Communication Networks:
Technology & Protocols

Stavros Tripakis (stavros@eecs)

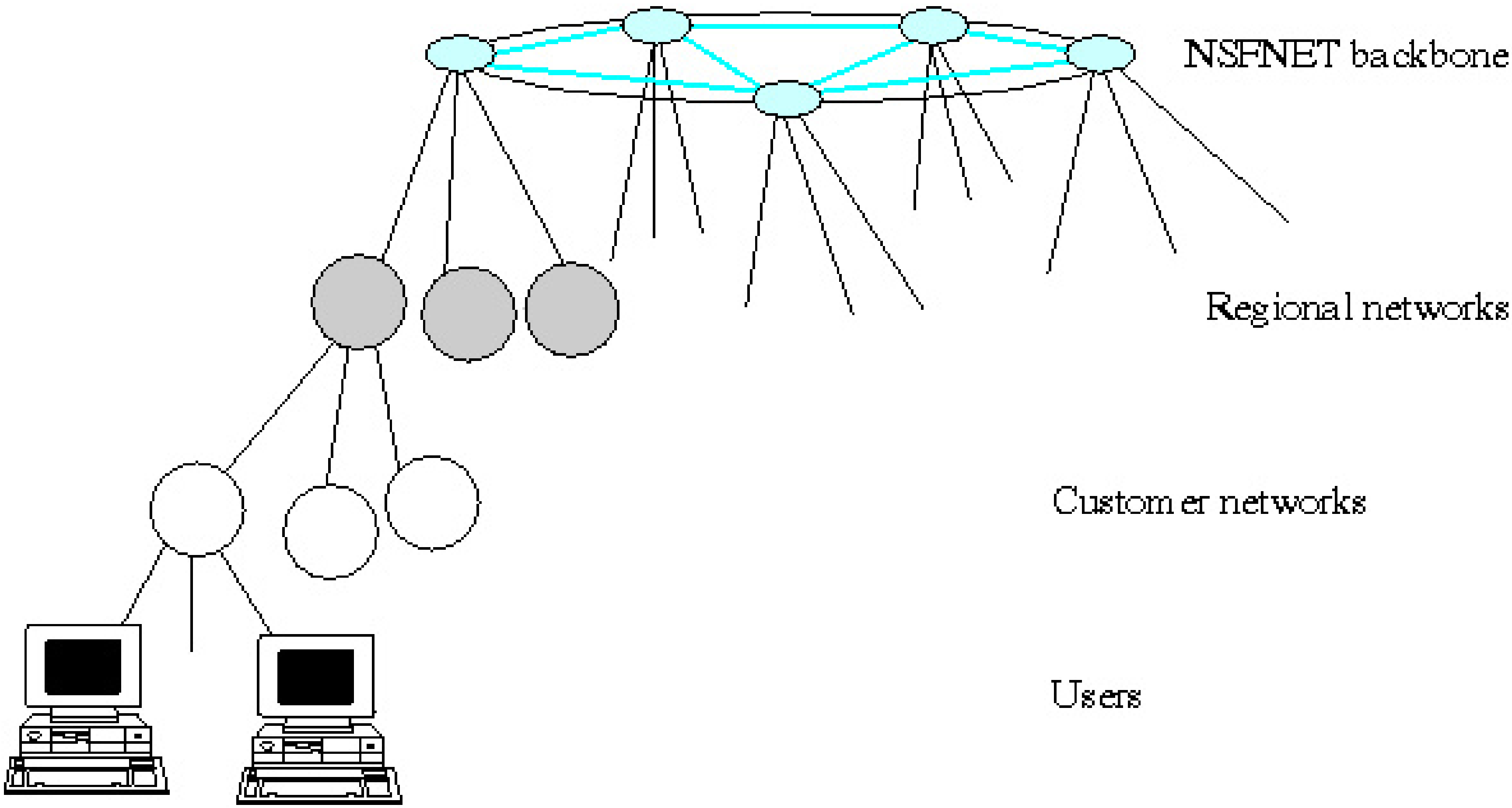
Lecture 8

September 10

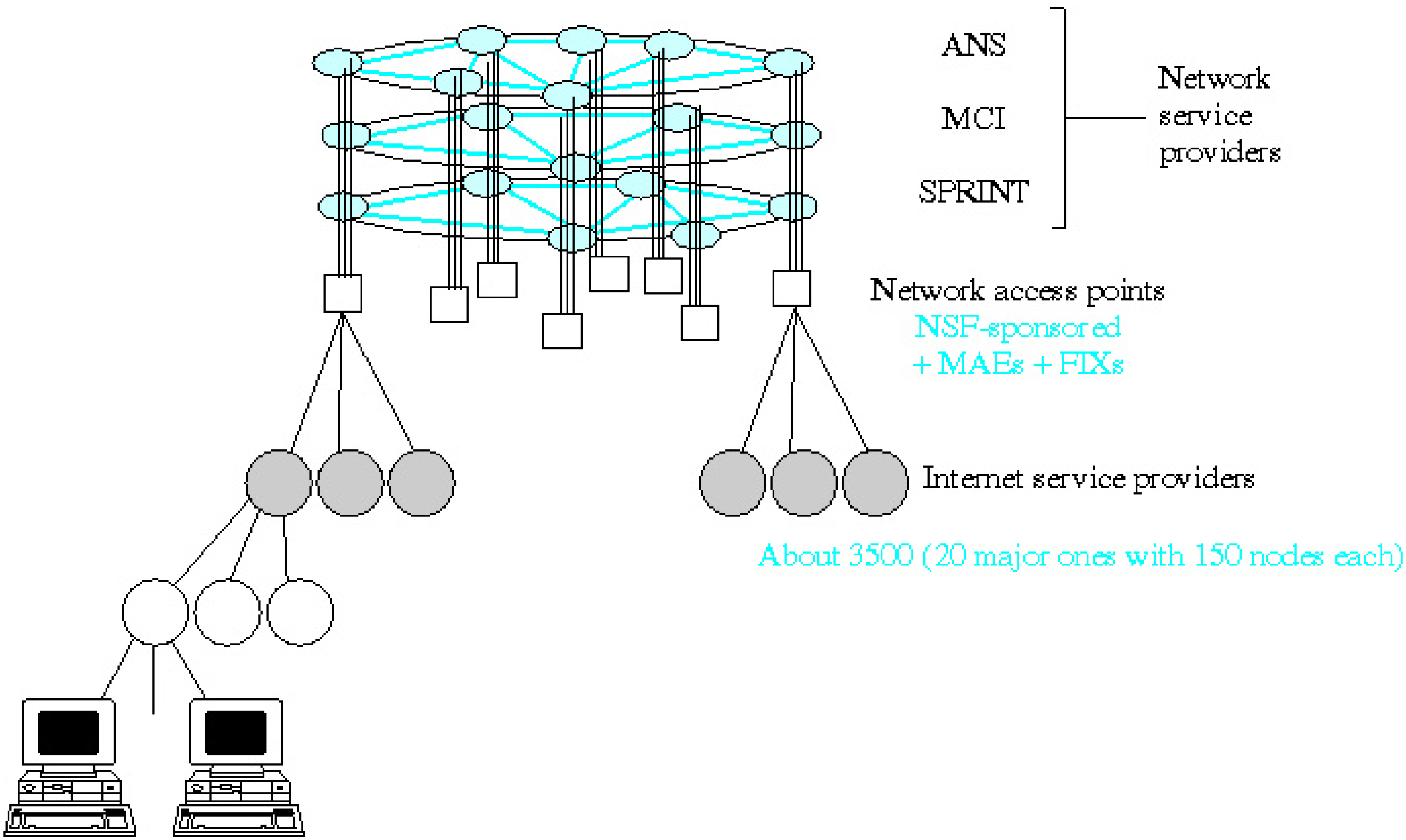
Logistics

- Web site:
  - www.cs.berkeley.edu/~amc/eecs122
- **Homework 3** (due Friday 9/17) will be available on web-site later today.
  - Homework 1 due today (**counts for enrollment**).
  - Homework 2 due Monday 9/13.

