1. Distinguish between switching, forwarding and routing.

   - Switching

     Switching is the method Switches use to determine where to send frames.

     The switching table is a mapping from MAC addresses to output ports on a switch.

     The algorithm for switching is as follows:

     – On receipt of a packet from MAC address x on port y add an entry into the switching table to send packets for MAC address x out on port y.

     – On receipt of a packet destined for MAC address x, search for it in the switching table. If there is an entry saying which port it should be sent on, send it on that port.

       Otherwise, flood the network asking if any nodes *have* MAC address x or have an entry for x in their switching table. If no response is received, x is not on the network so the packet is dropped. If a response is received, the entry is input into the forwarding table and x is sent out on the correct link.

   - Routing

     Routing is a network-wide operation which is done every time the network changes. Routing is done in the control plane and is used to construct the forwarding table.

   - Forwarding

     Forwarding is a per-packet operation performed by Routers in the data plane. The Forwarding Table is a mapping from ranges of IP addresses to output ports. A copy of this is stored on the Input Linecard. For each packet, the Input Linecard finds finds the longest prefix match (the most specific destination) in the forwarding table and sends the packet to the corresponding output port.

2. Describe the Link State and Distance Vector routing algorithms.

   - Link State Routing

     In Link State routing, each node is given total information about the whole network then computes the shortest path to the destination and stores the first-hop in the routing table. Initially, each node knows only its local link state (information about links it is directly connected to). Each node floods the network with this information (using eager reliable broadcast). Every node on the network now has the global link state.

     Each node now has a copy of the global link state. Every node then independently uses Dijkstra's Algorithm to work out a least-cost path to every router.

   - Distance Vector Routing

     In Distance Vector Routing, each node maintains a Distance Vector D, containing a triple of $(\text{address prefix}, \text{subnet mask}, \text{distance}, \text{router})$ where $(a, s, d, r) \in D$ means the subnet a/s can be reached in distance d through router r.

     Nodes continuously broadcast their distance vector to their neighbours. They then update their distance vectors to the elementwise minimum.

3. What are RIP, OSPF and BGP?

   RIP is "Router Information Protocol" – the most common implementation of Distance Vector Routing.

   OSPF is "Open Shortest Path First Protocol" – a common implementation of link-state routing.

BGP is "Border Gateway Protocol" – an inter-domain routing protocol which uses a variant of of Distance Vector Routing and is designed to allow border routers to hide what goes on inside their networks.

Simultaneous: when you send data on broadcast, all links get it at the same time.

Head of line blocking occurs in the input buffer. The head of the input queue can't be sent because the output port is receiving a lot of data. Other packets in the buffer are blocked even though they could go...

"Virtual output buffering": have a buffer for each output in every input buffer. This resolves HOL blocking.

Fragmentation IN SWITCHES: crossbar switches need to be clocked to be efficient. So they fragment the frames on the input and reassemble them on the output.

The same crossbar switching happens in routing. In order to function, a router needs a switching fabric within it. The physical devices are named by the highest layer of intelligence they contain.

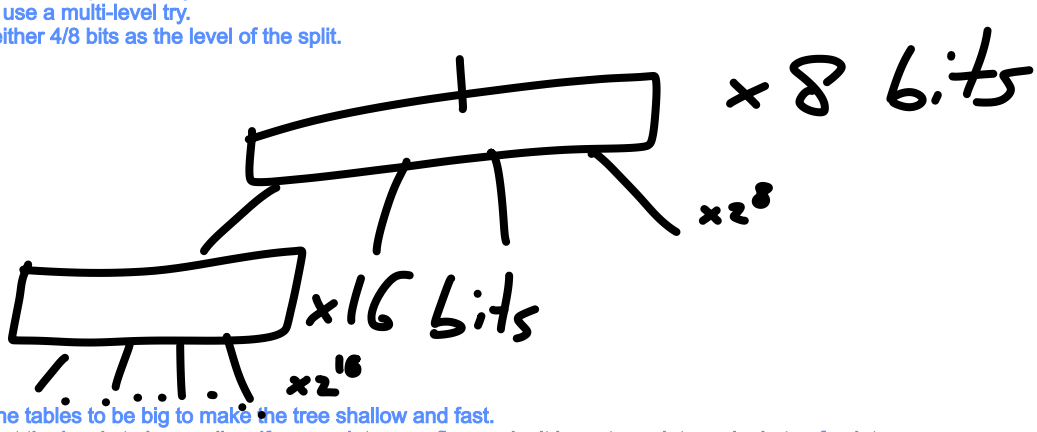
Switching tables are fully associative -- "a fully associative L2 cache"
Q: why isn't it set associative? A: It's faster

This scales linearly with the size of the network. i.e. if you broadcast a packet then every switch has to do a read and write to the switching table.

If the "link beat" drops then all entries for the switching table on that link are deleted.

The same idea doesn't scale from switching table to forwarding table (we can't store a cache...)

Forwarding table maps IP to layer 2 link on which to send data.
Instead we use a multi-level try.
This uses either 4/8 bits as the level of the split.



You want the tables to be big to make the tree shallow and fast.
But you want the levels to be small so if you update a prefix you don't have to update a gigabyte of pointers.
These are maintained in software but perform in hardware (control / data plane)

In distance vector, you send out data on a timer -- every 30s. This isn't based on whether you change.
This means data propagates slowly. But it's not too bad since most DV changes are pretty local.

A "solution" is triggered update. But this causes errors... i.e. dust in a fibre-optic cable causing it to turn on and off repeatedly...

"Count to infinity": I advertise a path. My path fails. I take up the offer from someone else. This other person increments my cost. Etc, etc. Count to infinity. You can count up to infinity (for RIPv1 this is 15, for RIPv2 this is 31... takes 15.5 seconds) or use a better path.

Split horizon: "don't avertise any path back to the person who advertised it"... problem: loops of length 3 are not resolved...

Split horizon with poison reverse: same worst-case

Holddown: "after my link to x fails, I refuse to accept that anyone else has a link to x for τs". τ=30*longest cycle... this is a guaranteed (hopefully shorter outage) instead of a longer outage.

Best solution is "Triggered update" with a clamp (of ~3s)

BGP is a "Path Vector" protocol.
BGP speakers represent their entire ISP.
At BGP, an entire ISP is a single hop... So one node can stretch from singapore to LA

OSPF is Link State
You have very fast convergence. You run everything in parallel on every router.
You could only do this within an ISP. You couldn't scale it up to the size of the internet.

OSPF is ubiquitous.
RIP is dead.
BGP happens between routers.

Routing is a background process in the data plane whose process is to reconfigure the forwarding table.