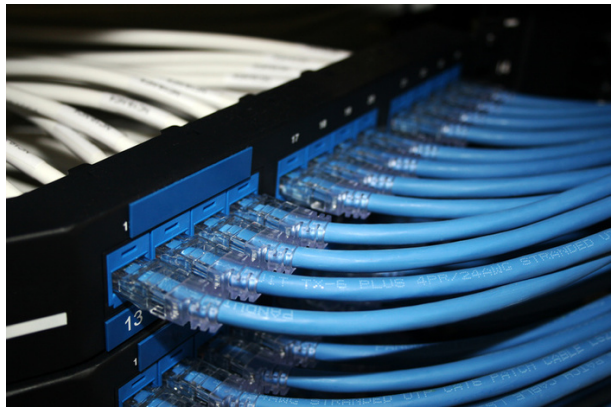
Data Center Ethernet



Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides and audio/video recordings of this class lecture are at:

http://www.cse.wustl.edu/~jain/cse570-21/

Student Questions

Overview

1. Residential vs. Data Center Ethernet
2. Review of Ethernet Addresses, devices, speeds, algorithms
3. Enhancements to Spanning Tree Protocol
4. Virtual LANs
5. Data Center Bridging Extensions

Student Questions

Quiz: True or False?

Which of the following statements are generally true?

T	F	
o	o	Ethernet is a local area network (Local ≤ 2km)
o	o	Token ring, Token Bus, and CSMA/CD are the three most common LAN access methods.
o	o	Ethernet uses CSMA/CD.
o	o	Ethernet bridges use spanning tree for packet forwarding.
o	o	Ethernet frames are 1518 bytes.
o	o	Ethernet does not provide any delay guarantees.
o	o	Ethernet has no congestion control.
o	o	Ethernet has strict priorities.

Student Questions

Residential vs. Data Center Ethernet

Residential	Data Center
❑ Distance: up to 200m	q No limit
❑ Scale: ➢ Few MAC addresses ➢ 4096 VLANs	 q Millions of MAC Addresses q Millions of VLANs Q-in-Q
❑ Protection: Spanning tree	q Rapid spanning tree, … (Gives 1s, need 50ms)
❑ Path determined by spanning tree	q Traffic engineered path
❑ Simple service	q Service Level Agreement. Rate Control.
❑ Priority ⇒ Aggregate QoS	q Need per-flow/per-class QoS
❑ No performance/Error monitoring (OAM)	q Need performance/BER

