

Networking Fundamentals (II)

Network Layer Routing

- We know that when a packet gets into a L2 network, it can be delivered by using ARP and MAC address. Then, **how does a network send a packet to the subnet that contains the packet's destination IP address?**
- It uses network layer routing.
- Network layer routing is **hop-by-hop routing**. This means that you don't have to figure out the entire routing path from a source to a destination when sending the packet. You can send the packet to the next hop and ask that hop for further direction.
- A router usually connects different subnets.
- A network node maintains a **routing table** that **mapping next hops with prefixes** (aka subnets). So, it can decide which node is the next hop based on the prefixes.

Routing Table

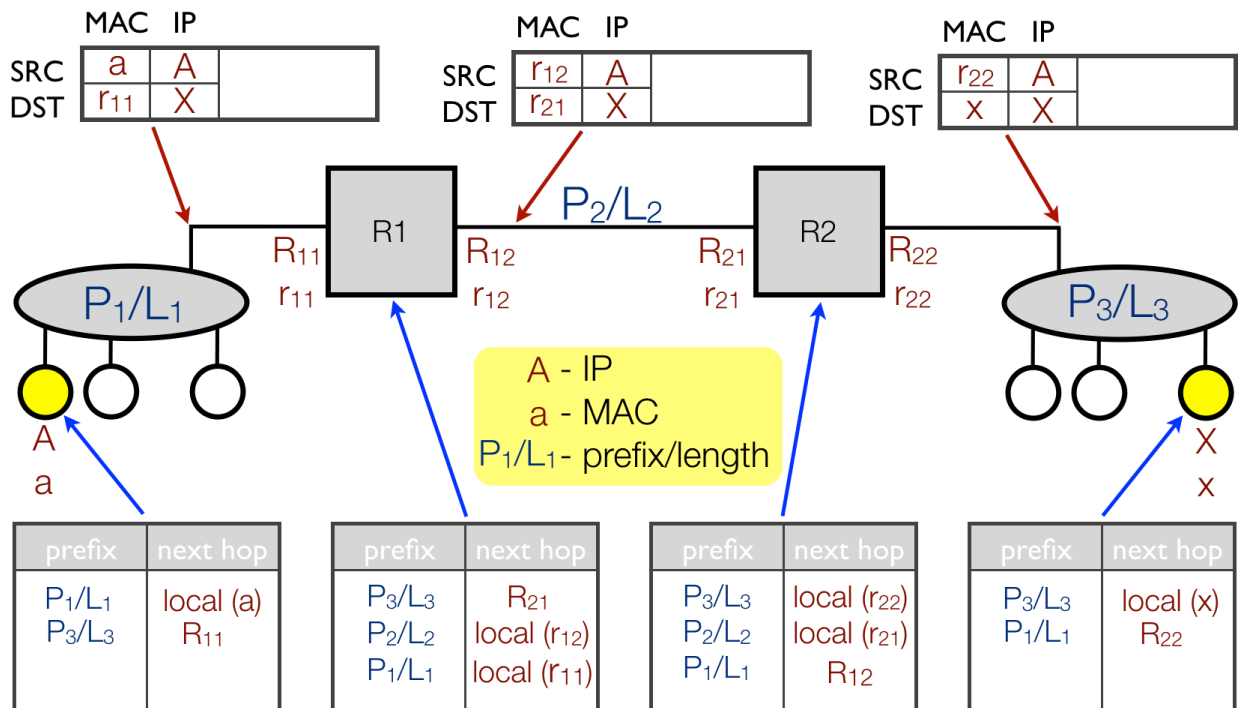
- What does a routing table look like? We can use the routing table on a computer for example.
 - The routing table of d122-01 below indicates:
 - **137.146.183.0/24**: local delivery via interface enp128s31f6 (subnet for CS department)
 - use ARP to find the MAC address for the destination IP
 - **0.0.0.0/0** (everything else): send to 137.146.183.1 (gateway of CS dept.) via enp128s31f6
 - use ARP to find the MAC address for 137.146.183.1

```
yingli@davis122-917fc6:~$ /sbin/route
Kernel IP routing table
Destination      Gateway         Genmask         Flags Metric Ref    Use Iface
default          _gateway       0.0.0.0         UG    100    0      0 enp128s31f6
137.146.183.0    0.0.0.0        255.255.255.0   U     100    0      0 enp128s31f6
yingli@davis122-917fc6:~$ /sbin/route -n [-n: do not resolve IP address]
Kernel IP routing table
Destination      Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0          137.146.183.1  0.0.0.0         UG    100    0      0 enp128s31f6
137.146.183.0    0.0.0.0        255.255.255.0   U     100    0      0 enp128s31f6
```

- If the destination is in the connected subnet, use ARP to find the MAC address for the destination.
- If the destination is in another subnet, use ARP to find out the MAC address of the next hop (gateway or router). Once the packet reaches at the router then update the dest MAC.
- Please note that network routers usually have multiple interfaces, which is different from computers. Each interface connects to a different subnet and has its distinct IP and MAC.

