Chapter 6 Link Layer and LANs

A note on the use of these ppt slides:

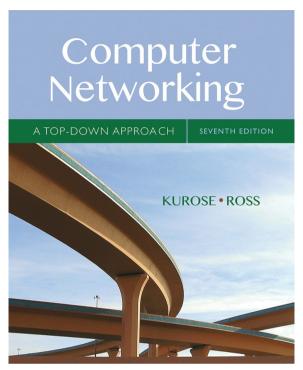
We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you see the animations; and can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a lot of work on our part. In return for use, we only ask the following:

- If you use these slides (e.g., in a class) that you mention their source (after all, we'd like people to use our book!)
- If you post any slides on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

All material copyright 1996-2016
J.F Kurose and K.W. Ross, All Rights Reserved

Nota: Conteúdo atualizado, UMinho, PMC, 2022.



Computer
Networking: A Top
Down Approach
6th and 7th edition
Jim Kurose, Keith Ross
Addison-Wesley
March 2012, 2017.

Chapter 6: Link layer and LANs

our goals:

- understand principles behind link layer services:
 - error detection, correction
 - sharing a broadcast channel: multiple access
 - link layer addressing
 - local area networks: Ethernet, VLANs
- instantiation, implementation of various link layer technologies

Link layer, LANs: outline

- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- **6.4 LANs**
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

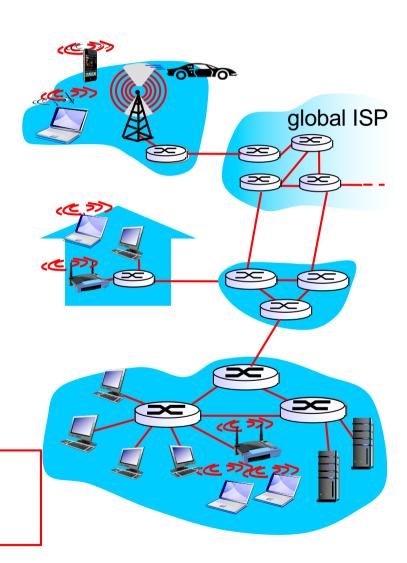
- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

Link layer: introduction

terminology:

- hosts and routers: nodes
- communication channels that connect adjacent nodes along communication path: links
 - wired links
 - wireless links
 - LANs
- layer-2 packet: frame, encapsulates datagram

data-link layer has responsibility of transferring datagram from one node to physically adjacent node over a link



Link layer: context

- datagram transferred by different link protocols over different links:
 - e.g., Ethernet on first link, frame relay on intermediate links, 802.11 on last link
- each link protocol provides different services
 - e.g., may or may not provide reliable data transfer (rdt) over each link

transportation analogy:

- trip from Princeton to Lausanne
 - car: Princeton to JFK
 - plane: JFK to Geneva
 - train: Geneva to Lausanne
- tourist = datagram
- transport segment = communication link
- transportation mode = link layer protocol
- travel agent = routing algorithm

Link layer services

- framing, link access:
 - encapsulate datagram into frame, adding header, trailer
 - obtain channel access if shared medium
 - "MAC" addresses used in frame headers to identify source, destination
 - different from IP address!
- reliable delivery between adjacent nodes
 - seldom used on low bit-error link (fiber, some twisted pair)
 - wireless links: high error ratios
 - Q: why both link-level and end-end reliability?

Link layer services (more)

flow control:

pacing between adjacent sending and receiving nodes

error detection:

- errors caused by signal attenuation, noise.
- receiver detects presence of errors:
 - signals sender for retransmission or drops frame (rdt vs. no rdt)

error correction:

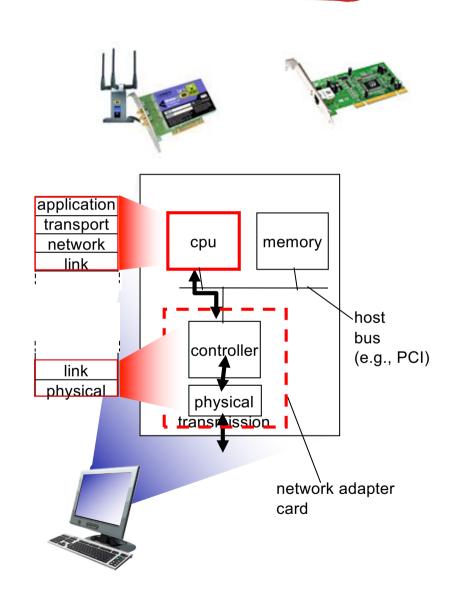
 receiver identifies and corrects bit error(s) without resorting to retransmission

half-duplex and full-duplex

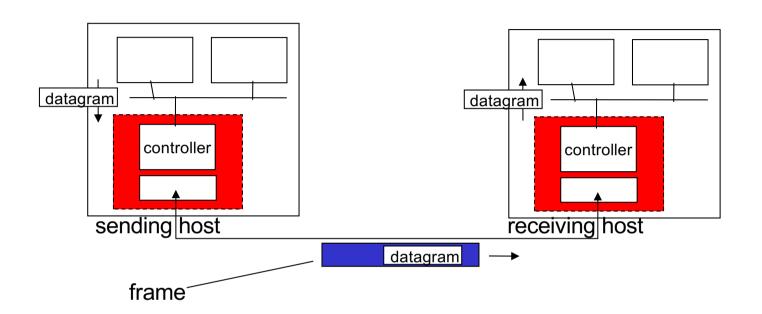
 with half duplex, nodes at both ends of link can transmit, but not at same time

Where is the link layer implemented?

- in each and every host
- link layer implemented in "adaptor" (aka network interface card NIC) or on a chip
 - Ethernet card, 802.11 card; Ethernet chipset
 - implements link, physical layer
- attaches into host's system buses
- combination of hardware, software, firmware



Adaptors communicating



- sending side:
 - encapsulates datagram in frame
 - adds (eventually) error checking bits, reliability (rdt), flow control, etc.

