Virtual LAN,
VXLAN and Ethernet VPN (EVPN)

Professor Chien-Chao Tseng

曾建超 教授

Department of Computer Science

National Yang Ming Chiao Tung University

cctseng@cs.nctu.edu.tw

- Virtual Local Area Network (VLAN)
- QinQ (IEEE 802.1ad)
- Virtual Extensible Local Area Network (VLAN (VXLAN)
- SIMPLIFIED: WHY EVPN/VXLAN? By Mike Bushong, Juniper
- Ethernet VPN/VXLAN

LAN, IP, MAC and ARP

LAN and WAN

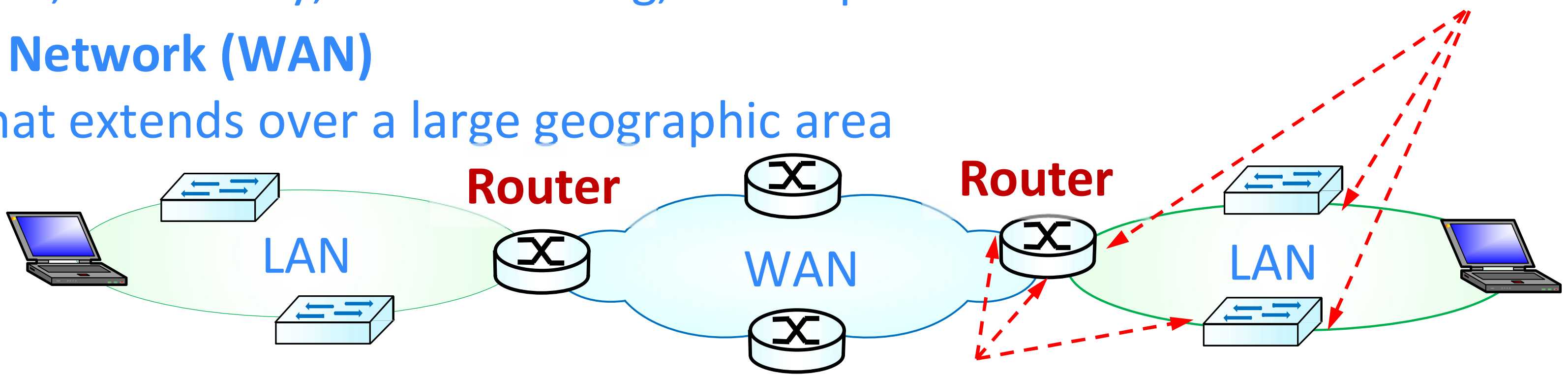
- **Local Area Network (LAN)**

  A network that interconnects computers within a limited area of **a single broadcast domain**

  - a residence, laboratory, office building, or campus

- **Wide Area Network (WAN)**

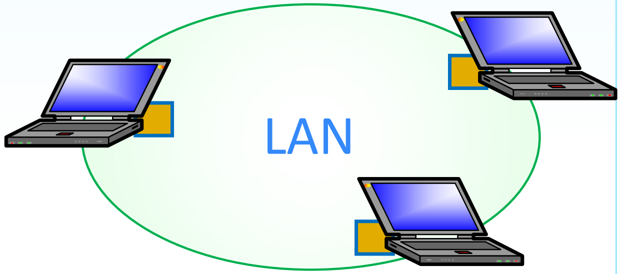
  Network that extends over a large geographic area

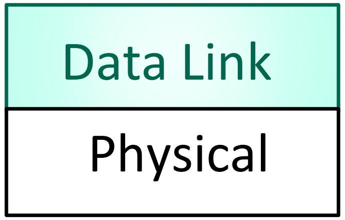


  - LAN Port: of a switch, router or WiFi AP
    a socket that allows you to connect your device to a local network,
  - WAN Port: of a switch, router or WiFi AP
    a socket that allow you to connect to the wider internet or other networks outside of your local network.

Network Interface Card and MAC Address

- Machines need Network Interface Cards (NICs) to connect to networks.
  - NIC aka Network/LAN Adapter, or Physical NW Interface
- Each NIC has a Media Access Control (MAC) Address
- MAC (aka LAN, Physical, or Ethernet) address:
  - 48 bits
  - Locally unique on a LAN
  - Flat address
    - Portable: can move NIC from one LAN to another
- Two machines on the same network
  - Can physically communicate with each other directly with MAC addresses





LAN, IP and MAC

- Host1 sends packets to Host2

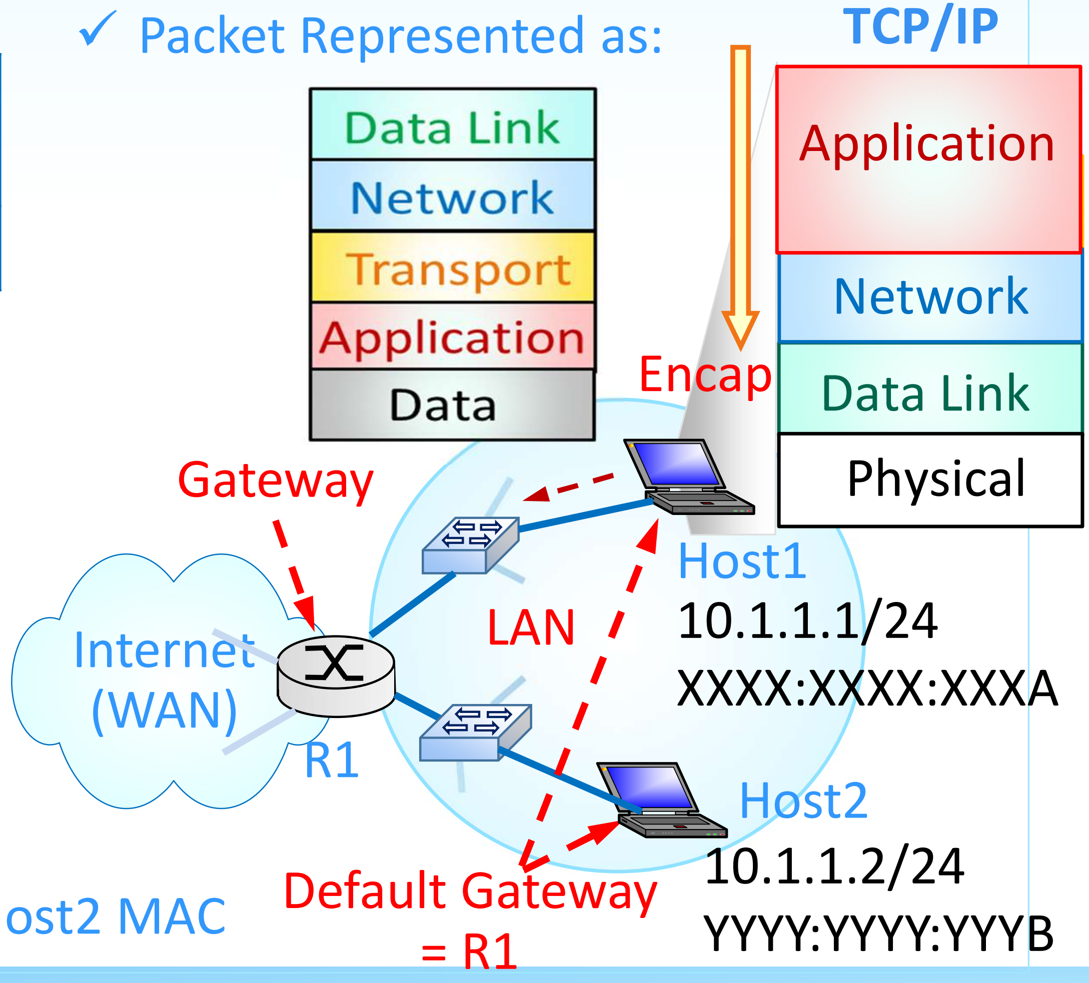
L3 Header	Src IP: 10.1.1.1 Dst IP: 10.1.1.2
Data	TCP ○ ○ ○

- After Data Link layer encapsulation

L2 Header	Src MAC: XXXX:XXXX:XXXA Dst MAC: YYYY:YYYY:YYYB
L3 Header	Src IP: 10.1.1.1 Dst IP: 10.1.1.2
Data	TCP ○ ○ ○

➢ Need a mechanism to map Host2 IP to Host2 MAC

✓ Packet Represented as:



Data Link
Network
Transport
Application
Data

TCP/IP
Application
Network
Data Link
Physical

Encap

Gateway

Internet (WAN)

R1

LAN

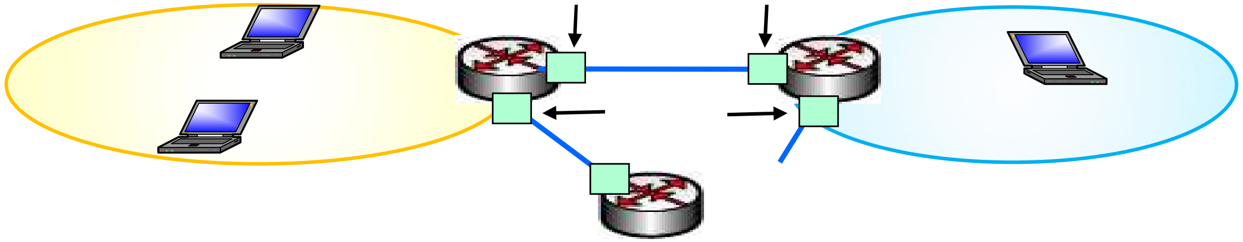
Host1
10.1.1.1/24
XXXX:XXXX:XXXA

Host2
10.1.1.2/24
YYYY:YYYY:YYYB

Default Gateway = R1

Address Resolution Protocol

- Address Resolution Protocol (ARP): mapping Internet addresses (IP) to physical addresses (MAC)
- Each IP node (host, router) has ARP table(s)
  - One table for each interface associated with a network (LAN)



- ARP table: IP/MAC mappings (for some nodes on the LAN)