Routing Example

- In the above example, node A in P1/L1 wants to send a packet to node X in P3/L3.

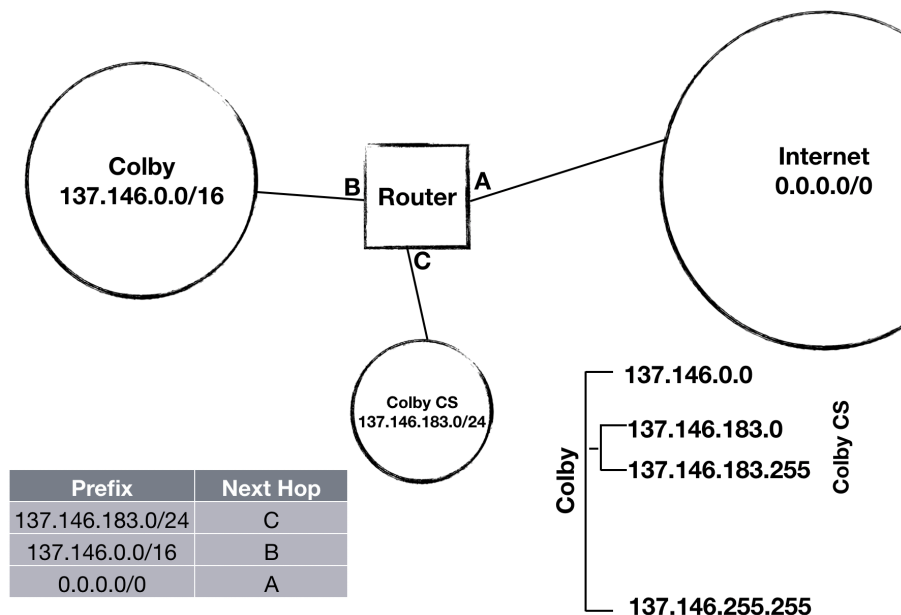


- When A sends the packet, it will look up the routing table first to find the IP for the next hop to X. After getting the IP, A will look into the ARP table to find the MAC address of next hop and use ARP if necessary. In this case, A will look up ARP table or use ARP for R11's MAC address, which is r11, the MAC address of the interface of the router R1 that connects P1/L1. A will then set the dest MAC address of the packet to r11.
- When the packet arrives at R1, R1 checks its routing table and finds packets for P3/L3 should go through R21. It then uses ARP if necessary, finds the MAC address of R21, and updates the dest MAC of the packet. Meanwhile, it also updates the src MAC to the MAC of its own interface R12. The reason is that it can provide hop-by-hop reliability if the packet is lost, which is more efficient than going back to the source and resend the packet.
- When the packet arrives at R2, R2 checks the routing table and finds that X belongs to P3/L3, its local network. It can then use ARP if necessary to find the MAC address of X, update the src and dest MAC of the packet, and sends the packet through R22.
- Please note that the src and dest IP are not changed during the entire process. The MAC addresses are updated every time when the packet is forward to the next hop.
- Please also note that it's important to **correctly configure the routing tables**. If the first entry in R2's routing table is R12, then the packet won't be delivered to X, and it will be tossed back

and forth between R1 and R2. This creates a **routing loop**, which will lead to significant packet loss and degrade the network performance.

Longest Prefix Match

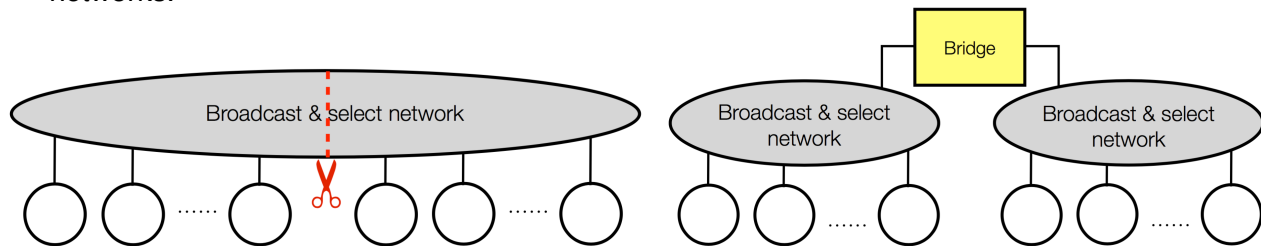
- We know that if the leftmost N bits of an IP address are the same as those in the subnet ID, the IP address belongs to the subnet specified by the prefix.
- However, the IP range of subnets are nested usually. For example, as shown in the figure below, the IP range for CS department is nested in the range for Colby.
- So, if the destination of a packet is an IP of CS department, how does the router route the packet to the appropriate subnet?
- It uses the **longest prefix match**, forwarding the packet to the hop that connects a subnet whose prefix is the most specific matching (aka **matching entry with the longest prefix length**).