- receiving side
 - if in use, looks for errors, rdt, flow control, etc.
 - extracts datagram, passes to upper layer at receiving side

Link layer, LANs: outline

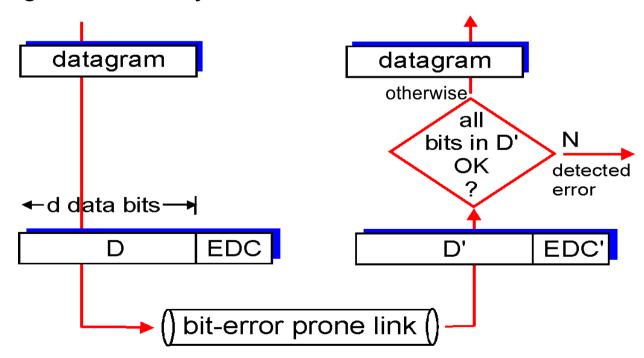
- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- **6.4 LANs**
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

Error detection

EDC= Error Detection and Correction bits (redundancy)

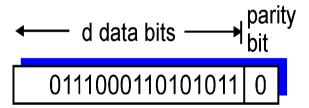
- D = Data protected by error checking, may include header fields
- Error detection not 100% reliable!
 - protocol may miss some errors, but rarely
 - larger EDC field yields better detection and correction



Parity checking

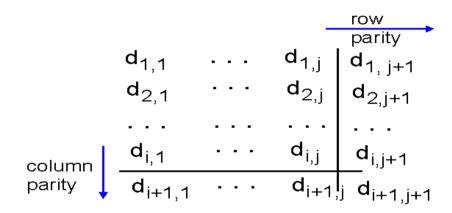
single bit parity:

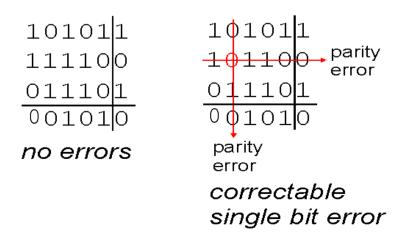
detect single bit errors



two-dimensional bit parity:

detect and correct single bit errors





Internet checksum (review)

goal: detect "errors" (e.g., flipped bits) in transmitted packet (note: used at transport layer only)

sender:

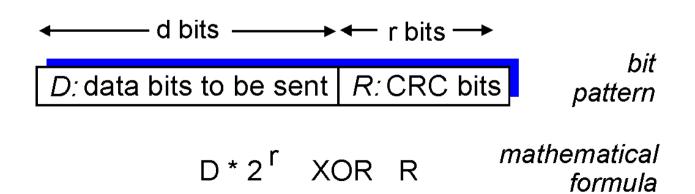
- treat segment contents as sequence of 16-bit integers
- checksum: addition (I's complement sum) of segment contents
- sender puts checksum value into UDP checksum field

receiver:

- compute checksum of received segment
- check if computed checksum equals checksum field value:
 - NO error detected
 - YES no error detected. But maybe errors nonetheless?

Cyclic redundancy check

- more powerful error-detection coding
- view data bits, D, as a binary number
- choose r+1 bit pattern (generator), G
- goal: choose r CRC bits, R, such that
 - <D,R> exactly divisible by G (modulo 2)
 - receiver knows G, divides <D,R> by G. If non-zero remainder: error detected!
 - can detect all burst errors less than r+1 bits
- widely used in practice (Ethernet, 802.11 WiFi, ATM)



CRC example

want:

 $D2^r XOR R = nG$

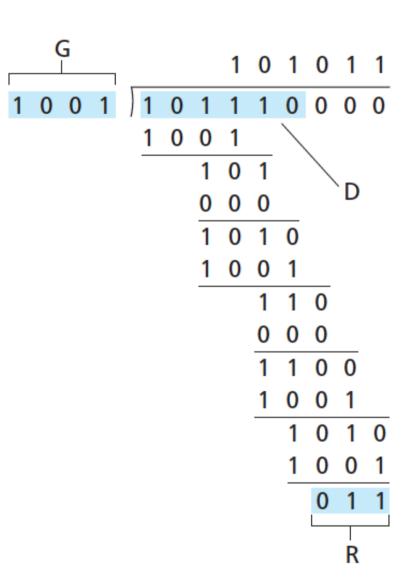
equivalently:

 $D2^r = nG XOR R$

equivalently:

if we divide D2^r by G, want remainder R to satisfy:

R = remainder[
$$\frac{D \cdot 2^r}{G}$$
]



Link layer, LANs: outline

- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- **6.4 LANs**
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

Multiple access links, protocols

two types of "links":

- point-to-point
 - (PPP for dial-up access; HDLC for point-to-point or point-to-multipoint)
 - point-to-point link between Ethernet switch, host
- broadcast (shared wire or medium)
 - old-fashioned Ethernet.
 - upstream HFC (Hybrid Fiber Coax)
 - 802.11 wireless LAN



shared wire (e.g., cabled Ethernet)



shared RF (e.g., 802.11 WiFi)



shared RF (satellite)



humans at a cocktail party (shared air, acoustical)

Multiple access protocols

- single shared broadcast channel
- two or more simultaneous transmissions by nodes: interference
 - collision if node receives two or more signals at the same time

multiple access protocol

- distributed algorithm that determines how nodes share channel, i.e., determine when node can transmit
- communication about channel sharing must use channel itself!
 - no out-of-band channel for coordination

An ideal multiple access protocol

given: broadcast channel of rate R bps desiderata: (seria desejável...)

- I. when one node wants to transmit, it can send at rate R.
- 2. when M nodes want to transmit, each can send at average rate R/M
- 3. fully decentralized:
 - no special node to coordinate transmissions
 - no synchronization of clocks, slots
- 4. simple

MAC protocols: taxonomy

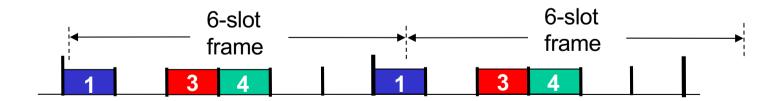
three broad classes:

- channel partitioning
 - divide channel into smaller "pieces" (time slots, frequency,...)
 - allocate piece to node for <u>exclusive</u> use
- random access
 - channel not divided, allow collisions
 - "recover" from collisions
- "taking turns"
 - nodes take turns, but nodes with more to send can take longer turns

Channel partitioning MAC protocols:TDMA

TDMA: time division multiple access

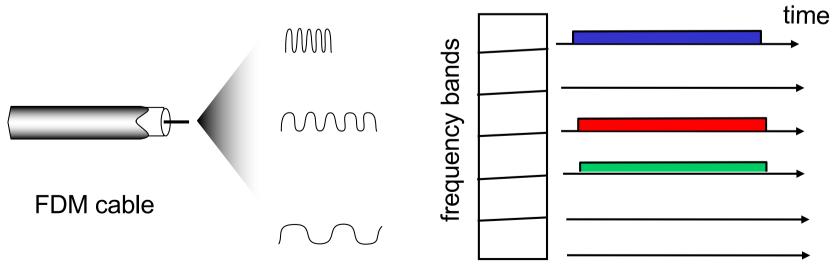
- channel capacity is divided in time frames, and then time frames in time slots
- access to channel in "rounds"
- each station gets fixed length slot (length = pkt transmission time) in each round (each transmitting at R/N)
- unused slots go idle
- example: 6-station LAN, 1,3,4 have pkt, slots 2,5,6 idle
- Pros: no collisions, fair. Cons: time wait, may waste capacity