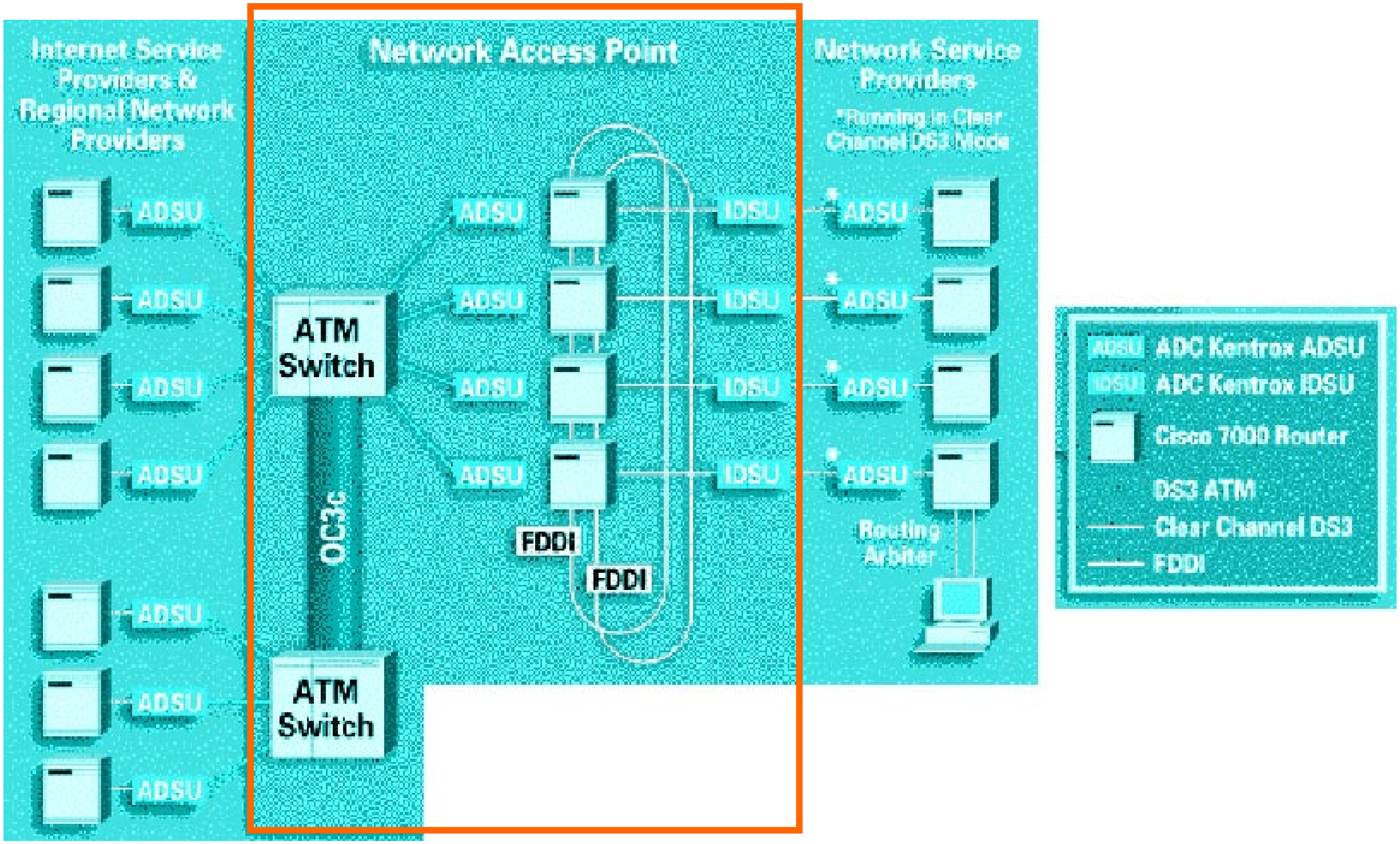
The Internet around 1990



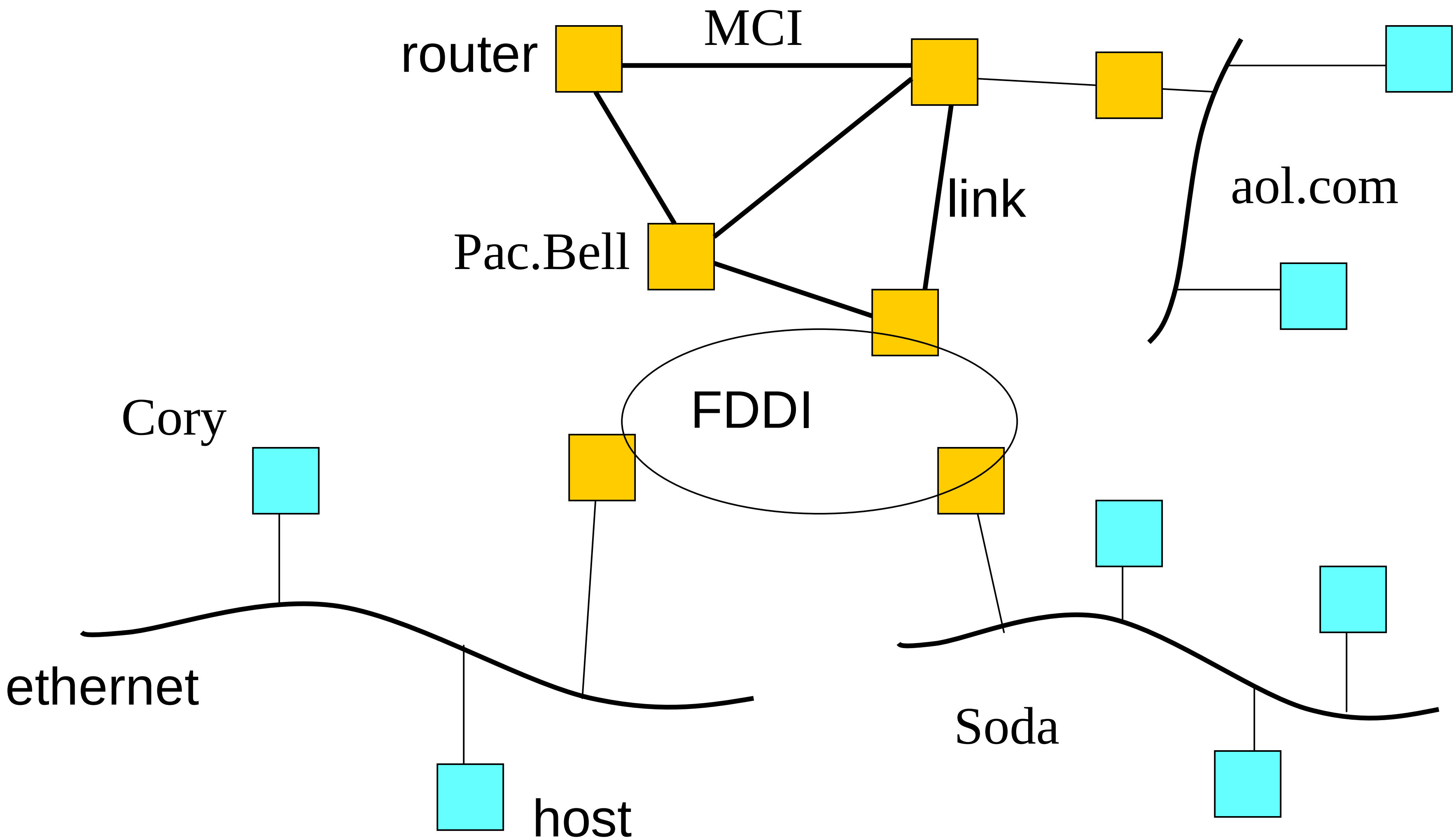
The Internet in 1997



A typical Network Access Point (NAP)