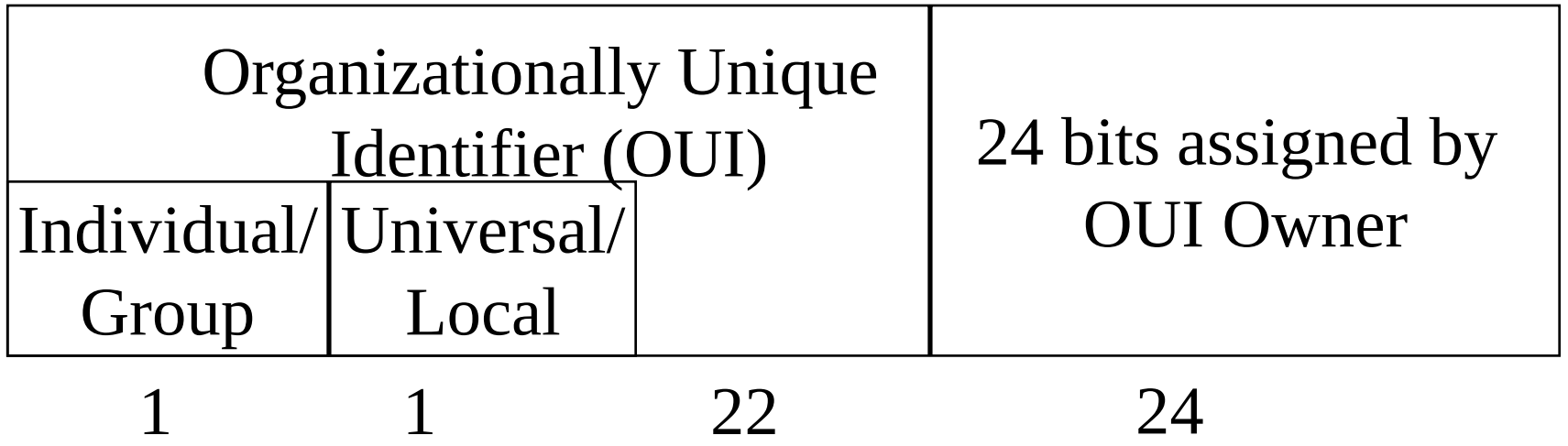
Student Questions

❑ Is the residential IP more trusted by the web server- and less likely to be noticed and blocked?

No. All residential IPs, when outside, are public IP addresses. They are from a block assigned to the carrier. You can not differentiate residential from non-residential IP addresses.

IEEE 802 Address Format

- 48-bit:1000 0000 : 0000 0001 : 0100 0011 : 0000 0000 : 1000 0000 : 0000 1100 = 80:01:43:00:80:0C



- Multicast = "To all bridges on this LAN"
- Broadcast = "To all stations" (Note: Local bit is set) = 111111....111 = FF:FF:FF:FF:FF:FF

Student Questions

- Can you explain the MAC emulation in more detail? Is that in order to "change" MAC addresses?

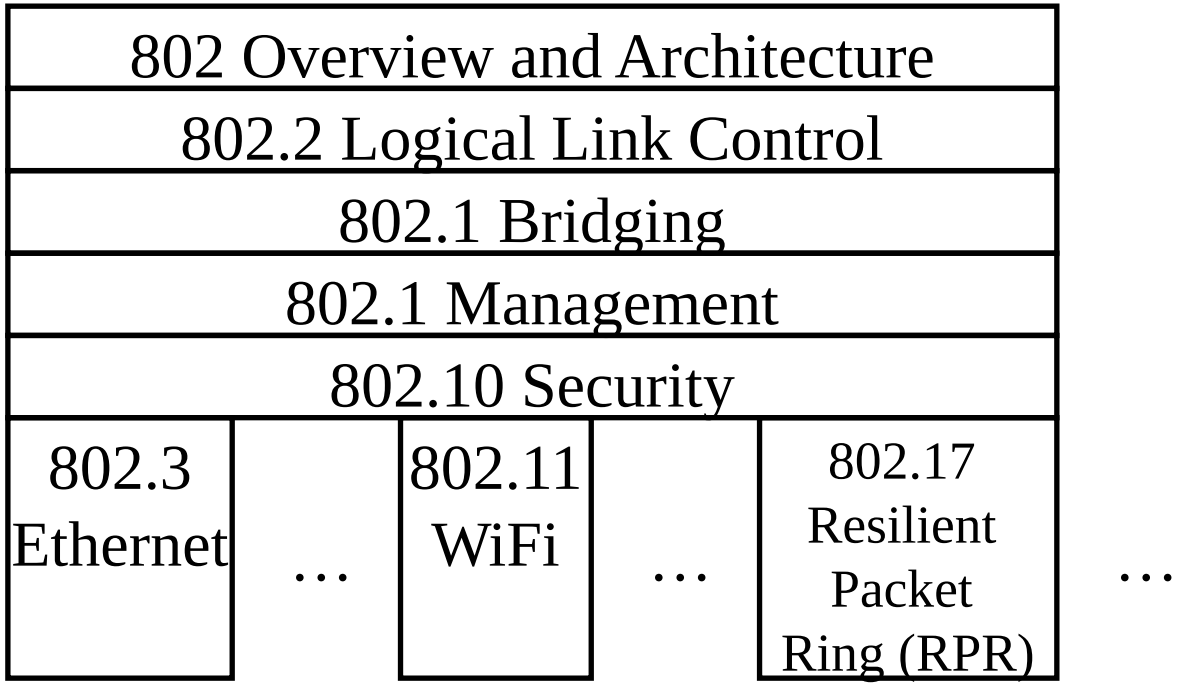
The Ethernet chip allows software to provide the MAC address in place of burnt-in MAC.