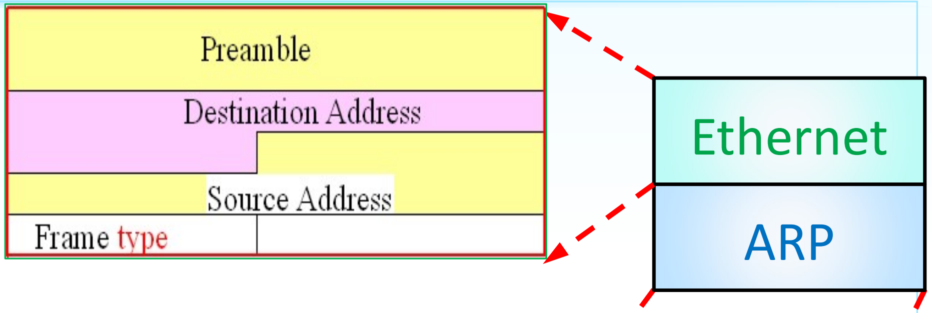
  < IP address; MAC address; TTL>

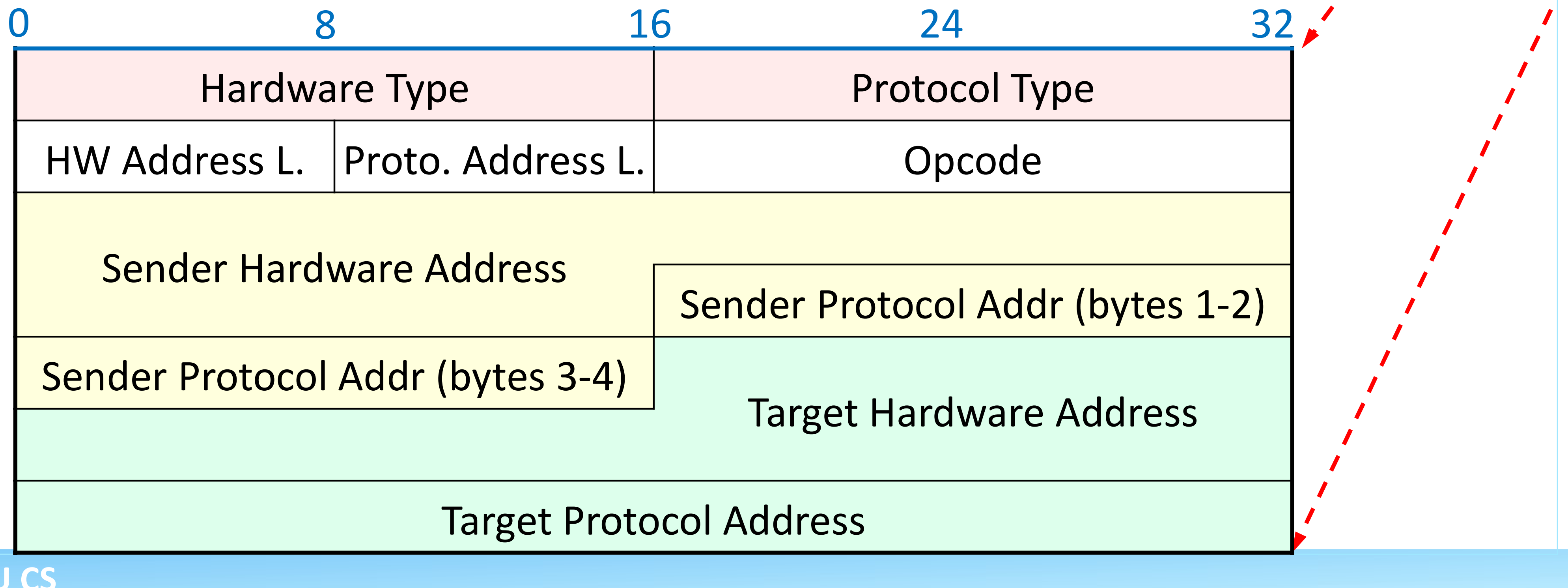
  - TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

- Soft State: information that times out (goes away) unless refreshed

ARP Message Format

- Used to resolve IP to MAC address mapping
- Ethernet frame type:
  - 0806$_{16}$ for ARP messages
- ARP Message Format:





Hardware Type | Protocol Type

HW Address L. | Proto. Address L. | Opcode

Sender Hardware Address

Sender Protocol Addr (bytes 1-2)

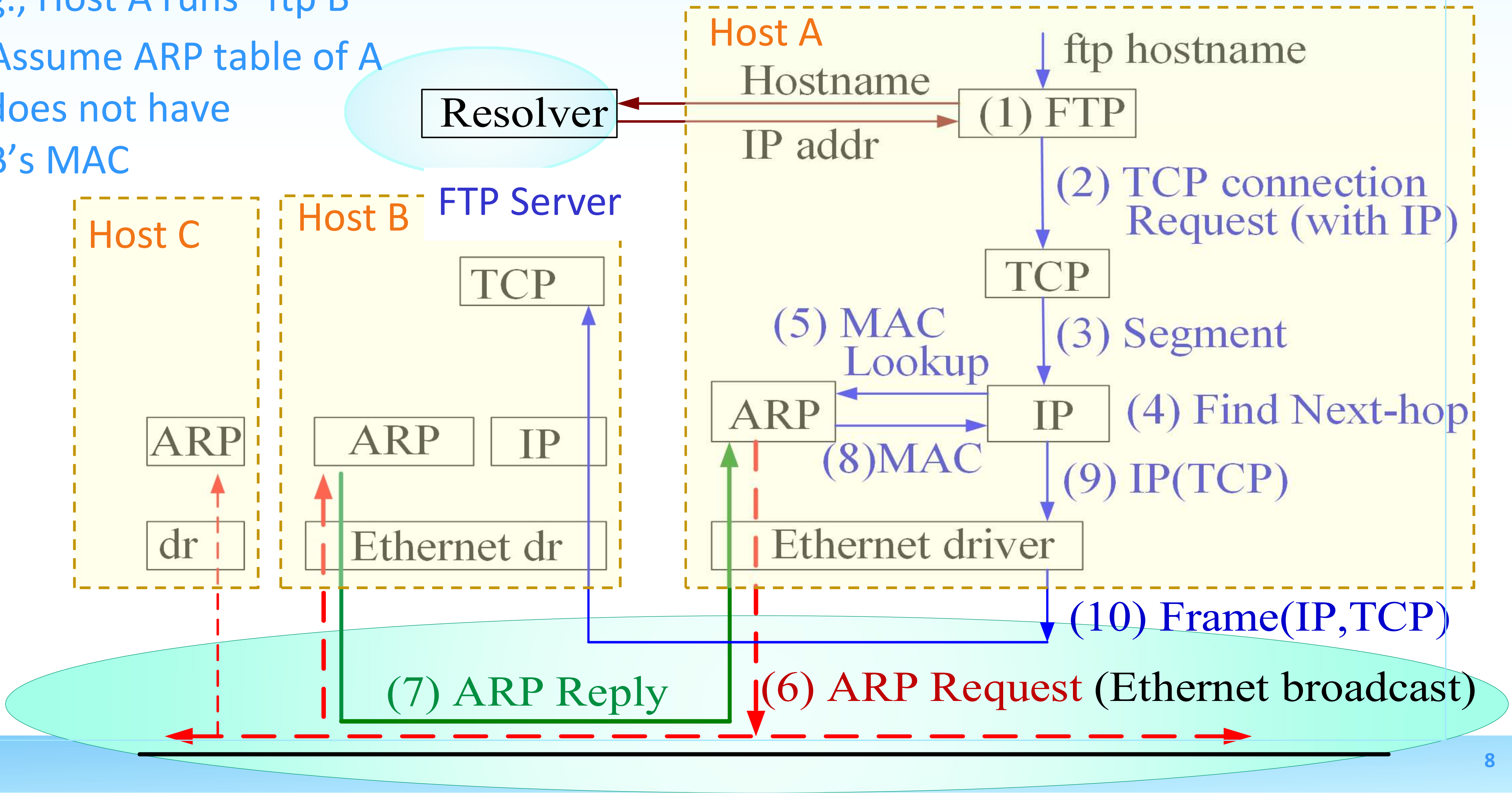
Sender Protocol Addr (bytes 3-4)

Target Hardware Address

Target Protocol Address

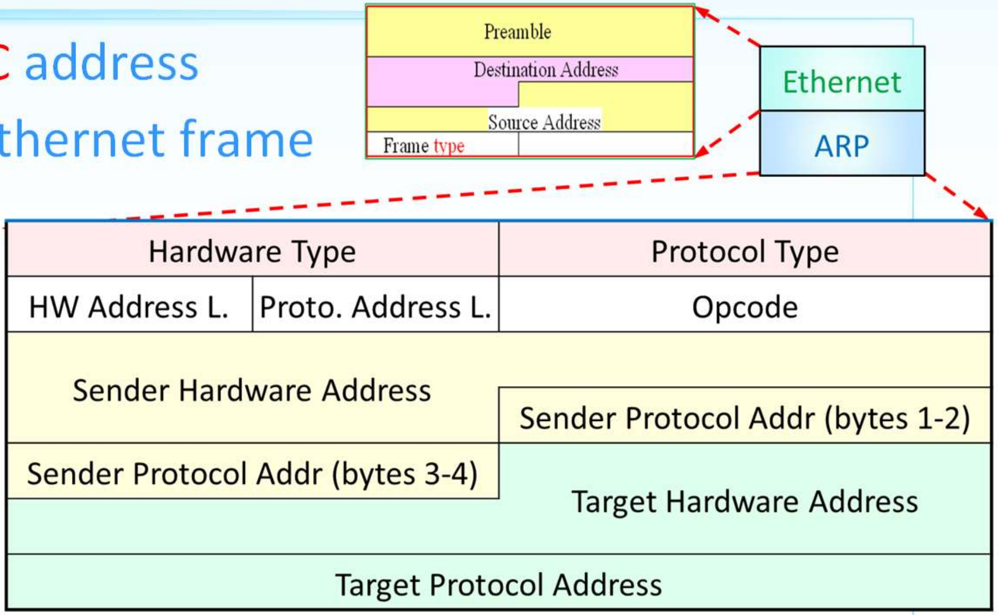
Procedure of ARP: Host A runs “ftp B”

- E.g., Host A runs “ftp B”
  - Assume ARP table of A does not have B’s MAC



Procedure of ARP: Host A runs “ftp B” (cont.)

