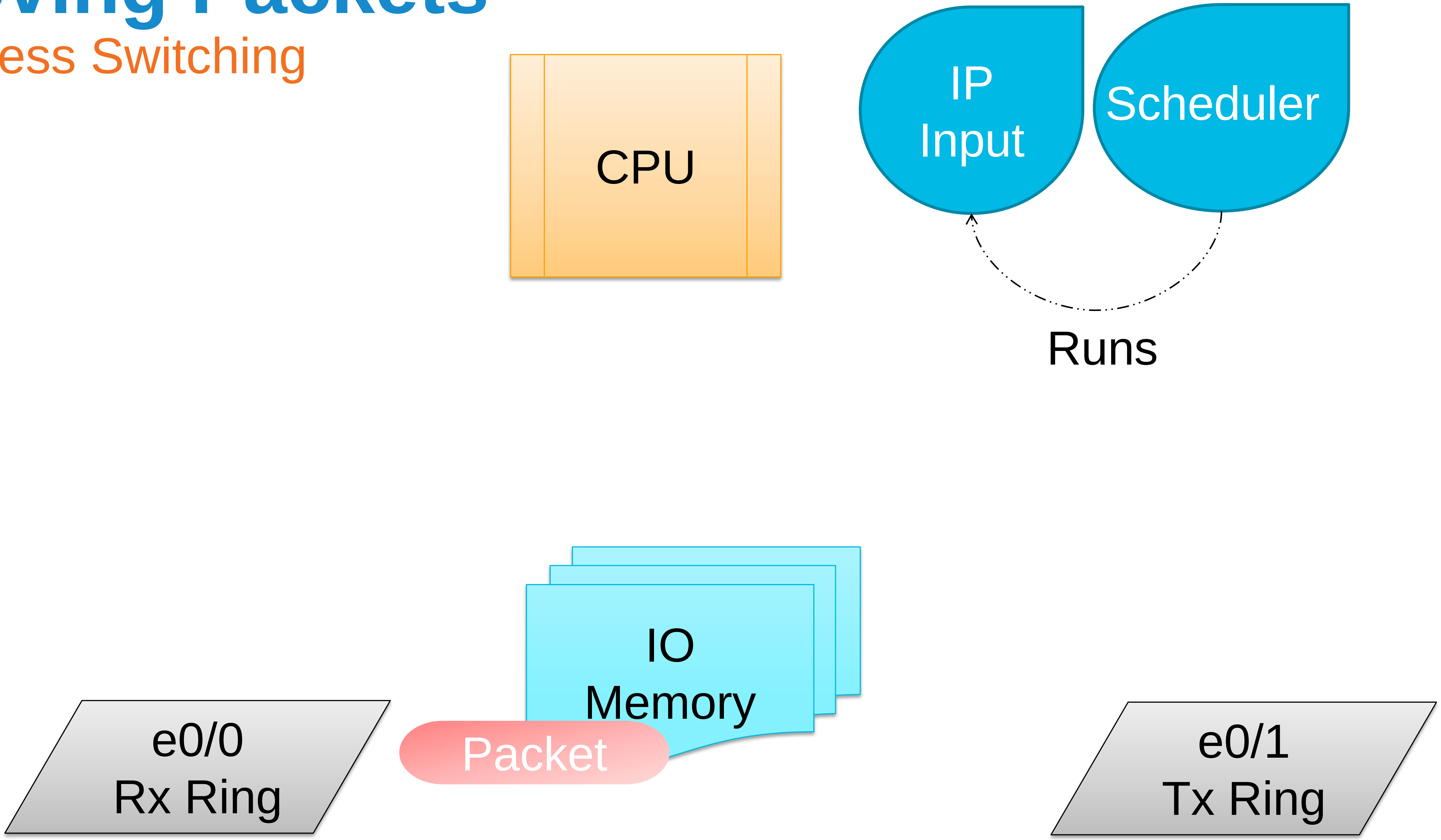


Moving Packets

Process Switching

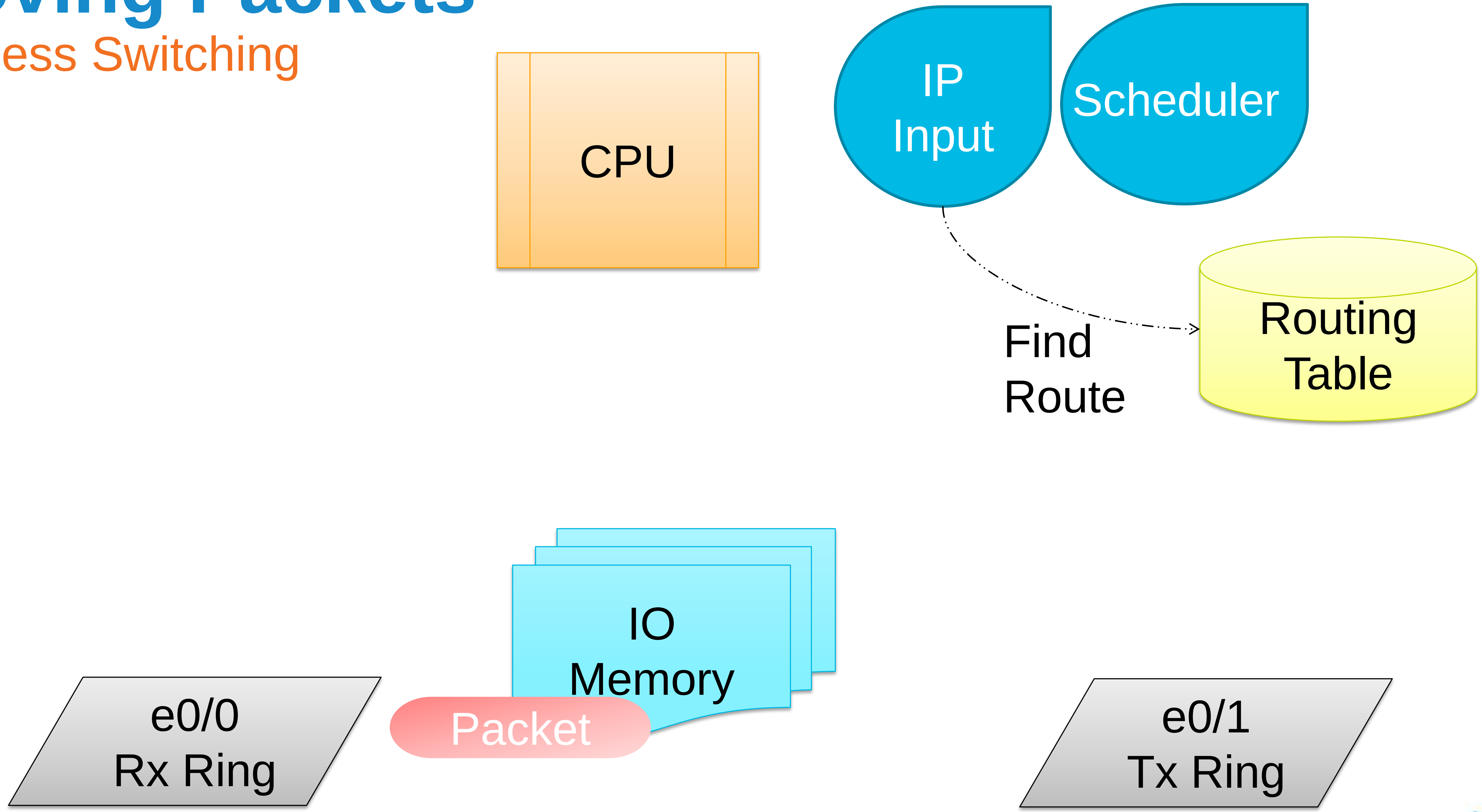


Moving Packets
Process Switching

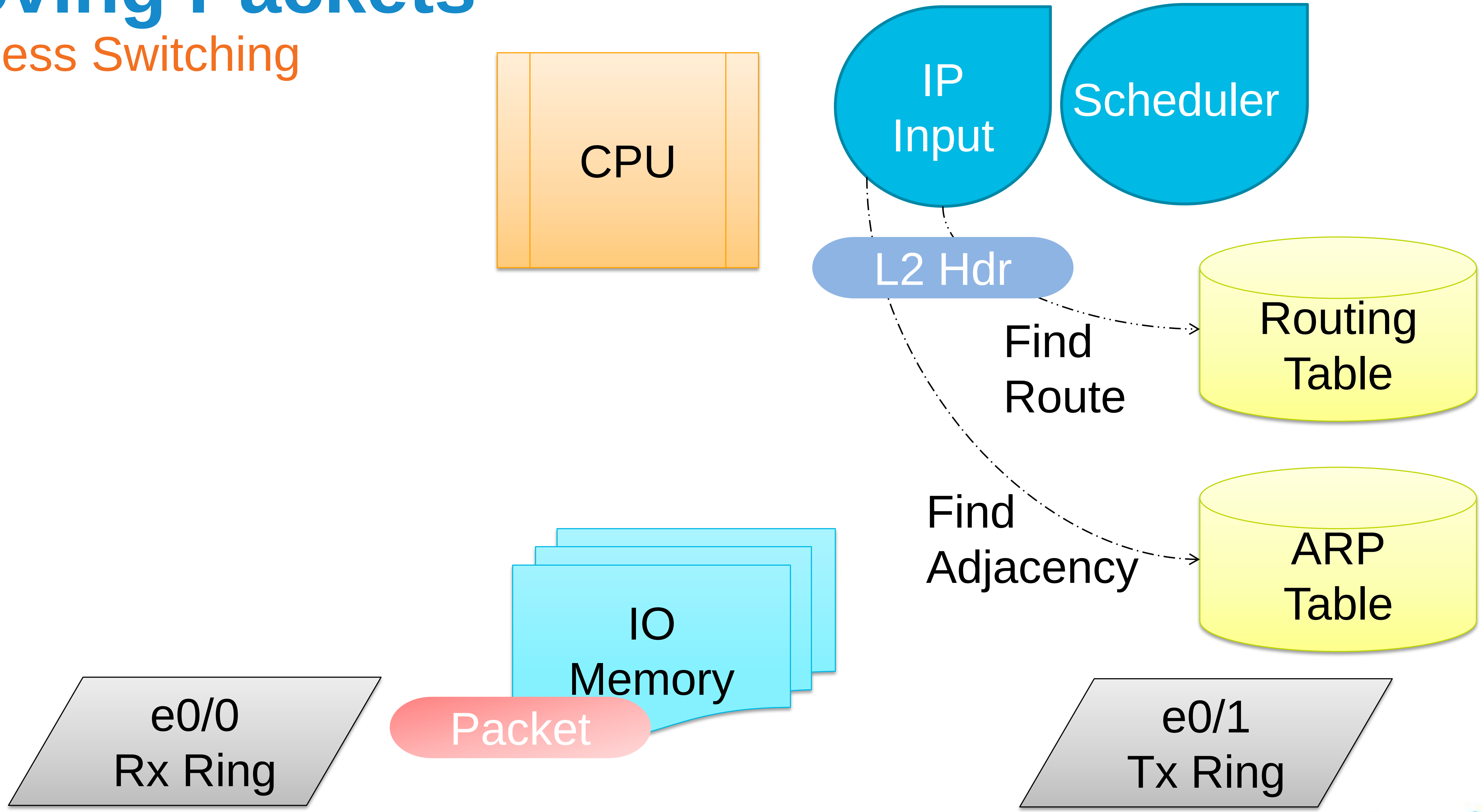


Moving Packets

Process Switching

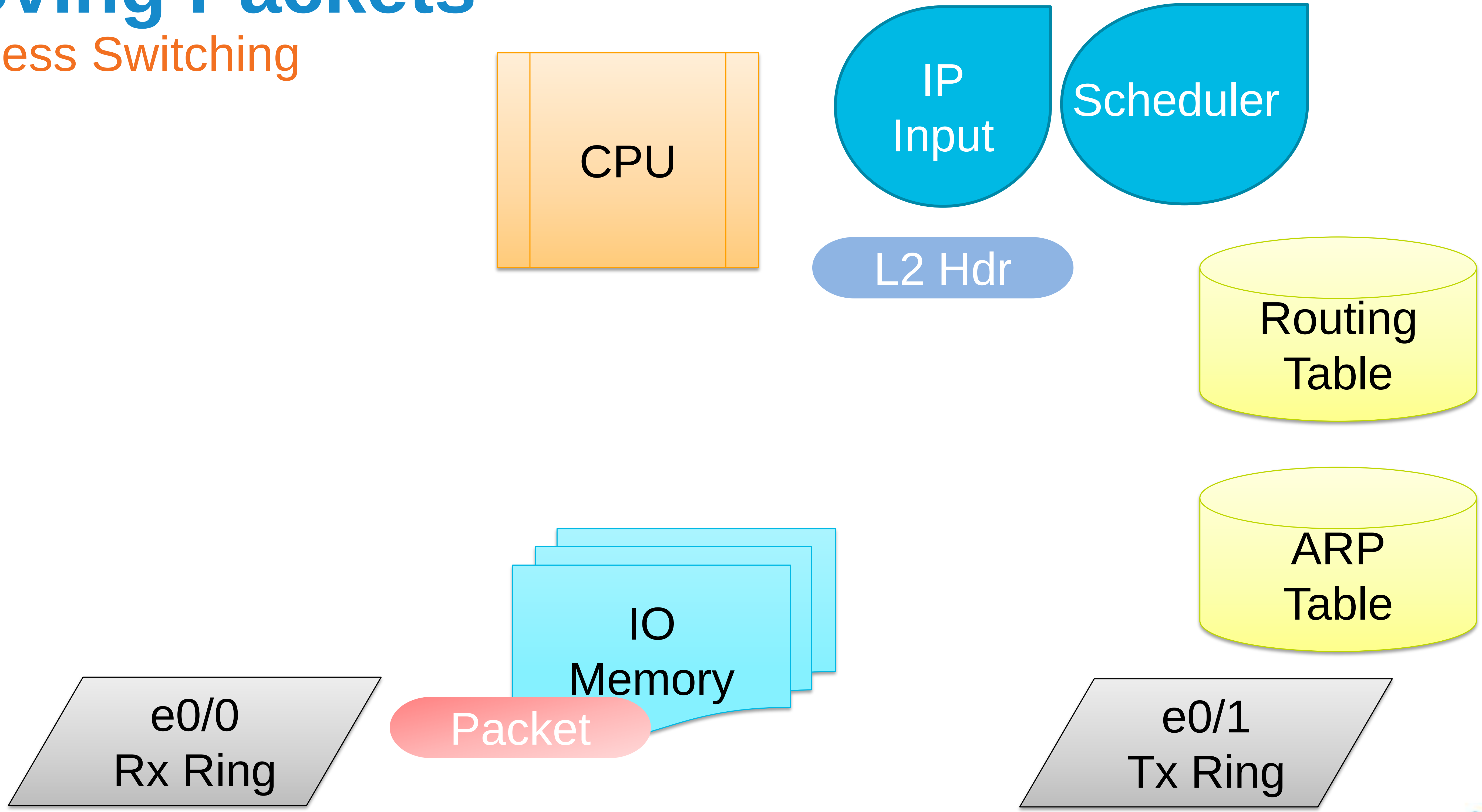


Moving Packets
Process Switching



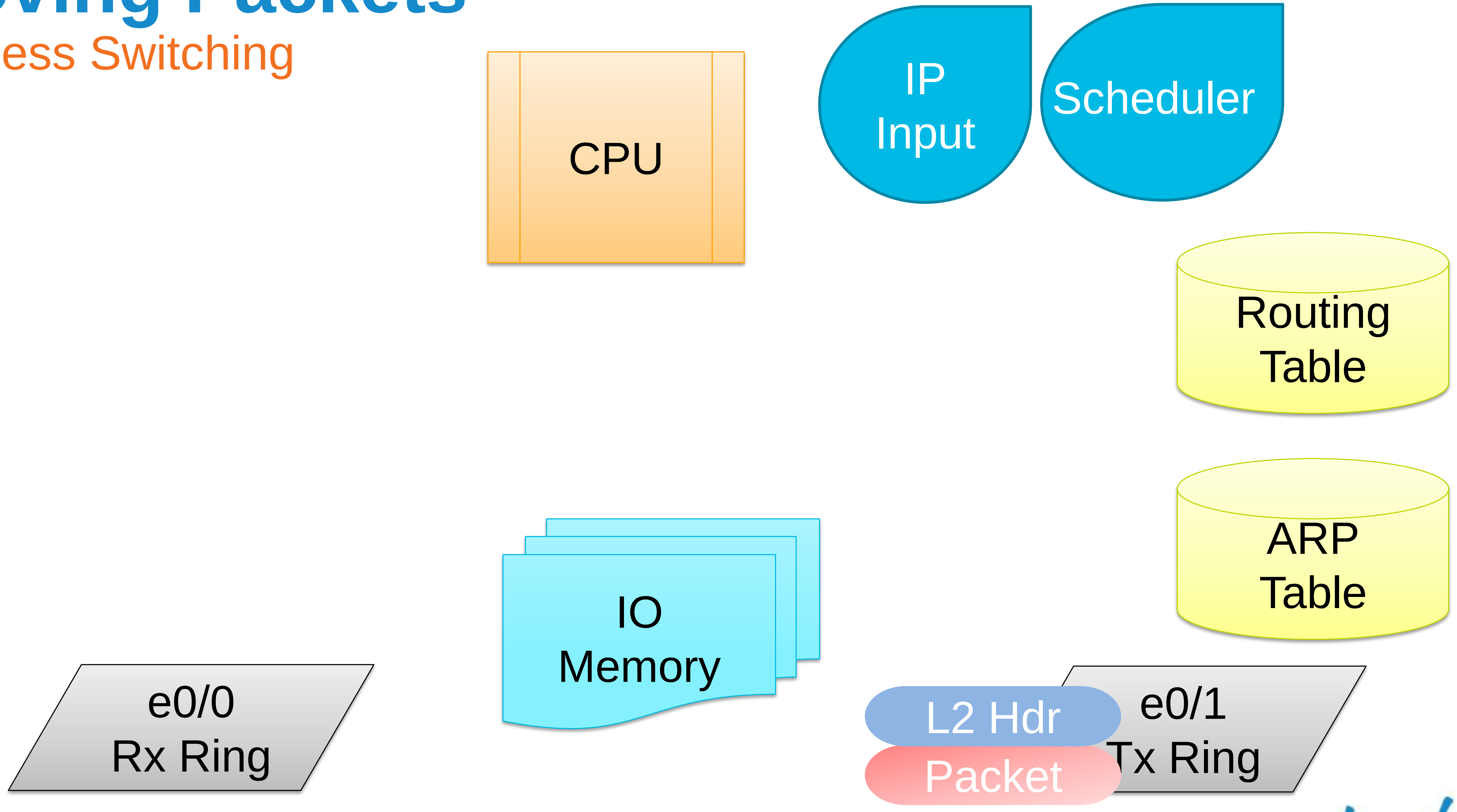
Moving Packets

Process Switching



Moving Packets

Process Switching



Moving Packets
Process Switching

- Process Switching is **BAD - CPU Intensive**
- Multiple lookups
- Inefficient data structures
- Process scheduling

- What can we do to improve?
  – Better data structures
  – Pre-compile forwarding information

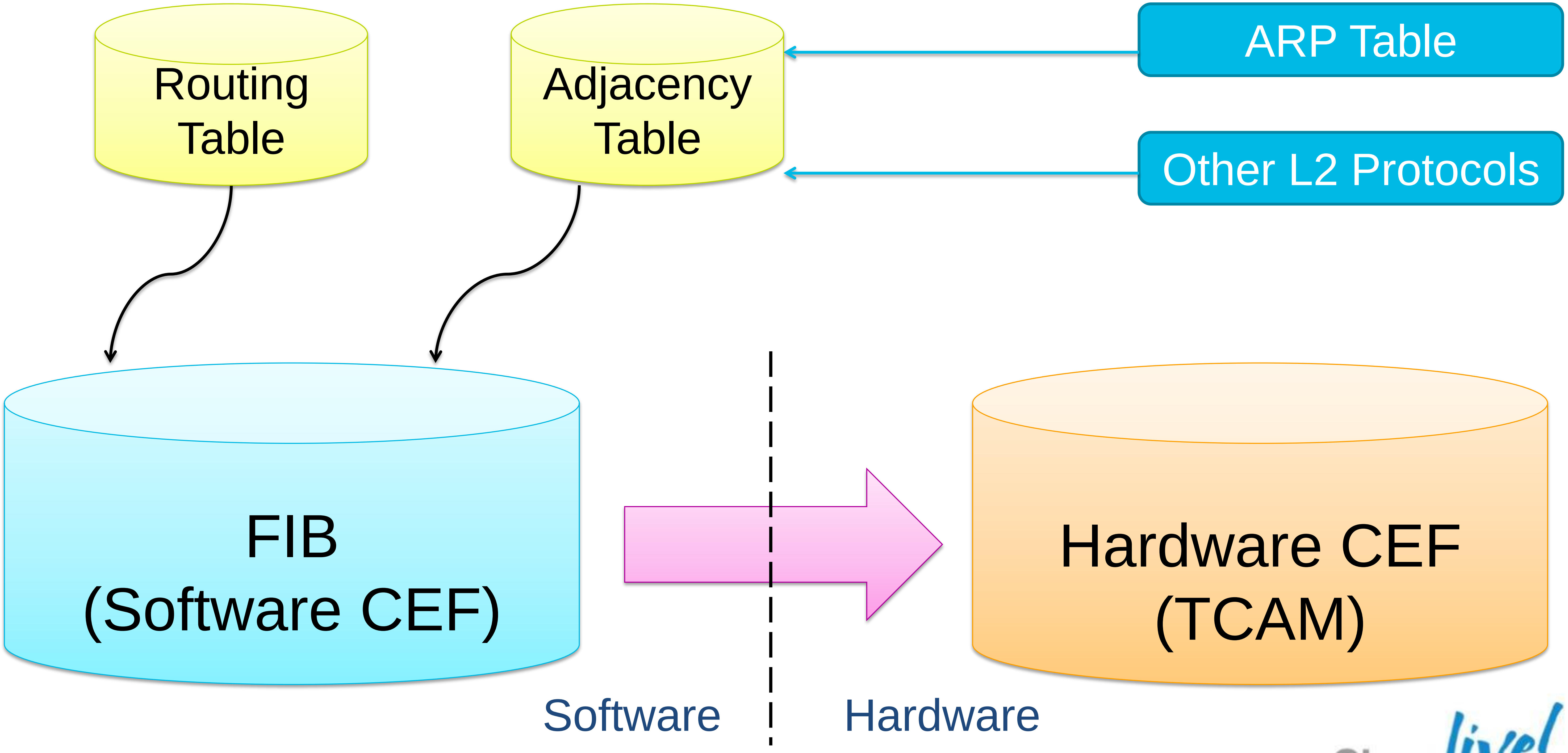
Moving Packets

CEF Switching

TOMORROW starts here.