- Since there are only universal addresses, does the second bit always be 0 except broadcast address?

No. Local MACs are still in use, e.g., private group addresses.

IEEE Standards Numbering System

- IEEE 802.* and IEEE 802.1* standards (e.g., IEEE 802.1Q-2011) apply to all IEEE 802 technologies:
  - IEEE 802.3 Ethernet
  - IEEE 802.11 WiFi
  - IEEE 802.16 WiMAX



Student Questions

IEEE Standards Numbering (Cont)

- IEEE 802.3* standards apply only to Ethernet, e.g., IEEE802.3ba-2010
- Standards with all upper case letters are base standards E.g., IEEE 802.1AB-2009
- Standards with lower case are additions/extensions/revisions. Merged with the base standard in its next revision. e.g., IEEE 802.1w-2001 was merged with IEEE 802.1D-2004
- Standards used to be numbered, sequentially, e.g., IEEE 802.1a, …, 802.1z, 802.1aa, 802.1ab, …
- Recently they started showing base standards in the additions, e.g., IEEE 802.1Qau-2010

Student Questions

Names, IDs, Locators

Name: John Smith

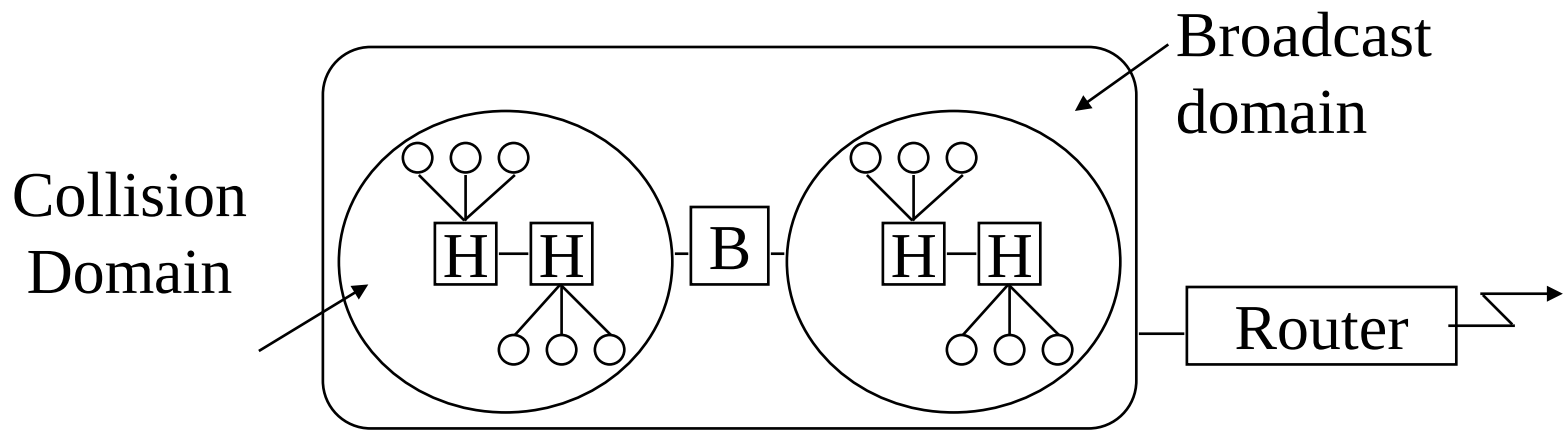
ID: 012-34-5678

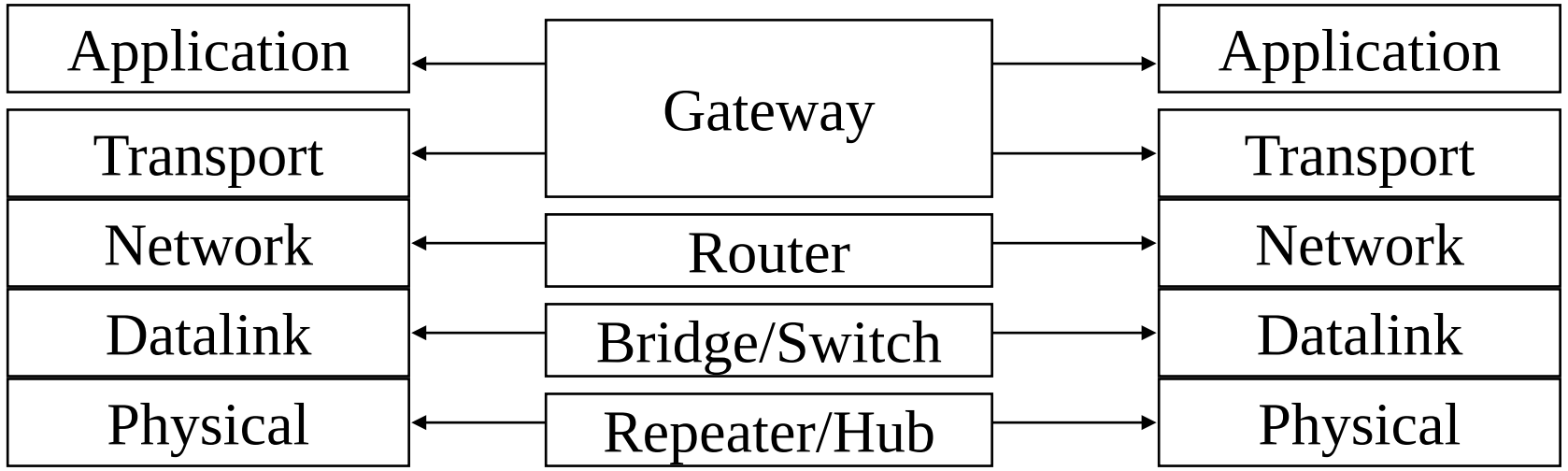
Locator:
1234 Main Street
Big City, MO 12345
USA

- Locator changes as you move, ID and Names remain the same.
- **Examples**:
  - Names: Company names, DNS names (Microsoft.com)
  - IDs: Cell phone numbers, 800-numbers, Ethernet addresses, Skype ID, VOIP Phone number
  - Locators: Wired phone numbers, IP addresses

Student Questions

Interconnection Devices





Student Questions

- What is the difference between switch and router?

