Switching models

- datagram / connectionless
  - packets can be sent immediately by a host
  - uses forwarding tables in switches
  - packets may traverse different routes from source to destination
  - robust
- virtual circuit
  - need to establish connection state in each switch on the source-destination path first
  - can be administratively configured or established by signalling
  - smaller per-packet overhead than datagram approach
- source routing
  - can be used in datagram & virtual circuit networks
  - information about source-destination route is placed in packet by source
  - strict or loose

Addressing & forwarding / routing

- **link layer LAN switching – extended LANs** (datagram)
  - MAC addresses used
    - every network adapter has a unique MAC address, usu. 6 bytes long
    - MAC addresses have a flat structure & do not change when an adapter changes its physical location (unlike IP addresses)
    - the Address Resolution Protocol (ARP) resolves IP addresses to MAC addresses
    - LAN switches are plug-&-play (unlike routers which need configuration of IP addresses)
  - forwarding tables are used by switches to decide which output(s) to direct a frame to, or whether to drop the frame
    - a forwarding table contains entries for some subset of nodes
    - a self-learning switch builds the table as follows:
      - when the switch first boots, the table is empty
      - for every frame received, the switch stores/updates a table entry w̄: (1) the frame's source MAC address, (2) the interface from which the frame arrived, & (3) the current time

- each entry is deleted after some specified timeout period

If a frame arrives at a switch on an interface x,

- if its destination address is not in the forwarding table, the switch broadcasts the frame on all interfaces except x
- if its destination address is associated with interface x in the table, it is dropped (filtering)
- if its destination address is associated with an interface $y \neq x$ in the table, it is forwarded on interface y

• To prevent cycling of broadcast frames in a cyclic topology, switches run a <u>spanning tree protocol</u> to decide on a spanning tree over which frames are forwarded

- dynamic, distributed algorithm — if a bridge fails or topology changes, a new spanning tree is formed (allows redundancy in network topology)
- each switch has a unique ID, which is composed of a configurable priority & a fixed MAC address; each of its interfaces is similarly ID'd
- results:
  - the switch with the smallest ID is elected as the root switch
  - each switch determines which of its

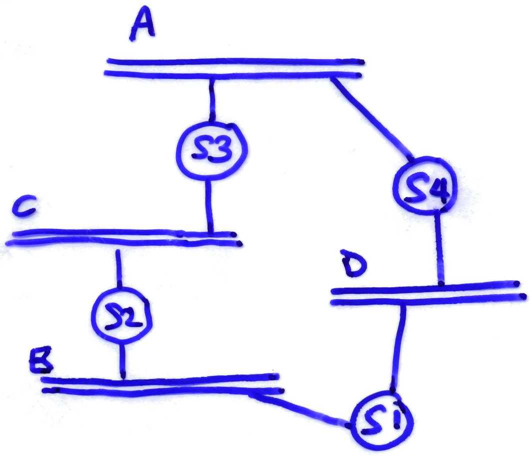
interfaces provides the shortest path to the root (chosen as root port), and the length of the path

- each LAN segment elects the switch closest to the root as the designated switch for that segment (ties broken by switch IDs)
- each switch selects interfaces connected to segments for which it is the designated switch, & places remaining interfaces in an inactive state (do not send/receive data frames)

- protocol details:
  - switches exchange configuration messages containing: (1) the sender's switch & port IDs, (2) the ID of the switch that the sender believes is the root, & (3) the cost of the path to the root from the sender
  - initially each switch thinks it is the root, & sends configuration messages out on each of its ports, identifying itself as the root & specifying distance to root = 0
  - when a switch receives a config message indicating a root with a smaller ID, it stops initiating new config messages & only propagates received config messages (with the distance to root ↑'d by the length of the link on which the message is received)

- when a switch receives a config message from a sending switch that is closer to the root (or as close but with a smaller ID), it knows it is not the designated switch for that interface, & stops sending config messages on that interface
- when the system stabilizes, only the root switch continues to generate config messages, & all other switches propagate these messages only on interfaces for which they are the designated switch

- Eg.



- Denote a config message as (sender, believed root, distance)
- Suppose all switches boot around the same time

- focusing on S3,
  - S3 receives (S2, S2, 0) & (S4, S4, 0), thinks S2 = root
  - S3 sends (S3, S2, 1) to S4
  - S3 receives (S2, S1, 1) & (S4, S1, 1), thinks S1 = root & both S2 & S4 are closer too root
  - S3 places both its ports in an inactive state

- the original Spanning Tree Protocol (Perlman, DEC) detects topology changes when bridges stop receiving periodic config messages from the root

- the Rapid Spanning Tree Protocol, which replaced STP in IEEE 802.1D-2004, improves recovery time by having faster link change detection & having switches automatically determine an <u>alternate port</u> that is a backup to the root port

- <u>limitations</u> of extended LANs
  - scalability: use of broadcast is not scalable
    - scalability can be improved by partitioning an extended LAN into several <u>virtual LANs</u> (VLANs)
    - each VLAN is assigned an identifier; packets are broadcast only among VLANs with the same identifier
    - requires each switch port to be configured with appropriate VLAN identifiers
  - heterogeneity: can only interconnect networks using the same frame header / address format eg. 802.5 rings, various Ethernet technologies