Channel partitioning MAC protocols: FDMA

FDMA: frequency division multiple access

- channel spectrum divided into frequency bands
- each station assigned fixed frequency band
- unused transmission time in frequency bands goes idle
- example: 6-station LAN, 1,3,4 have pkt, frequency bands 2,5,6 idle



Random access protocols

- when node has packet to send
 - transmit at full channel data rate R.
 - no a priori coordination among nodes
- ❖ two or more transmitting nodes → "collision"
- random access MAC protocol specifies:
 - how to detect collisions
 - how to recover from collisions (e.g., via delayed retransmissions)
- examples of random access MAC protocols:
 - ALOHA; slotted ALOHA (not covered in SCR)
 - CSMA, CSMA/CD, CSMA/CA

CSMA (carrier sense multiple access)

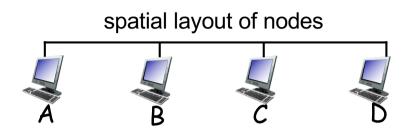
CSMA: listen before transmit:

- * if channel sensed idle: transmit entire frame
- if channel sensed busy, defer transmission

human analogy: don't interrupt others!

CSMA collisions

- collisions can still occur: propagation delay means two nodes may not hear each other's transmission
- collision: entire packet transmission time wasted
 - distance & propagation delay play role in in determining collision probability





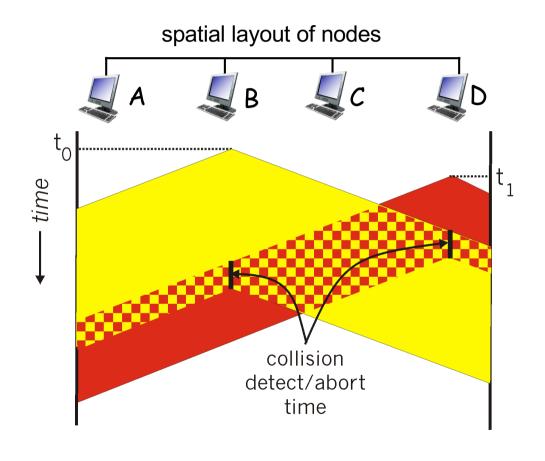
 t_1

CSMA/CD (collision detection)

CSMA/CD: carrier sensing, deferral as in CSMA

- collisions detected within short time
- colliding transmissions aborted, reducing channel wastage
- collision detection:
 - easy in wired LANs: measure signal strengths, compare transmitted, received signals
 - difficult in wireless LANs: received signal strength overwhelmed by local transmission strength
- human analogy: the polite conversationalist

CSMA/CD (collision detection)



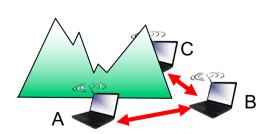
Ethernet CSMA/CD algorithm

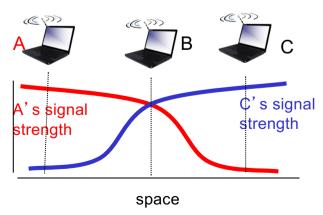
- I. NIC receives datagram from network layer, creates frame
- 2. If NIC senses channel idle, starts frame transmission. If NIC senses channel busy, waits until channel idle, then transmits.
- 3. If NIC transmits entire frame without detecting another transmission, NIC is done with frame!

- 4. If NIC detects another transmission while transmitting, aborts and sends jam signal
- 5. After aborting, NIC enters binary (exponential) backoff:
 - after mth collision, NIC chooses K at random from {0,1,2,..., 2^m-1}. NIC waits K·512 bit times, returns to Step 2
 - longer backoff interval with more collisions

CSMA/CD fails in WLANs

- In WLANs collision detection does not work properly!
 - difficult to receive (sense collisions) when transmitting due to weak received signals (fading)
 - can't sense all collisions in any case: hidden terminal, fading
- Goal: avoid collisions CSMA/C(ollision)A(voidance)
- CSMA sense before transmitting
 - don't collide with ongoing transmission by other node
- ❖ CA scheme based on proper timing and ACKs (see IEEE 802.11)





"Taking turns" MAC protocols

channel partitioning MAC protocols:

- share channel efficiently and fairly at high load
- inefficient at low load: delay in channel access, I/N bandwidth allocated even if only I active node!

random access MAC protocols

- efficient at low load: single node can fully utilize channel
- high load: collision overhead

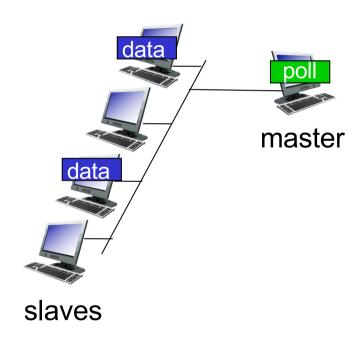
"taking turns" protocols

look for best of both worlds!

"Taking turns" MAC protocols

polling:

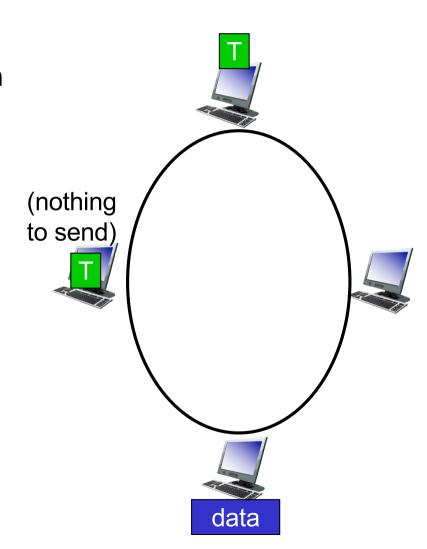
- master node "invites" slave nodes to transmit in turn
- typically used with "dumb" slave devices
- concerns:
 - polling overhead
 - latency
 - single point of failure (master)



"Taking turns" MAC protocols

token passing:

- control token passed from one node to next sequentially.
- token message
- concerns:
 - token overhead
 - latency
 - single point of failure (token)



Summary of MAC protocols

- channel partitioning, by time, frequency (or code)
 - Time Division, Frequency Division
- random access (dynamic),
 - ALOHA, S-ALOHA, CSMA, CSMA/CD
 - carrier sensing: easy in some technologies (wire), hard in others (wireless)
 - CSMA/CD used in Ethernet
 - CSMA/CA used in 802.11
- taking turns
 - polling from central site
 - token passing: bluetooth, FDDI, token ring

