Network Layer

Lecture 15 Layer 3 Switching

How do we go from LAN to a much larger network?
Why doesn't ethernet switching scale?

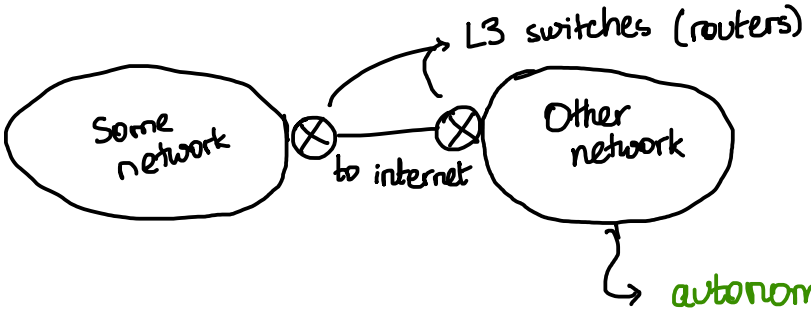
→ In the spanning tree, the path between two nodes could be long.
(potentially very unoptimal because we are not using all links)

→ The forwarding table, whose size can be as large as the number of hosts, can be very cumbersome to use.
This is a result of flat addressing.
To fix this, we use hierarchical addressing in IP.

→ If a switch in the tree goes down, we reconstruct the spanning tree.
There are periodic "Hello" messages to ensure that the tree is intact.
(if not received by someone, we reconstruct)
In a large network this could happen often, thus wasting resources frequently.

→ Earlier, there were no common addressing scheme or communication protocols across the globe.

L3 switches forward based on the IP address.



autonomous system (AS)

Each AS can choose its own internal routing protocol.
(the distance heuristic mentioned at the end of the prev. section)

There is intra-domain routing (within AS) and inter-domain routing (between AS)
In the internet, inter-domain routing is done using BGP — the Border

Gateway Protocol.

Let us start with intra-domain routing. It is broadly of two types: distance vector and link-state routing

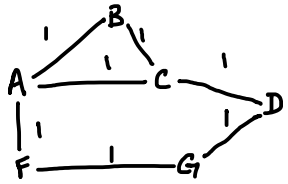
- distance vector → RIP (Routing Information Protocol)
- link-state → OSPF (Open Shortest Path First), IS-IS (Intermediate System to Intermediate System)

They are essentially just algorithms to:

- → find shortest path (each hop is assigned a weight by the admin)
- → avoid cycles.

A router does not need to know the entire route, only the next hop in a shortest path.

Distance vector routing uses a distributed version of the Bellman Ford algorithm.



A (and each node) first sends out (A, 0)

its IP ↵ ↳ distance to itself

It then updates its forwarding table after hearing each message.

Destination	Next hop	Cost
A	-	0
B	B	1
C	C	1
F	F	1

Next, A sends its table to its own neighbours.

(A,0), (B,1), (C,1), (F,1)

From C, it hears (C,0), (A,1), (B,1), (D,1)

It then updates its table as

Destination	Next hop	Cost
A	-	0
B	B	1
C	C	1
F	F	1
D	C	2
G	F	2

Proceeding, it builds up a forwarding table, choosing the neighbour closest to a destination at each step

What happens if a link fails?

A node X recognizes that the link has failed and sends this information to its neighbours, saying that its distance to that node is now ∞.

Similarly, if a neighbour's next hop for that destination is X, it updates its own cost as ∞ as well

This spreads until we reach a node with a different next hop.

If we receive a packet for that node in the intermittent period, it is discarded.

How often does this occur?

→ Triggered update: An event triggers a routing update.
(We try to send on a link and we fail)

→ Periodic update: Periodically, give neighbours information about routing table.

No particular node knows the topology of the entire network.