- ✓ Assume A’s ARP table does not have B’s MAC address
- A broadcasts ARP request, encapsulated in Ethernet frame
  - – Ethernet Header: containing
    - Source MAC: A’s MAC
    - Destination MAC: FF-FF-FF-FF-FF-FF
  - – ARP Request message: containing
    - Sender A’s MAC, A’s IP address
    - Target B’s MAC all 0s, B's IP address,
- All machines on the LAN receive ARP Request
  - – B replies (unicast) to A with B's MAC address
  - ➢ A caches (saves) IP-to-MAC address pair in ARP Table
- ✓ ARP is “plug-and-play”:
  Create table automatically without intervention from administrator

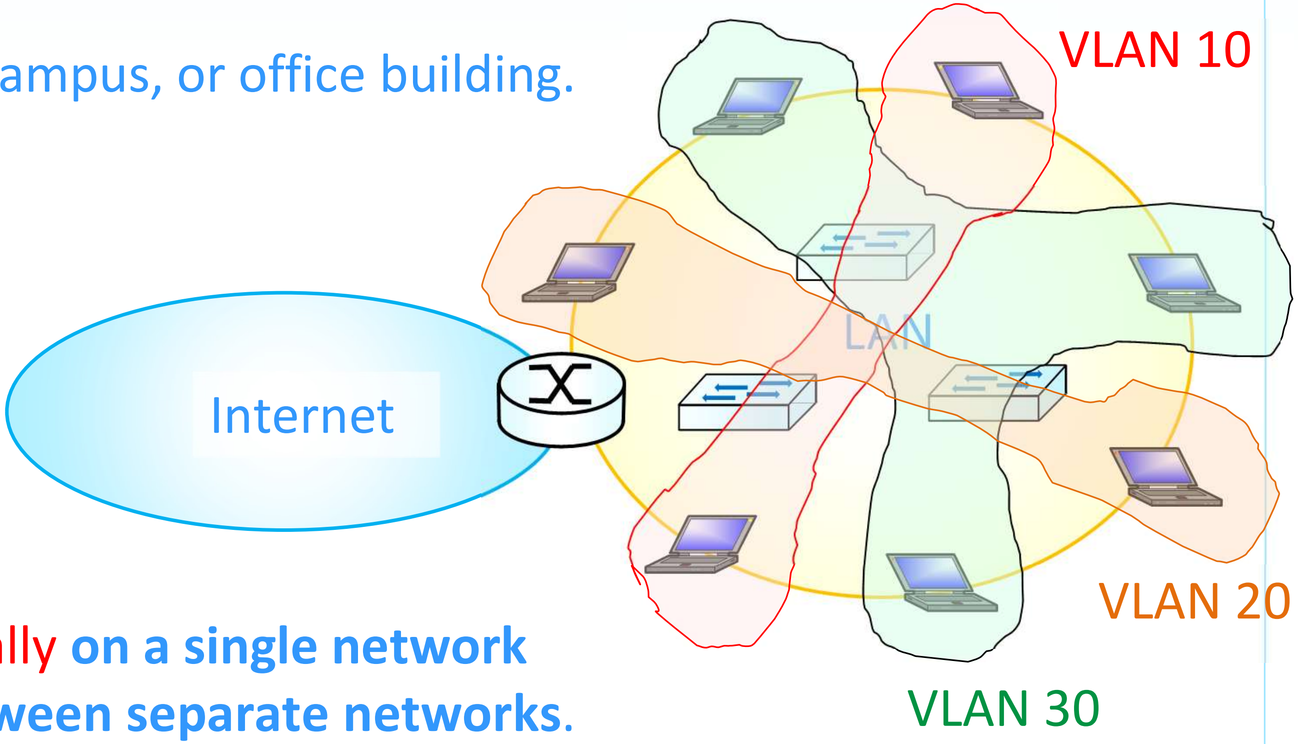


Hardware Type		Protocol Type
HW Address L.	Proto. Address L.	Opcode
Sender Hardware Address		Sender Protocol Addr (bytes 1-2)
Sender Protocol Addr (bytes 3-4)		Target Hardware Address
Target Protocol Address		

Virtual LAN (VLAN)

Virtual LAN (VLAN)

- Local area network (LAN)
  - a computer network that interconnects computers within a limited area of **a single broadcast domain**.
    - a residence, school, lab, campus, or office building.
- **Virtual LAN** (**VLAN**)
  - A broadcast domain that is partitioned and isolated in a computer network at the **data link layer** (OSI layer 2).

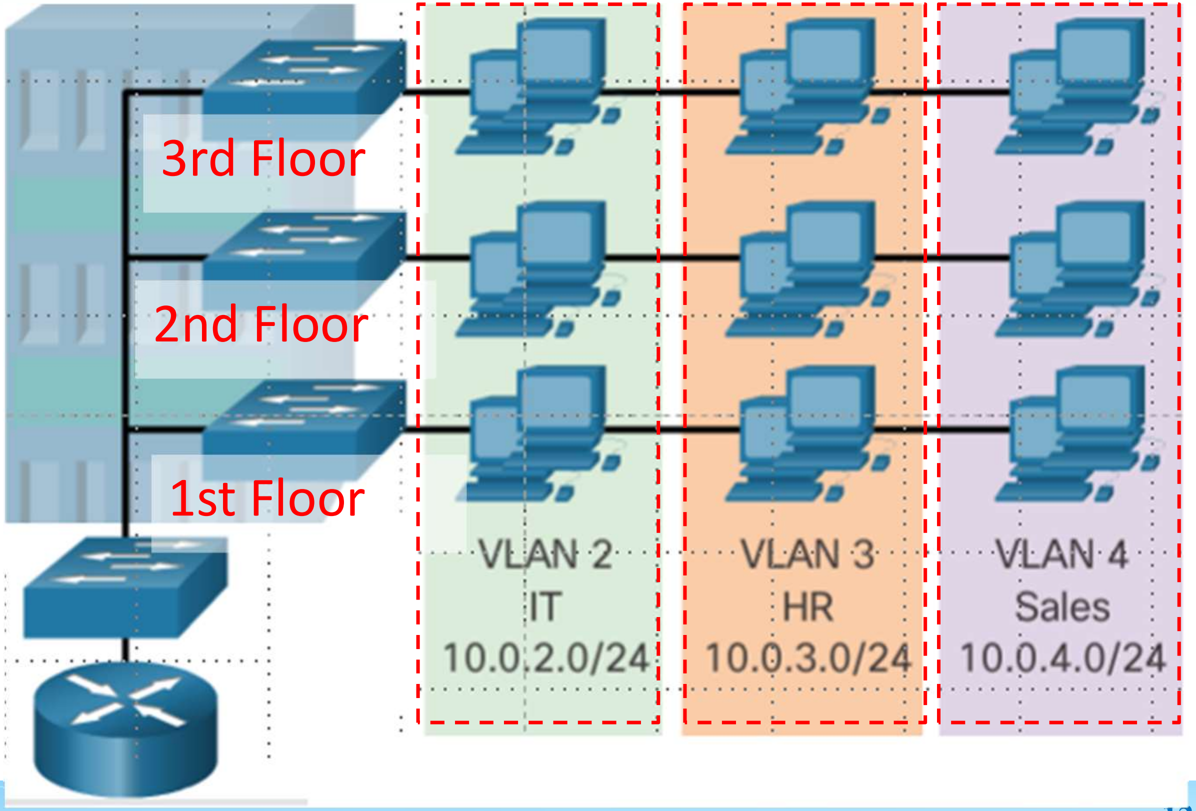


- VLAN-enabled LAN isolation

  Network traffic that is physically **on a single network** but acts as if it were **split between separate networks**.

VLAN Usages

- VLANs provide a way to group devices within a LAN
  - To communicate as if they were attached to the same cable
    - via logical connections, instead of physical connections
- May segment networks based on
    - Functions,
    - Project teams,
    - Applications, or
    - …
  - Without considering physical location of users or devices



- Reference: Cisco

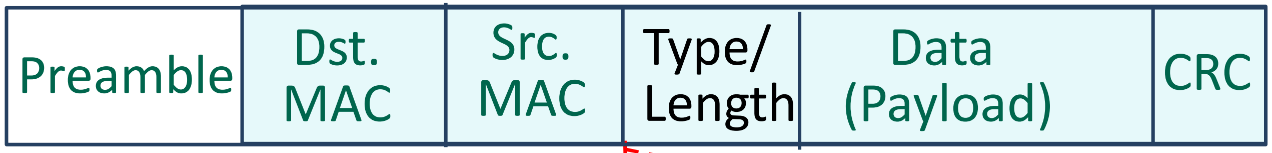
IEEE 802.1Q VLAN

- **IEEE 802.1Q** often referred to as **Dot1q**
- A networking standard that supports virtual LANs (VLANs) on an IEEE 802.3 Ethernet network.
- It defines
  - **A system** of **VLAN tagging** for **Ethernet frames** and
  - **Accompanying procedures** to be used by **bridges** and **switches** in handling such frames.

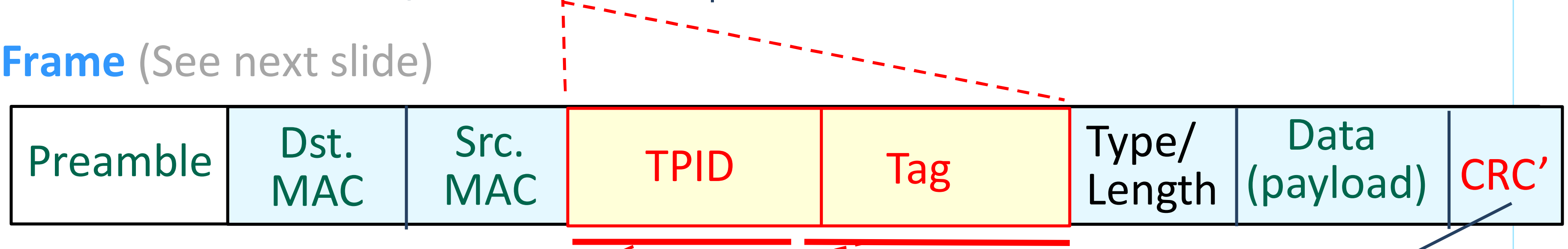
802.1Q VLAN Frame Format (1/2)

- Extending Ethernet Frame Format for **VLAN Trunk**
  - Adding a 4-byte field between source MAC and EtherType

- **802.3 Ethernet Frame**



- **802.1Q Frame** (See next slide)



Recomputed CRC
(FCS for Ethernet)

- TPID: Tag Protocol Identifier, 2 Bytes; Value: 0x8100 (next slide)
- Tags: Tag Control Information, 2 Bytes
  - VLAN IDentification Number (**VID**): 12 bits, up to 4096 VLAN IDs
- CRC: Cyclic Redundancy Check
  - FCS: Frame Check Sequence (for Ethernet)