The FIB (Forwarding Information Base)

"Show IP CEF"



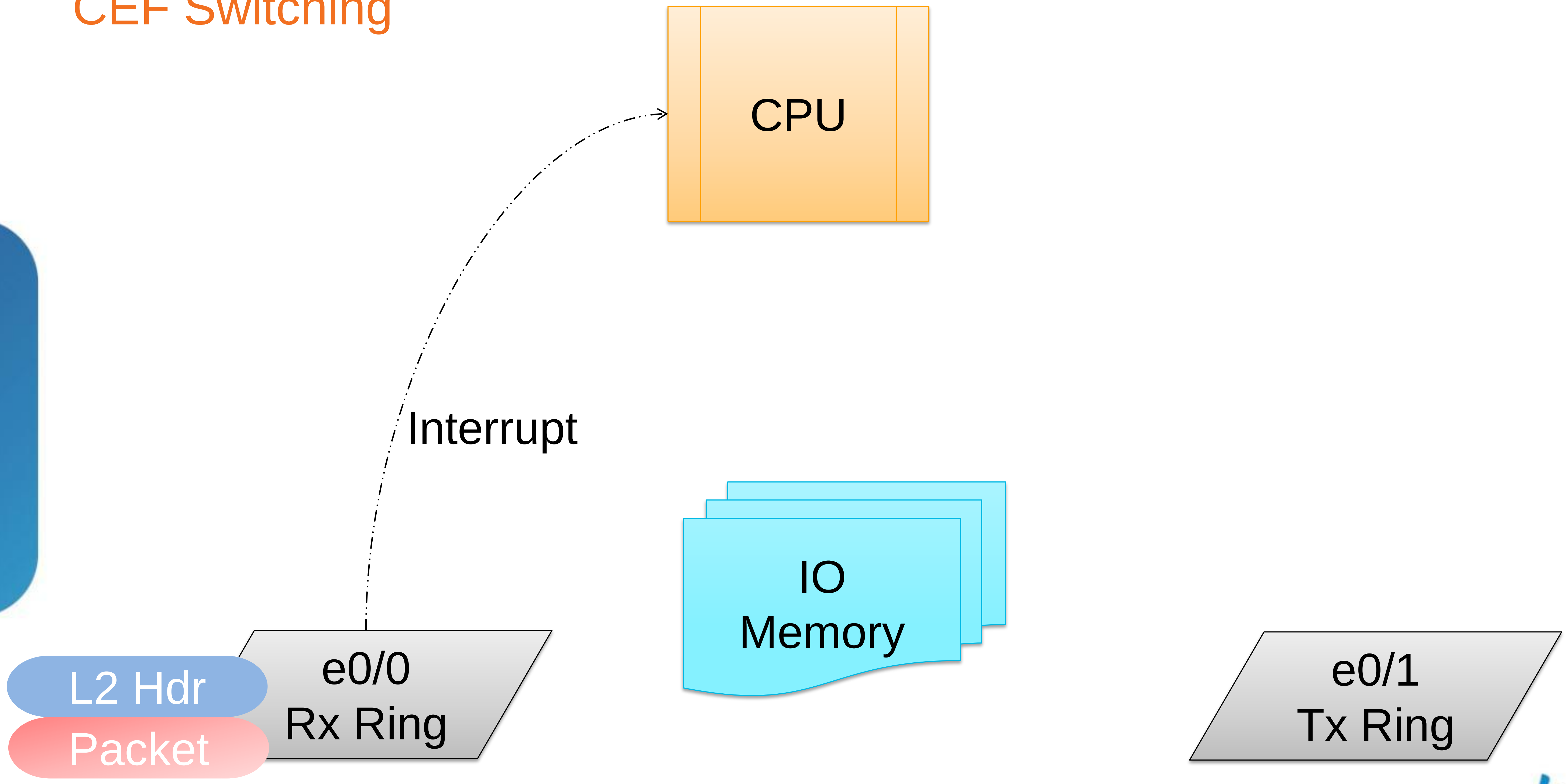
Moving Packets

CEF Overview

- CEF Table = Route + Egress Interface + L2 Destination
- Single lookup (and faster too!)
- No process scheduling

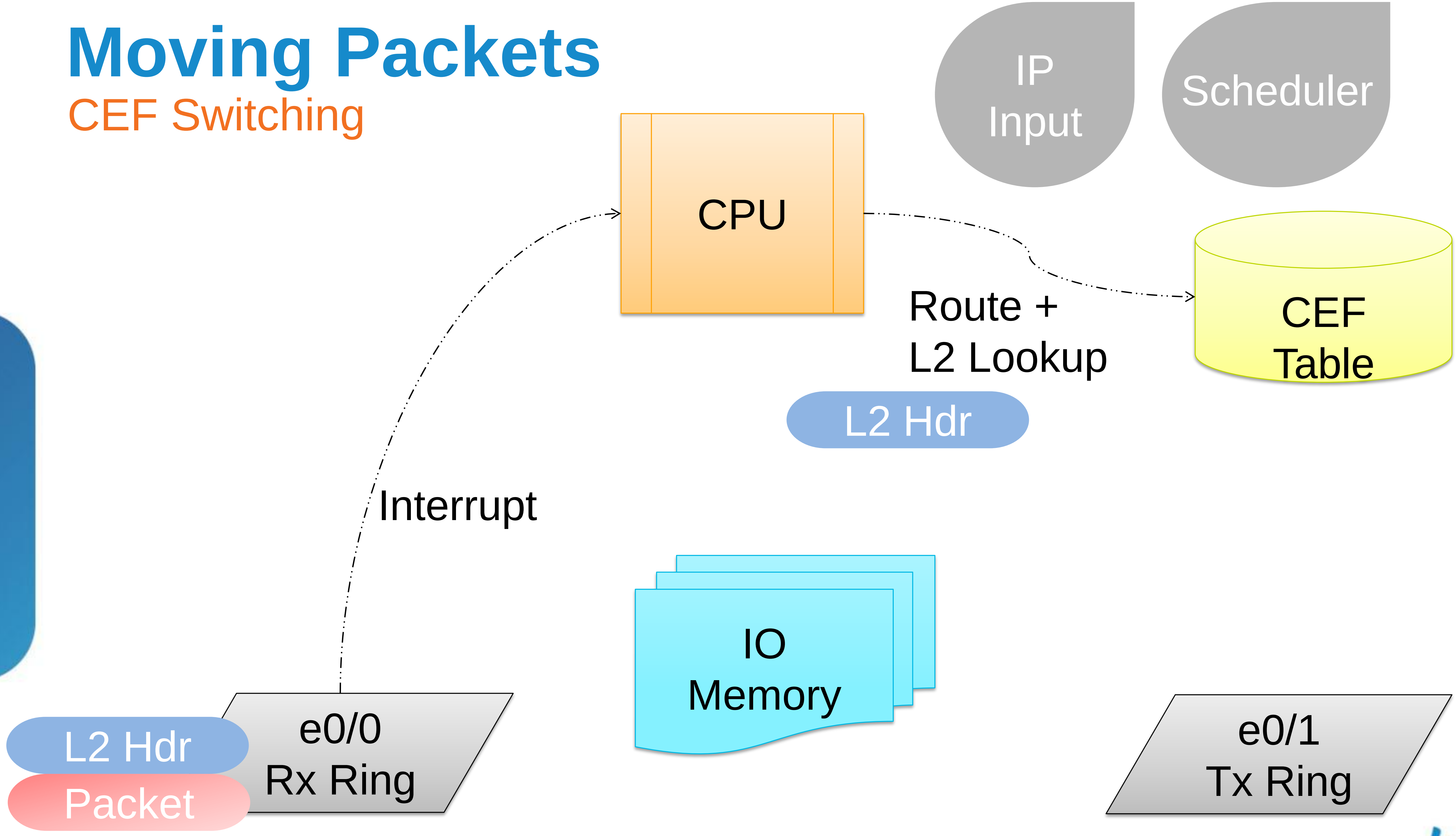
Moving Packets

CEF Switching



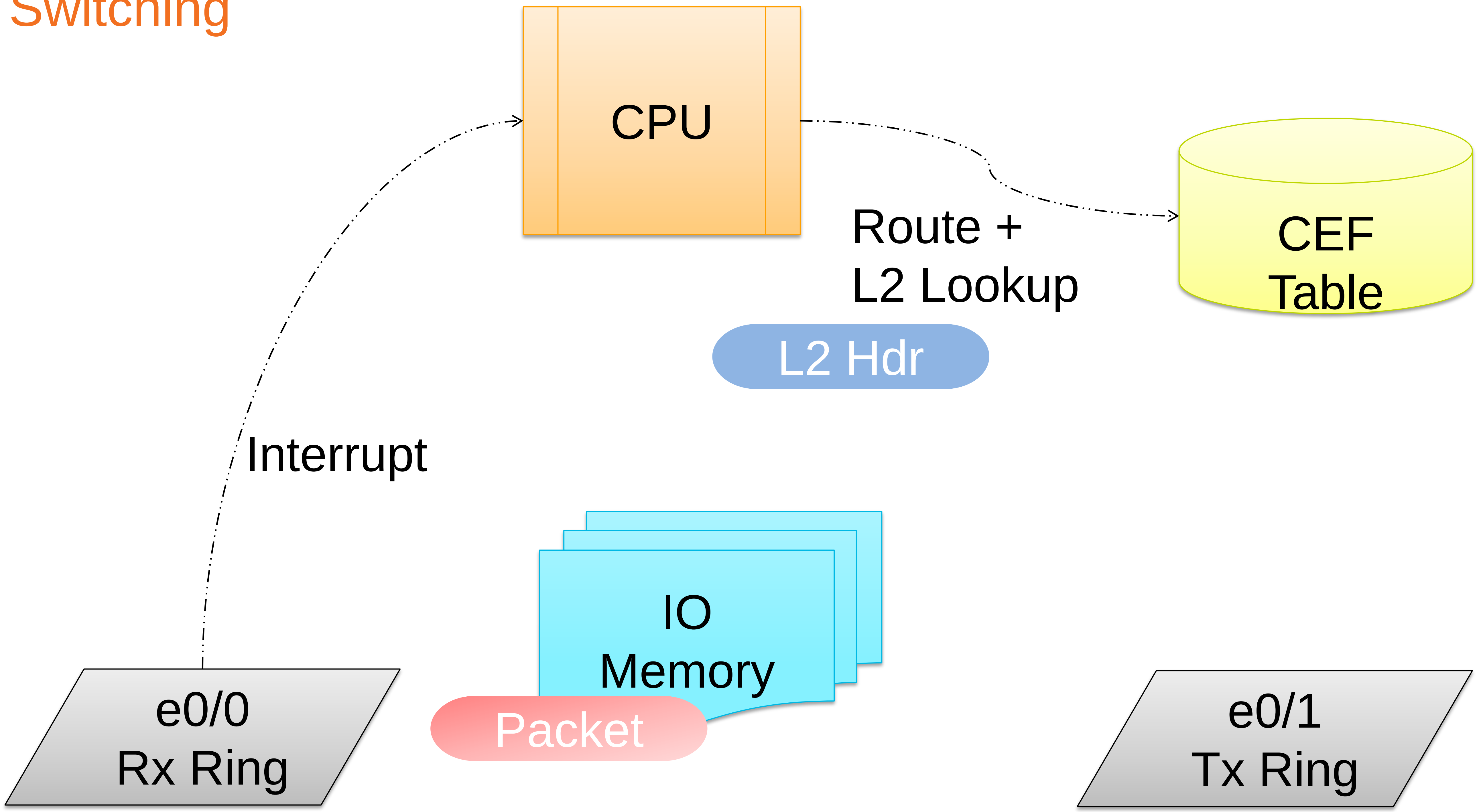
Moving Packets

CEF Switching



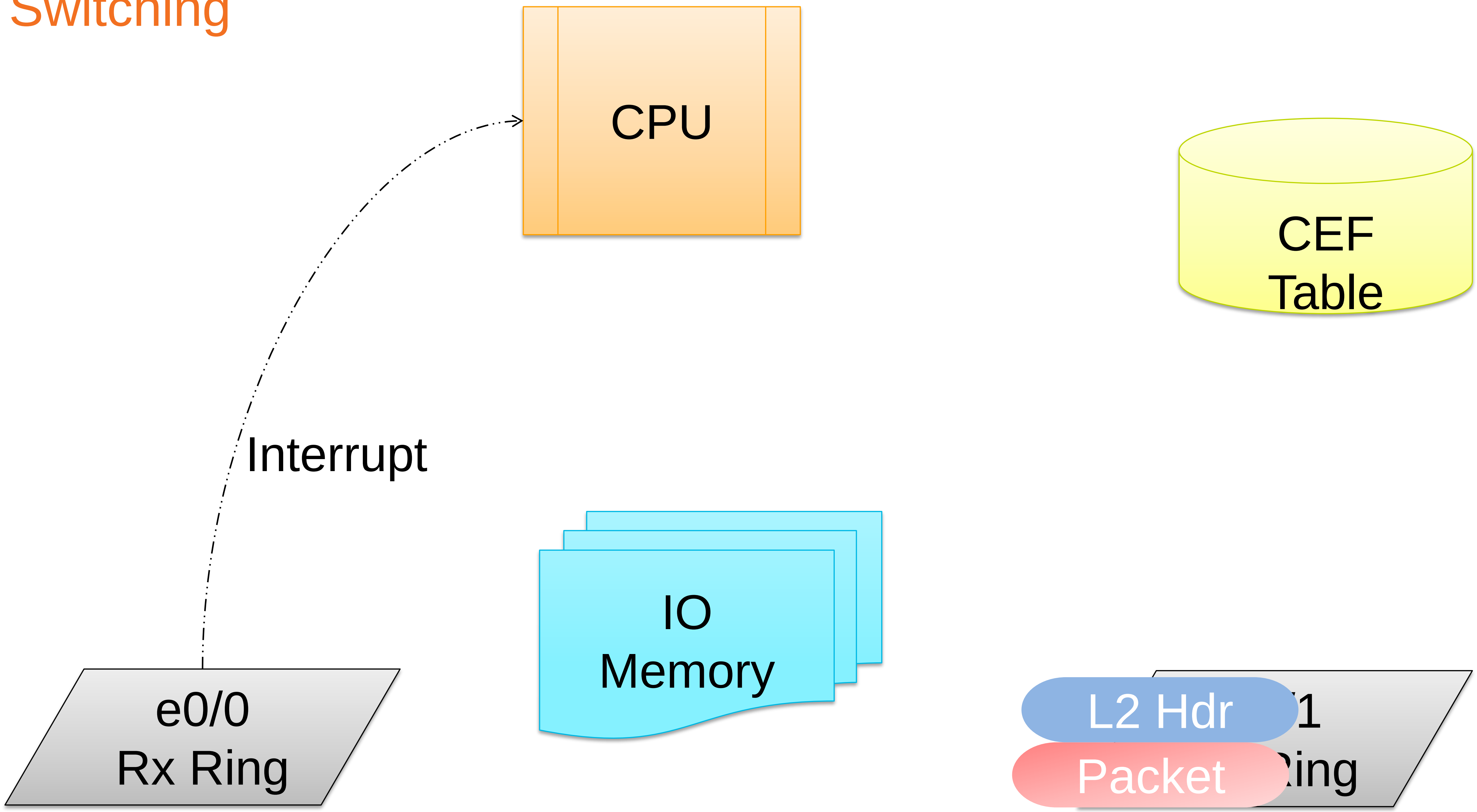
Moving Packets

CEF Switching



Moving Packets

CEF Switching



Moving Packets

CEF Switching - Summary

- Interrupt removes process scheduling
- Pre-compiled Interface + L2 information (cache)
- CEF table data structure improvement
  - RIB is a hash
  - CEF is a mtrie
- **Single lookup** for all necessary forwarding information

Moving Packets

Features and Switching Paths

- Supported in CEF
  - QoS
  - ACL
  - Zone Based Firewall
  - NAT
  - Netflow
  - IPSec
  - GRE
  - PBR
  - Many more!
- Process Switching Only
  - ACL Logging (ie deny any any log)
  - Packets destined to the router (BGP,OSPF ping etc)
  - No L2 Adjacency

Routing Operations in Cisco IOS Routers

Agenda

- Router Components
- Moving Packets
- CEF, CPU and Memory
- Outbound Load Sharing
- Routing Convergence Improvements

CEF and CPU Utilisation

CEF and CPU Utilisation

- CPU does everything
- Processes vs. Interrupts
  - SPF
  - BGP
  - Routed Packets

```
CPU utilization for five seconds: 5%/2%; one minute: 3%; five minutes: 2%
  PID  Runtime (ms)     Invoked    uSecs      5Sec    1Min     5Min  TTY     Process
...
    2            68         585      116     1.00%   1.00%       0%    0     IP Input
   17            88        4232       20     0.20%   1.00%       0%    0     BGP Router
   18           152       14650       10        0%      0%       0%    0     BGP Scanner
...
```

CEF and CPU Utilisation

CPU Utilisation Examples

1. CPU Utilisation due to moderate traffic rates

```
CPU utilisation for five seconds: 47%/46%; one minute: 40%; five minutes: 39%
```

CEF and CPU Utilisation

CPU Utilisation Examples

1. CPU Utilisation due to moderate traffic rates

```
CPU utilization for five seconds: 47%/46%; one minute: 40%; five minutes: 39%
```

2. High CPU due to OSPF Reconvergence

```
CPU utilization for five seconds: 99%/3%; one minute: 53%; five minutes: 49%
 PID  Runtime(ms)  Invoked   uSecs     5Sec     1Min    5Min   TTY Process
 357       319932   138750   21039   88.32%   41.18%  36.78%     0 OSPF-1 Router
```

CEF and CPU Utilisation

CPU Utilisation Examples

1. CPU Utilisation due to moderate traffic rates

```
CPU utilization for five seconds: 47%/46%; one minute: 40%; five minutes: 39%
```

2. High CPU due to OSPF Reconvergence

```
CPU utilization for five seconds: 99%/3%; one minute: 53%; five minutes: 49%
 PID   Runtime(ms)   Invoked    uSecs       5Sec      1Min    5Min    TTY Process
 357        319932    138750    21039     88.32%    41.18%  36.78%      0 OSPF-1 Router
```

3. High CPU due to multiple Virtual Exec Processes

```
CPU utilization for five seconds: 99%/3%; one minute: 99%; five minutes: 99%
 PID Runtime(ms)   Invoked       uSecs   5Sec   1Min   5Min TTY Process
  3     24871276   47622133        522 30.62% 31.62% 31.57%   2 Virtual Exec
122     24812452   47528825        522 30.53% 31.62% 31.60%   3 Virtual Exec
131     24790280   47490842        522 32.84% 31.88% 31.31%   4 Virtual Exec
```

CEF and CPU Utilisation

Process Priority

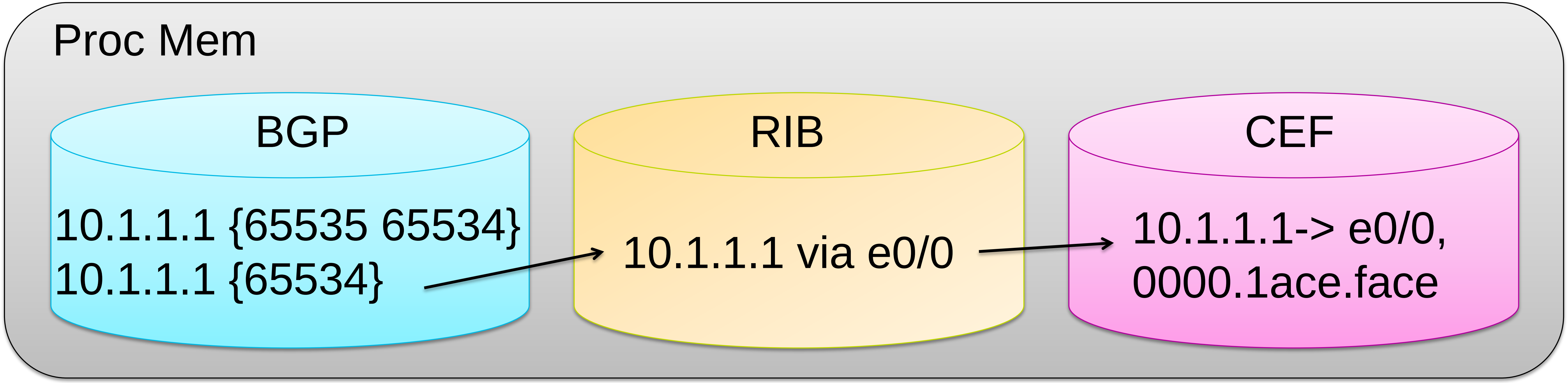
- Processes assigned priority
  – Critical/High/Medium/Low
- Priority Scheduler
- Run to Completion Model
  – Processes choose to suspend
- Interrupts break the rules

Routing Protocols, CEF and Memory

Routing Protocols, CEF and Memory

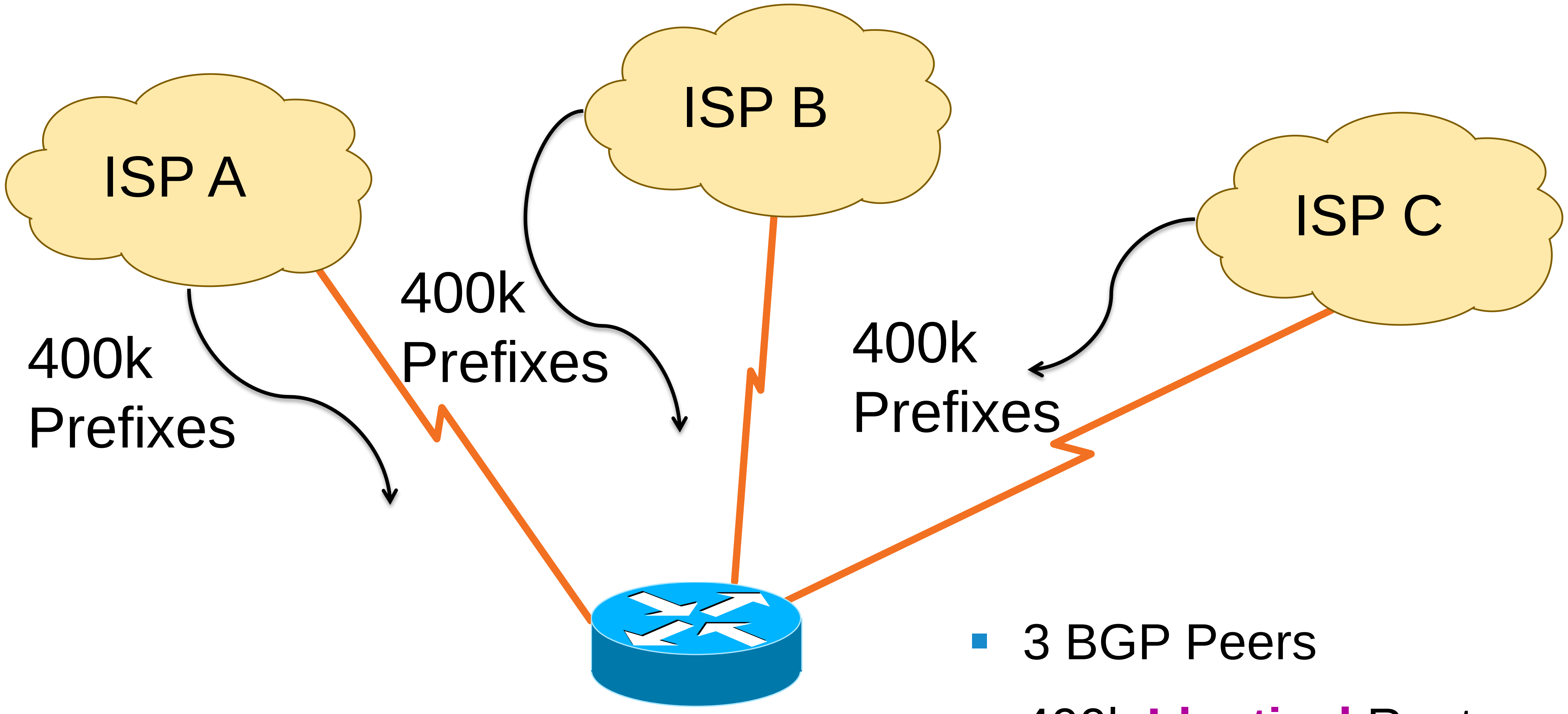
Process Memory

- Routing Protocol, RIB, and CEF each take their own memory
- RIB built from Routing Protocols
- CEF built from RIB