A small Internet



Routing

- The Internet is a network of **heterogeneous** networks:
  - using different technologies;
  - belonging to different administrative authorities.
- Goal of routing: interconnect all these networks.
- Routers, switches, bridges.
- Routing protocols.

Routing requirements

- Scalability.
- Robustness to network changes (link failures, full buffers, etc).
- Efficiency: "good" routes.
- Security, administration issues.

Internet routing: hierarchical

- Routers only know about networks, not hosts.
- Routing table:

Destination Network	Next hop IP address
Net1	Router 1
Net2	Router 2
Net3	Router 3

- Network addresses = pairs (IP address, mask)
  - E.g., (128.96.34.0, 255.255.255.128)
  - Can encode class-based, subnetting, supernetting.

Internet routing: hierarchical

- Class-based:
  - class A, e.g., network 4 has pair (4.0.0.0, 255.0.0.0)
  - class B, e.g., network 128.29: (128.29.0.0, 255.255,0,0)
  - class C, e.g., net 192.44.3: (192.44.3.0, 255.255.255.0)
- Subnetting:
  - E.g., Berkeley class B address: 128.29 and EECS subnet address: 128.29.4: (128.29.3, 255.255.255.0)
- Supernetting:
  - Supernet of 16 class C networks: 192.44.16-31 has pair: (192.44.16.0, 255.255.240.0)

Internet routing: hierarchical

- When a packet with destination IP address X arrives at the router:
  - the router goes over the entries in its routing table
  - for each entry ((Y, M), Z), it checks whether:

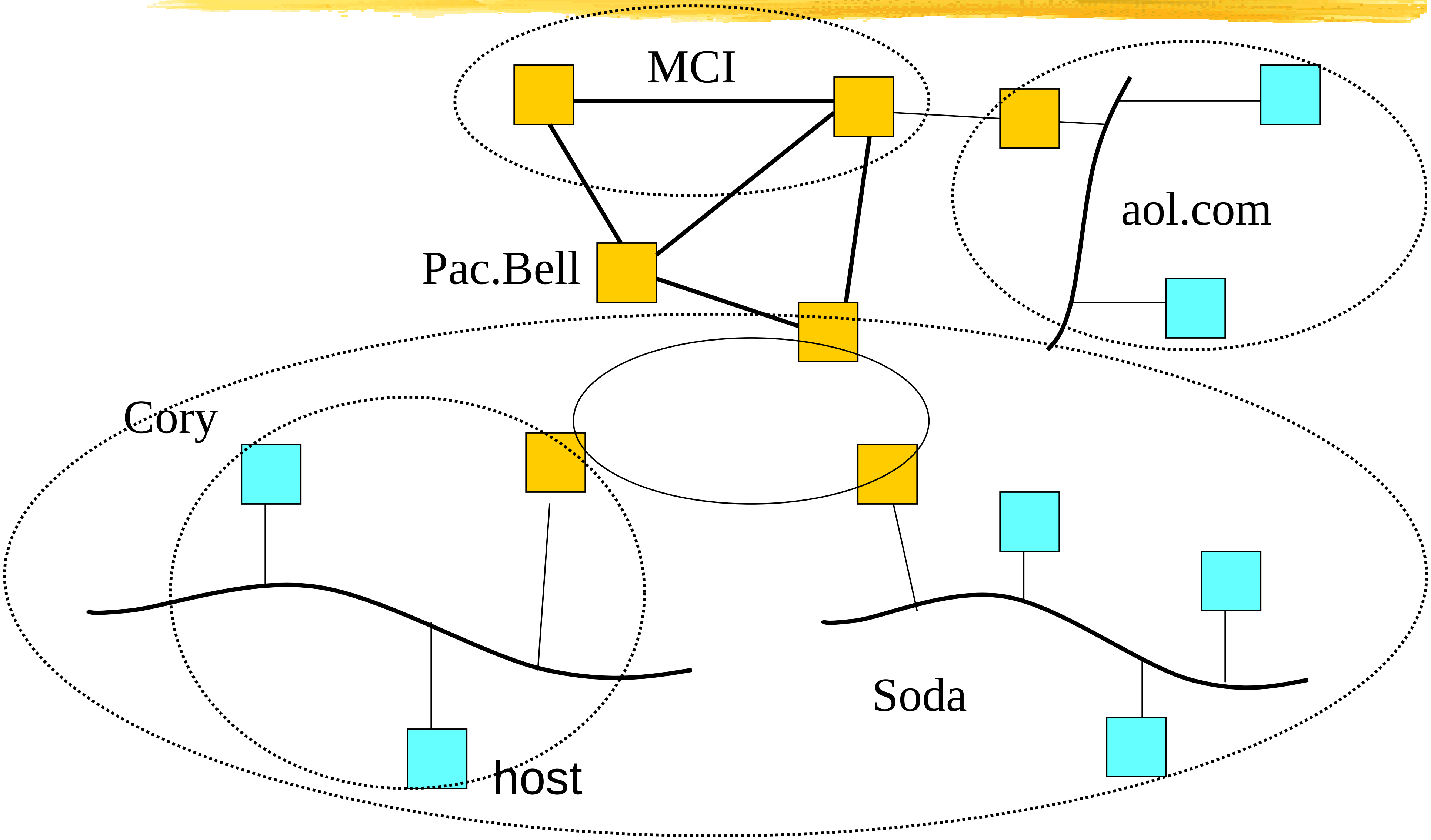
    $$X \wedge M = Y$$

  - if so, then Y is the (sub-/super-) network where X belongs, and the packet is forwarded to Z.
- What if multiple entries match ?
  - Pick the one with the longest match.

Internet routing: hierarchical

- **Autonomous system** (AS): a set of networks subject to a common authority. E.g.,
  - Berkeley campus network.
  - Company network.
  - Regional network (1990), ISP network (today).
  - NSFNET backbone (1990), one of the corporation backbones (today).

A small Internet: ASs



traceroute to ormelune.imag.fr (129.88.43.35): 1-30 hops, 38 byte packets

```
 1  cnr239net.EECS.Berkeley.EDU (128.32.239.1)  0.737 ms  0.503 ms  0.405 ms
 2  fast4-0-0.inr-110-cory.Berkeley.EDU (169.229.1.41)  0.871 ms  0.796 ms  0.752 ms
 3  f4-0.inr-107-eva.Berkeley.EDU (128.32.120.107)  1.49 ms  1.56 ms  1.8 ms
 4  f8-0.inr-666-eva.Berkeley.EDU (128.32.2.2)  2.63 ms  2.70 ms  2.3 ms
 5  fast0-0-0.inr-002-eva.Berkeley.EDU (128.32.0.82)  1.70 ms  2.8 ms  1.57 ms
 6  pos0-2.inr-000-eva.Berkeley.EDU (128.32.0.73)  2.24 ms  1.45 ms  1.59 ms
 7  c2-berk-gsr.calren2.net (128.32.0.90)  1.93 ms  1.58 ms  1.57 ms
 8  Abilene-BERK.POS.calren2.net (198.32.249.42)  5.19 ms  4.71 ms  5.37 ms
 9  denv-scrm.abilene.ucaid.edu (198.32.8.2)  26.9 ms  27.2 ms  26.3 ms
10  kscy-denv.abilene.ucaid.edu (198.32.8.14)  36.6 ms  36.8 ms  37.0 ms
11  ipls-kscy.abilene.ucaid.edu (198.32.8.6)  45.6 ms  47.2 ms  46.4 ms
12  chicago-atm40.1.opentransit.net (193.251.128.169)  151 ms  51.0 ms  50.5 ms
13  bagnolet1-atm30.5.opentransit.net (193.55.152.77)  153 ms  153 ms  152 ms