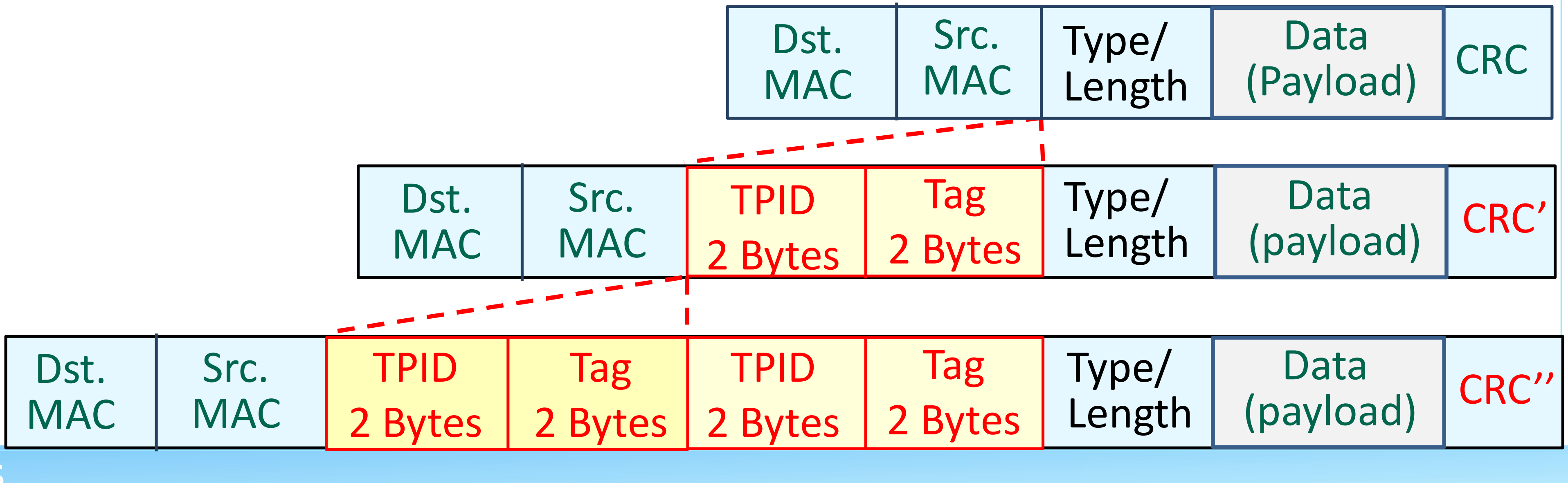
802.1Q VLAN Frame Format (2/2)

Preamble	Dst. MAC	Src. MAC	TPID	PRI	CFI	VID	Type/ Length	Data (payload)	CRC’

- VLAN Tag Fields:
  - **Tag Protocol ID (TPID)**: 2-byte value
    - **Type compliant with Ether type**, Value of 802.1Q Header: 0x8100
  - **User Priority:** 3-bit
    - Level or service implementation (like IP TOS)
  - **Canonical Format Identifier (CFI)**: 1-bit
    - Identifier that enables Token Ring frames to be carried across Ethernet links
  - **VLAN ID (VID)**: 12-bit
    - VLAN identification number, up to 4096 VLAN IDs
- Switch Inserts or removes TPID and Tag Control Information fields
  - Need to recalculate FCS/CRC values and inserts new FCS into the frame
    - FCS/CRC: Fault Check Sequence/Cyclic Redundant Code

QinQ (IEEE 802.1ad**)**

- Increasing number of users in networks or VMs in Datacenter
  - Require a large number of VLAN IDs.
- QinQ (also known as **Stacking VLAN** or **Double VLAN**)
  - Encapsulates packets with **two layers** of **VLAN Tags**
    - An **802.11Q tagged packet** is encapsulated in **another 802.1Q tag**
  - to extend the VLAN numbers up to 4096 × 4096



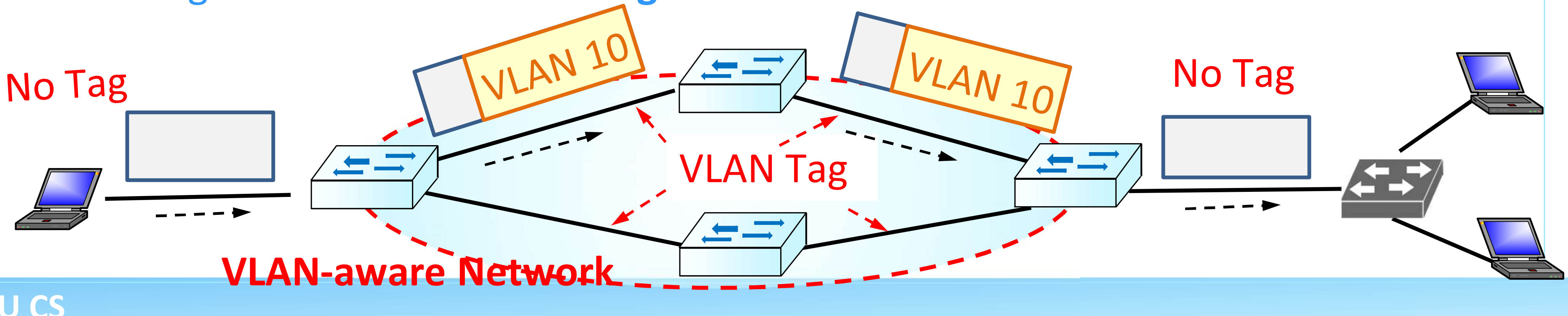
Dst. MAC	Src. MAC	Type/ Length	Data (Payload)	CRC

Dst. MAC	Src. MAC	TPID 2 Bytes	Tag 2 Bytes	Type/ Length	Data (payload)	CRC’

Dst. MAC	Src. MAC	TPID 2 Bytes	Tag 2 Bytes	TPID 2 Bytes	Tag 2 Bytes	Type/ Length	Data (payload)	CRC’’

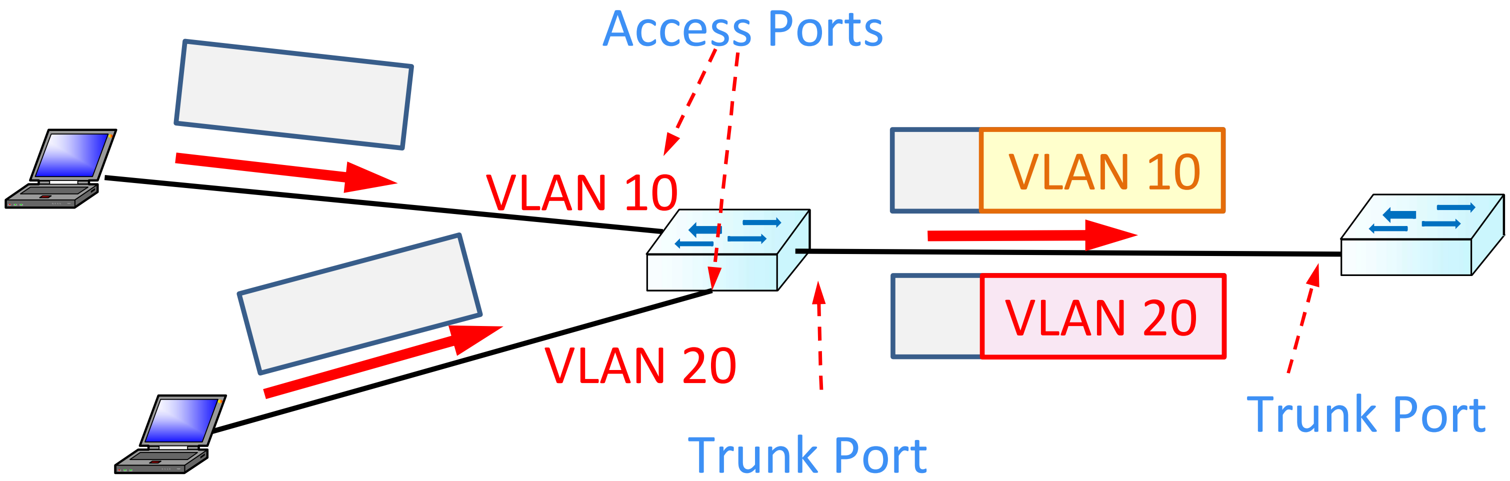
VLAN-aware Network

- **Networks** that are IEEE 802.1Q Conformant
  - Can add or remove VLAN tags.
- **VLAN-aware network** must **distinguish** each frame as being **within exactly one VLAN**.
  - When a frame **enters** a VLAN-aware network, the edge switch **adds** a tag to the frame
    - to represent the VLAN membership
  - When a frame **is leaving** a VLAN-aware network, the edge switch **removes the tag** on the frame



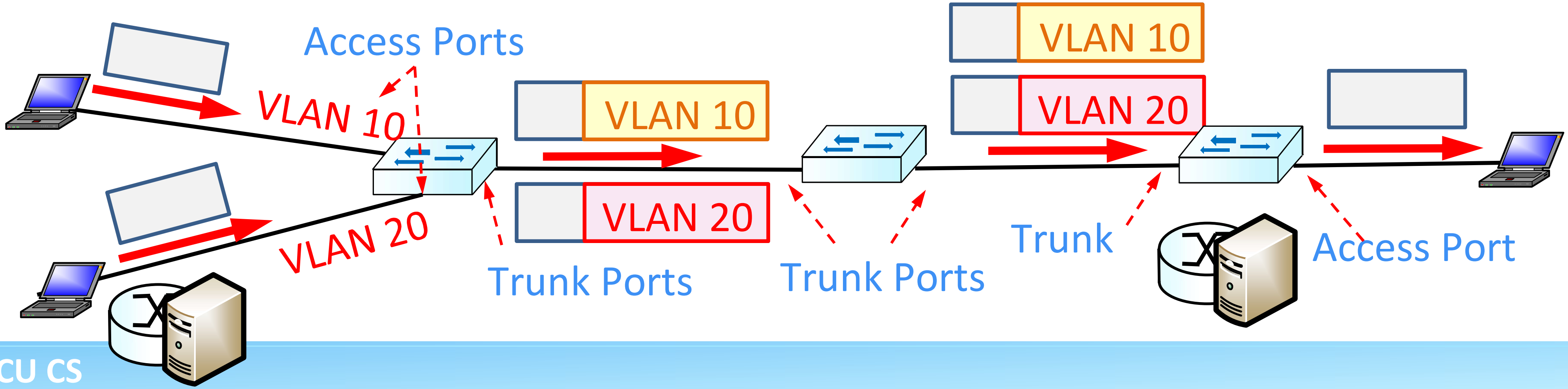
Access Ports

- **Access Port**
  - A connection on a switch that transmits data to and from a **specific VLAN**.
  - Only assigned to **a single VLAN**
    - When a host is plugged into the port it is in the VLAN
  - Sends and receives frames that **aren’t tagged**
    - Drop a frame with VLAN tags