Routing Protocols, CEF and Memory

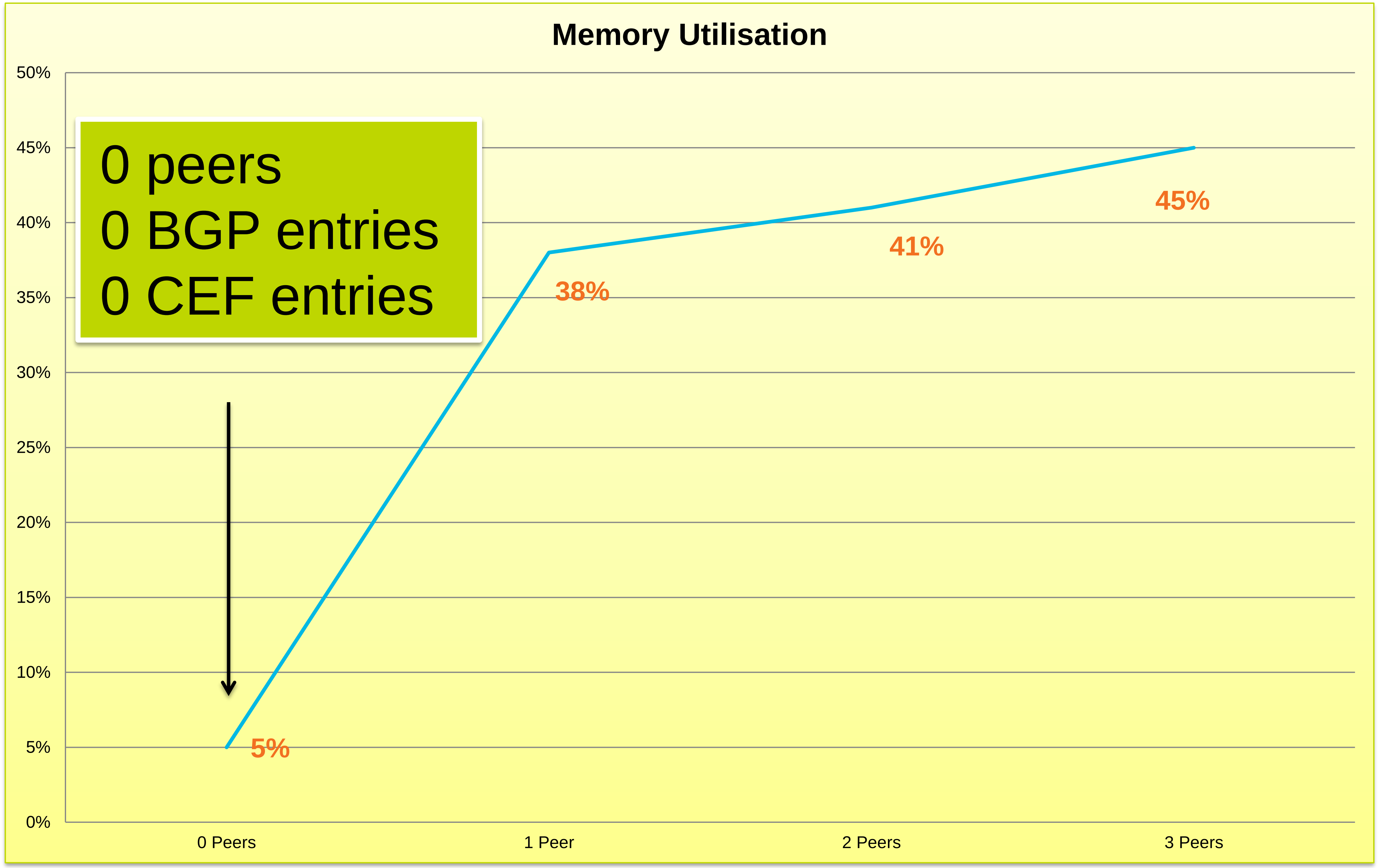
Memory Utilisation – Multiple Internet Tables

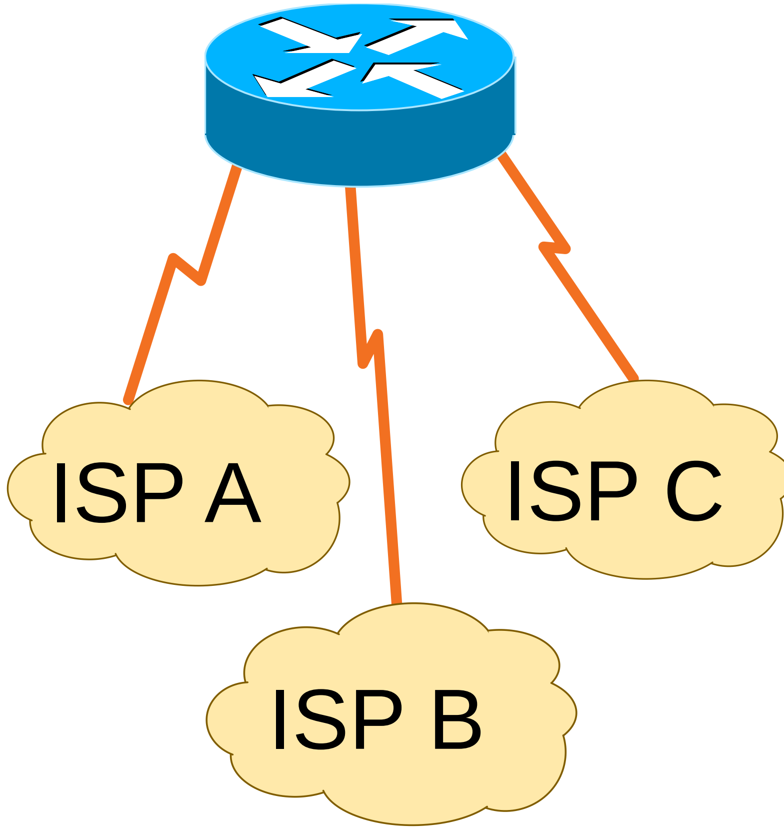


- 3 BGP Peers
- 400k **Identical** Routes
- 15.2(2)T

Routing Protocols, CEF and Memory

Memory Utilisation – Multiple Internet Tables

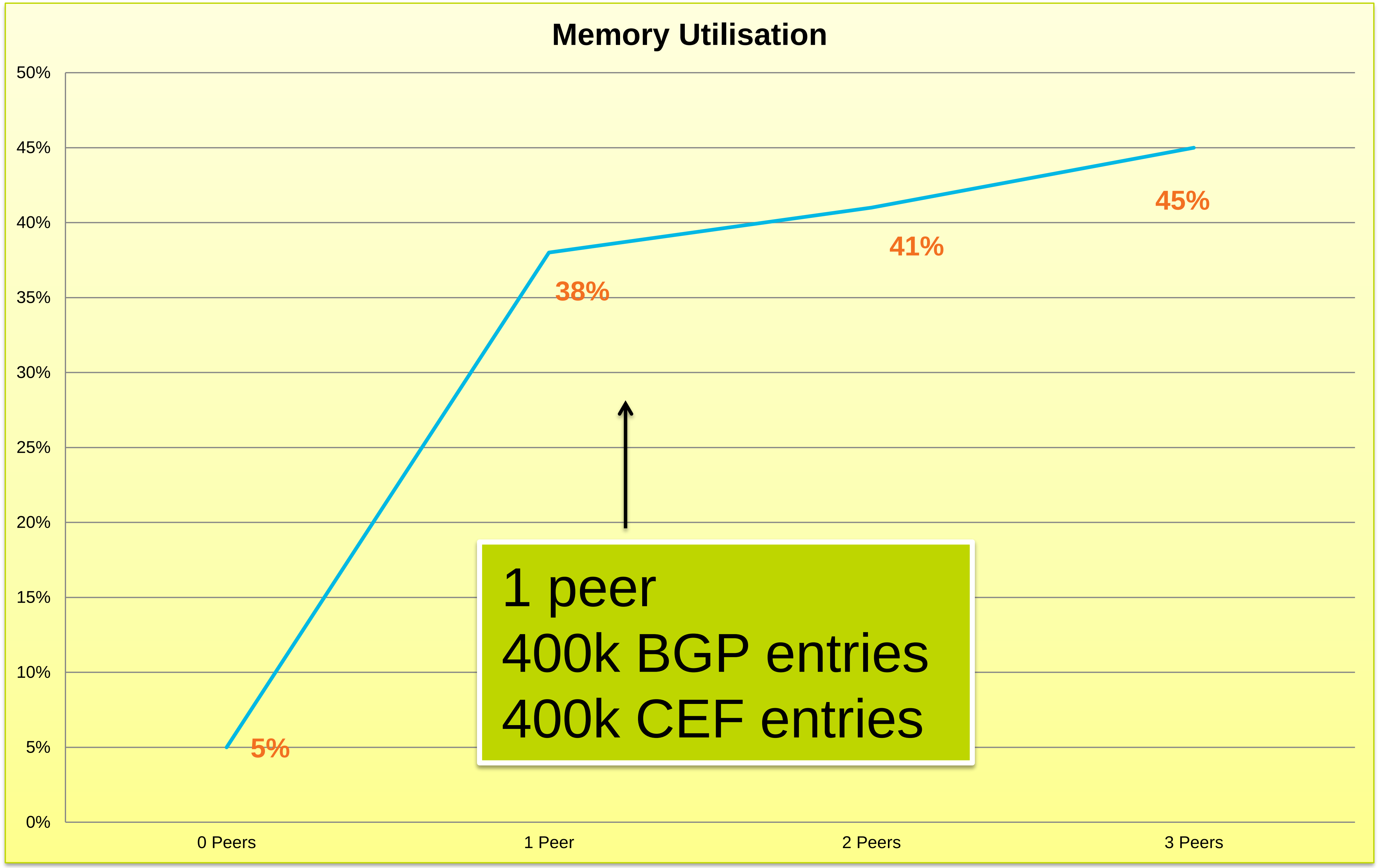


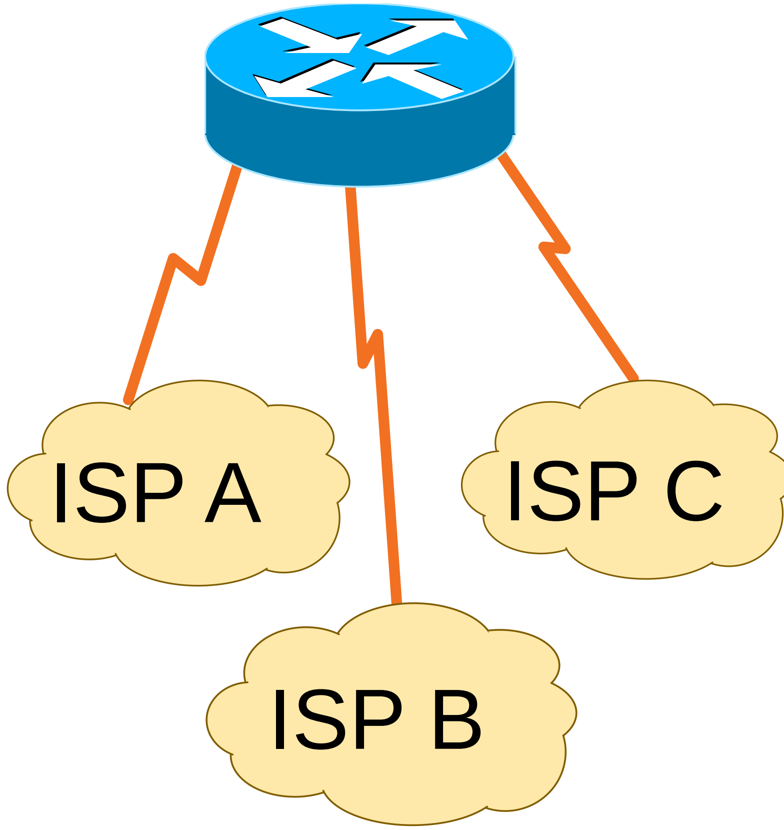


- 3 BGP Peers
- 400k Identical Routes
- 15.2(2)T

Routing Protocols, CEF and Memory

Memory Utilisation – Multiple Internet Tables

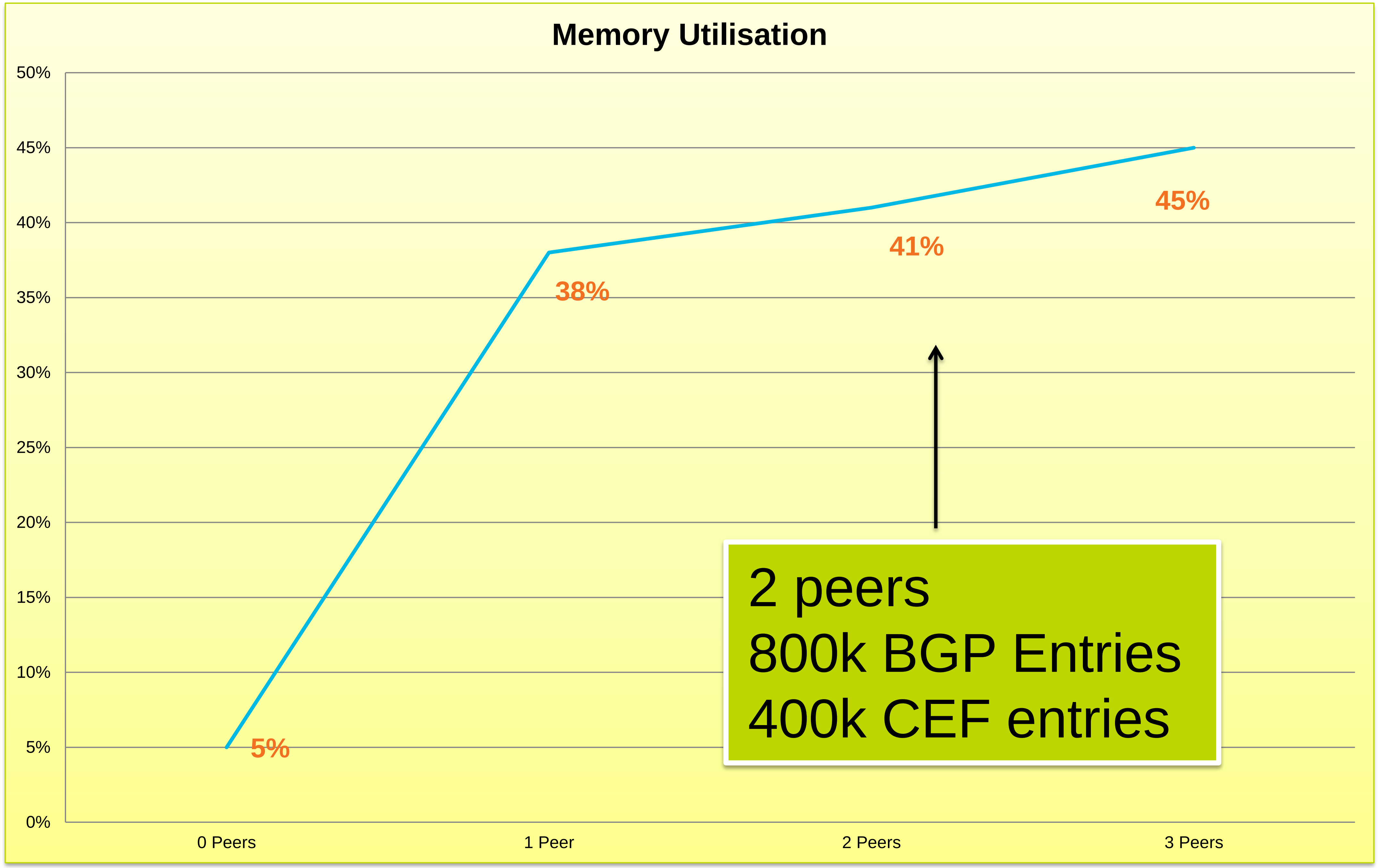


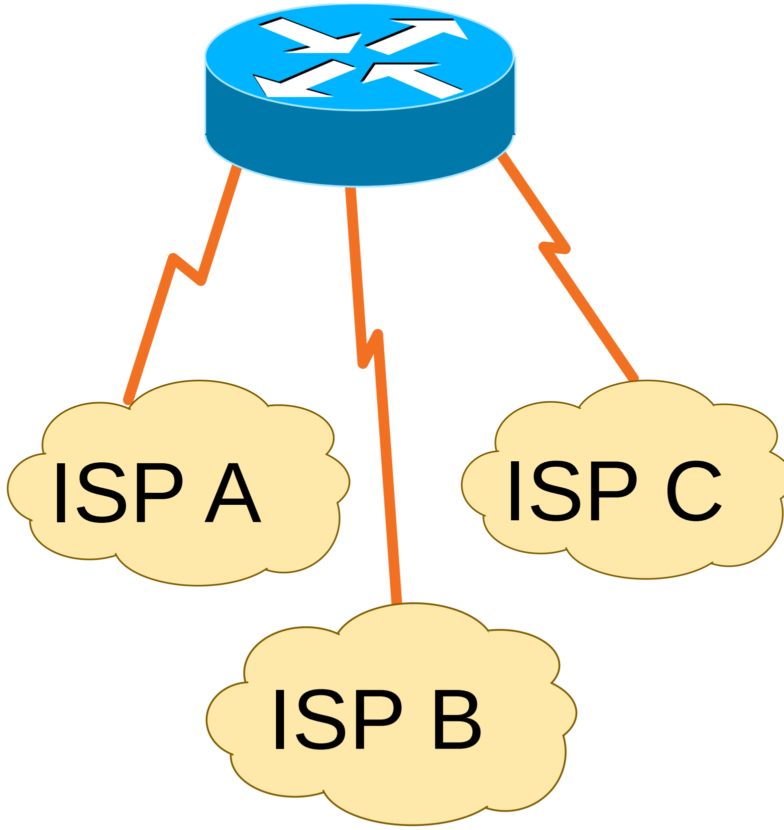


- 3 BGP Peers
- 400k Identical Routes
- 15.2(2)T

Routing Protocols, CEF and Memory

Memory Utilisation – Multiple Internet Tables

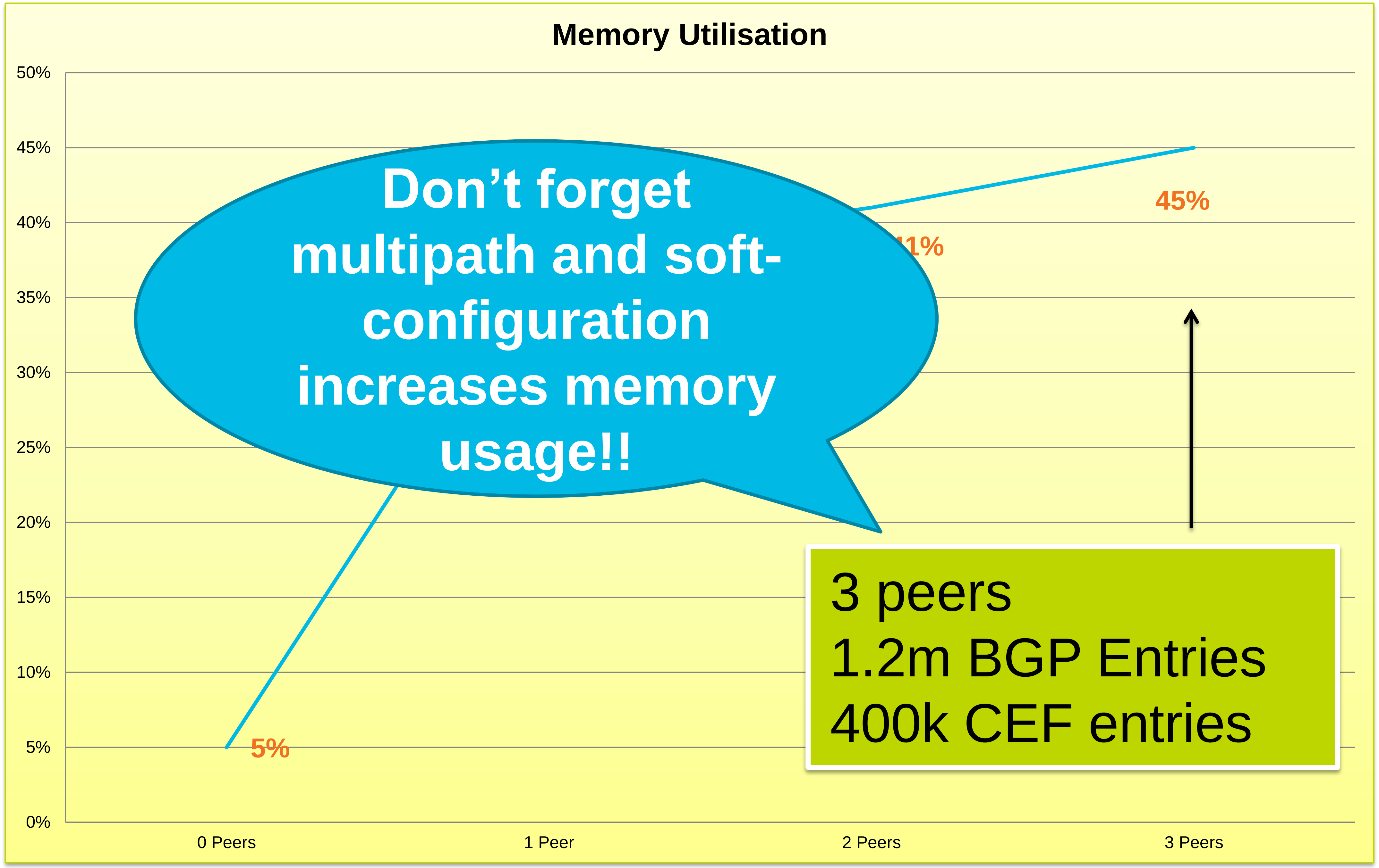


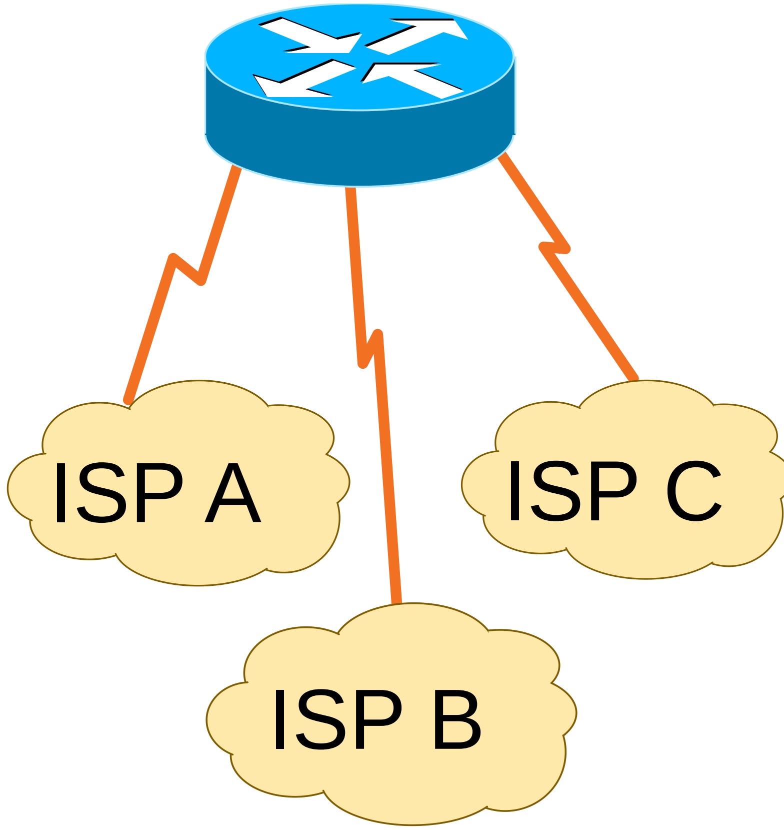


- 3 BGP Peers
- 400k Identical Routes
- 15.2(2)T

Routing Protocols, CEF and Memory

Memory Utilisation – Multiple Internet Tables





- 3 BGP Peers
- 400k Identical Routes
- 15.2(2)T

Routing Operations in Cisco IOS Routers

Agenda

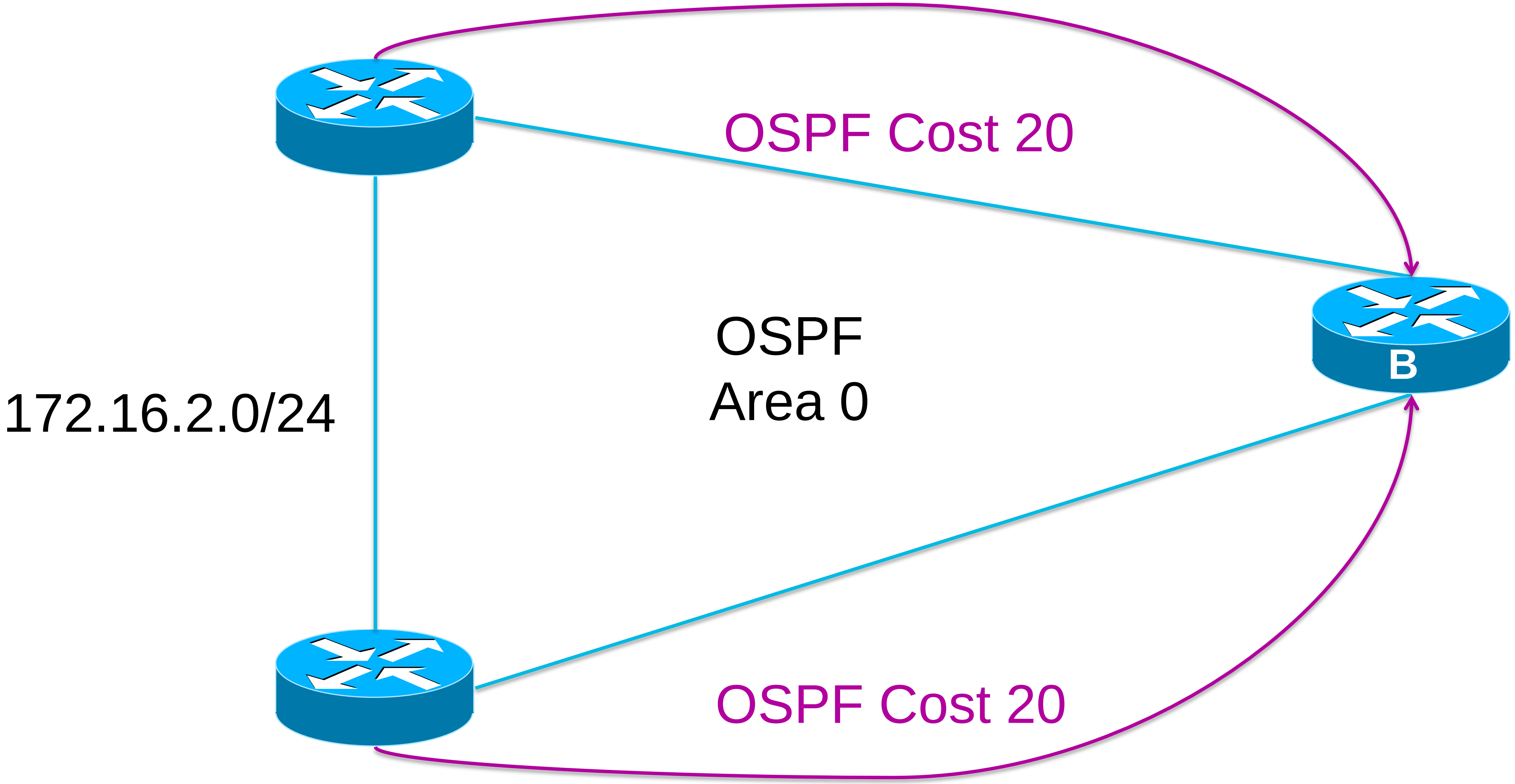
- Router Components
- Moving Packets
- CEF, CPU and Memory
- Outbound Load Sharing
- Routing Convergence Improvements