Link layer, LANs: outline

- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- 6.4 LANs
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

MAC addresses and ARP

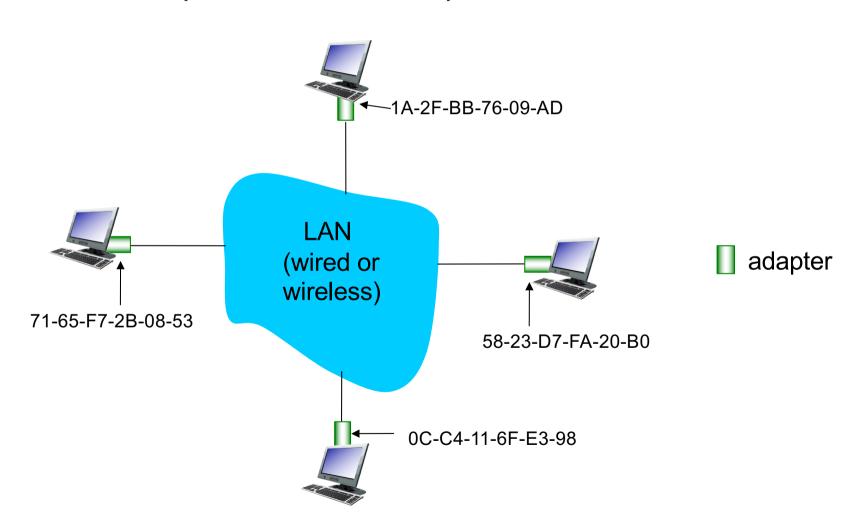
- 32-bit IP address:
 - network-layer address for interface
 - used for layer 3 (network layer) forwarding
- MAC (or LAN or physical or Ethernet) address:
 - function: used locally to get frame from one interface to another physically-connected interface (same network, in IP-addressing sense)
 - 48 bit MAC address (for most LANs) burned in NIC ROM, also sometimes software settable
 - e.g.: IA-2F-BB-76-09-AD, also as company_ID-76-09-AD

hexadecimal (base 16) notation (each "number" represents 4 bits)

1st 3 octets represent company_ID (e.g. D-Link, Cisco, HP, Intel, etc.)

LAN addresses and ARP

each adapter on LAN has unique LAN address

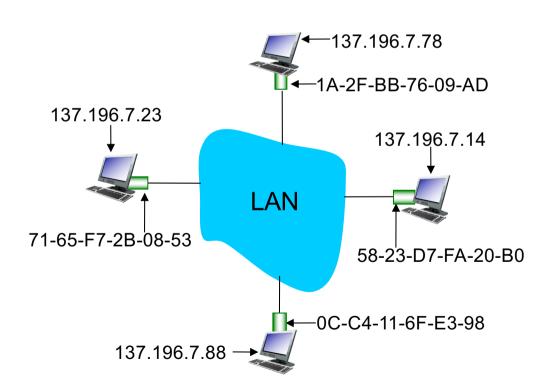


LAN addresses (more)

- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)
- analogy:
 - MAC address: like Social Security Number
 - IP address: like postal address
- ❖ MAC flat address → portability
 - can move LAN card from one LAN to another
- IP hierarchical address not portable
 - address depends on IP subnet to which node is attached

ARP: address resolution protocol

Question: how to determine interface's MAC address, knowing its IP address?



ARP table: each IP node (host, router) on LAN has table

- IP/MAC address mappings for some LAN nodes:
- < IP address; MAC address; TTL>
- TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

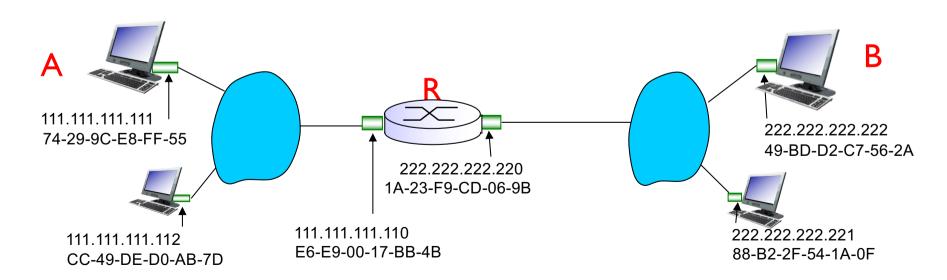
ARP protocol: same LAN

- A wants to send datagram to B
 - B's MAC address not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address
 - destination MAC address = FF-FF-FF-FF-FF
 - all nodes on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC address
 - frame sent to A's MAC address (unicast)

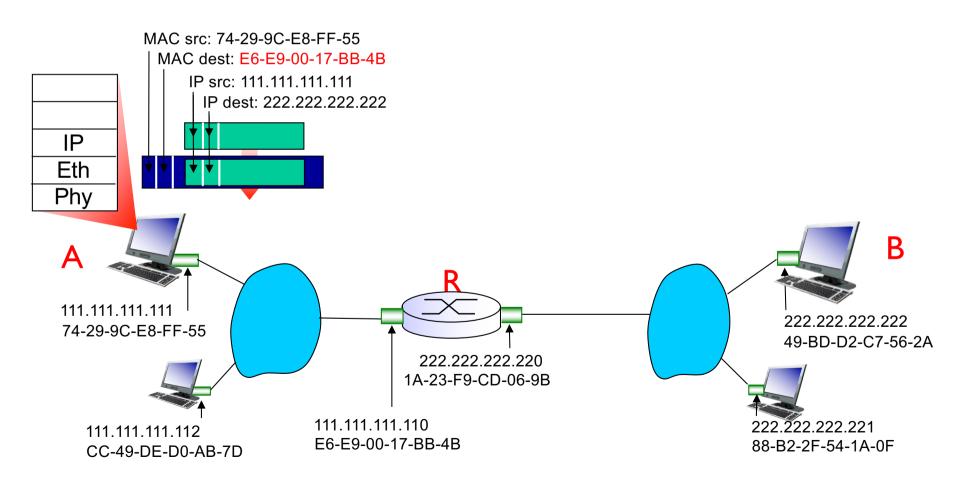
- A caches (saves) IP-to-MAC address pair in its ARP table until information becomes old (times out)
 - soft state: information that times out (goes away) unless refreshed
- ARP is "plug-and-play":
 - nodes create their ARP tables without intervention from net administrator