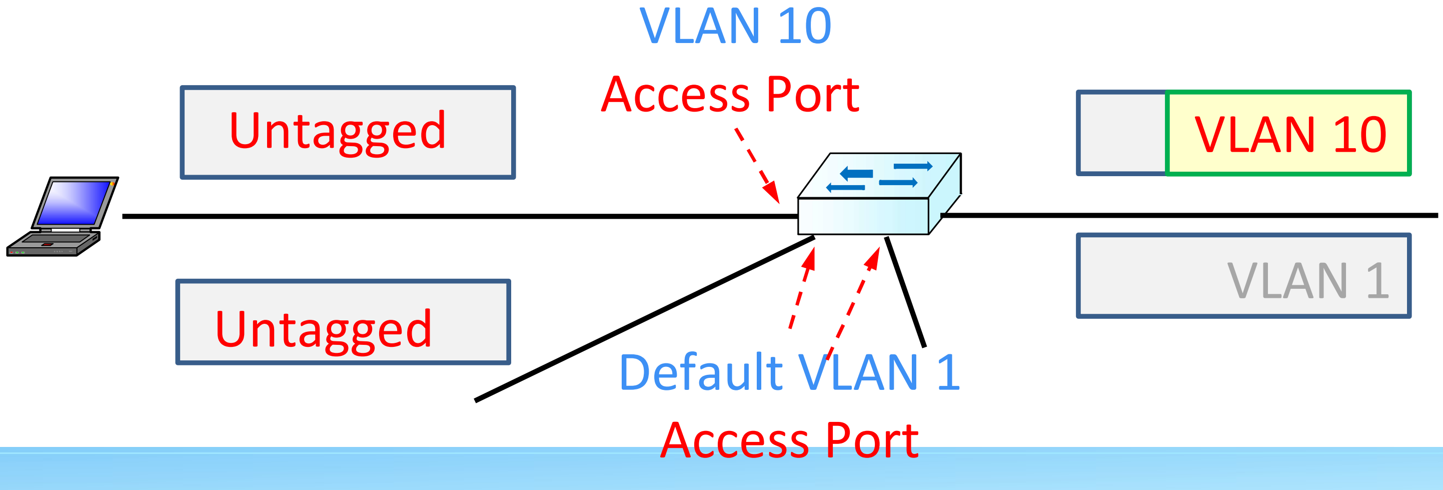
Trunk Ports

- **Trunk Port**
  - A connection on a switch that can transmit data from **multiple VLANs**
    - A conduit for **multiple VLANs** between switches, routers, or servers.
  - Route traffic to and from multiple VLANs using a tagging protocol
    - Allows for multiple VLANs to co-exist on a single network.
    - Sends and receives frames that are **tagged**, in general
  - Typically offer higher bandwidth and lower latency than access ports.



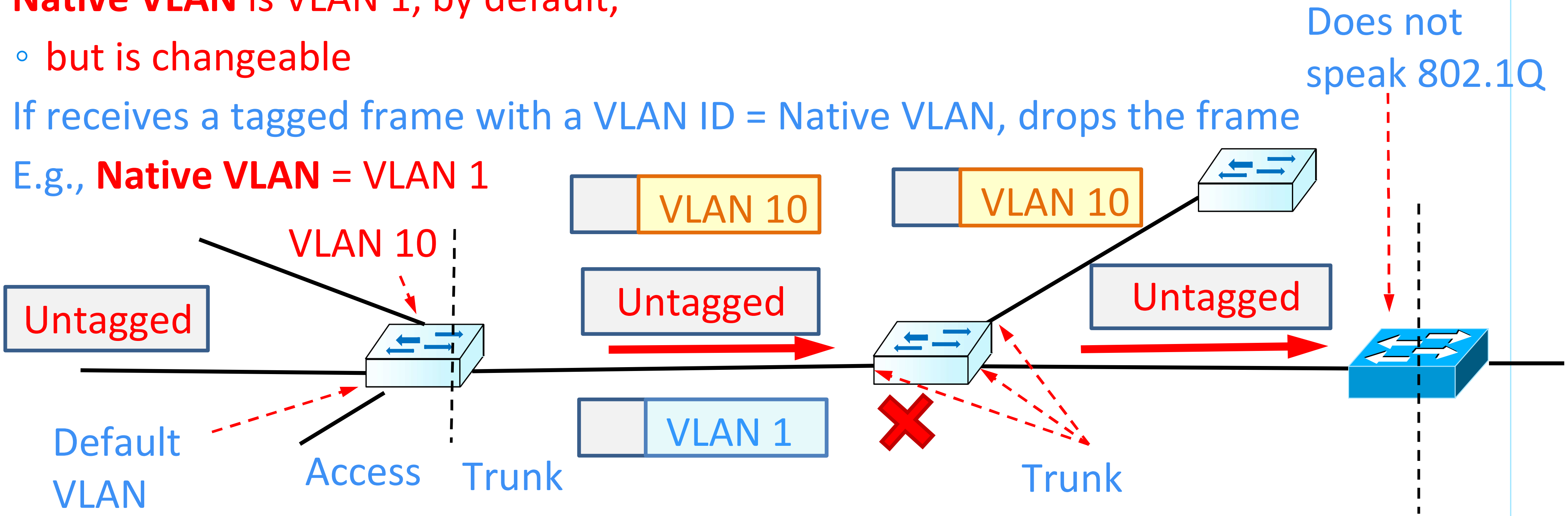
Default VLAN

- The VLAN which all Access Ports are assigned to until they are explicitly placed in another VLAN.
  - Only relevant on an **Access port**
    - The VLAN **all access ports** are assigned initially
    - Sends and receives traffic **without VLAN tags**
- Default VLAN is **VLAN 1**, in general
  - By default, **Layer 2 Control Traffic** is associated with **VLAN 1**
- **E.g.,**



Native VLAN

- Used to define which VLAN will be **untagged** on a trunk
- By default, **trunk ports** are in the **Native VLAN**
  - **Trunk ports** will not tag outgoing frames that belong to the Native VLAN
- **Native VLAN** is VLAN 1, by default,
  - but is changeable
- If receives a tagged frame with a VLAN ID = Native VLAN, drops the frame
- E.g., **Native VLAN** = VLAN 1



- If both sides of a link do not agree on Native VLAN, trunk will not operate properly

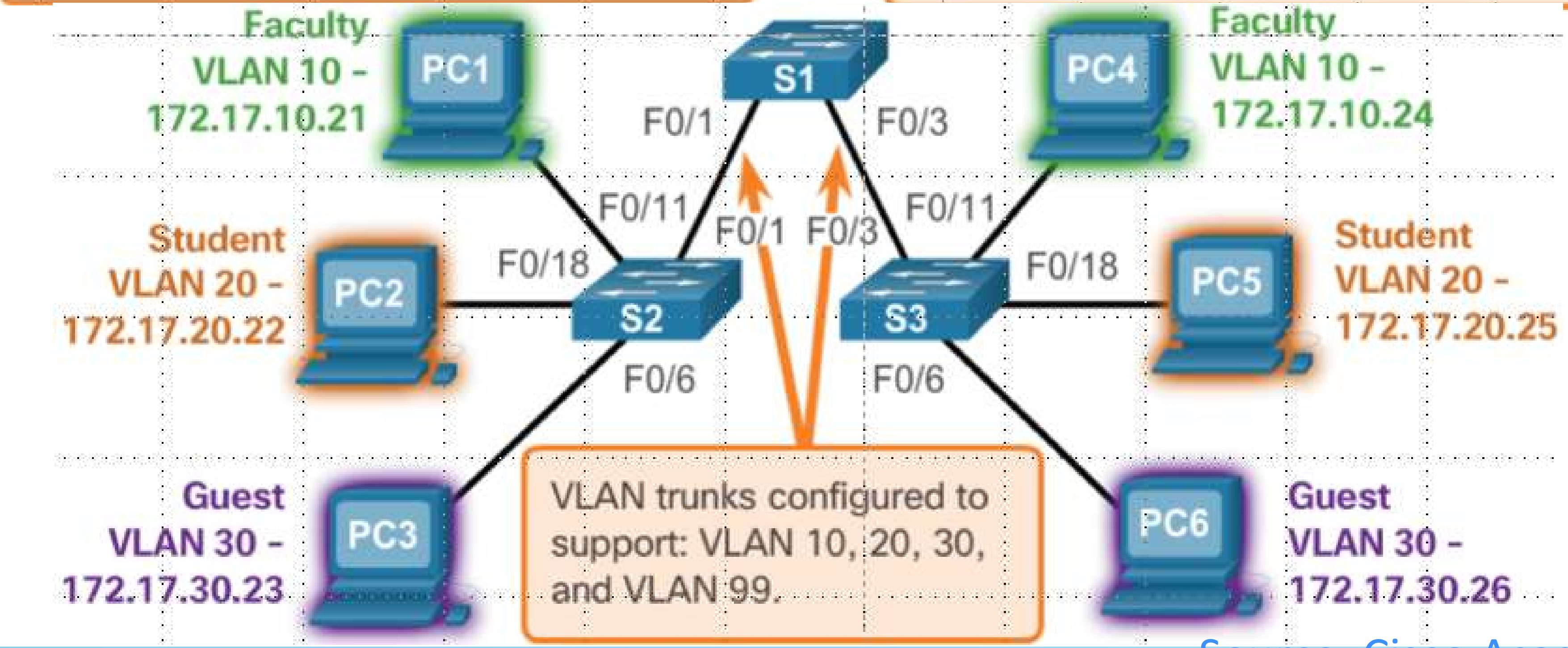
Example of VLAN Trunk

VLAN 10 Faculty/Staff - 172.17.10.0/24
VLAN 20 Students - 172.17.20.0/24
VLAN 30 Guest - 172.17.30.0/24
VLAN 99 Management and Native - 172.17.99.0/24

F0/1-5 are 802.1Q trunk interfaces with native VLAN 99.

F0/11-17 are in VLAN 10.
F0/18-24 are in VLAN 20.
F0/6-10 are in VLAN 30.

Access Ports



VLAN trunks configured to support: VLAN 10, 20, 30, and VLAN 99.

Source: Cisco Academy

Management VLAN and Switch Virtual Interface

- **Management VLAN**

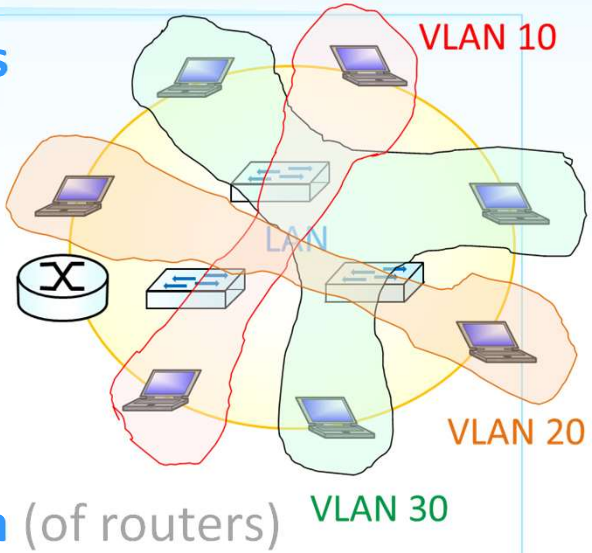
  A VLAN configured to access the management capabilities of a switch

- By default, VLAN 1 is the management VLAN
  - But VLAN 1 is a bad choice because Default VLAN is VLAN 1 too.
    - Recommend to assign other VLAN ID for the management VLAN
- **Switch Virtual Interface (SVI)**
  - a logical layer-3 interface on a switch.
  - Can be used to connect and manage switch remotely
- Managing switch remotely:
  - Create a SVI for the management VLAN
  - Assign **an IP address** and **a subnet mask** to the **SVI**

  ➢Allowing the switch to be managed via HTTP, Telnet, SSH, or SNMP

Inter-VLAN Routing

What is Inter-VLAN Routing?