Load Sharing

CEF Load Sharing

Equal Cost Multipath (ECMP)

Equal Cost Overview



Equal Cost Multipath

The Routing Table (Equal Cost)

```
RouterB#show ip route 172.16.2.0

Routing entry for 172.16.2.0/24
  Known via "ospf 1", distance 110, metric 20, type intra area
  Last update from 172.16.1.1 on Ethernet0/0, 1d02h ago
  Routing Descriptor Blocks:

  * 192.168.100.1, from 192.168.200.1, 1d02h ago, via Ethernet0/1
      Route metric is 20, traffic share count is 1

    172.16.1.1, from 192.168.200.1, 1d02h ago, via Ethernet0/0
      Route metric is 20, traffic share count is 1
```

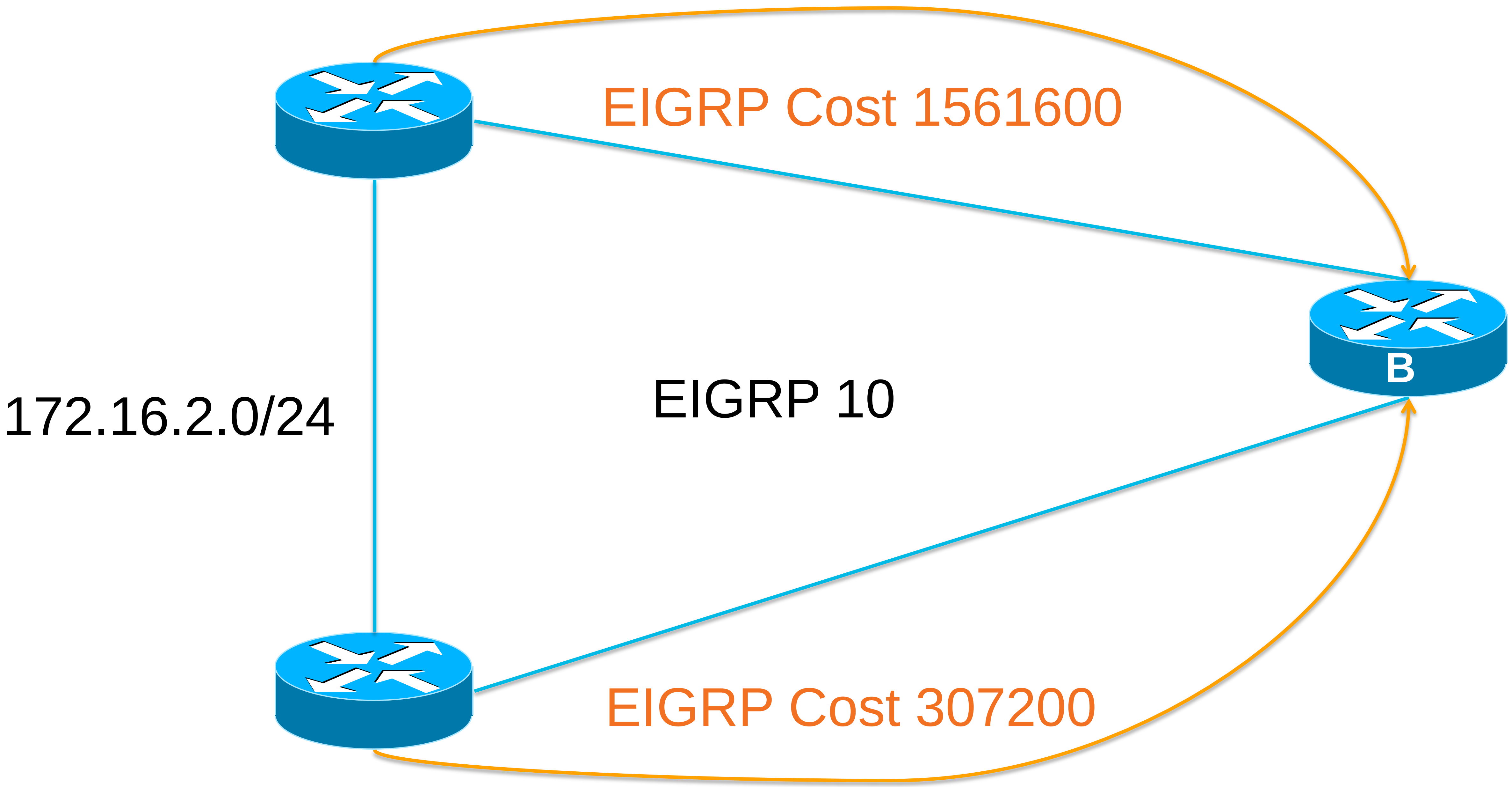
Equal Cost Multipath

The Routing Table (Equal Cost)

```
RouterB#show ip route 172.16.2.0

Routing entry for 172.16.2.0/24
  Known via "ospf 1", distance 110, metric 20, type intra area
  Last update from 172.16.1.1 on Ethernet0/0, 1d02h ago
  Routing Descriptor Blocks:

  * 192.168.100.1, from 192.168.200.1, 1d02h ago, via Ethernet0/1
      Route metric is 20, traffic share count is 1

    172.16.1.1, from 192.168.200.1, 1d02h ago, via Ethernet0/0
      Route metric is 20, traffic share count is 1
```

Equal Cost Multipath

Unequal Cost Overview



Equal Cost Multipath

The Routing Table (Unequal Cost)

```
RouterB#show ip route 172.16.2.0

Routing entry for 172.16.2.0/24
  Known via "eigrp 10", distance 90, metric 307200, type internal
  Last update from 172.16.1.1 on Ethernet0/0, 1d02h ago
  Routing Descriptor Blocks:

    192.168.100.1, from 192.168.200.1, 1d02h ago, via Ethernet0/1
      Route metric is 1561600, traffic share count is 47

      ...

  * 172.16.1.1, from 172.16.1.1, 00:00:16 ago, via Ethernet0/0
      Route metric is 307200, traffic share count is 240
```

Unequal Metrics

Equal Cost Multipath

The Routing Table (Unequal Cost)

```
RouterB#show ip route 172.16.2.0

Routing entry for 172.16.2.0/24
  Known via "eigrp 10", distance 90, metric 307200, type internal
  Last update from 172.16.1.1 on Ethernet0/0, 1d02h ago
  Routing Descriptor Blocks:

    192.168.100.1, from 192.168.200.1, 1d02h ago, via Ethernet0/1
      Route metric is 1561600, traffic share count is 47

      ...

  * 172.16.1.1, from 172.16.1.1, 00:00:16 ago, via Ethernet0/0
      Route metric is 307200, traffic share count is 240
```

Unequal traffic share count

Equal Cost Multipath

The Routing Table (Unequal Cost)

```
RouterB#show ip route 172.16.2.0

Routing entry for 172.16.2.0/24
  Known via "eigrp 10", distance 90, metric 307200, type internal
  Last update from 172.16.1.1 on Ethernet0/0, 1d02h ago
  Routing Descriptor Blocks:

    192.168.100.1, from 192.168.200.1, 1d02h ago, via Ethernet0/1
      Route metric is 1561600, traffic share count is 47

      ...

  * 172.16.1.1, from 172.16.1.1, 00:00:16 ago, via Ethernet0/0
      Route metric is 307200, traffic share count is 240
```

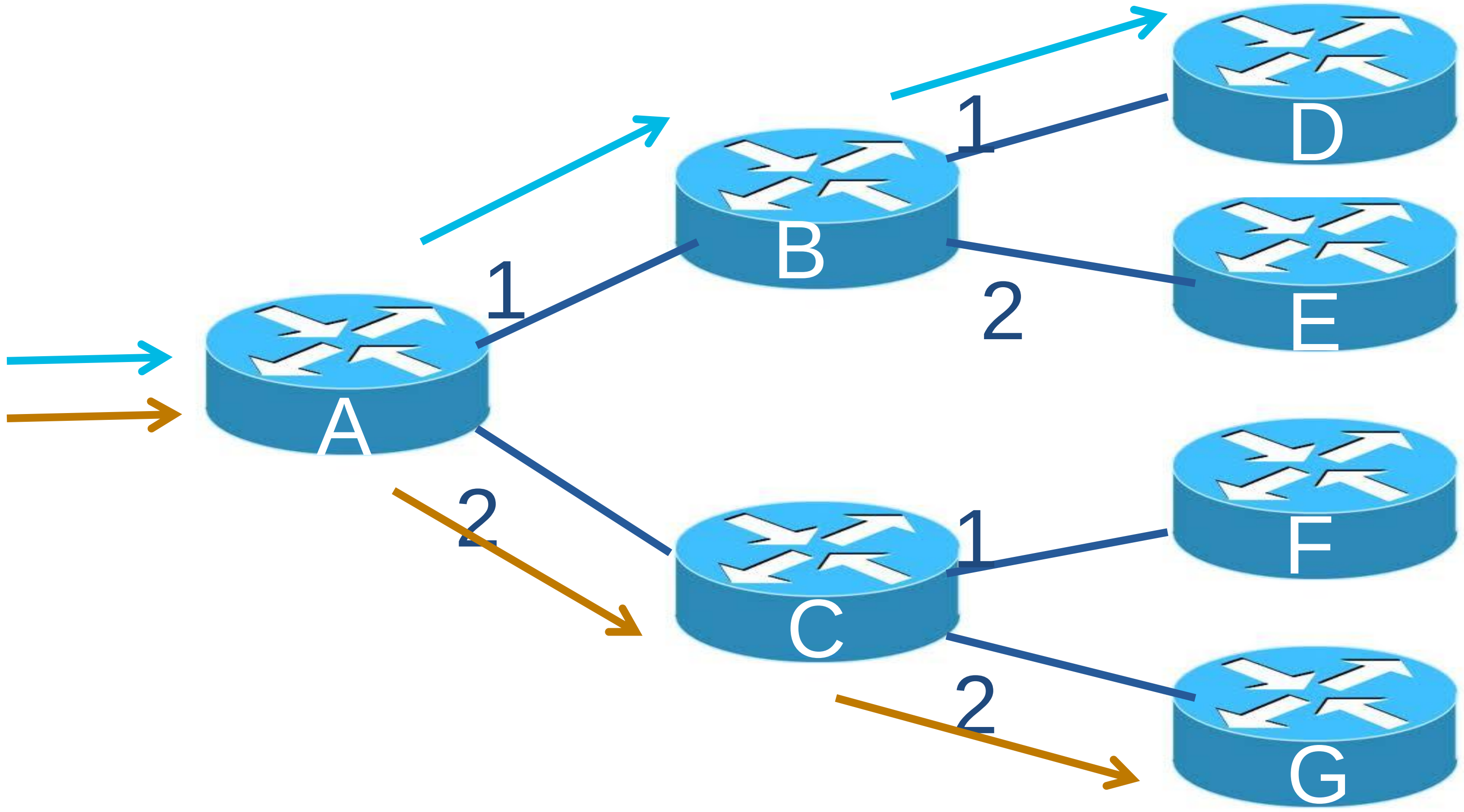
Only accomplished with EIGRP **`variance`**

Equal Cost Multipath

CEF Polarisation

- Hash is deterministic
  - Same input always provides the same output

Packet 1 = src 10.1.1.1 dst 10.2.2.2
Packet 2 = src 10.1.1.1 dst 10.3.3.3



- Without randomisation every router makes the same decision
- Downstream routers never loadshare

Equal Cost Multipath

CEF Hashing Algorithm

- Default hash is “Universal”

  Source IP + Destination IP + Universal Identifier

- Universal ID prevents polarisation
- Other hashes can be used for fixing unequal load sharing

```
RouterB#show cef state
CEF Status:
…
 universal per-destination load sharing algorithm,
id 0F33353C
```

Equal Cost Multipath

CEF Load Sharing Options

- Per-Packet
  - More even load sharing
  - Jitter
  - Out of Order packets (bad for lots of applications)
- Per-Destination (default)
  - Can be less even load sharing
  - Ordered delivery
  - Hashing challenges

Equal Cost Multipath

CEF Hashing

```
RouterB#show ip CEF 172.16.2.1 internal

172.16.2.0/24, epoch 0, RIB[I], refcount 5, per-destination sharing
…
  ifnums:
   Ethernet0/0(3): 172.16.1.1
   Ethernet0/1(4): 192.168.200.1

  path 08172748, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 172.16.1.1 Eth0/0, adj IP adj out Eth0/0, addr 172.16.1.1 081E35A0

  path 08172898, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 192.168.200.1 Eth0/1, adj IP adj out Eth0/1, addr 192.168.200.1 0F75D9F8

  flags: Per-session, for-rx-IPv4, 2buckets
    2 hash buckets

      < 0 > IP adj out of Ethernet0/0, addr 172.16.1.1 081E35A0
      < 1 > IP adj out of Ethernet0/1, addr 192.168.200.1 0F75D9F8
```

Equal Cost Multipath

CEF Hashing

```
RouterB#show ip CEF 172.16.2.1 internal

172.16.2.0/24, epoch 0, RIB[I], refcount 5, per-destination sharing
…
  ifnums:
   Ethernet0/0(3): 172.16.1.1
   Ethernet0/1(4): 192.168.200.1

  path 08172748, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 172.16.1.1 Eth0/0, adj IP adj out Eth0/0, addr 172.16.1.1 081E35A0

  path 08172898, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 192.168.200.1 Eth0/1, adj IP adj out Eth0/1, addr 192.168.200.1 0F75D9F8

  flags: Per-session, for-rx-IPv4, 2buckets
    2 hash buckets

      < 0 > IP adj out of Ethernet0/0, addr 172.16.1.1 081E35A0
      < 1 > IP adj out of Ethernet0/1, addr 192.168.200.1 0F75D9F8
```

Equal Cost Multipath

CEF Hashing

```
RouterB#show ip CEF 172.16.2.1 internal

172.16.2.0/24, epoch 0, RIB[I], refcount 5, per-destination sharing
…
  ifnums:
   Ethernet0/0(3): 172.16.1.1
   Ethernet0/1(4): 192.168.200.1

  path 08172748, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 172.16.1.1 Eth0/0, adj IP adj out Eth0/0, addr 172.16.1.1 081E35A0

  path 08172898, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 192.168.200.1 Eth0/1, adj IP adj out Eth0/1, addr 192.168.200.1 0F75D9F8

  flags: Per-session, for-rx-IPv4, 2buckets
    2 hash buckets

      < 0 > IP adj out of Ethernet0/0, addr 172.16.1.1 081E35A0
      < 1 > IP adj out of Ethernet0/1, addr 192.168.200.1 0F75D9F8
```

Equal Cost Multipath
CEF Hashing

```
RouterB#show ip CEF 172.16.2.1 internal

172.16.2.0/24, epoch 0, RIB[I], refcount 5, per-destination sharing
…
  ifnums:
   Ethernet0/0(3): 172.16.1.1
   Ethernet0/1(4): 192.168.200.1

  path 08172748, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 172.16.1.1 Eth0/0, adj IP adj out Eth0/0, addr 172.16.1.1 081E35A0

  path 08172898, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 192.168.200.1 Eth0/1, adj IP adj out Eth0/1, addr 192.168.200.1 0F75D9F8

  flags: Per-session, for-rx-IPv4, 2buckets
    2 hash buckets

      < 0 > IP adj out of Ethernet0/0, addr 172.16.1.1 081E35A0
      < 1 > IP adj out of Ethernet0/1, addr 192.168.200.1 0F75D9F8
```

Equal Cost Multipath

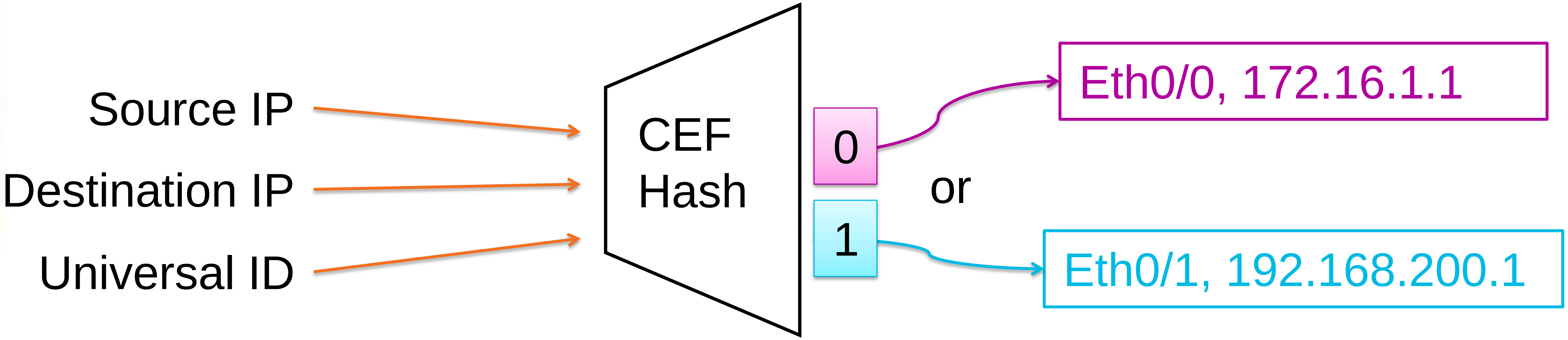
CEF Hashing

```
RouterB#show ip CEF 172.16.2.1 internal

172.16.2.0/24, epoch 0, RIB[I], refcount 5, per-destination sharing
…
  ifnums:
   Ethernet0/0(3): 172.16.1.1
   Ethernet0/1(4): 192.168.200.1

  path 08172748, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 172.16.1.1 Eth0/0, adj IP adj out Eth0/0, addr 172.16.1.1 081E35A0

  path 08172898, path list 100071A8, share 1/1, type attached nexthop, for IPv4
  nexthop 192.168.200.1 Eth0/1, adj IP adj out Eth0/1, addr 192.168.200.1 0F75D9F8

  flags: Per-session, for-rx-IPv4, 2buckets
    2 hash buckets

      < 0 > IP adj out of Ethernet0/0, addr 172.16.1.1 081E35A0
      < 1 > IP adj out of Ethernet0/1, addr 192.168.200.1 0F75D9F8
```

Equal Cost Multipath

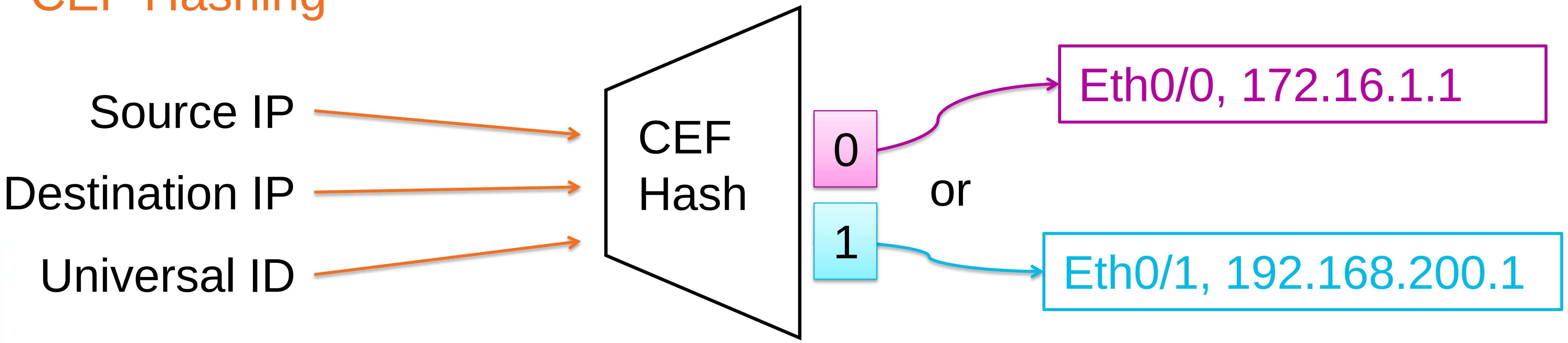
CEF Hashing

```
2 hash buckets
     < 0 > IP adj out Ethernet0/0, addr 172.16.1.1 081E35A0
     < 1 > IP adj out Ethernet0/1, addr 192.168.200.1 0F75D9F8
```



Equal Cost Multipath

CEF Hashing




```
RouterB#show ip CEF exact-route 192.168.2.38 172.16.2.24
192.168.2.38 -> 172.16.2.24 => IP adj out Ethernet0/1, addr 192.168.200.1

RouterB#show ip CEF exact-route 192.168.2.40 172.16.2.24
192.168.2.40 -> 172.16.2.24 => IP adj out Ethernet0/0, addr 172.16.1.1
```

Equal Cost Multipath

Summary

- CEF is built from the routing table
- Load sharing is part of routing decision
- Not 100% equal (60/40)
- Based on Source/Destination IP + Universal ID
- Only one router

How do I load share on more than one router?