walkthrough: send datagram from A to B via R

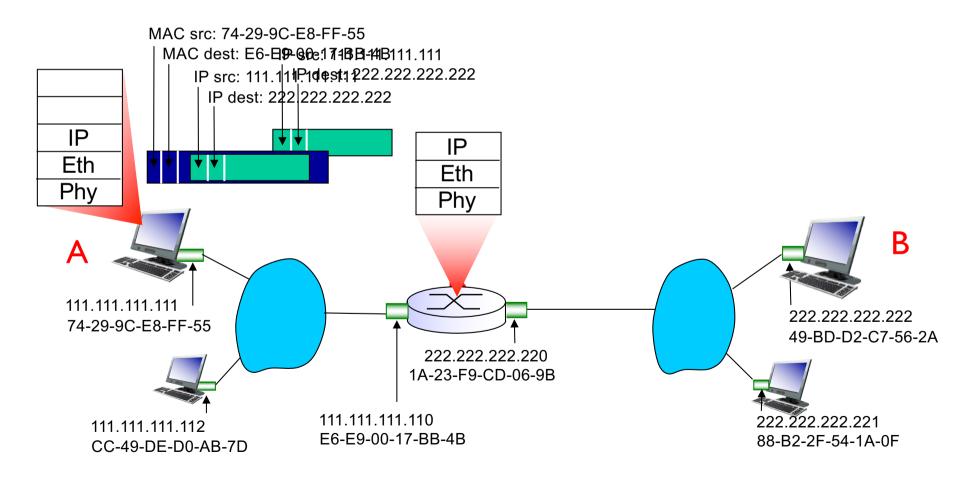
- focus on addressing at IP (datagram) and MAC layer (frame)
- assume A knows B's IP address
- assume A knows IP address of first hop router, R (how?)
- assume A knows R's MAC address (how?)



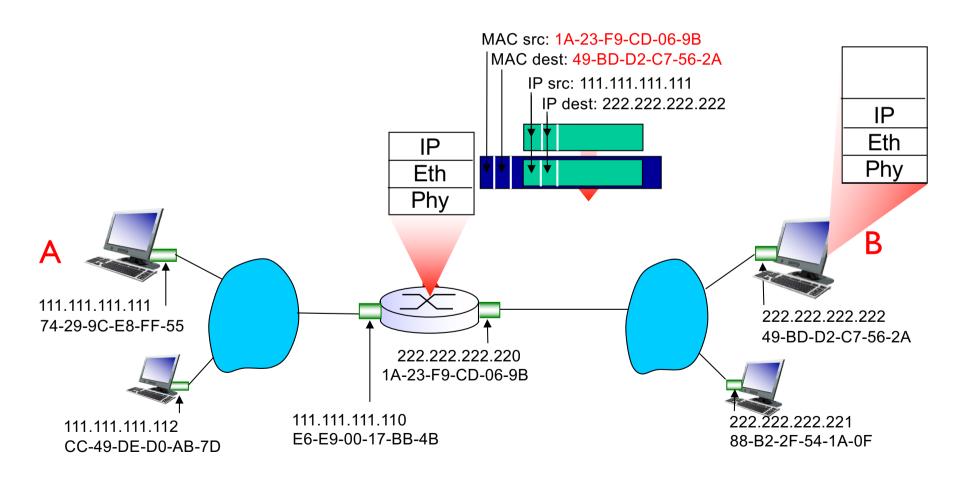
- A creates IP datagram with IP source A, destination B
- A creates link-layer frame with R's MAC address as destination address, frame contains A-to-B IP datagram



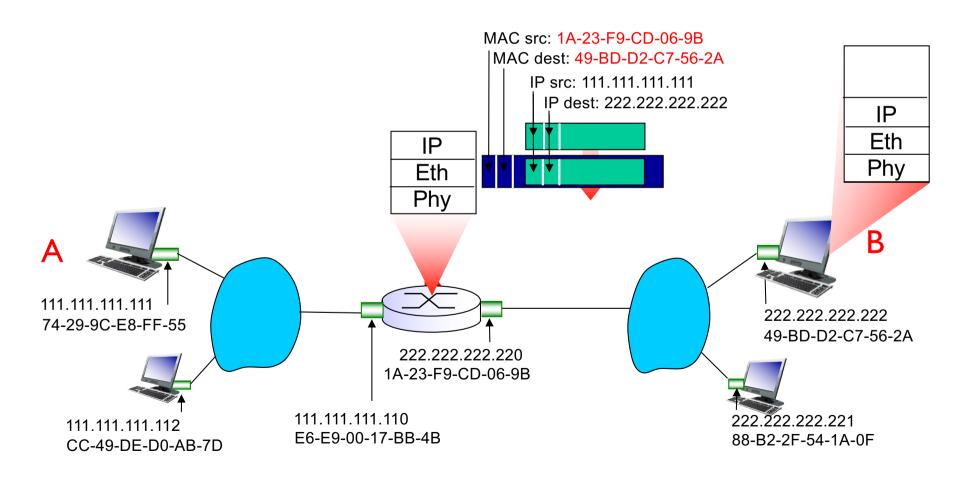
- frame sent from A to R
- frame received at R, datagram removed, passed up to IP



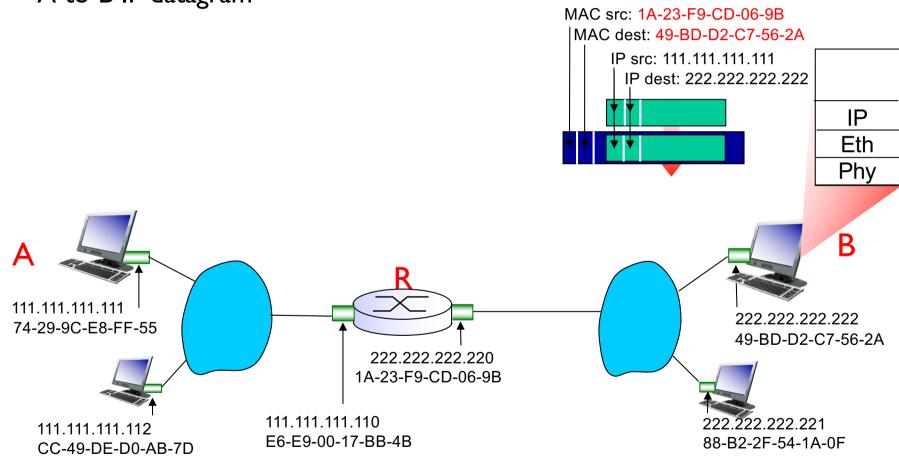
- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as destination address, frame contains A-to-B IP datagram



- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as destination address, frame contains A-to-B IP datagram



- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram



^{*} Check out the online interactive exercises for more examples: http://gaia.cs.umass.edu/kurose_ross/interactive/