14  bagnolet2-fddi000.opentransit.net (193.55.152.178)  153 ms  153 ms  153 ms
15  rbs2.renater.ft.net (195.220.180.30)  156 ms  154 ms  153 ms
16  stamand1.renater.ft.net (195.220.180.17)  154 ms  154 ms  154 ms
17  grenoble.renater.ft.net (195.220.180.6)  166 ms  166 ms  166 ms
18  194.199.224.113 (194.199.224.113)  180 ms  207 ms  216 ms
19  194.199.224.122 (194.199.224.122)  210 ms  179 ms  168 ms
20  aramis.grenet.fr (193.54.184.1)  170 ms  167 ms  167 ms
21  r-ujf.grenet.fr (193.54.185.124)  170 ms  168 ms  168 ms
22  bio-gate.ujf-grenoble.fr (193.54.238.9)  169 ms  169 ms  170 ms
23  ormelune.imag.fr (129.88.43.35)  171 ms  *  172 ms
```

Internet routing: hierarchical

- *Intradomain routing* : routing inside an AS.
  - Packet forwarding in LANs.
  - Distance-vector routing (Bellman-Ford's shortest-path algorithm, RIP protocol).
  - Link-state routing (Dijkstra's shortest-path algorithm, OSPF protocol).
- *Interdomain routing* : routing across many ASs.
  - EGP and BGP.

Packet-forwarding in LANs

- At node A, given destination IP address B, is B directly connected to A ?
  - Is B in the same LAN as A (e.g., Ethernet, FDDI) or is there a point-to-point link between A and B ?
  - If yes, what is the LAN/link interface address ?
  - If not, send to (default) router.
- Address Resolution Protocol (ARP).

Packet-forwarding in LANs

- ARP table (or cache):

IP address	LAN address
C	x, Ethernet
D	y, Ethernet
E	z, FDDI
...	

- Basic operations:
  - If A doesn't have an entry for B, it **broadcasts** message "B, are you on my LAN? If yes, give me your interface address".
  - If B is in the LAN, it replies, and A adds an entry in its ARP cache.