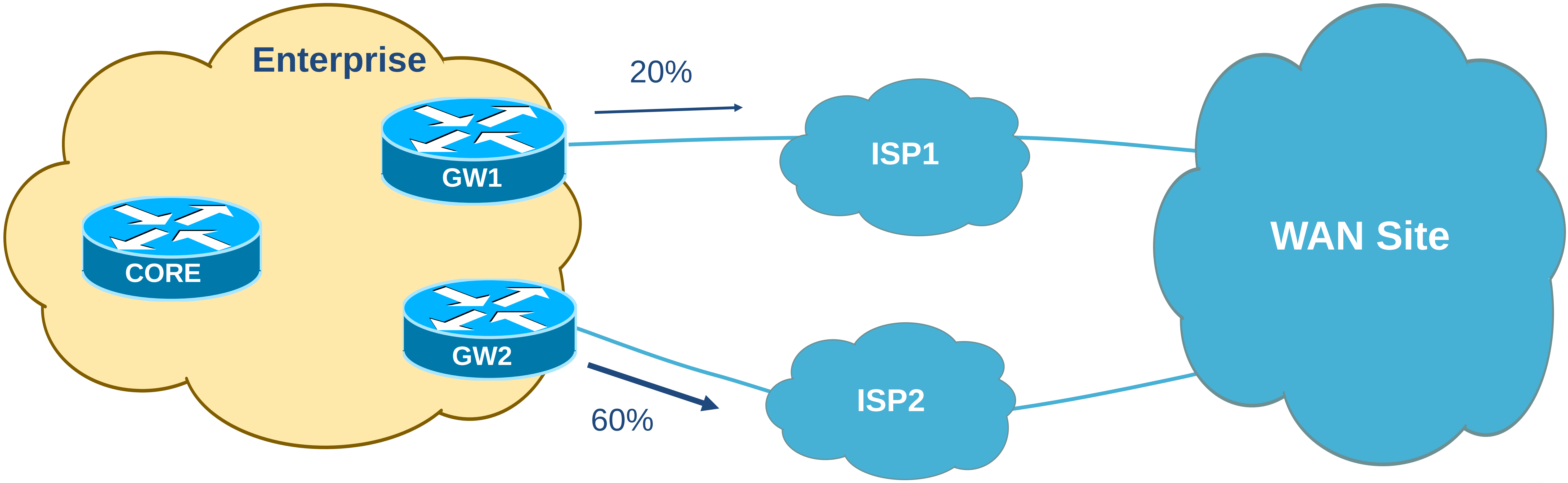
Load Sharing

Outbound Load Sharing with PfR

Load Sharing with PfR

PfR Introduction

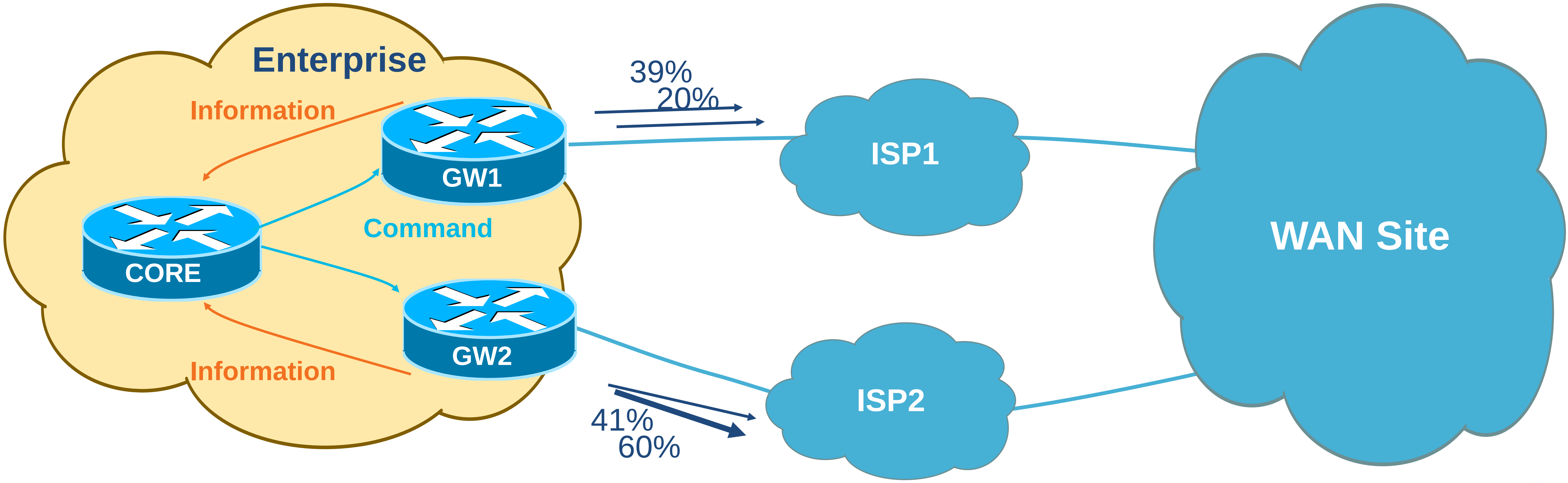
- CEF ECMP works per-router
- No dynamic way to get even sharing across routers



Load Sharing with PfR

PfR Introduction

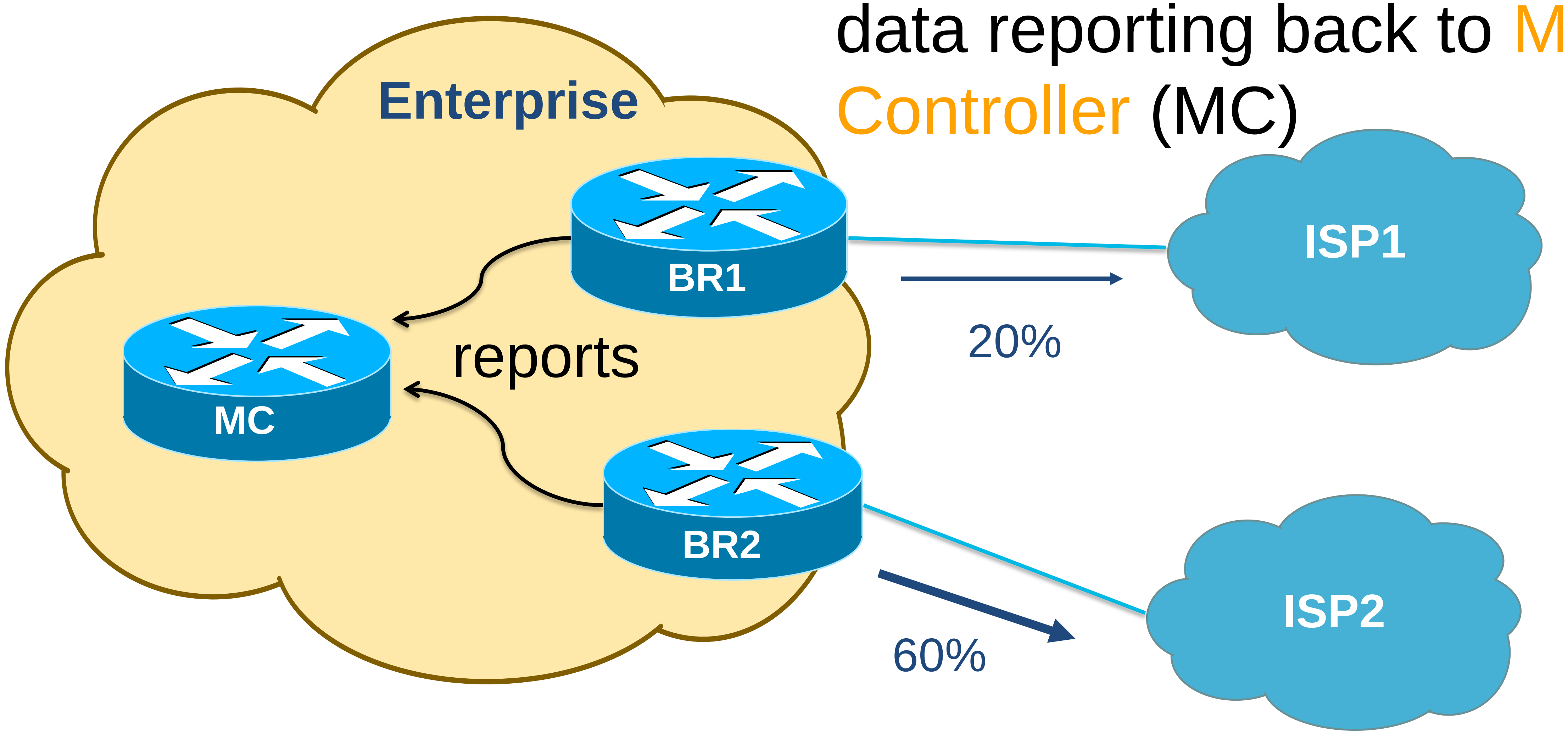
- PfR defines polices and modifies routing
- Allows for dynamic routing changes based on load or delay



Load Sharing with PfR

Operations

Border Routers(BR) provide constant data reporting back to Master Controller (MC)

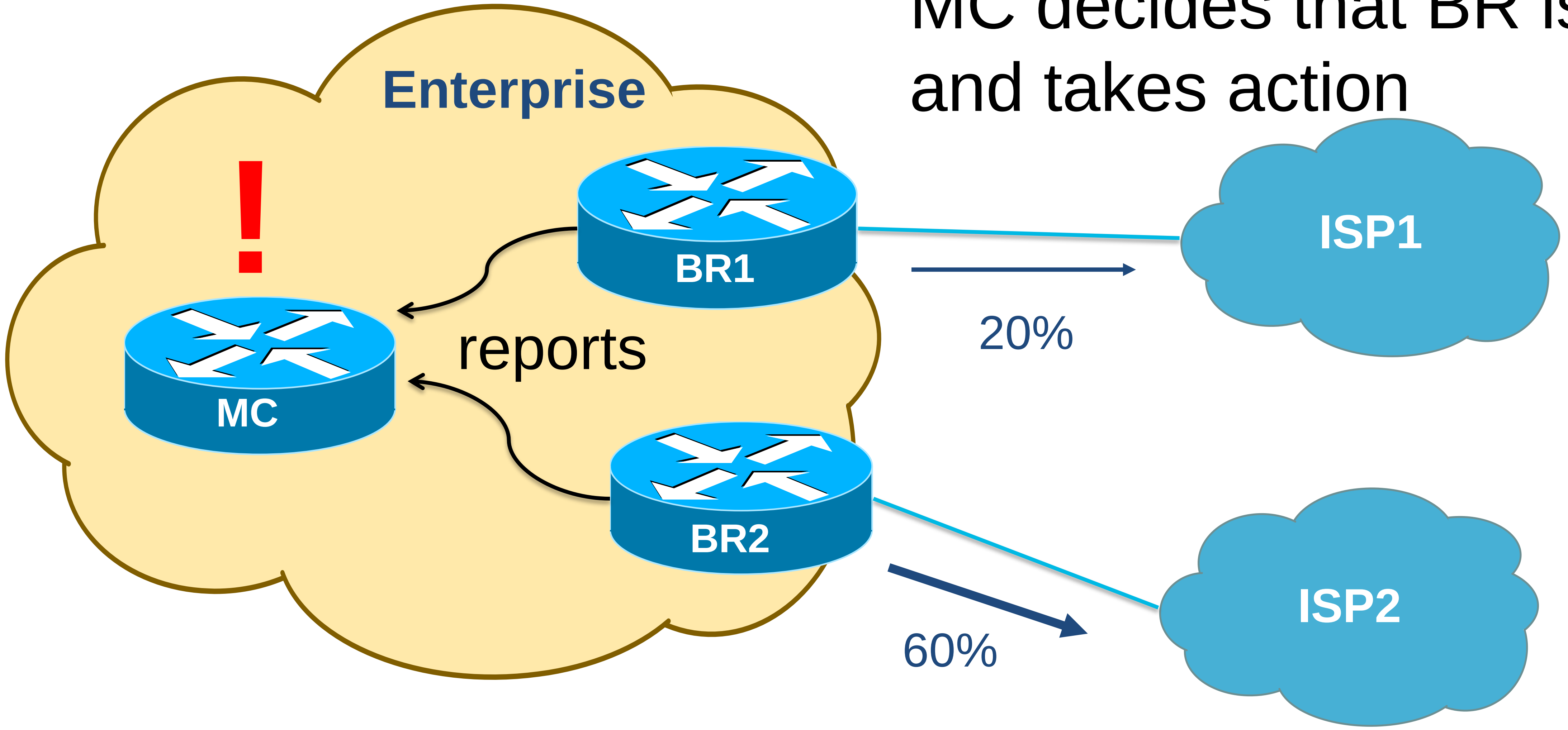


BRs can monitor loss, delay, BW, and more

Load Sharing with PfR

Operations

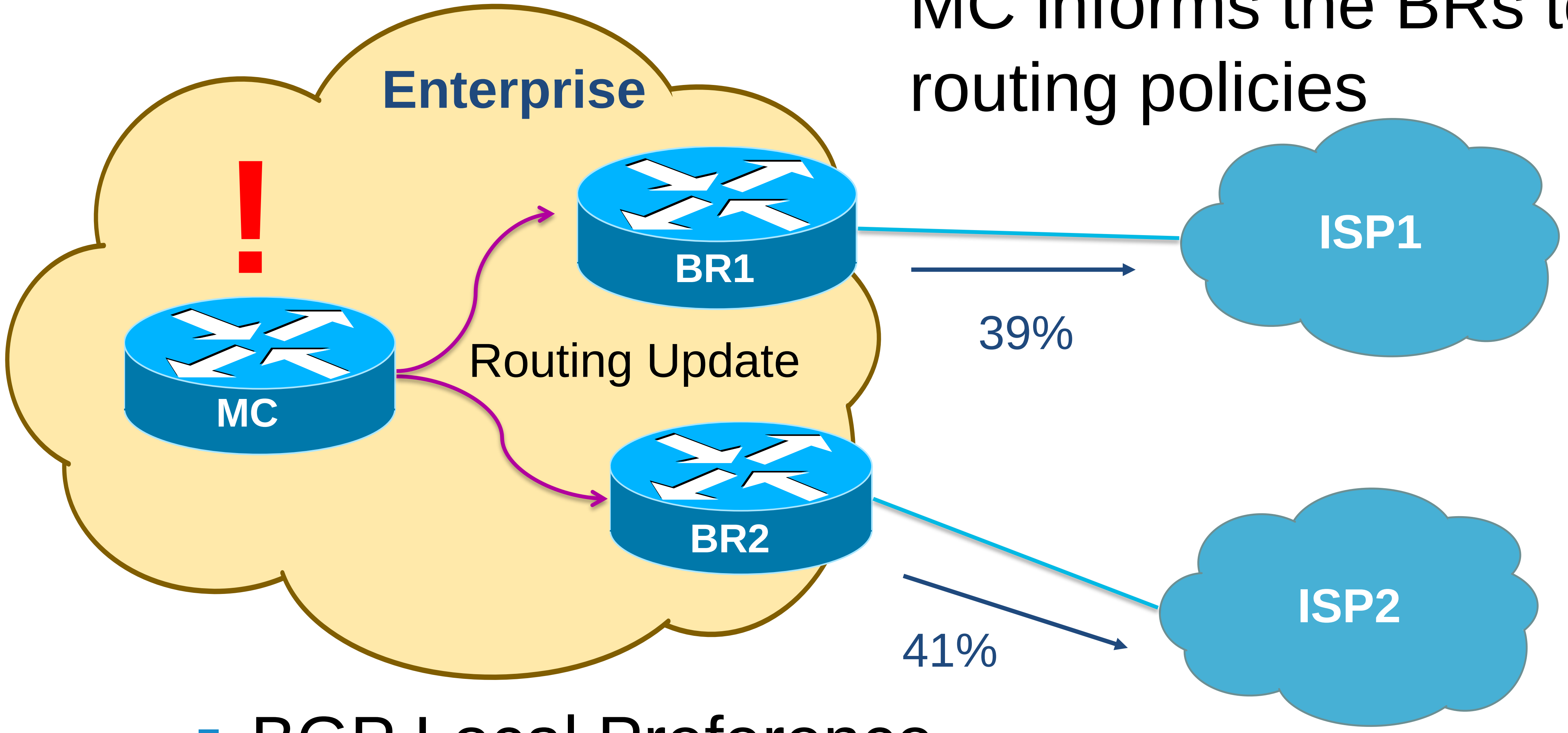
MC decides that BR is out of policy and takes action



Load Sharing with PfR

Operations

MC informs the BRs to change their routing policies

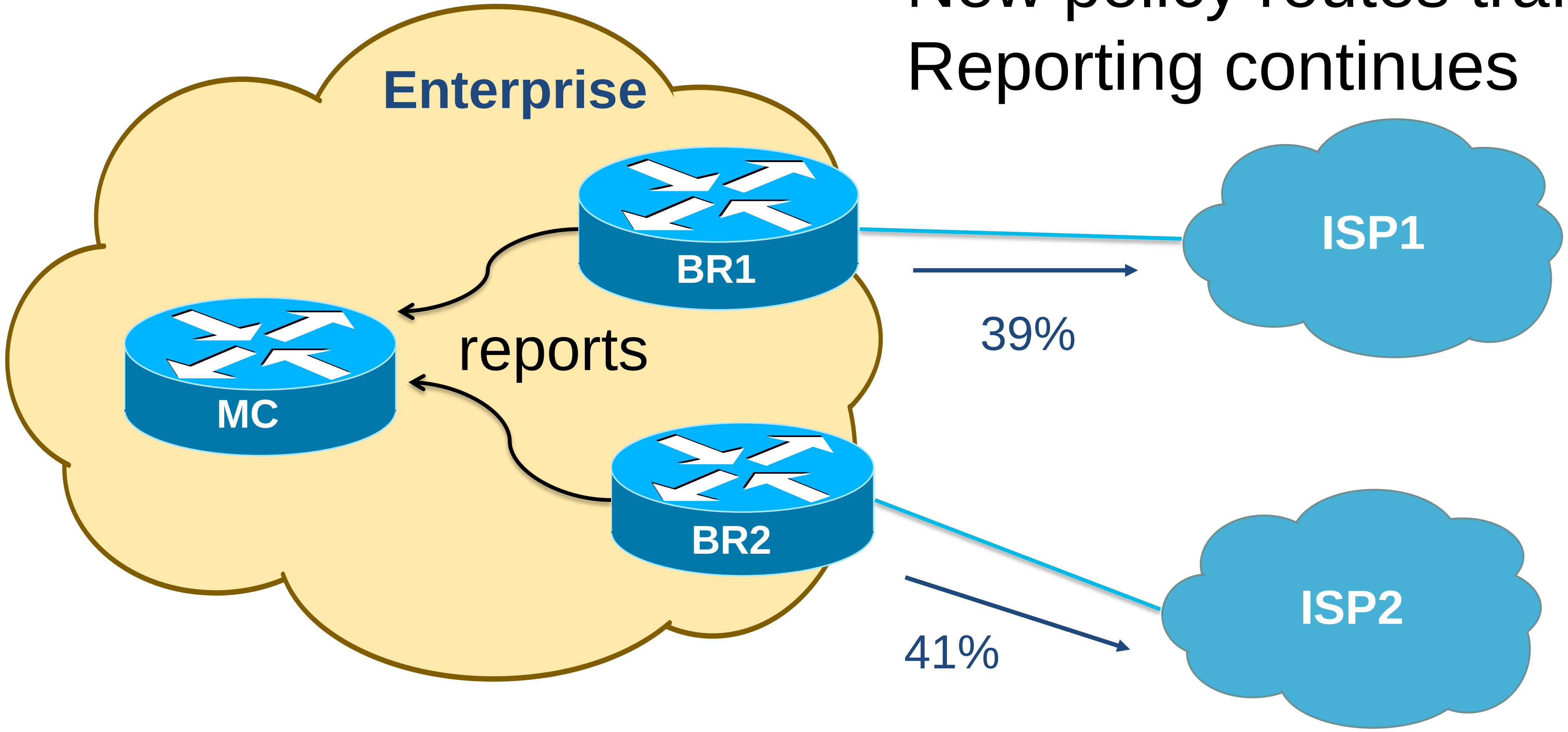


- BGP Local Preference
- Static Route Injection
- PBR

Load Sharing with PfR

Operations

New policy routes traffic to other BR.
Reporting continues



CEF and RIB behaviours are the same!

PfR simply provides more information

Routing Operations in Cisco IOS Routers

Agenda

- Router Components
- Moving Packets
- CEF, CPU and Memory
- Outbound Load Sharing
- Routing Convergence Improvements

Routing Convergence Improvements
What's to Improve?

- Routing changes are **bad**
- Small changes can require (potentially) large recalculation
- Routing Protocols are slow
  - Failure detection is fast
  - Event propagation + calculation is the bottleneck
- Chain Reaction
  - Protocol Change -> RIB Change -> CEF Change
- Protocol can already know what to do before failure

Routing Convergence Improvements

Failure Detection with BFD

- Bidirectional Forwarding Detection
- **VERY** fast (50ms hello/150ms dead)
- Lightweight
  – 24 bytes BFD Hello vs. 56 byte OSPF Hello
- Handled in Interrupt
- Protocols are BFD clients
- Offloaded to hardware* (HWO)

*12k, 7600 with ES+

Routing Convergence Improvements

OSPF Loop Free Alternate

TOMORROW starts here.