- **VLANs** divide a LAN into multiple, separate **broadcast domains**
  - ➢Computers on separate VLANs are unable to communicate (without the intervention of a routing device)
- ■ Inter-VLAN routing
  - ◦ forwarding network traffic from one **VLAN** to another **VLAN**
- Recall:

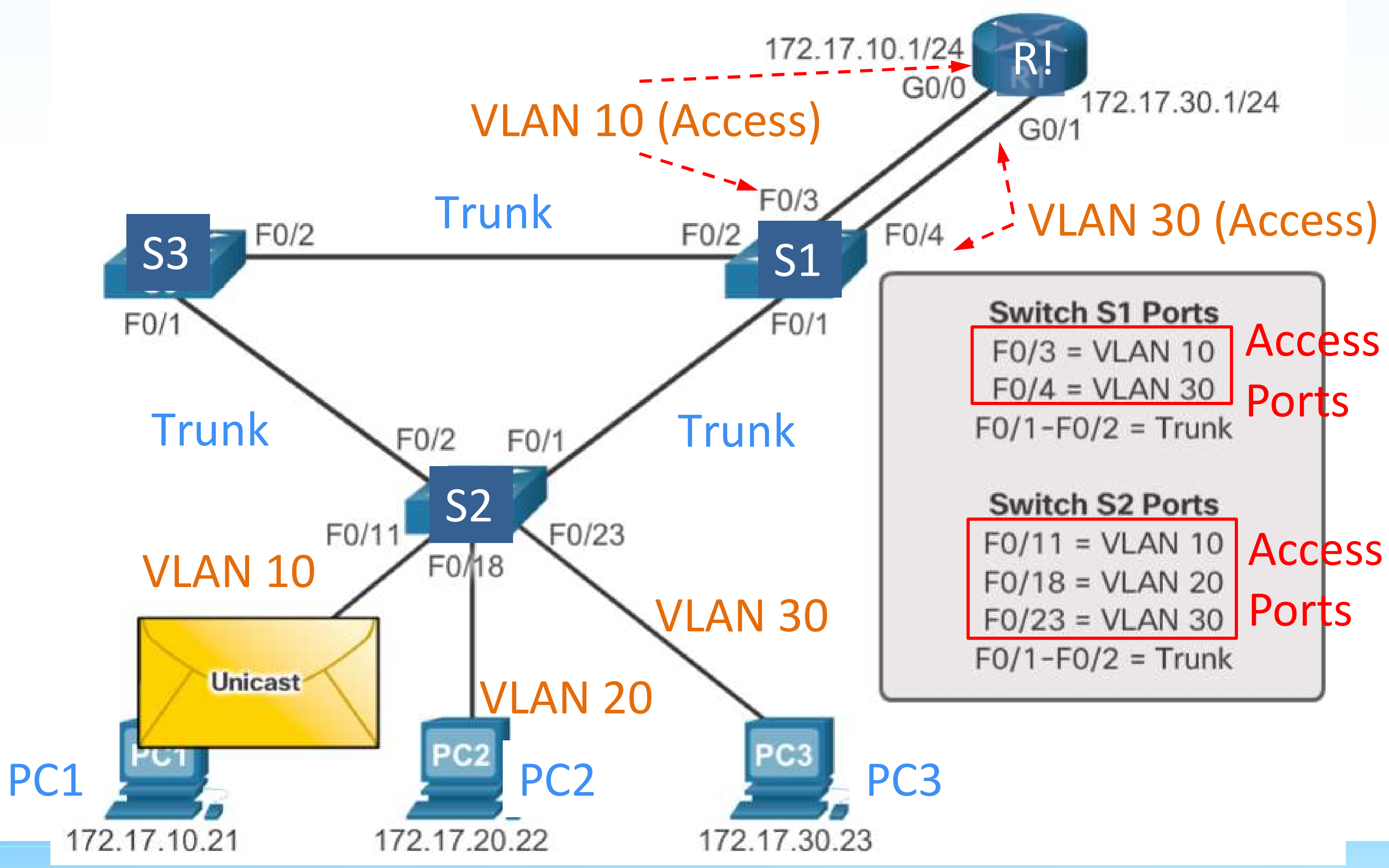
  Layer 2 switches cannot perform the dynamic **routing function** (of routers)
- Three inter-VLAN routing mechanisms:
  - ◦ Legacy inter-VLAN routing
  - ◦ Router-on-a-Stick
  - ◦ Layer 3 switching **via SVIs**

Legacy Inter-VLAN Routing

- **Legacy inter-VLAN routing** relies on **routers** with multiple physical interfaces.
- **Each router's physical interface** is connected to **a unique VLAN**.
- Configuration
  - Connecting different physical router interfaces to different physical switch ports
    - Each physical interface is assigned to a different VLAN
    - **Switch ports** connected to the router are placed in **access mode**
  - Each router interface can then
    - Accept traffic from the VLAN associated with the switch interface
    - Route traffic to the other VLANs connected to the other interfaces

Example of legacy inter-VLAN routing (1/2)

- Router R1 has a separate interface configured for each of VLANs.
- E.g., PC1 PC3
  - PC1 on VLAN 10
  - PC3 on VLAN 30
  - Communicate
    through router R1.

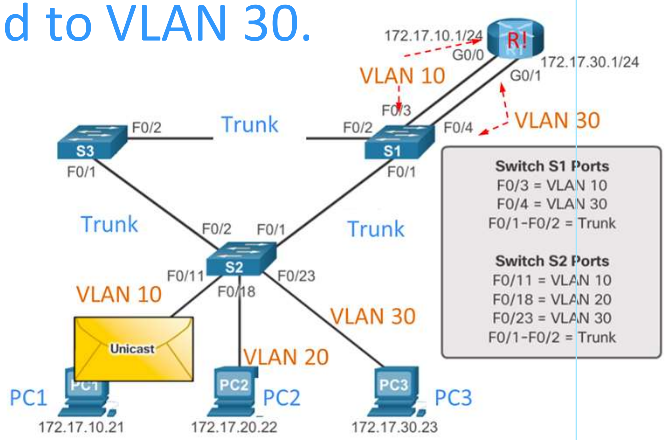


Switch S1 Ports
F0/3 = VLAN 10
F0/4 = VLAN 30
F0/1-F0/2 = Trunk

Switch S2 Ports
F0/11 = VLAN 10
F0/18 = VLAN 20
F0/23 = VLAN 30
F0/1-F0/2 = Trunk

Example of legacy inter-VLAN routing (2/2)

1. PC1 sends unicast traffic destined for PC3 to switch S2 on VLAN 10,
2. S2 forwards traffic out the trunk interface to S1.
3. S1 forwards unicast traffic, through interface F0/3, to interface G0/0 on R1.
4. R1 routes traffic, through interface G0/1, which is connected to VLAN 30.
   - ➢Unicast traffic forwarded to S1 on VLAN 30.
5. S1 forwards the unicast traffic to S2 through the trunk link,
6. S2 can then forward the unicast traffic to PC3 on VLAN 30.



- **Note:** The router was configured with two separate physical interfaces to interact with the different VLANs and performed the routing.
  - ➢Not efficient and is generally **no longer implemented in switched networks**.

Inter-VLAN Routing – Router-on-a-Stick

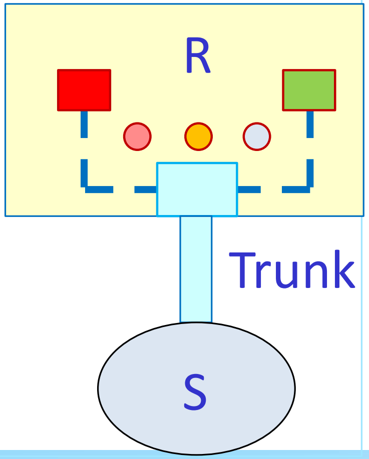
- **Sub-interfaces:**

  Software-based virtual interfaces, associated with a single physical interface
  - Each **sub-interface** is independently configured with
    - A VLAN assignment and
    - **An IP address in the subnet** corresponding to the VLAN assignment