Packet-forwarding in LANs

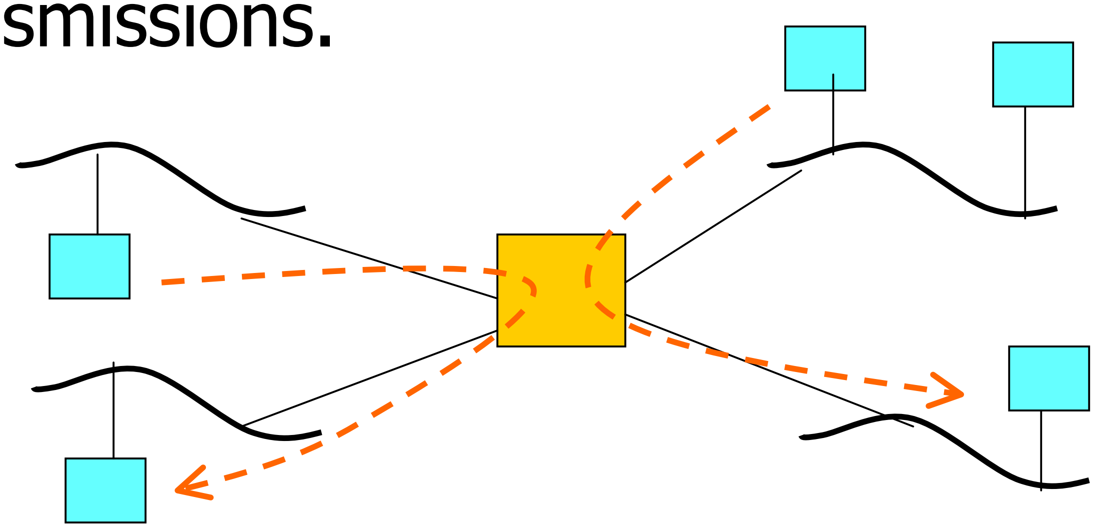
- ARP table (or cache):

IP address	LAN address
C	x, Ethernet
D	y, Ethernet
E	z, FDDI
...	

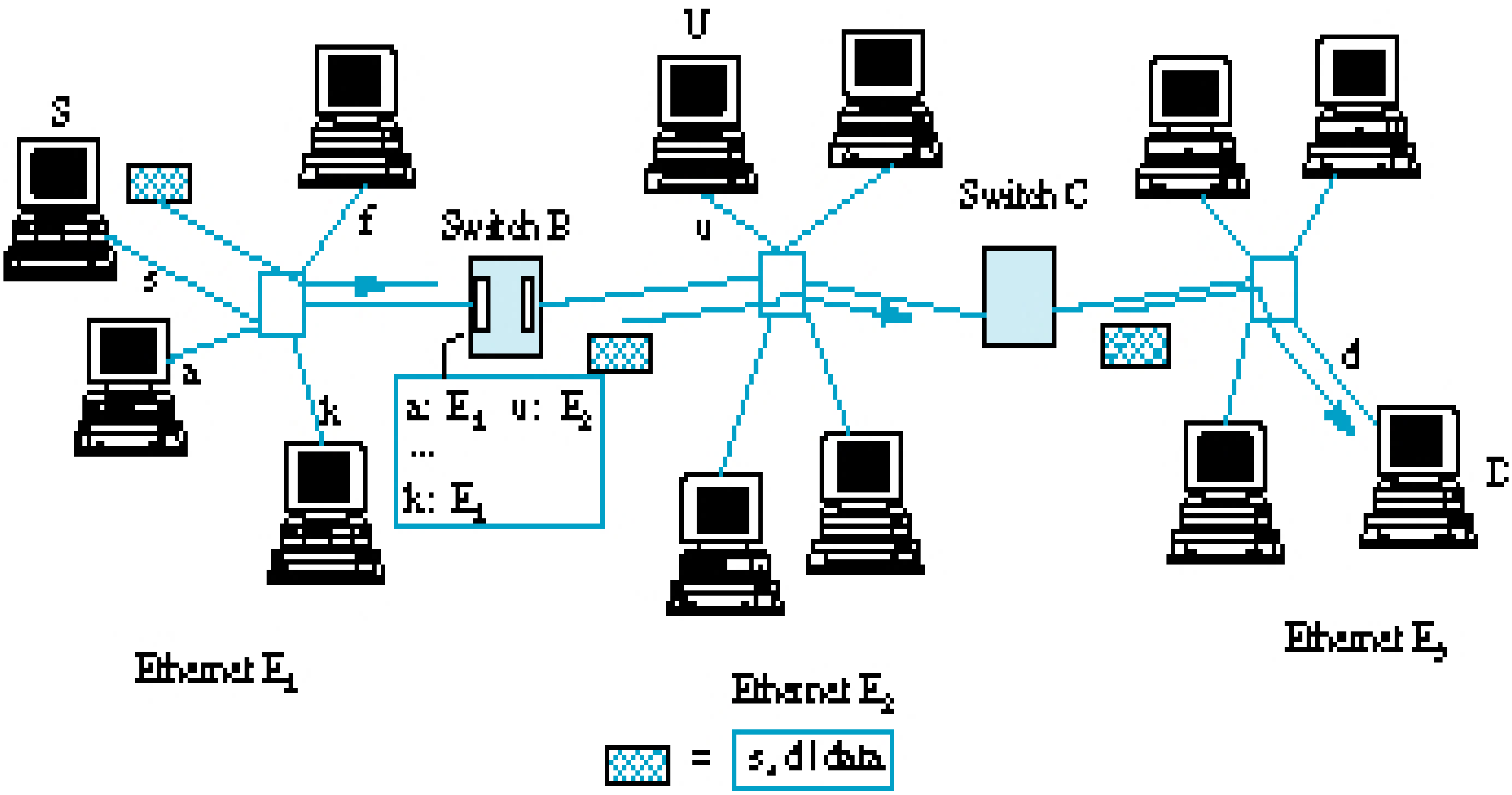
- Notes:
  - LAN has to support broadcasting (special broadcast address used).
  - Point-to-point link addresses statically configured.