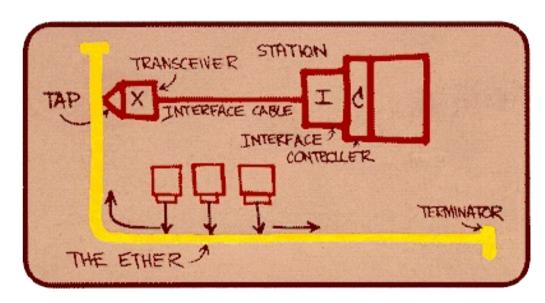
Link layer, LANs: outline

- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- 6.4 LANs
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

Ethernet

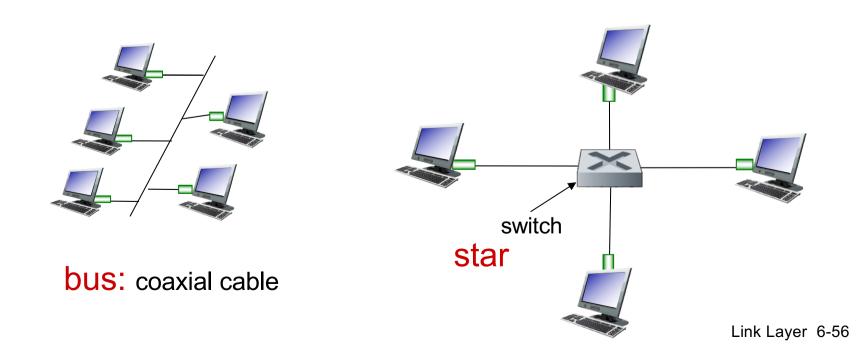
- "dominant" wired LAN technology:
- first widely used LAN technology
- simple, cheap
- single chip, multiple speeds (e.g., Broadcom BCM5761)
- ★ kept up with speed race: I0 Mbps 40 Gbps



Metcalfe's Ethernet hand sketch, 1976

Ethernet: physical topology

- bus: popular through mid 90s
 - all nodes in same collision domain (can collide with each other)
- star: prevails today
 - active switch in center
 - each "spoke" runs a (separate) Ethernet protocol (nodes do not collide with each other)



Ethernet frame structure

sending adapter encapsulates IP datagram (or other network layer protocol packet) in Ethernet frame



preamble:

- 7 bytes with pattern 10101010 followed by one byte with pattern 10101011
- used to synchronize receiver, sender clock rates

Ethernet frame structure (more)

- addresses: 6 byte source, destination MAC addresses
 - if adapter receives frame with matching destination address, or with broadcast address (e.g. ARP packet), it passes data in frame to network layer protocol
 - otherwise, adapter discards frame
- type: indicates higher layer protocol (mostly IP but others possible, e.g., ARP, Novell IPX, AppleTalk)
- CRC: cyclic redundancy check at receiver
 - aka Frame Check Sequence (FCS)
 - error detected: frame is dropped

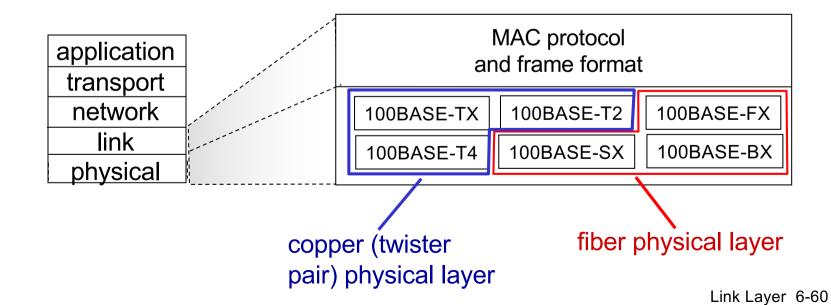


Ethernet: unreliable, connectionless

- connectionless: no handshaking between sending and receiving NICs
- unreliable: receiving NIC doesn't send acks or negative acks (nacks) to sending NIC
 - data in dropped frames recovered only if initial sender uses higher layer reliable data transfer (rdt) protocol (e.g., TCP), otherwise dropped data is lost
- Ethernet's MAC protocol: unslotted CSMA/CD (Carrier Sense Multiple Access / Collision Detection) with binary backoff.

802.3 Ethernet standards: link & physical layers

- many different Ethernet standards
 - common MAC protocol and frame format
 - different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10Gbps, 40 Gbps
 - different physical layer media: fiber, cable



Link layer, LANs: outline

- 6. I introduction, services
- 6.2 error detection, correction
- 6.3 multiple access protocols
- 6.4 LANs
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

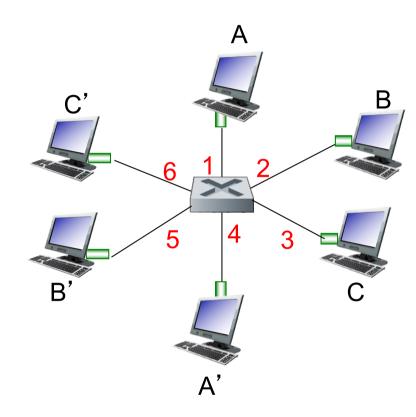
- 6.5 link virtualization: MPLS
- 6.6 data center networking
- 6.7 a day in the life of a web request

Ethernet switch

- link-layer device: takes an active role
 - store, forward Ethernet frames
 - examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links when frame is to be forwarded on segment, uses CSMA/CD to access segment
- transparent
 - hosts are unaware of presence of switches
- plug-and-play, self-learning
 - switches do not need to be configured

Switch: multiple simultaneous transmissions

- hosts have dedicated, direct connection to switch
- switches buffer packets
- Ethernet protocol used on each incoming link, but no collisions; full duplex
 - each link is its own collision domain
- switching: A-to-A' and B-to-B' can transmit simultaneously, without collisions



switch with six interfaces (1,2,3,4,5,6)

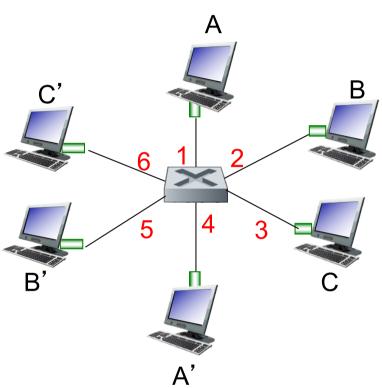
Switch forwarding table

Q: how does switch know A' reachable via interface 4, B' reachable via interface 5?

- * A: each switch has a switch table, each entry:
 - (MAC address of host, interface to reach host, timestamp)
 - looks like a routing table!

Q: how are entries created, maintained in switch table?