Switch is a L2 device. Router is an L3 device. L2 devices do not look at L3-L7 headers. L3 devices do not have access to L1-L2 headers.

- What is the thing connecting one hub to another hub in the same collision domain?

It is just another wire or link.

Interconnection Devices (Cont)

- **Repeater**: PHY device that restores data and collision signals
- **Hub**: Multiport repeater + fault detection and recovery
- **Bridge**: Datalink layer device connecting two or more collision domains. MAC multicasts are propagated throughout the LAN.
- **Router**: Network layer device. IP, IPX, AppleTalk. Does not propagate MAC multicasts.
- **Switch**: Multiport bridge with parallel paths
- These are functions. Packaging varies.

Student Questions

- How does the hub provide fault detection and recovery? I thought this is provided by L4.

Hub cannot provide fault recovery. But they can detect L1 faults, such as a wire cut.

Ethernet Speeds

- IEEE 802.3ba-2010 (40G/100G) standard
- 10Mbps, 100 Mbps, 1 Gbps versions have both CSMA/CD and Full-duplex versions
- No CSMA/CD in 10G and up
- No CSMA/CD in practice now even at home or at 10 Mbps
- 1 Gbps in residential, enterprise offices
- 1 Gbps in Data centers, moving to 10 Gbps and 40 Gbps
- 100G in some carrier core networks
  100G is still more expensive than 10×10G
- Note: only decimal **bit** rates are used in networking
  No cheating like binary byte values used in storage
  1 Gbps = 10^9 b/s, Buy 256 GB Disk = 238.4 GB storage