➢ Router-on-a-stick: a type of router configuration **utilizing sub-interfaces**

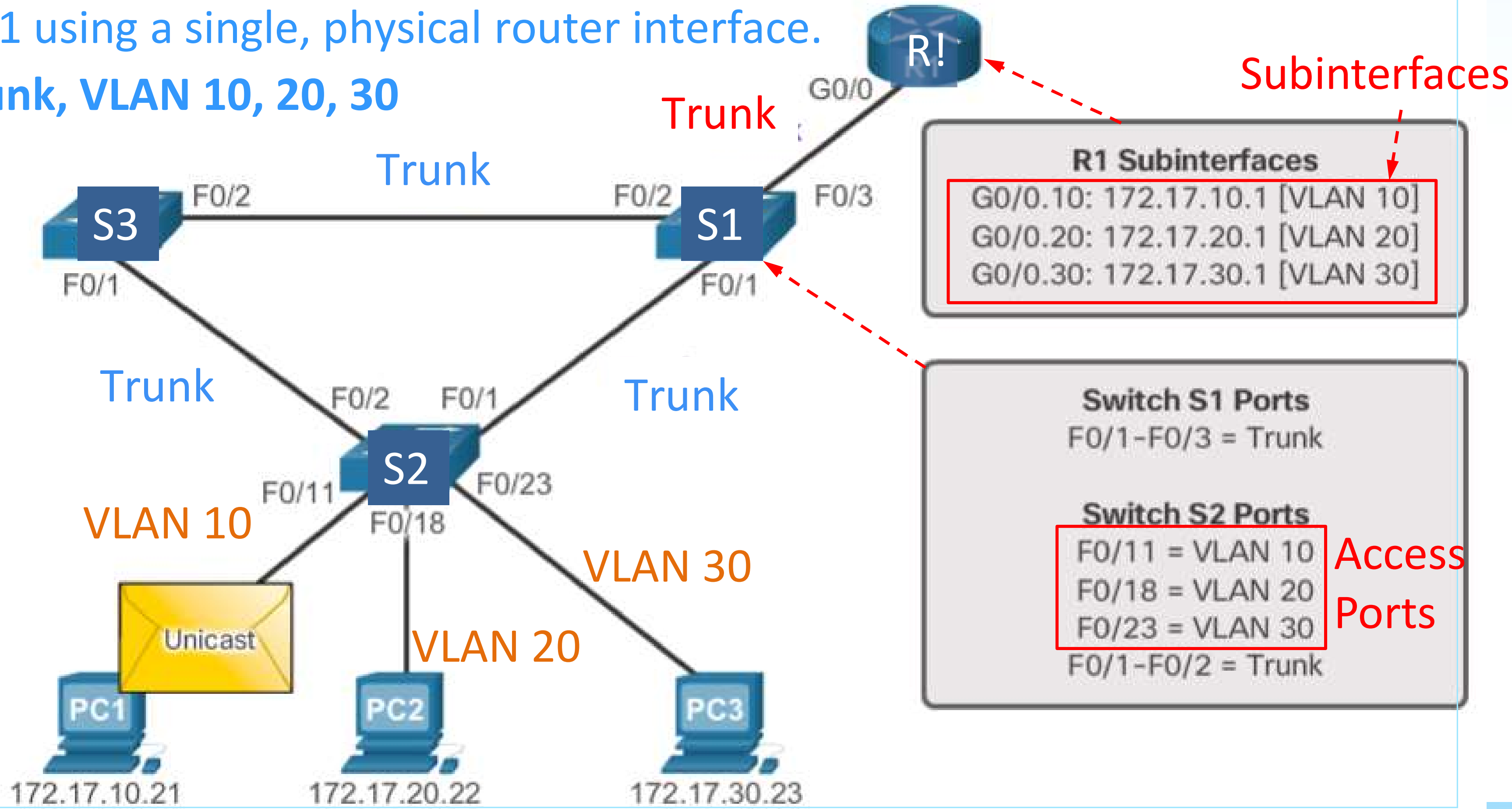
  a single physical interface routes traffic between **multiple VLANs** on a network
  - Configured to operate as a **trunk link** and
  - Connected to a switch port in **trunk** mode

- Router accepts **VLAN-tagged** traffic on the **trunk interface**
- Router internally **route traffic between VLANs** using **sub-interfaces**
  - **VLAN-tagged** for the **destination VLAN** and
  - Forward the traffic out **the same physical interface**



Example of Router-on-a-Stick Inter-VLAN Routing

- PC1 on VLAN 10 communicating with PC3 on VLAN 30
  - Through R1 using a single, physical router interface.
- **R1 G0/0: trunk, VLAN 10, 20, 30**



R1 Subinterfaces

G0/0.10: 172.17.10.1 [VLAN 10]
G0/0.20: 172.17.20.1 [VLAN 20]
G0/0.30: 172.17.30.1 [VLAN 30]

Switch S1 Ports

F0/1-F0/3 = Trunk

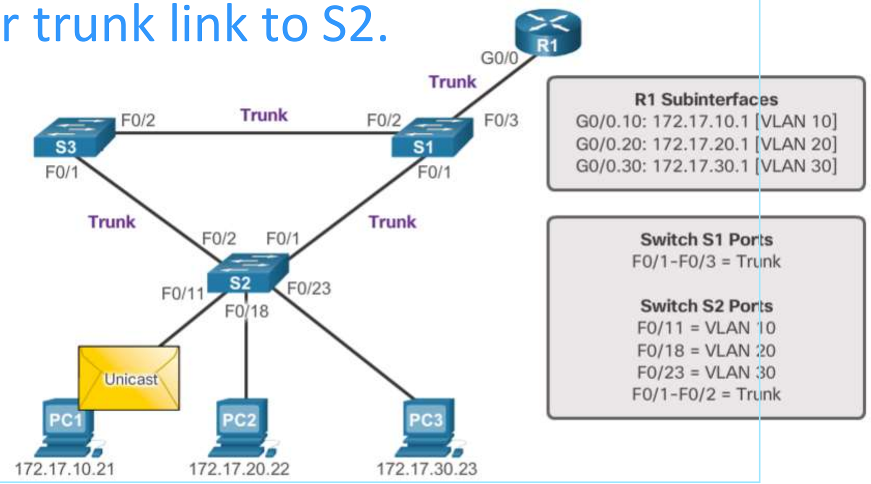
Switch S2 Ports

F0/11 = VLAN 10
F0/18 = VLAN 20
F0/23 = VLAN 30
F0/1-F0/2 = Trunk

Example of Router-on-a-Stick Inter-VLAN Routing (2/2)

1. PC1 sends its unicast traffic to switch S2.
2. S2 tags unicast traffic with **VLAN 10** and forwards traffic out its trunk link to S1.
3. S1 forwards **tagged traffic** out the other trunk interface on port F0/3 to R1.
4. R1 accepts tagged unicast traffic on **VLAN 10** and
   - Routes traffic to **VLAN 30** using its configured subinterfaces
     - Tag unicast traffic with **VLAN 30**
   - Sends tagged traffic out its interface to S1.
5. S1 forwards the tagged unicast traffic out the other trunk link to S2.
6. S2:
   - Removes VLAN tag of unicast frame and
   - Forwards the frame out to PC3 on port F0/23.

Note: can not scale beyond 50 VLANs



R1 Subinterfaces
G0/0.10: 172.17.10.1 [VLAN 10]
G0/0.20: 172.17.20.1 [VLAN 20]
G0/0.30: 172.17.30.1 [VLAN 30]

Switch S1 Ports
F0/1-F0/3 = Trunk

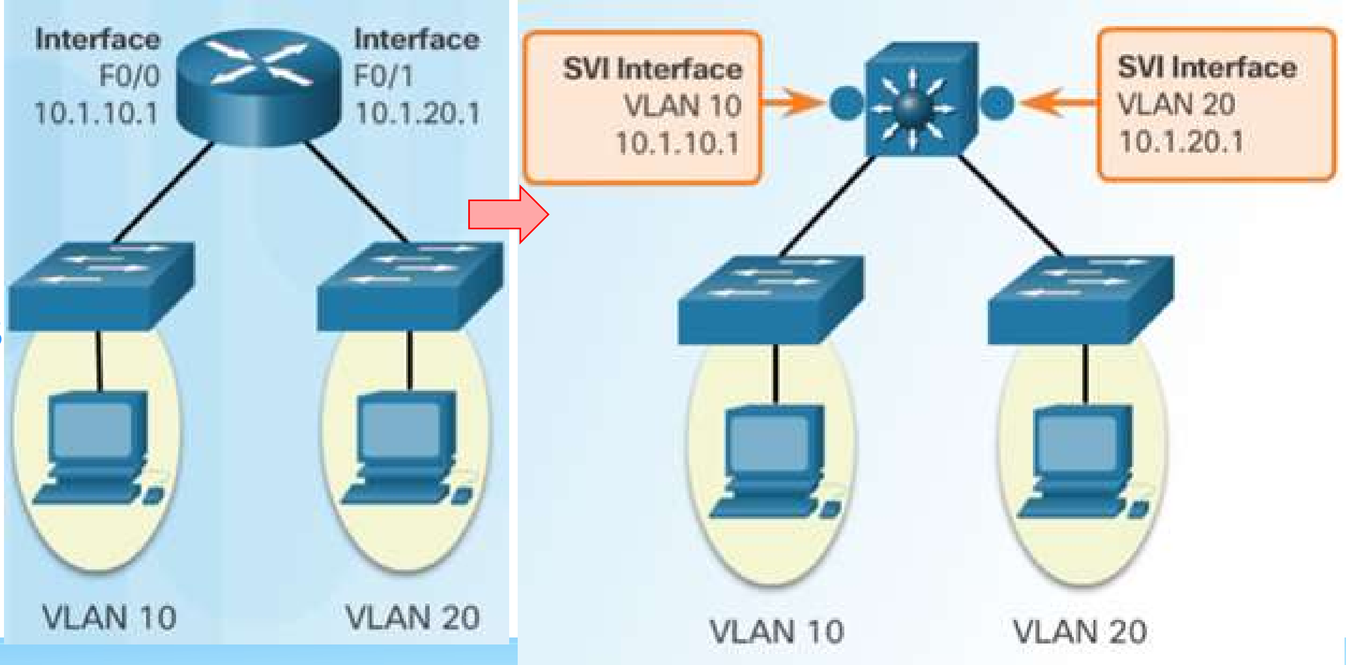
Switch S2 Ports
F0/11 = VLAN 10
F0/18 = VLAN 20
F0/23 = VLAN 30
F0/1-F0/2 = Trunk

Inter-VLAN Routing – Multilayer Switching

- Router-on-a-Stick still requires a dedicated router.
- A **multilayer switch (MLS)** is a networking device that
  - **Switches** frames **on OSI layer 2** (like an ordinary network switch) and
  - Provides **extra functions** on **higher OSI layers**.
- **Layer 3 Switches** can perform **Layer 2** and **Layer 3** functions,
  - May **replace the dedicated routers** to perform **basic routing** on a network.
- **Hardware-based switching** can achieve **high-packet processing rates**
  - **Packet-switching throughputs**
    - **Layer 3 switches: millions** of packets per second (pps),
    - **Traditional routers**: from **100,000** pps to more than **1 million** pps