Routing Convergence Improvements
OSPF Overview

- Link State Algorithm
  - – LSDB provides a view of the entire network
- Network changes exchanged via LSA (Link State Advertisement)
  - – Multiple events cause throttling (5000ms default)
- SPF algorithm determines best path
  - – Runs on receipt of LSA, delayed 5000ms (default)

Routing Convergence Improvements

OSPF Convergence Times

- Convergence =

  Failure Detection + Event Propagation + SPF + FIB Update

  Failure Detection ↑ Neighbour Down; Event Propagation ↑ LSA generation; FIB Update ↑ RIB + CEF + Hardware

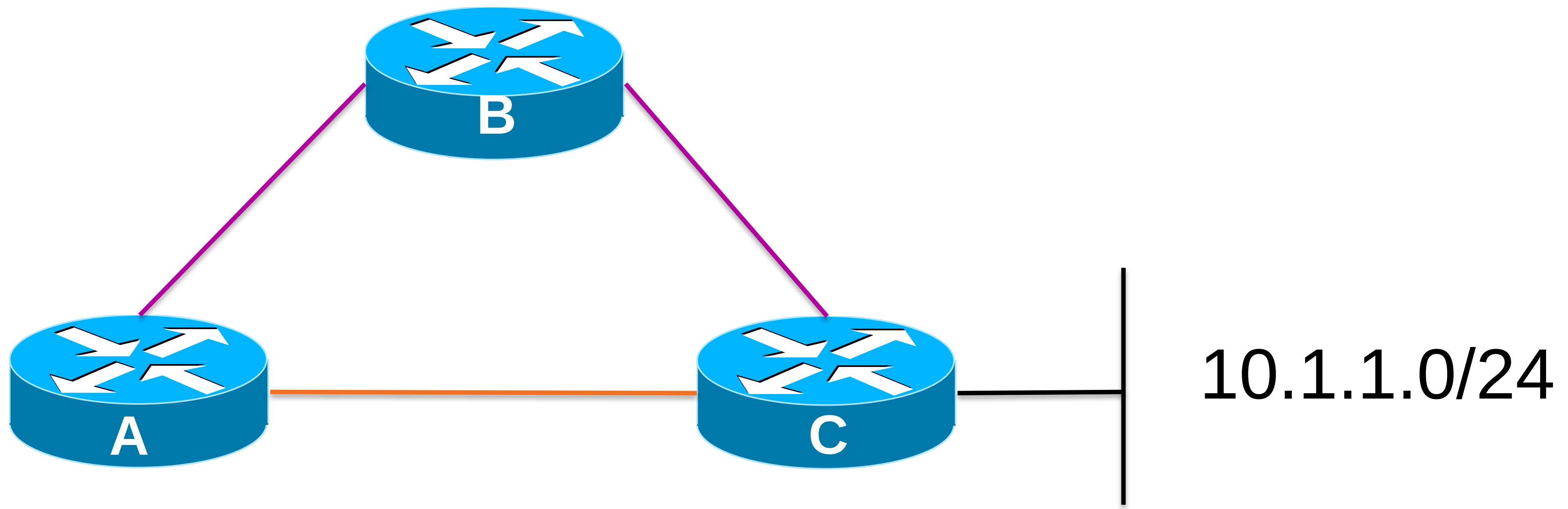
- Best case: ~160ms (SPF Tuning + BFD)
- Worst case: ~50 seconds

  (Dead Time + LSA throttle + SPF defaults)

- Failure Detection is **easy** (hardware)
- Control plane is **difficult** (software)

Routing Convergence Improvement

OSPF Loop Free Alternate



- A has a primary (A-C) and secondary (A-B-C) path to 10.1.1.0/24
- Link State allows A to know entire topology
- A should know that B is an alternative path
- **Loop Free Alternate** (LFA)

Routing Convergence Improvements

OSPF Loop Free Alternate

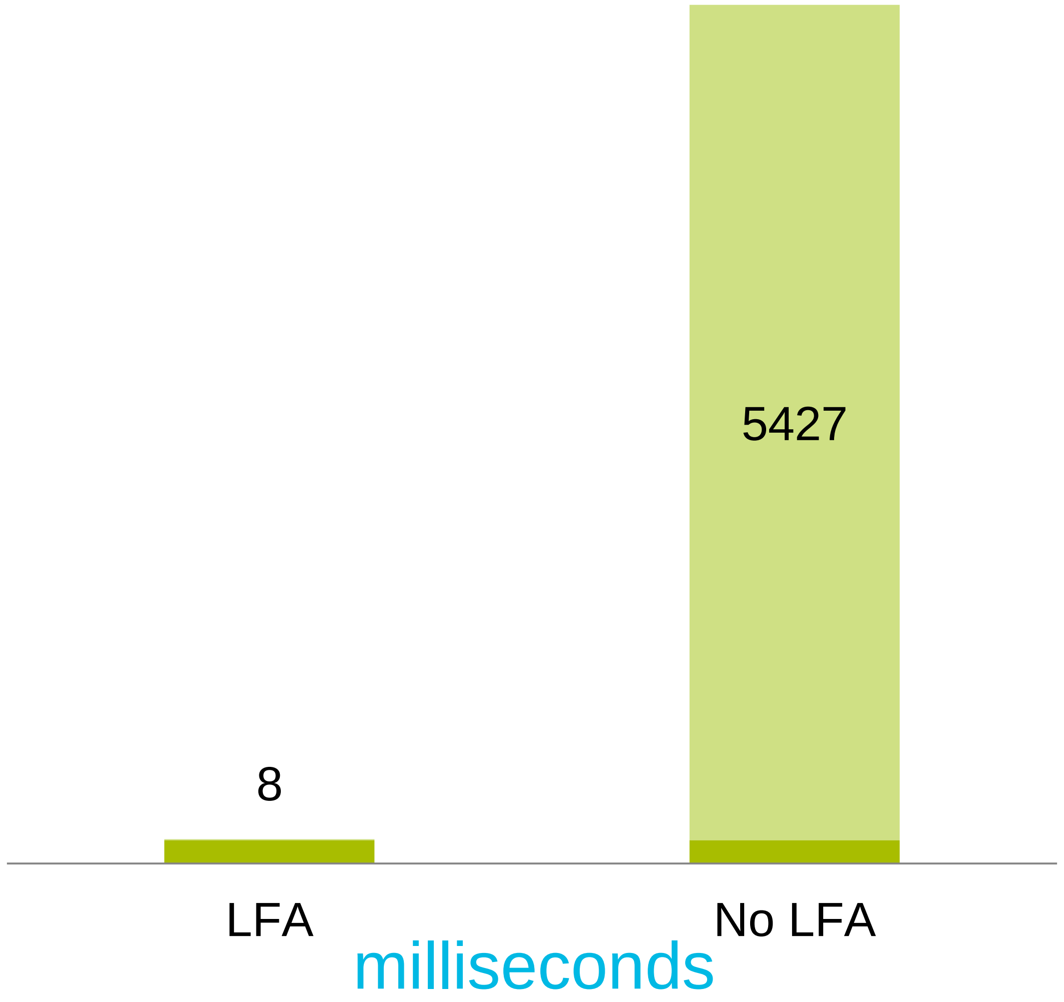
- OSPF presents a primary and backup to CEF
  - Backup calculated from secondary SPF run

```
RouterA# show ip route 10.1.1.0
Routing Descriptor Blocks:
 * 172.16.0.1, from 192.168.255.1, 00:01:57 ago, via Ethernet4/1/0
      Route metric is 2, traffic share count is 1
      Repair Path: 192.168.0.2, via Ethernet4/2/0

RouterA#show ip CEF 10.1.1.0
10.1.1.0/24
  nexthop 172.16.0.1 Ethernet4/1/0
    repair: attached-nexthop 192.168.0.2 Ethernet4/2/0
```

Routing Convergence Improvements

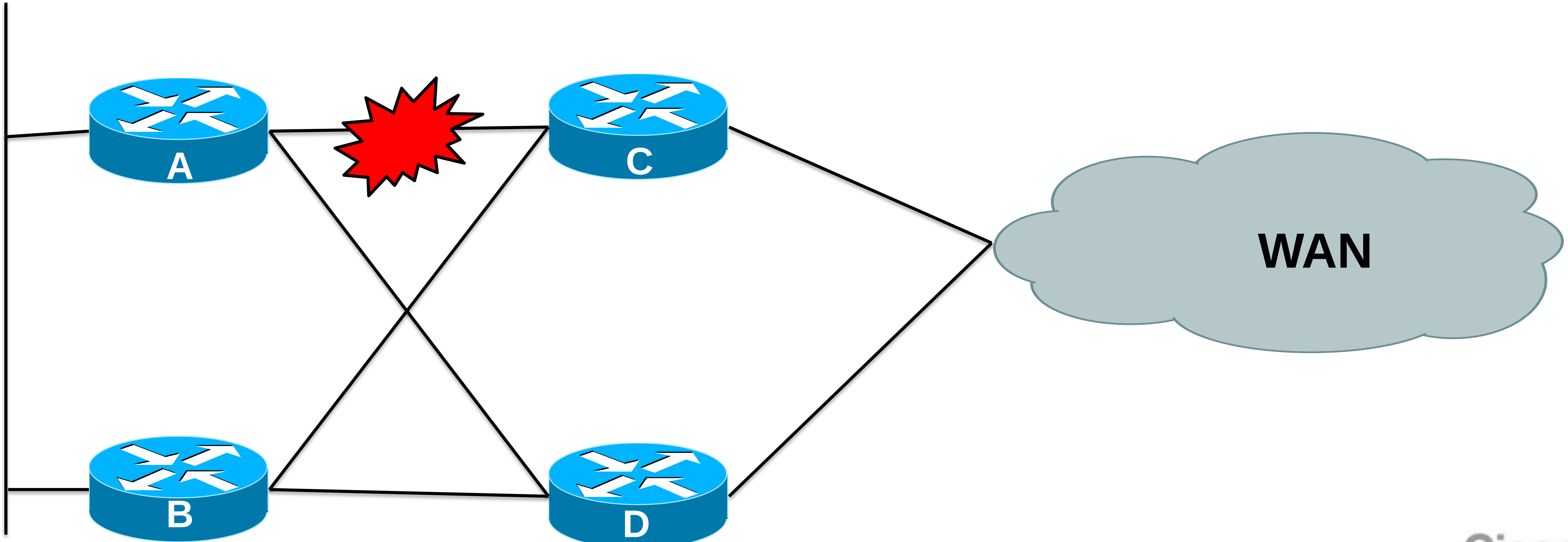
OSPF Loop Free Alternate

- Aims for <50ms reconvergence
- Triggers as soon as the failure is detected
  - NO fast hellos
  - Use BFD!
- Added to 7600/ASR1000 in 15.1(3)S
  - Not enabled by default



Routing Convergence Improvements
OSPF Loop Free Alternate

- Fast failure detection is key!
- Single Box
- Not a replacement for SPF Tuning



Routing Convergence Improvements

EIGRP Feasible Successor

TOMORROW starts here.

Routing Convergence Improvements

EIGRP Overview

- Distance Vector Protocol
  - Doesn't see the entire network like OSPF
- Based on QUERY and ACK messages for convergence
  - QUERY sent to determine best path for failed route
  - ACK sent when alternative path found or no other paths
- DUAL algorithm determines best path
  - Runs as soon as all outstanding QUERIES are received
- Query domain size can effect convergence time

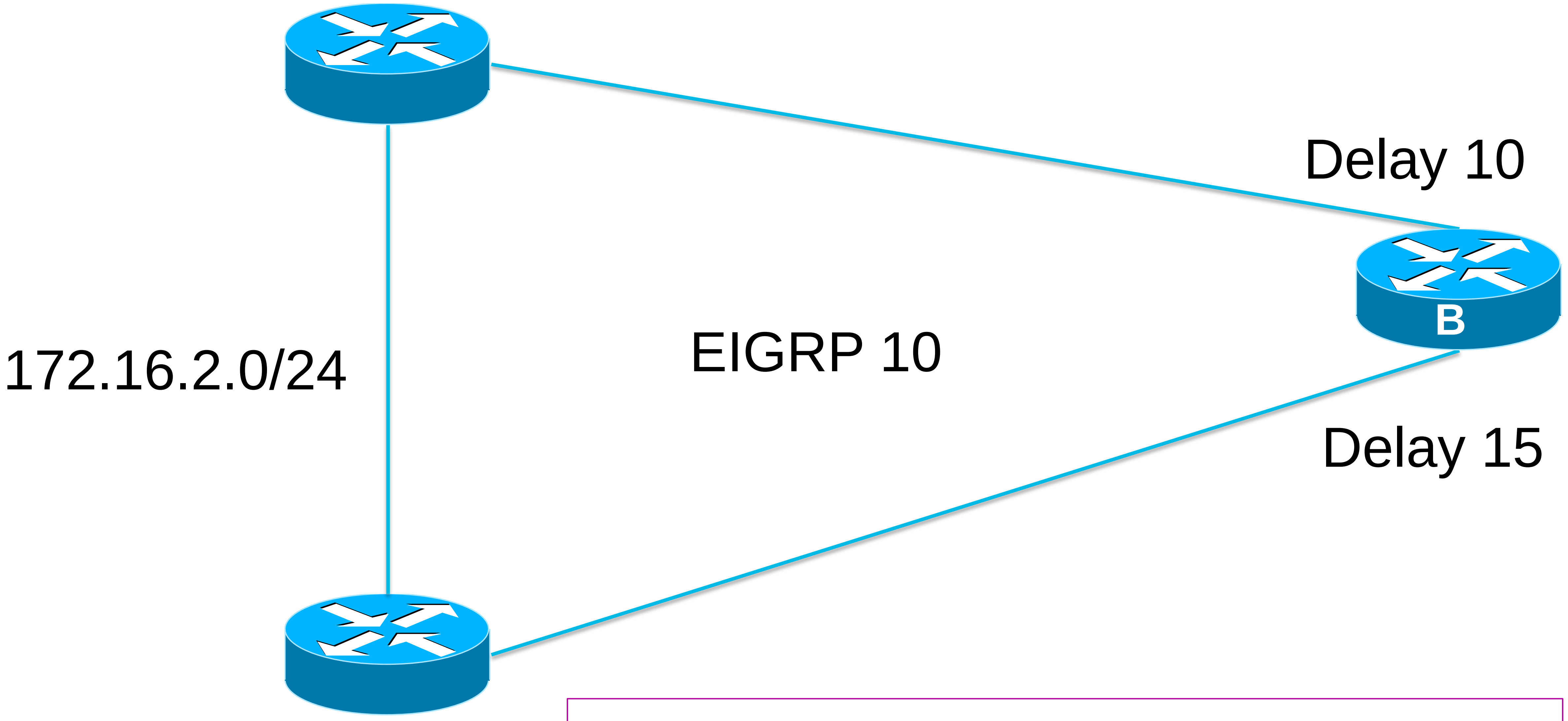
Routing Convergence Improvements

EIGRP Feasible Successors

- EIGRP selects Successor and Feasible Successor (FS)
- Successor is the best route
- FS is 2nd best route
- Must be mathematically loop-free (meets feasibility condition)
- FS acts as a “backup route”
- Kept in topology table (not routing table)
- Up to 6 Feasible Successors
- Built into the protocol, nothing to enable

Routing Convergence Improvements

EIGRP Feasible Successors

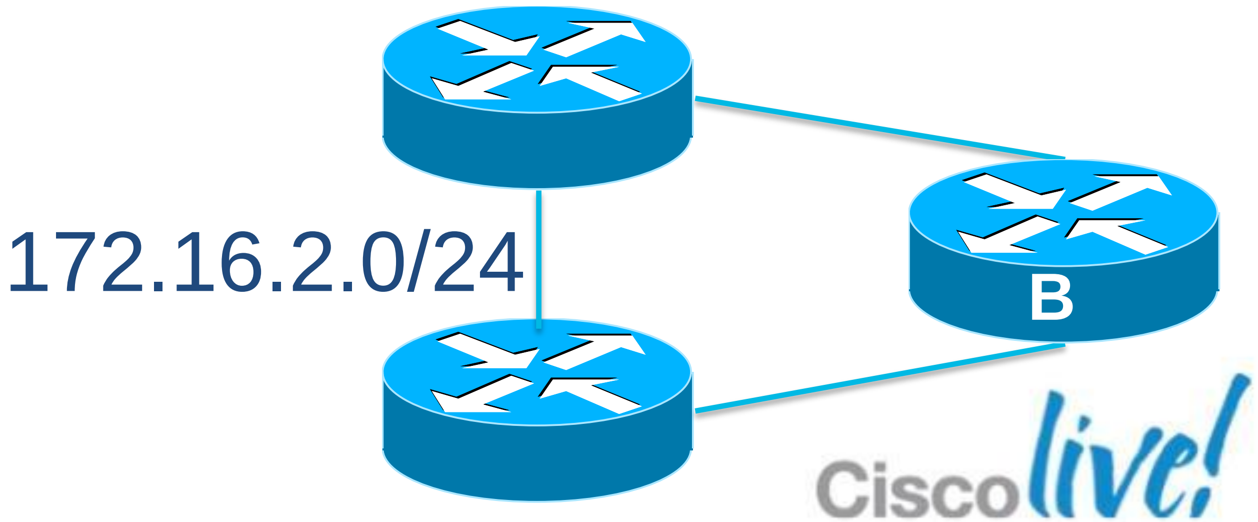


Metric based on bandwidth and delay

Routing Convergence Improvements

EIGRP Feasible Successors

```
RouterB# show ip route 172.16.2.0
Routing entry for 172.16.2.0/24
  Known via "eigrp 10", distance 90, metric 285440, type internal
  Routing Descriptor Blocks:
  * 192.168.200.1, from 192.168.200.1, 00:34:19 ago, via Eth0/1
      Route metric is 285440, traffic share count is 1
```



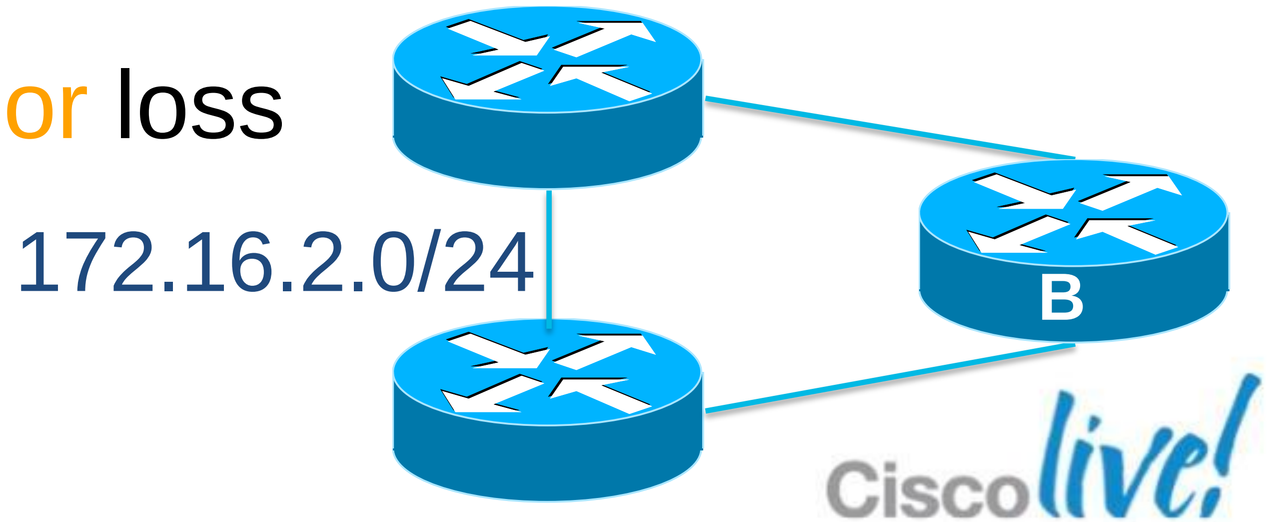
Routing Convergence Improvements

EIGRP Feasible Successors

```
RouterB#show ip eigrp topology
P 172.16.2.0/24, 1 successors, FD is 285440
        via 192.168.200.1 (285440/281600), Ethernet0/1
        via 172.16.1.1 (307200/281600), Ethernet0/0
```

Feasible Successor reported distance (281600)
is less than Successor feasible distance (285440)

- Feasibility Condition met
- Instant convergence after Successor loss



LFA-like repair path coming in 15.2(4)S/15.2(4)M

Routing Convergence Improvements

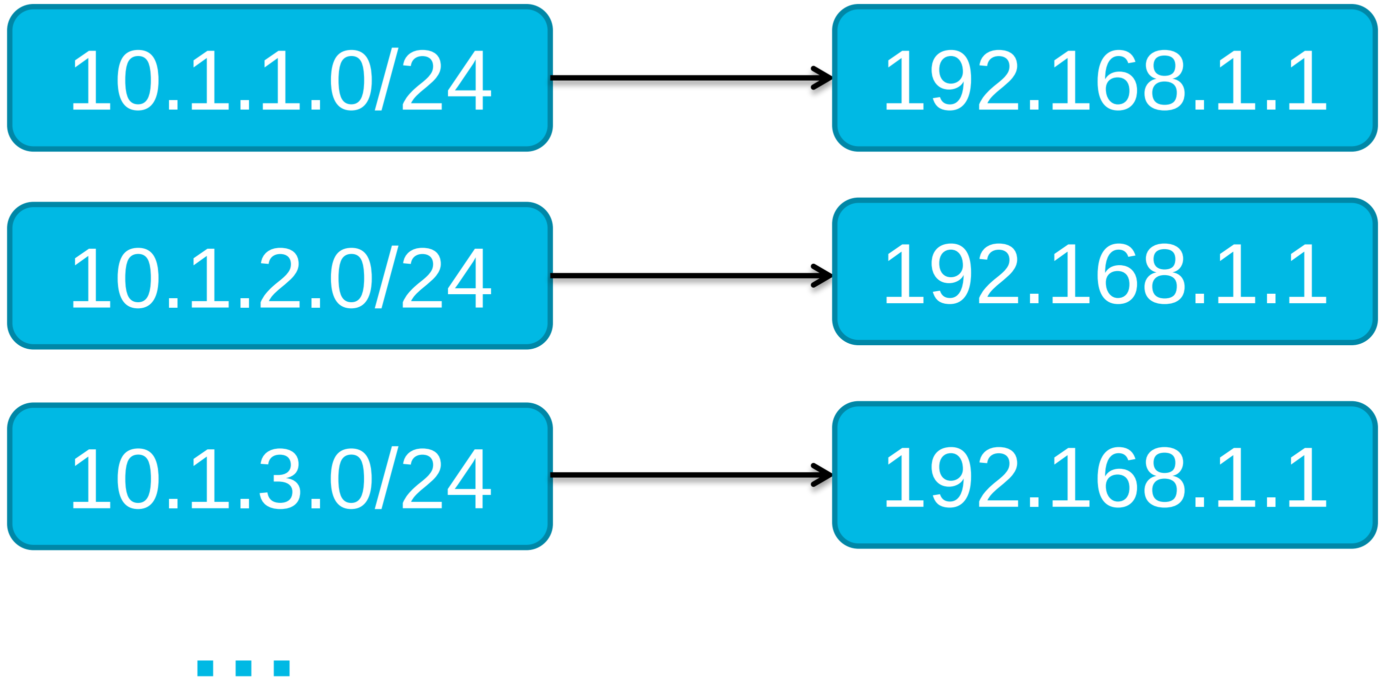
BGP Prefix Independent Convergence (Core)

TOMORROW starts here.

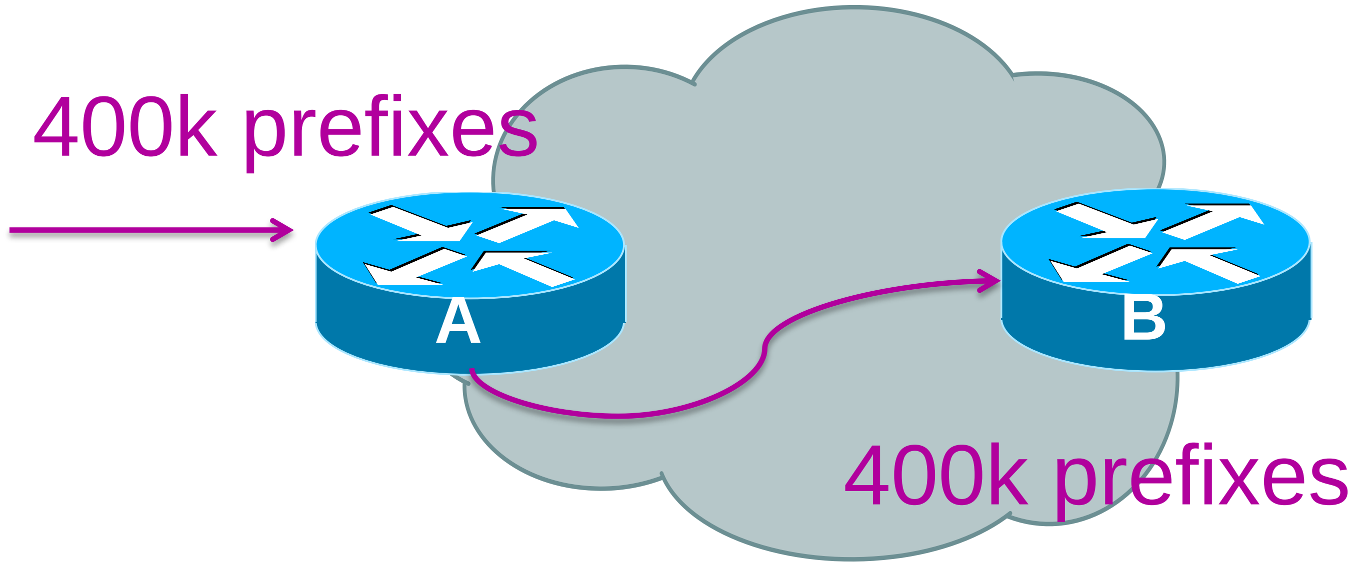
Routing Convergence Improvements

BGP Prefix Independent Convergence

- Today's RIB is flat



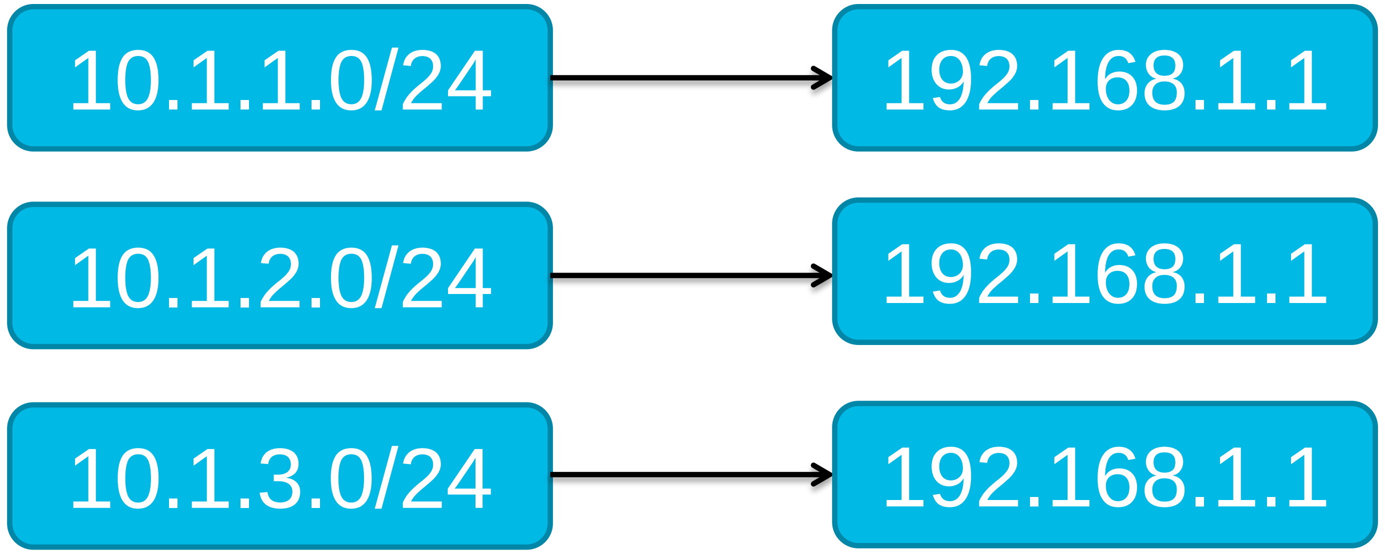
- 400k routes -> 400k updates
- BGP often has same next hop
- We can do better!