Switched Ethernets

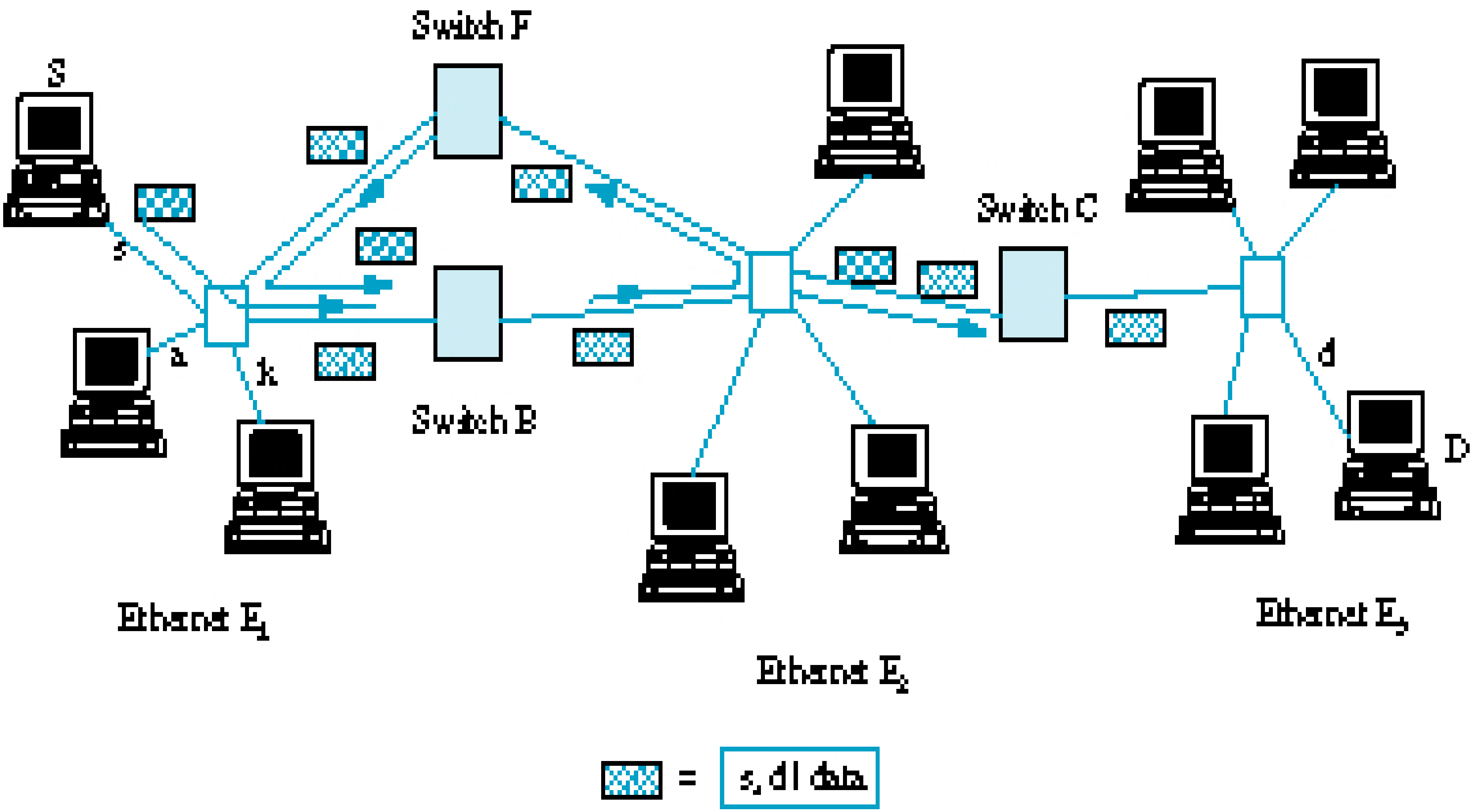
- In a single Ethernet two hosts cannot transmit simultaneously (collisions).
- A **switch** can break-up an Ethernet into many Ethernets, allowing a number of simultaneous transmissions.