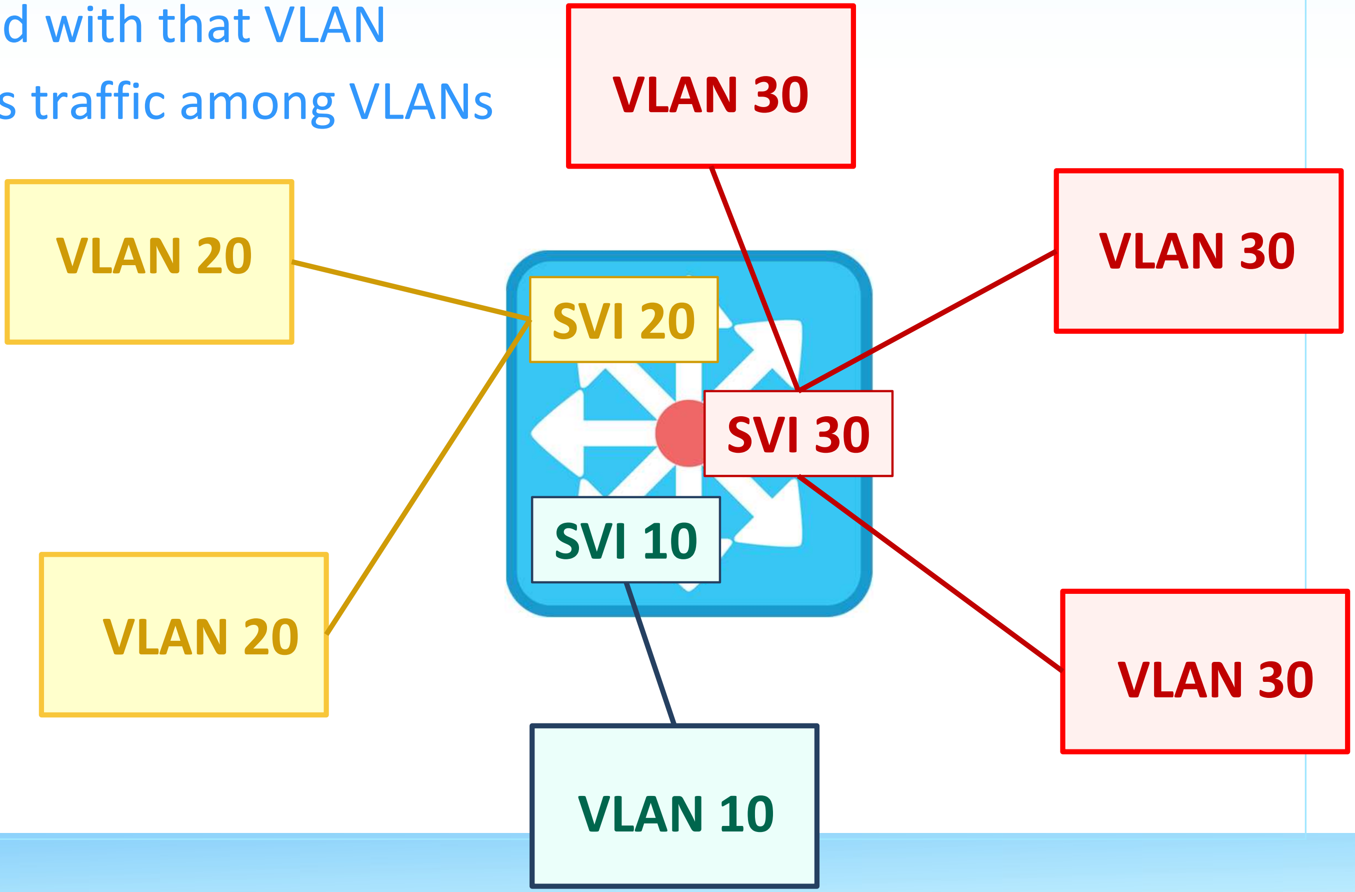
Switch Virtual Interfaces

- **Switch virtual interface (SVI)**
  a virtual interface configured within a multilayer switch
  - Can be created for any VLAN on the switch
  - Is virtual because there is **no physical port** dedicated to the interface
  - Can perform **the same functions** for the VLAN **as a router interface** would, and
  - Can be configured in much the same way as a router interface
    - IP address,
    - Inbound/Outbound ACLs
    - etc.



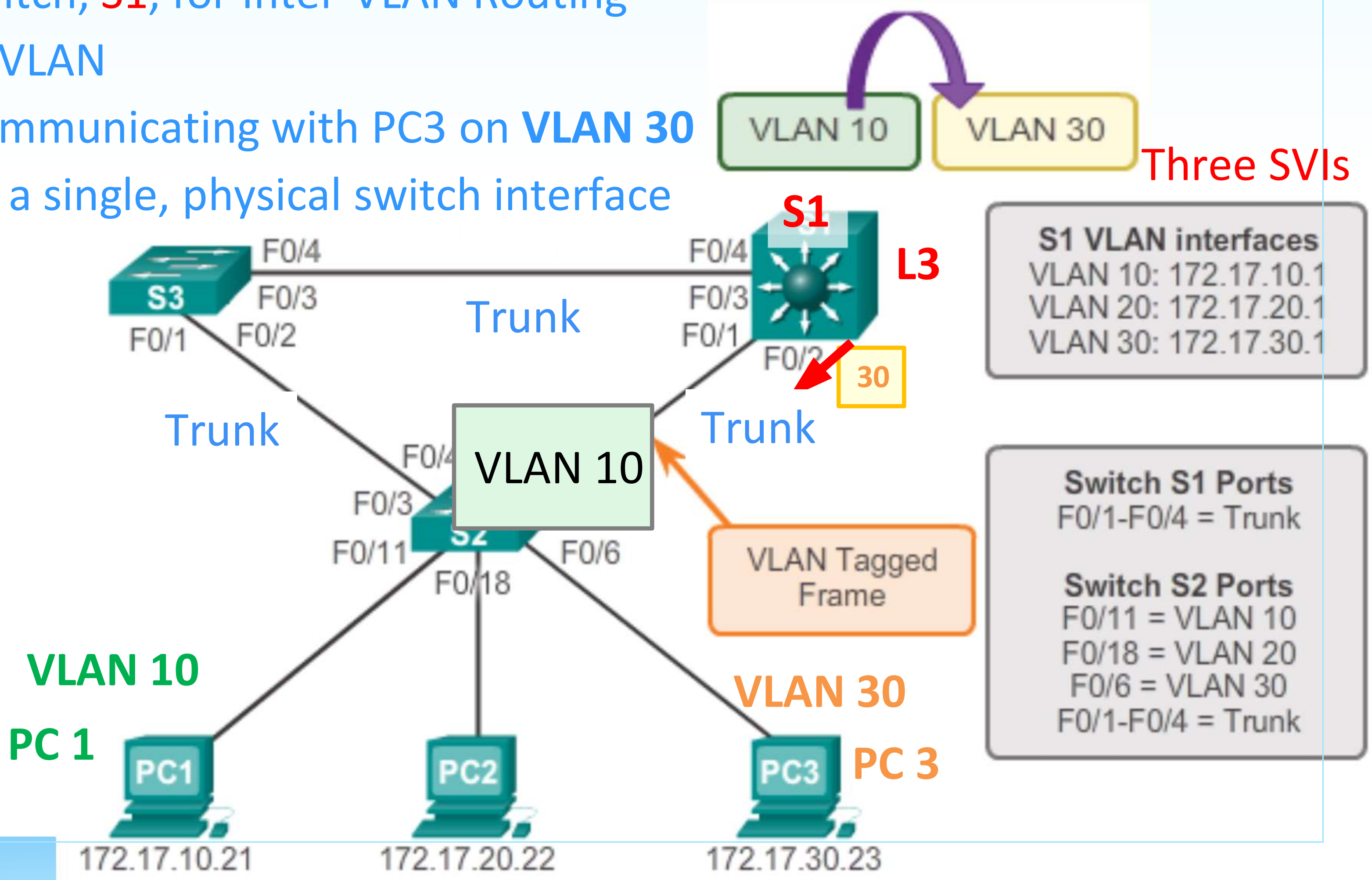
SVI as Default Gateway of VLAN

- An SVI served as the **default gate** for a **VLAN**
  - Provides Layer 3 processing for packets to or from **all switch ports** associated with that VLAN
- Multilayer (L3) Switch routes traffic among VLANs



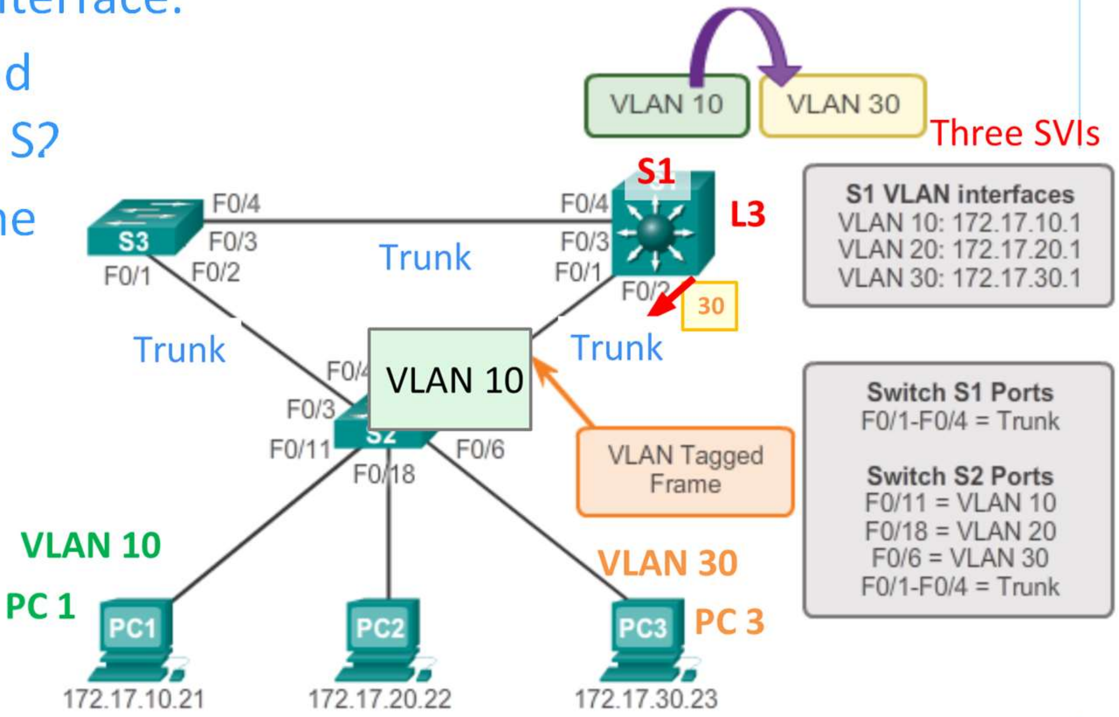
Example of Inter-VLAN Routing – Multilayer Switching (1/2)

- Deploying an L3 Switch, S1, for Inter-VLAN Routing
  - One SVI for each VLAN
- PC1 on **VLAN 10** communicating with PC3 on **VLAN 30**
  - Through S1 using a single, physical switch interface



Example of Inter-VLAN Routing – Multilayer Switching (2/2)

1. PC1 sends its unicast traffic to switch S2.
2. S2 tags unicast traffic with VLAN 10 and forwards traffic out its trunk link to S1.
3. S1 removes VLAN tag and forwards traffic to VLAN 10 interface.
4. S1 routes the traffic to its VLAN 30 interface.
5. S1 retags the traffic with VLAN 30 and forwards it out the trunk link back to S2
6. S2 removes VLAN tag of unicast frame
7. S2 forwards the frame out to PC3 on port F0/23



S1 VLAN interfaces
VLAN 10: 172.17.10.1
VLAN 20: 172.17.20.1
VLAN 30: 172.17.30.1

Switch S1 Ports
F0/1-F0/4 = Trunk

Switch S2 Ports
F0/11 = VLAN 10
F0/18 = VLAN 20
F0/6 = VLAN 30
F0/1-F0/4 = Trunk