Routing Convergence Improvements

BGP Prefix Independent Convergence

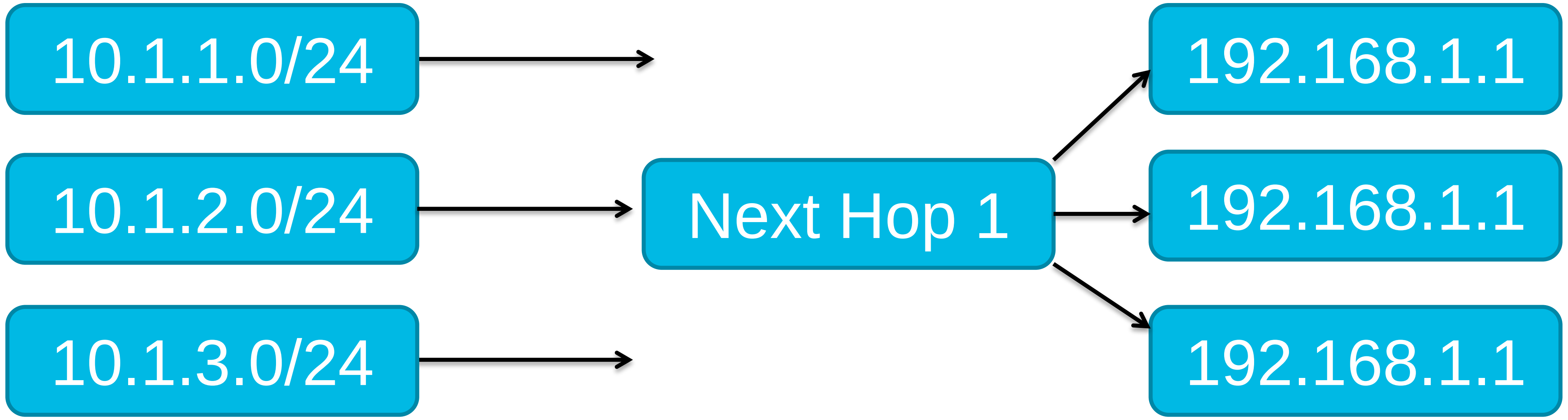
- Instead of flat FIB, Hierarchical



Routing Convergence Improvements

BGP Prefix Independent Convergence

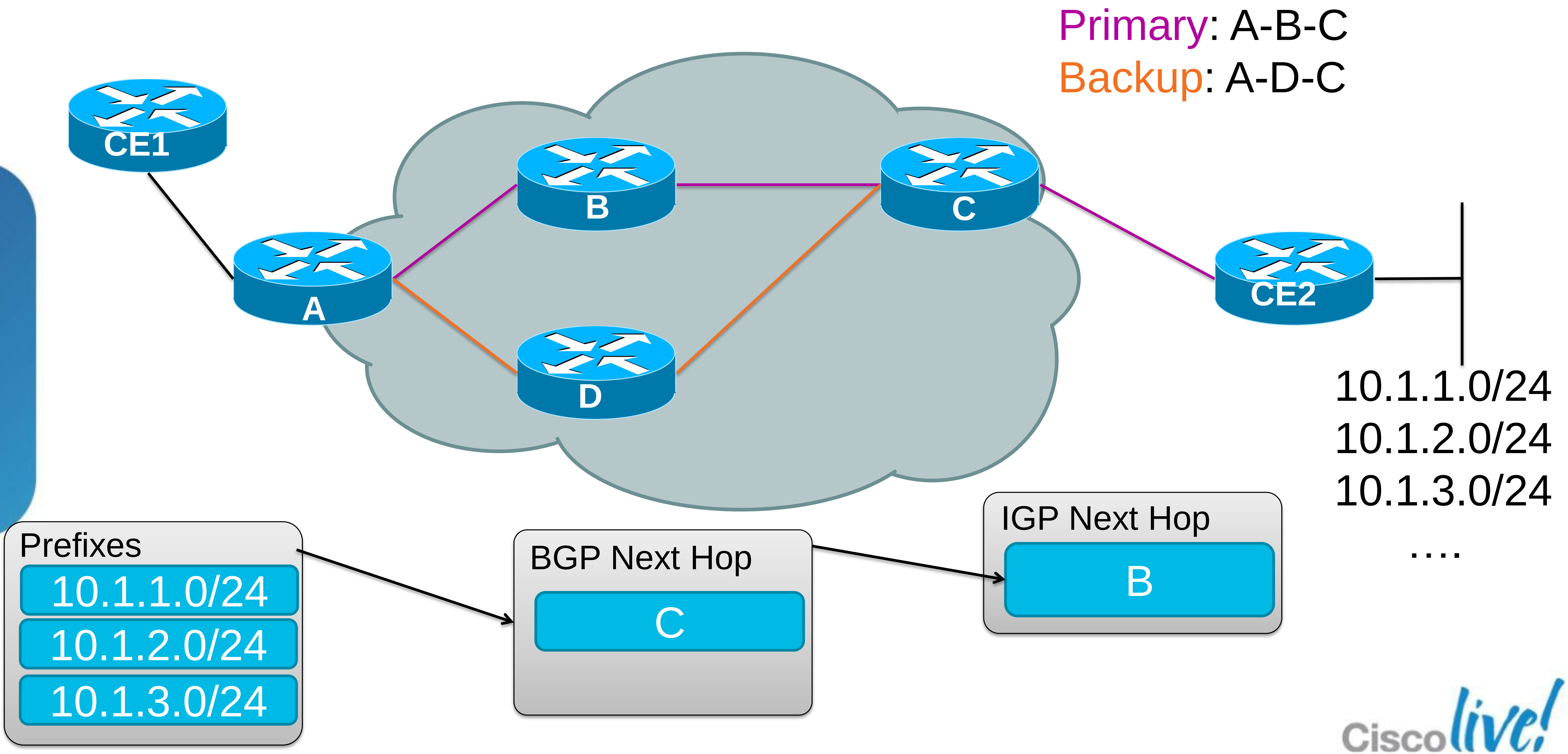
- Instead of flat FIB, Hierarchical



- Single change updates multiple entries
- Convergence time independent from prefix count

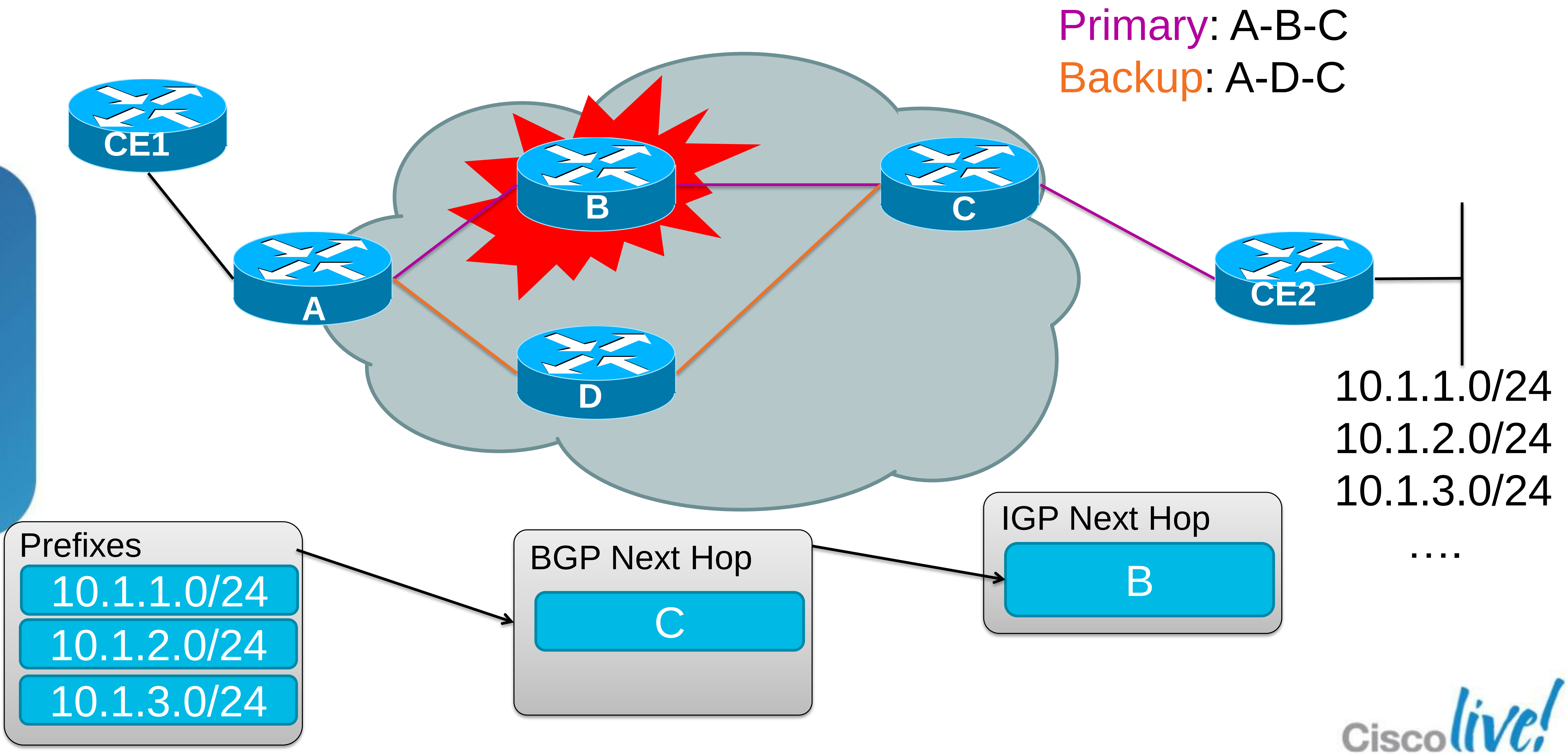
Routing Convergence Improvements

BGP PIC Core



Routing Convergence Improvements

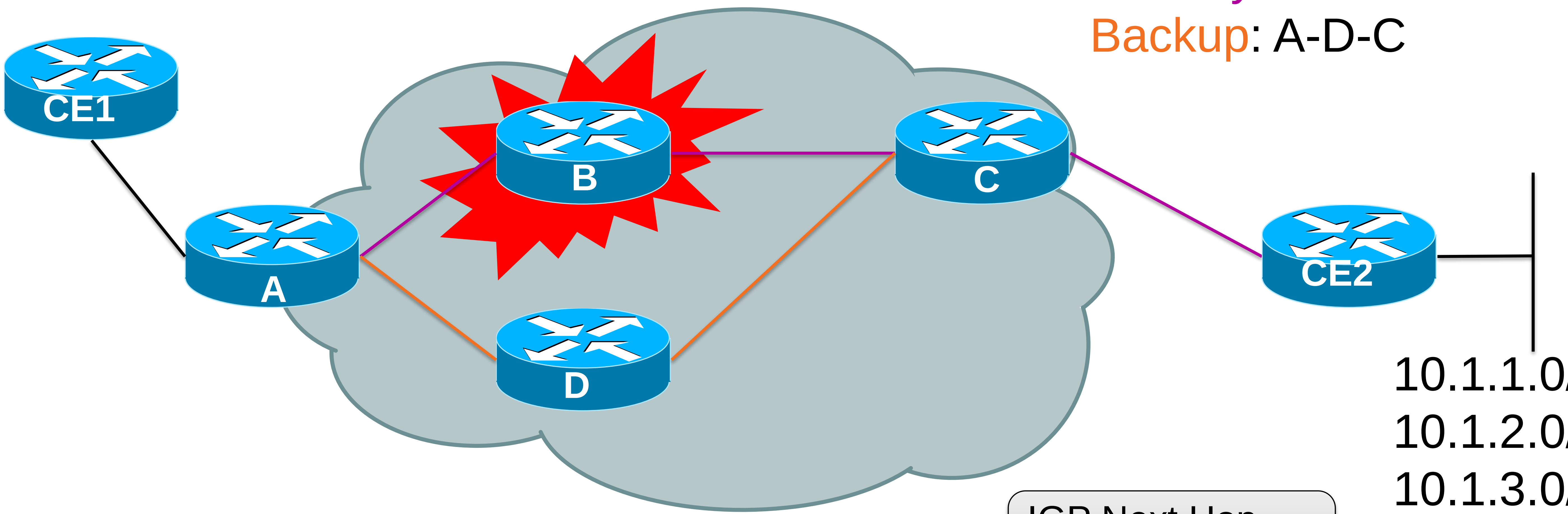
BGP PIC Core



Routing Convergence Improvements

BGP PIC Core

Primary: A-B-C
Backup: A-D-C



10.1.1.0/24
10.1.2.0/24
10.1.3.0/24
….

Prefixes
- 10.1.1.0/24
- 10.1.2.0/24
- 10.1.3.0/24

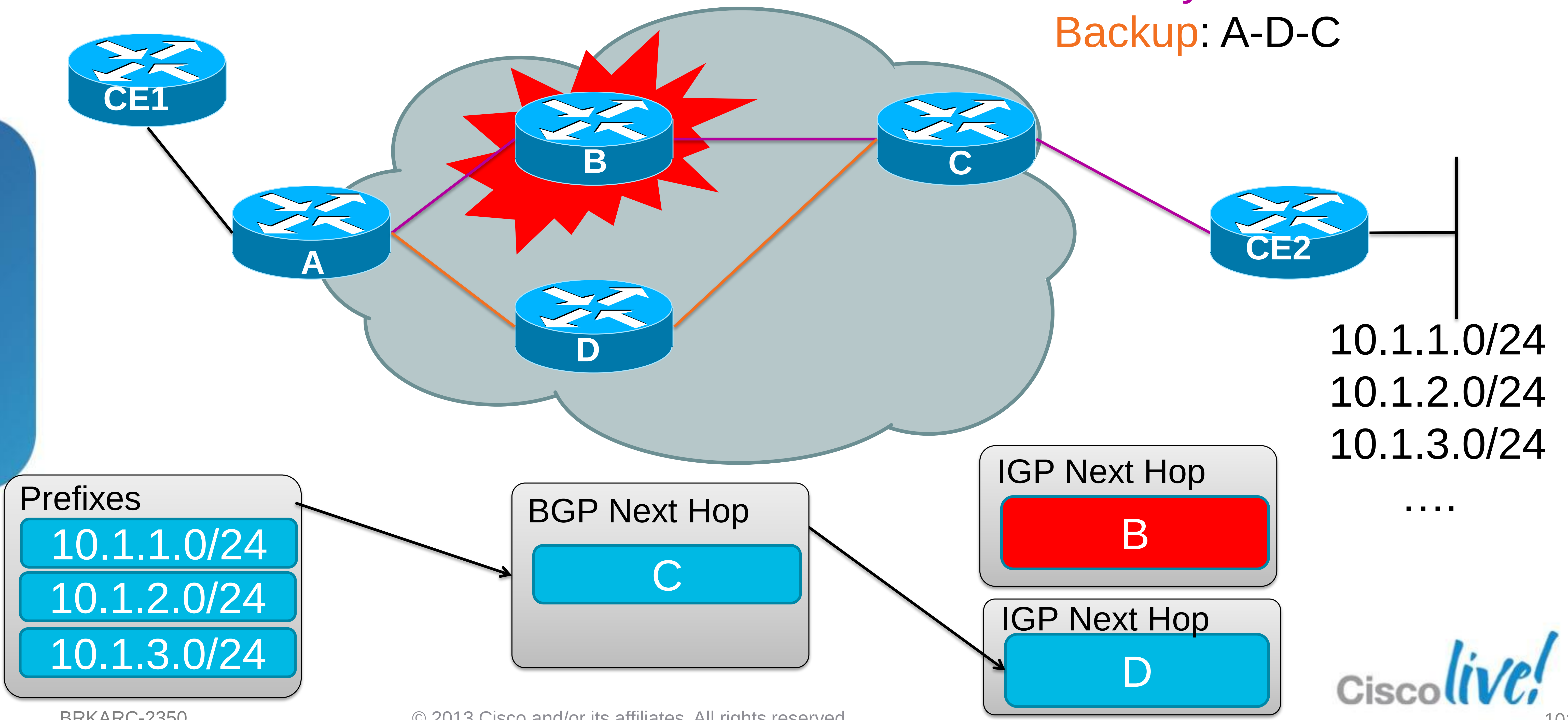
BGP Next Hop
- C

IGP Next Hop
- B

Routing Convergence Improvements

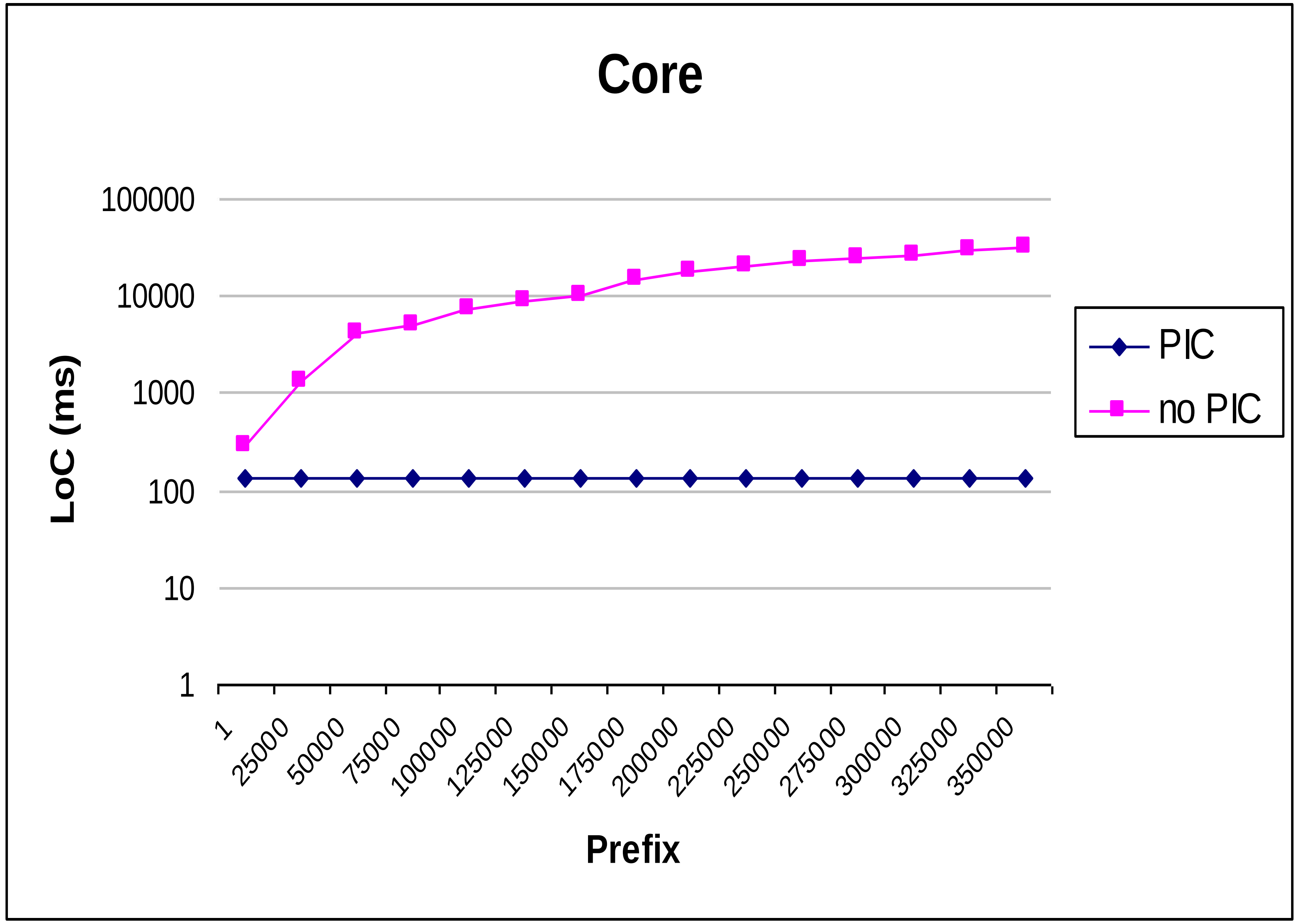
BGP PIC Core

Primary: A-B-C
Backup: A-D-C



Routing Convergence Improvements

BGP PIC Core



- BGP convergences starts after IGP convergence

Routing Convergence Improvements
BGP PIC Core Recap

- PIC Core part of migration to hierarchical FIB
- Still requires SPF convergence
  - OSPF LFA
  - EIGRP FS
- PIC Edge
  - Mainly for MPLS/VPN environments
  - Fast convergence for edge node failures
  - Beyond the scope of today's talk