Packet forwarding in switched Ethernets



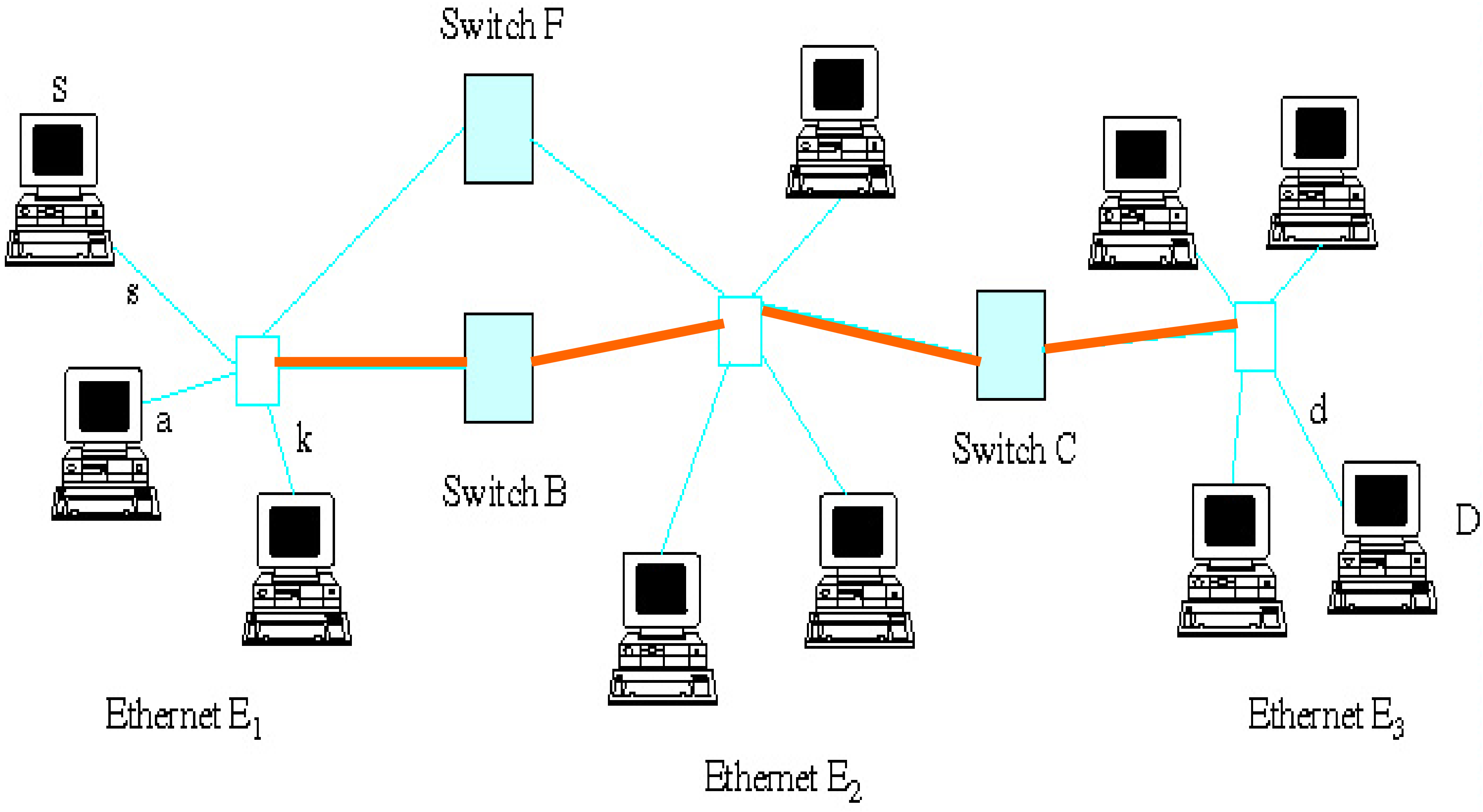
Packet forwarding in switched Ethernets: loops



Packet forwarding in switched Ethernets: loops

- Networks with loops desirable for reliability.
- However, loops should be avoided in forwarding:
  - Record all forwarded packets, do not re-forward already forwarded packet. Problem: too much bookkeeping.
  - Temporarily disable some of the links to break the loop: form **spanning-tree** (network without loops, where all hosts are connected).

Packet forwarding in switched Ethernets: loops



Packet forwarding in switched Ethernets: loops

Minimum spanning-tree algorithm:

- Network represented as a **graph**:
  - Nodes of the graph are Ethernets or switches.
  - Edge means a switch is connected to an Ethernet.
- Spanning tree: a sub-graph connecting all nodes (actually all “ethernet” nodes).
- Minimum spanning tree: a spanning tree with minimum number of edges.

Packet forwarding in switched Ethernets: loops

- Spanning-tree algorithm:
  - Centralized: O(m log(m)) complexity, where m is number of edges in the graph.
  - **Distributed** algorithm: non-trivial (IEEE 802.1 standard).
    - Has to be implemented among switches.
    - Switches are "blind": they only communicate by messages.
    - Steps in the algorithm to be "agreed upon" by all switches, e.g. election of root switch.
    - Final distributed knowledge has to be consistent.