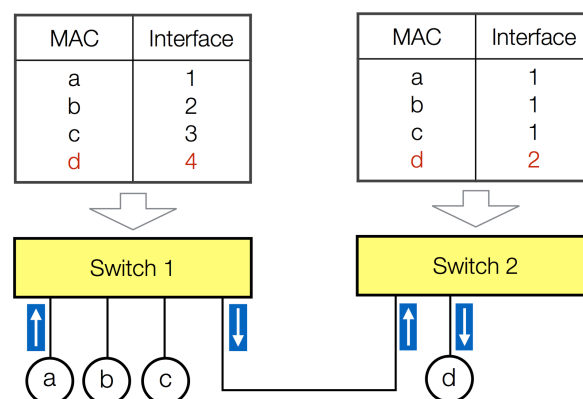
Dest Address	Matches	Used
137.146.3.15	B, A	B
137.146.183.35	C, B, A	C
200.100.0.15	A	A

Bridging

- Due to the broadcast nature of the underlying medium of L2 networks, it is difficult to scale up the L2 networks. The signal attenuate as the distance between two nodes increases, so as for the interference. No to mention the limited number of ports on a L2 devices (e.g., hub and switch).
- However, introducing more routers are expensive in terms of monetary cost and labor cost.
- Therefore, bridging is introduced to limit the scope of broadcasting and extend the reach of L2 networks.



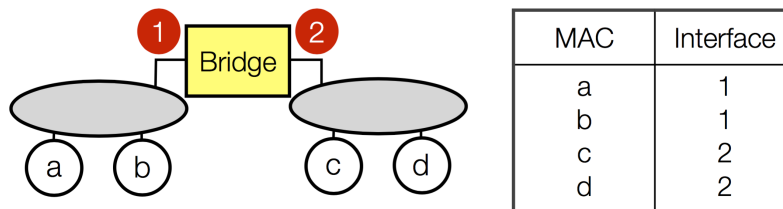
- The bridge sits between the sections of a subnet.
- In a simple subnet, it can be a wire connecting two sections of a subnet. In a complex subnet, it can be a switch connecting different sections.
- Please note that each **section** of the subnet **can** be a L2 network **connecting hosts via a switch**.
- Let's use a simple scenario to explain how a bridge works when connect L2 networks connecting hosts via switches.
- In a simple scenario, a bridge can be just a wire connecting two sections.
- If a node a in section 1 wants to send a packet to a node d in section 2, switch 1 will use its switching table to forward the packet.



- Once switches build up the switching table, they can tell the interface to the desired node. When node a pass the packet to switch 1, switch 1 will look up the table and forward it to

interface 4, as that interface can direct the packet to d. Once the packet arrives at switch 2, switch 2 will look up its table and forward the packet to interface 2, as that interface connects node d.

- In a more complex scenario, a bridging switch connects two sections of the same subnet. It **learns** node locations from passing traffic and stores them in its switching table.



- As shown in the above figure, the bridge connects two subnet sections and its switching table maps the MAC addresses of hosts in different sections to the interface that directly connects to that section.

Special Use IPv4 Addresses

- **10.0.0.0/8**, **172.16.0.0/12**, **192.168.0.0/16** - Private network
 - Range: 10.0.0.0 ~ 10.255.255.255, 172.16.0.0 ~ 172.31.255.255, 192.168.0.0 ~ 192.168.255.255
 - Scope: Private network
 - Purpose: Used for **local communications within a private network**
 - Commonly used for home, office, and enterprise networks
 - To **delay IPv4 address exhaustion**
 - May **use without approval from regional Internet registry**
 - Are **not allocated to any specific organization**
 - IP **packets addressed from these** addresses **cannot be transmitted through the public Internet**, and so if such a private network needs to connect to the Internet, it must do so via a network address translator (NAT) gateway, or a proxy server.
- **127.0.0.0/8** - Loopback
 - Range: 127.0.0.0 ~ 127.255.255.255
 - Scope: Host
 - Purpose: Used for **loopback address** to the local host
- REF: https://en.wikipedia.org/wiki/Reserved_IP_addresses
- A complete lists from RFC 5735

Disclaimer: Notes adapted from the textbook and online resources.