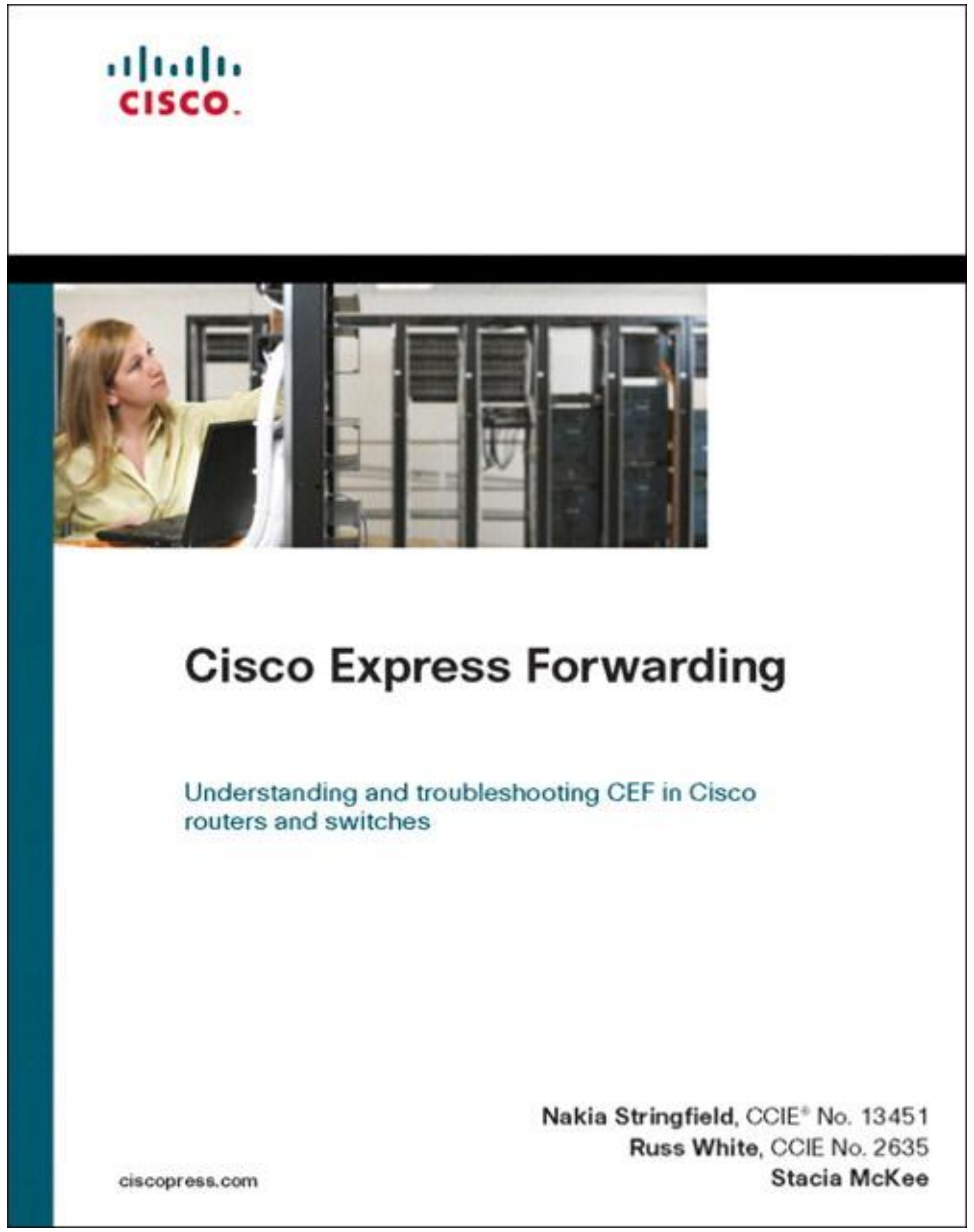
Review

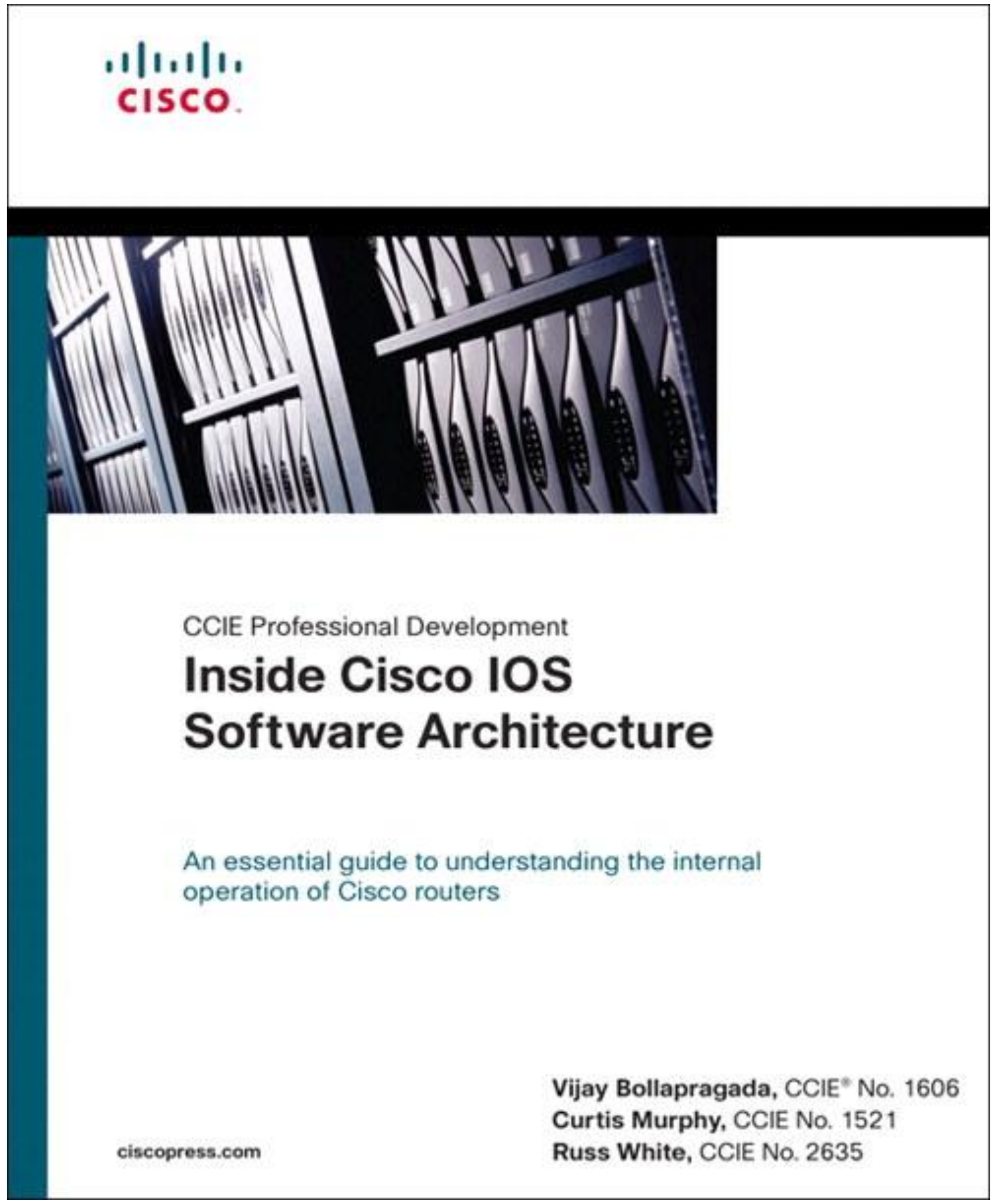
TOMORROW starts here.

Review

- Router Components
  - Control vs. Data plane
  - Software vs. Hardware based routers
- CPU and Memory
  - Interrupt (CEF) vs. Process (Routing Protocol)
  - Memory concerns for multiple routes
- Load Sharing
  - CEF and PfR
- Routing Enhancements
  - OSPF LFA/EIGRP Feasible Successors/BGP PIC

Further Reading





Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2013 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 8 March 12:00pm-2:00pm

Cisco live! 365

Don't forget to activate your Cisco Live 365 account for access to all session material, communities, and on-demand and live activities throughout the year. Log into your Cisco Live portal and click the "Enter Cisco Live 365" button.
www.ciscoliveaustralia.com/portal/login.ww

Appendix

Reading and Populating the Routing Table

The Routing Table

TOMORROW starts here.

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area


O     10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
      172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C        172.16.1.0/24 is directly connected, Ethernet0/0
L        172.16.1.2/32 is directly connected, Ethernet0/0
O        172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
      192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.200.0/24 is directly connected, Ethernet0/1
L        192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area


O     10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
      172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C        172.16.1.0/24 is directly connected, Ethernet0/0
L        172.16.1.2/32 is directly connected, Ethernet0/0
O        172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
      192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.200.0/24 is directly connected, Ethernet0/1
L        192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area



O     10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
      172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C        172.16.1.0/24 is directly connected, Ethernet0/0
L        172.16.1.2/32 is directly connected, Ethernet0/0
O        172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
      192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.200.0/24 is directly connected, Ethernet0/1
L        192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area


O      10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
       172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C         172.16.1.0/24 is directly connected, Ethernet0/0
L         172.16.1.2/32 is directly connected, Ethernet0/0
O         172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
       192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C         192.168.200.0/24 is directly connected, Ethernet0/1
L         192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
```

Classful Network

```
O       10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
        172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C          172.16.1.0/24 is directly connected, Ethernet0/0
L          172.16.1.2/32 is directly connected, Ethernet0/0
O          172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
        192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C          192.168.200.0/24 is directly connected, Ethernet0/1
L          192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area

O     10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
      172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C        172.16.1.0/24 is directly connected, Ethernet0/0
L        172.16.1.2/32 is directly connected, Ethernet0/0
O        172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
      192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.200.0/24 is directly connected, Ethernet0/1
L        192.168.200.2/32 is directly connected, Ethernet0/1
```

Subnetted Network

Reading the Routing Table

```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
```

Subnet

```
O     10.0.0.0/8 [110/20] via 172.16.1.1, 00:01:58, Ethernet0/0
      172.16.0.0/16 is variably subnetted, 4 subnets, 3 masks
C        172.16.1.0/24 is directly connected, Ethernet0/0
L        172.16.1.2/32 is directly connected, Ethernet0/0
O        172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
      192.168.200.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.200.0/24 is directly connected, Ethernet0/1
L        192.168.200.2/32 is directly connected, Ethernet0/1
```

Reading a Route Entry

```
O      172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
```

- 110 – Administrative Distance
  - – Locally significant

Reading a Route Entry

```
O    172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
```

- 110 – Administrative Distance
  - – Locally significant
  - – Lower is better (more trusted)
- 11 – Routing protocol metric
  - – Determined by the routing protocol
  - – Metrics of different protocols are not comparable

Reading a Route Entry

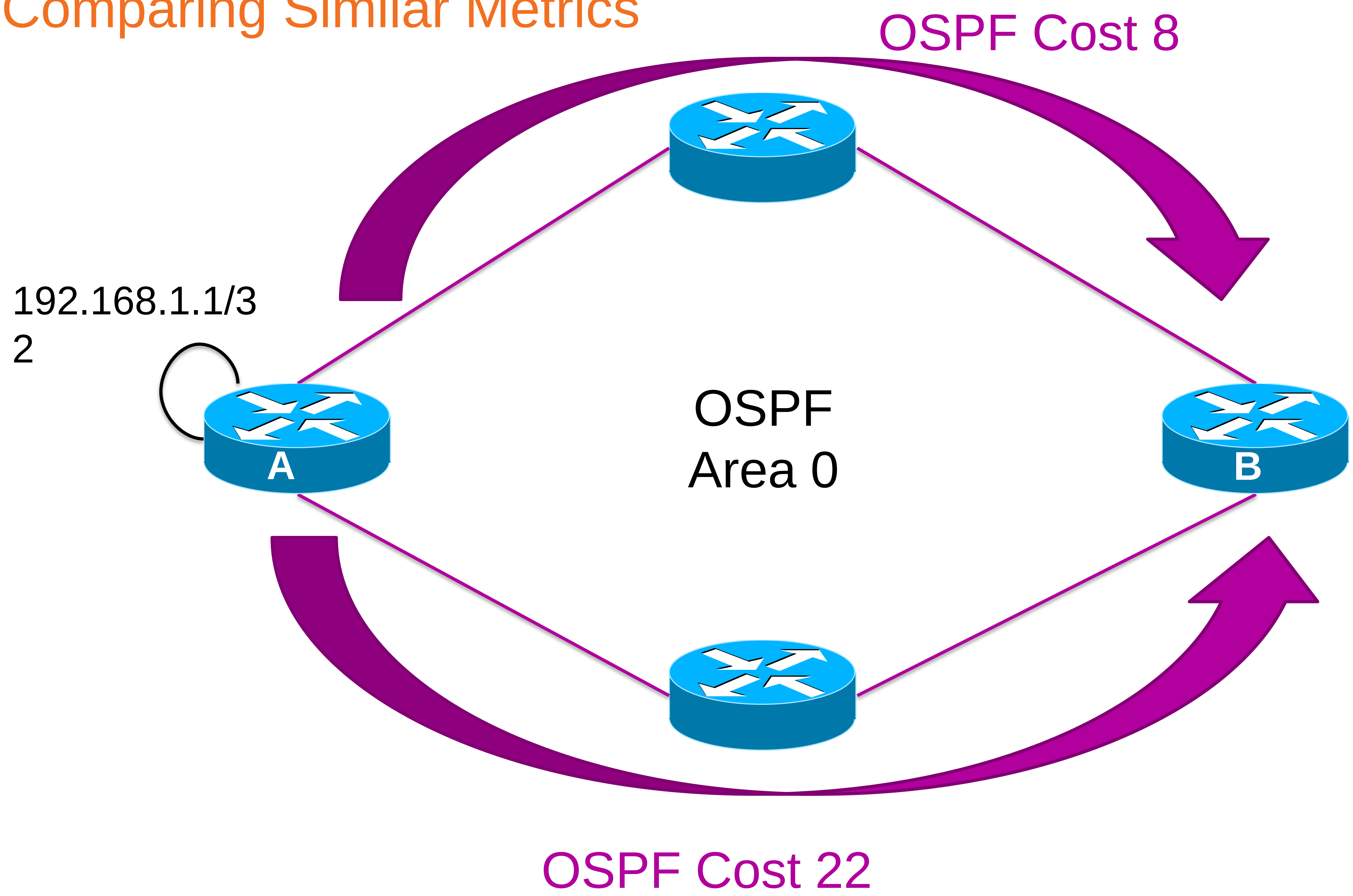
```
O     172.16.2.128/25 [110/11] via 172.16.1.1, 00:01:58, Ethernet0/0
```

- 110 – Administrative Distance
  - Locally significant
  - Lower is better (more trusted)
- 11 – Routing protocol metric
  - Determined by the routing protocol
  - Metrics of different protocols are not comparable
- 00:01:58 – Last modified time
  - EIGRP resets with any recalculation
  - OSPF/ISIS reset with SPF run

Route Selection

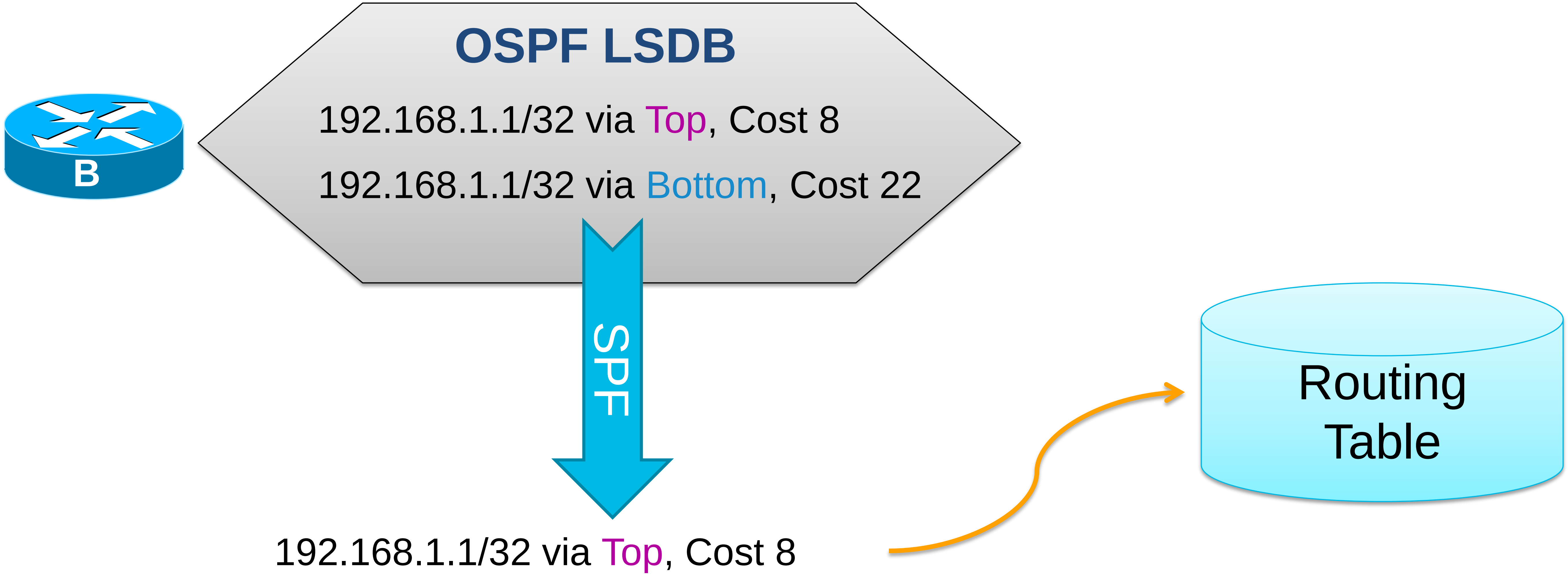
Apples to Apples

Comparing Similar Metrics



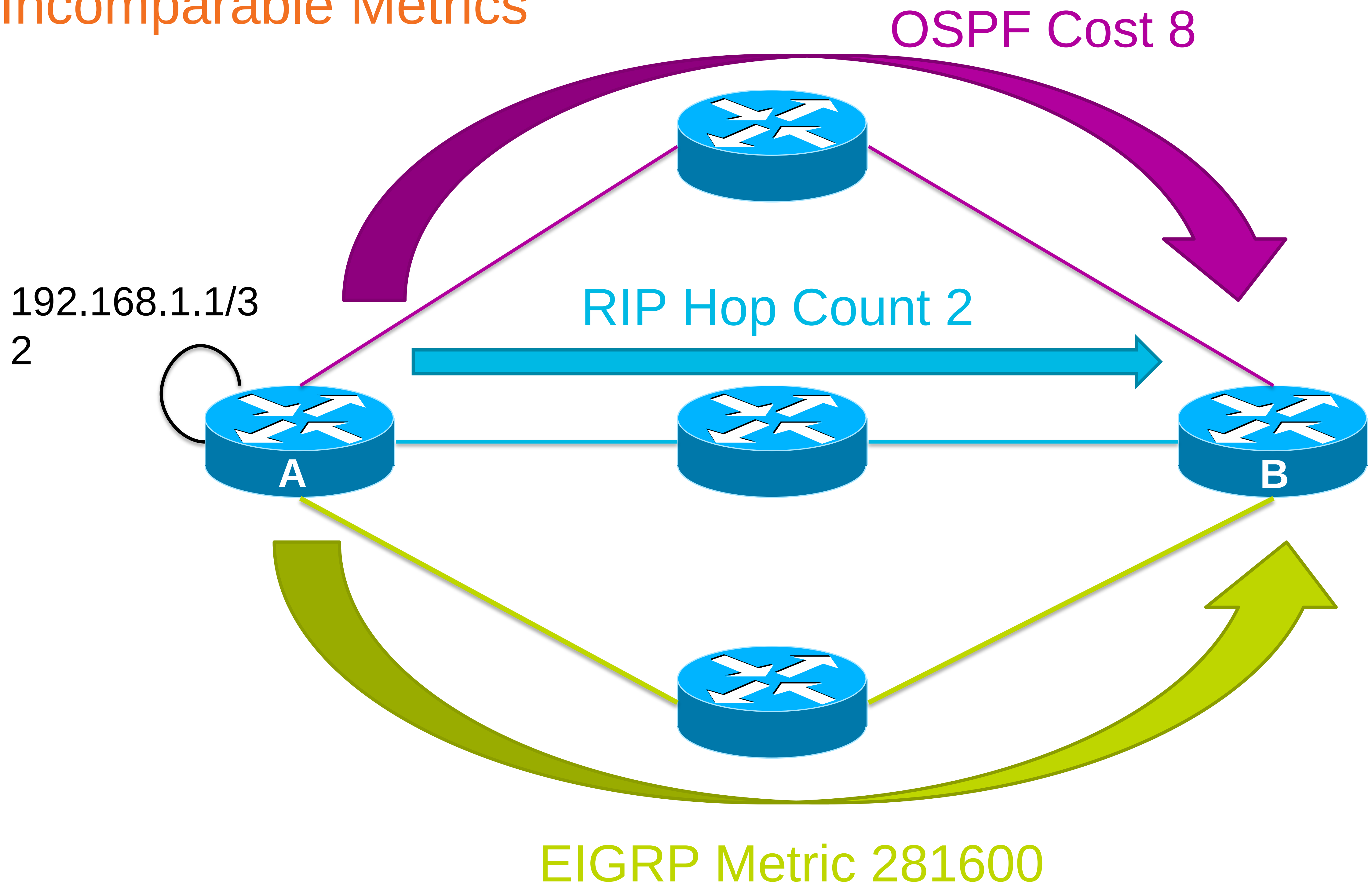
Apples to Apples

Comparing Similar Metrics

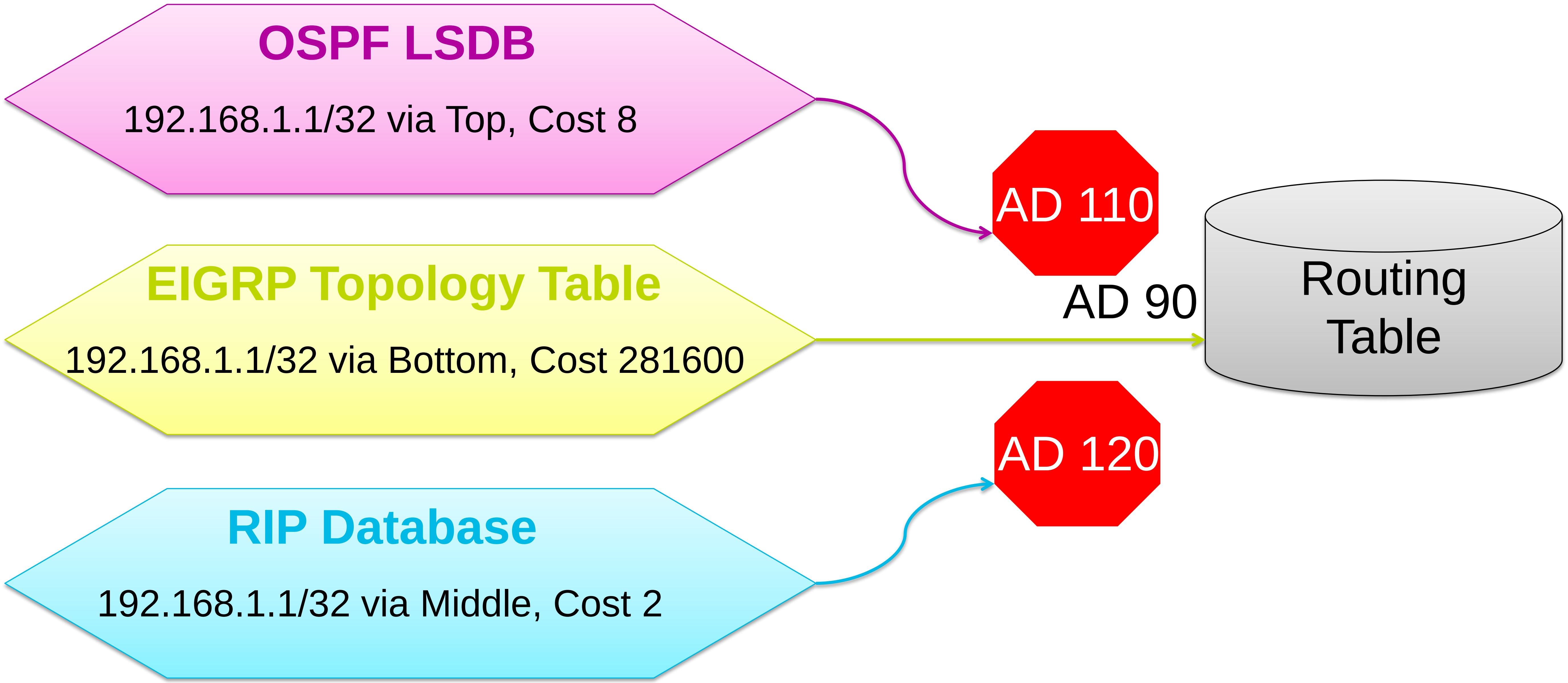


Apples to Oranges

Incomparable Metrics



Apples to Oranges
Incomparable Metrics



Presenting Routes

You Can't Install What Isn't There

- The routing protocol must "present" the route to the RIB for AD to have any effect
  - OSPF "Routing Bit" must be set
  - BGP must be able to reach the next hop
- If the routing protocol does not select the route, the AD never matters
- Possible for the "wrong" route to be presented
  - If BGP selects iBGP route as best, it will lose to any IGP route

Tie Breaking

Equal AD Behaviour

- If the AD of two protocols is the same the behaviour is undefined
  - Testing shows that the protocol that is the default will win

    (if RIP is changed to 90, the EIGRP Internal AD, the EIGRP route will win)
- OSPF – Oldest wins
- EIGRP – lowest AS
- Same Protocol – install both! (load balance)