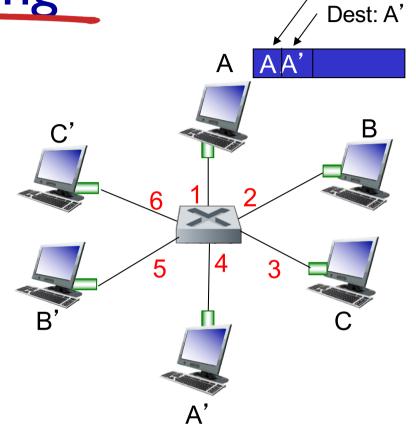
something like a routing protocol?



switch with six interfaces (1,2,3,4,5,6)

Switch: self-learning

- switch learns which hosts can be reached through which interfaces
 - when frame received, switch "learns" location of sender: incoming LAN segment
 - records sender/location pair in switch table



MAC addr	interface	TTL
Α	1	60

Switch table (initially empty)

Source: A

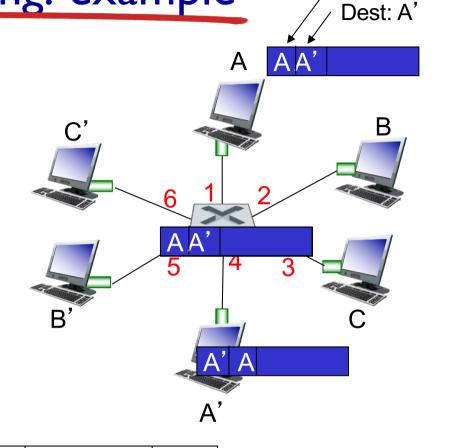
Switch: frame filtering/forwarding

when frame received at switch:

```
I. record incoming link, MAC address of sending host
2. index switch table using MAC destination address
3. if entry found for destination then {
    if destination on segment from which frame arrived then drop frame
    else forward frame on interface indicated by entry
    }
    else flood /* forward on all interfaces except arriving interface */
```

Self-learning, forwarding: example

- frame destination, A', location unknown: flood
- destination A location known: selectively send on just one link



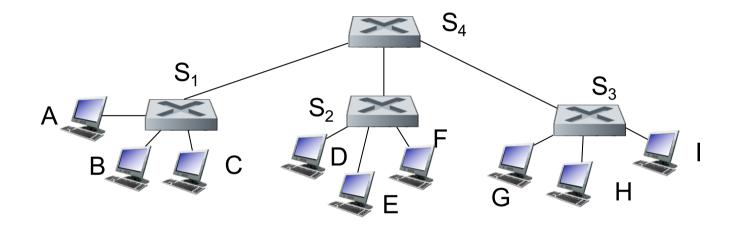
MAC addr	interface	TTL
Α	1	60
Α'	4	60

switch table (initially empty)

Source: A

Interconnecting switches

switches can be connected together

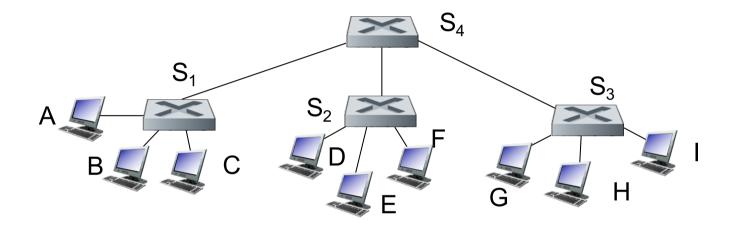


Q: sending from A to G - how does S_1 know to forward frame destined to G via S_4 and S_3 ?

*A: self learning! (works exactly the same as in single-switch case!)

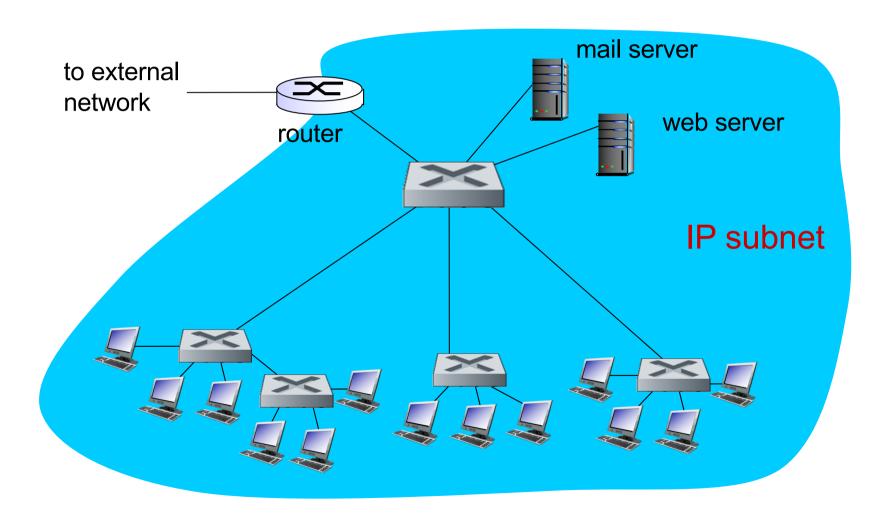
Self-learning multi-switch example

Suppose C sends frame to I, I responds to C



* Q: show switch tables and packet forwarding in S_1 , S_2 , S_3 , S_4

Institutional network



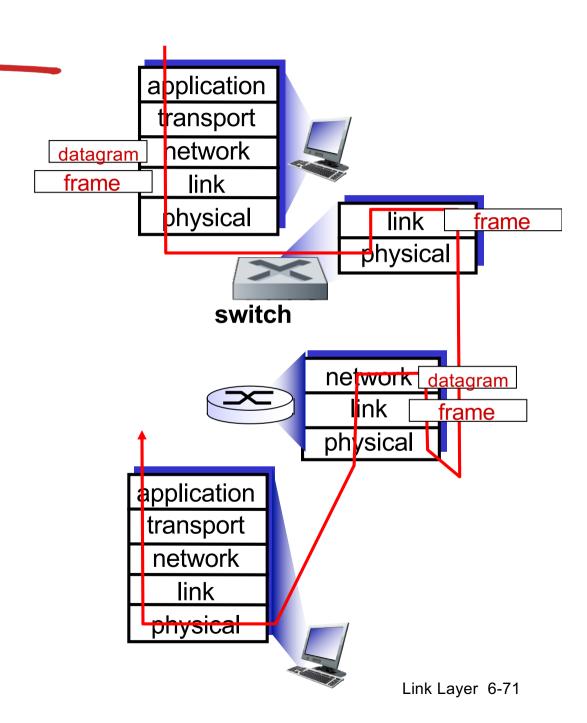
Switches vs. routers

both are store-and-forward:

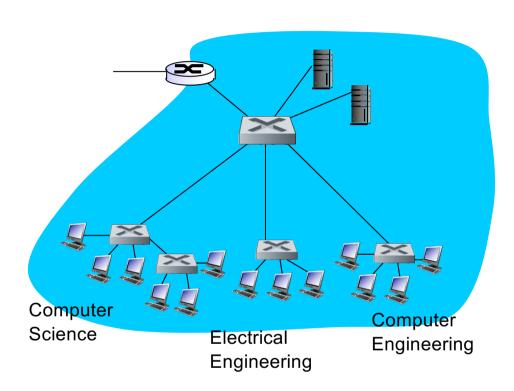
- routers: network-layer devices (examine network-layer headers)
- switches: link-layer devices (examine link-layer headers)

both have forwarding tables:

- routers: compute tables using routing algorithms, IP addresses
- *switches: learn forwarding table using flooding, learning, MAC addresses



VLANs: motivation



consider:

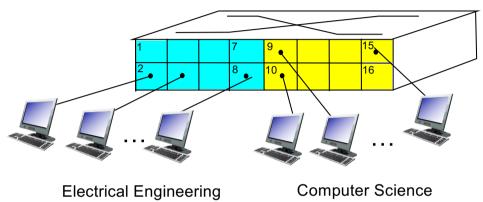
- CS user moves office to EE, but wants connect to CS switch?
- single broadcast domain:
 - all layer-2 broadcast traffic (ARP, DHCP, unknown location of destination MAC address) must cross entire LAN
 - security/privacy, efficiency issues

VLANs

Virtual Local Area Network

switch(es) supporting VLAN capabilities can be configured to define multiple virtual LANS over single physical LAN infrastructure.

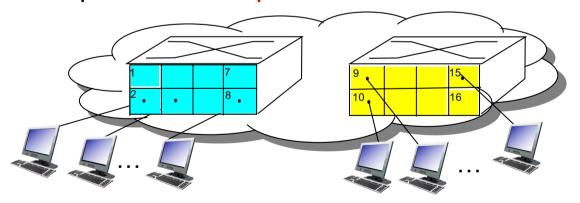
port-based VLAN: switch ports grouped (by switch management software) so that single physical switch



Electrical Engineering (VLAN ports 1-8)

Computer Science (VLAN ports 9-15)

... operates as multiple virtual switches



Electrical Engineering (VLAN ports 1-8)

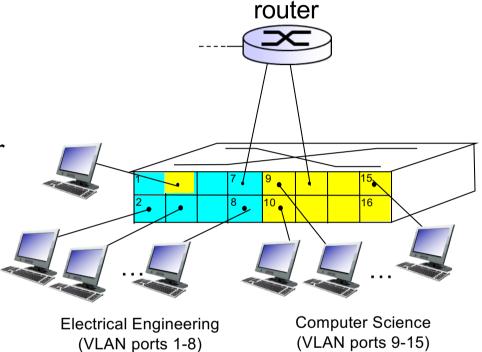
Computer Science (VLAN ports 9-16)

Port-based VLAN

traffic isolation: frames to/from ports I-8 can only reach ports
 I-8

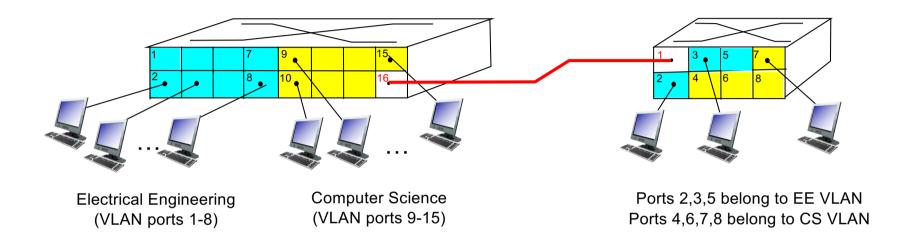
 can also define VLAN based on MAC addresses of endpoints, rather than switch port

dynamic membership: ports can be dynamically assigned among VLANs



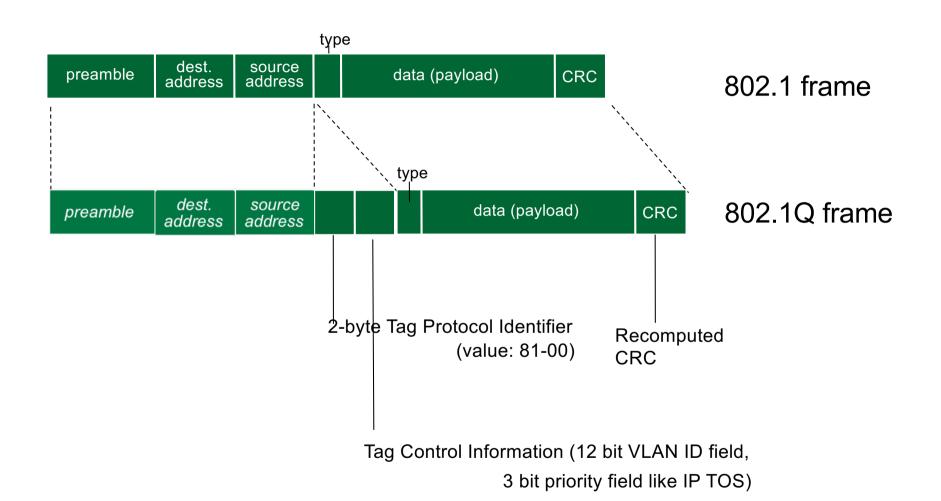
- forwarding between VLANS: done via routing (just as with separate switches)
 - in practice vendors sell combined switches plus routers

VLANS spanning multiple switches



- trunk port: carries frames between VLANS defined over multiple physical switches
 - frames forwarded within VLAN between switches can't be vanilla 802.1 frames (must carry VLAN ID info)
 - 802. I q protocol adds/removed additional header fields for frames forwarded between trunk ports

802. I Q VLAN frame format



Link layer, LANs: outline

- 5.1 introduction, services
- 5.2 error detection, correction
- 5.3 multiple access protocols
- **5.4** LANs
 - addressing, ARP
 - Ethernet
 - switches
 - VLANS

- 5.5 link virtualization: MPLS
- 5.6 data center networking
- 5.7 a day in the life of a web request

Fim do capítulo para RC

Chapter 5: Summary

- principles behind data link layer services:
 - error detection, correction
 - sharing a broadcast channel: multiple access
 - link layer addressing
- instantiation and implementation of various link layer technologies
 - Ethernet
 - switched LANS, VLANs
 - virtualized networks as a link layer: MPLS
- synthesis: a day in the life of a web request