Ref: http://en.wikipedia.org/wiki/100_Gigabit_Ethernet

Student Questions

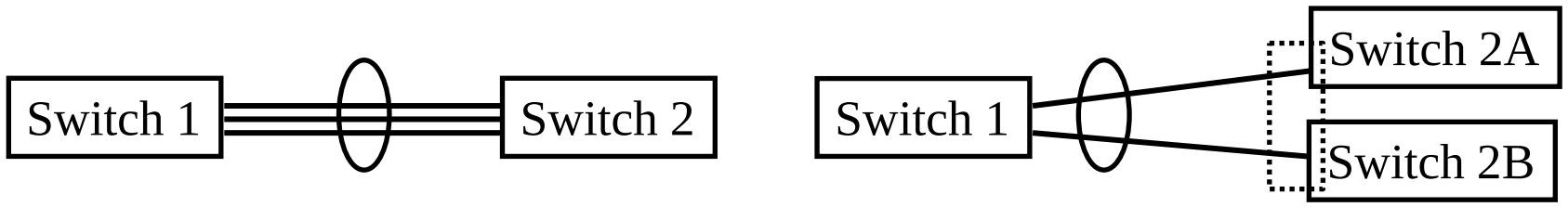
- Is Ethernet speed decided by Ethernet switching throughput?

No. Ethernet nominal speed is the bit rate on the wire. A 10 Gbps Ethernet link has 10 G bits per second on the wire.

- What exactly is CSMA/CD?

CSMA/CD=Carrier Sense Multiple Access with Collision Detection = Listen before you speak and stop if you hear some one else.

Link Aggregation Control Protocol (LACP)



- IEEE 802.1AX-2008/IEEE 802.3ad-2000
- Allows several parallel links to be combined as one link
  3×1Gbps = 3 Gbps
- Allows any speed links to be formed
- Allows fault tolerance
  ⇒ Combined Link remains connected even if one of the member links fails
- Several proprietary extensions. E.g., aggregate links to two switches which act as one switch.

Ref: Enterasys, "Enterasys Design Center Networking – Connectivity and Topology Design Guide," 2013, http://www.enterasys.com/company/literature/datacenter-design-guide-wp.pdf

Student Questions

- Is this essentially just redundant links between switches on the left side?

They are not redundant links. They are parallel links. You send different packets on different links at the same time.

- links are aggregated using LACP, but switches are daisy-chained?

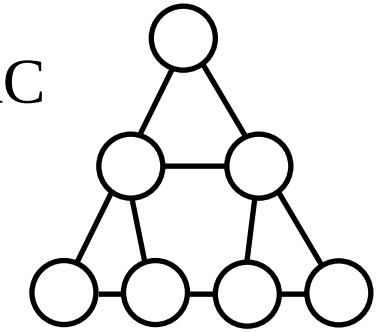
There is no daisy-chaining. The figures have been redrawn to remove this confusion.

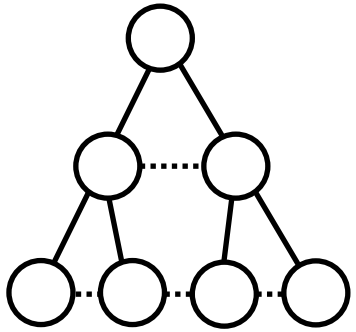
- When we use two links to two distinct switches (Switch 2A and 2B), do we send frames to both of them? or, we just use one of them and the other one is actually standby (like HA solutions or HSRP protocol

We send packets to just one. That's how we double the data rate. If we send to both, the data rate is equal to the minimum of the two.

Spanning Tree Algorithm

- Helps form a tree out of a mesh topology
- All bridges multicast to "All bridges"
  - My ID. 64-bit ID = 16-bit priority + 48-bit MAC address.
  - Root ID
  - My cost to root
- The bridges update their info using Dijkstra's algorithm and rebroadcast
- Initially all bridges are roots but eventually converge to one root as they find out the lowest Bridge ID.
- On each LAN, the bridge with minimum cost to the root becomes the Designated bridge
- All ports of all non-designated bridges are blocked.





Student Questions

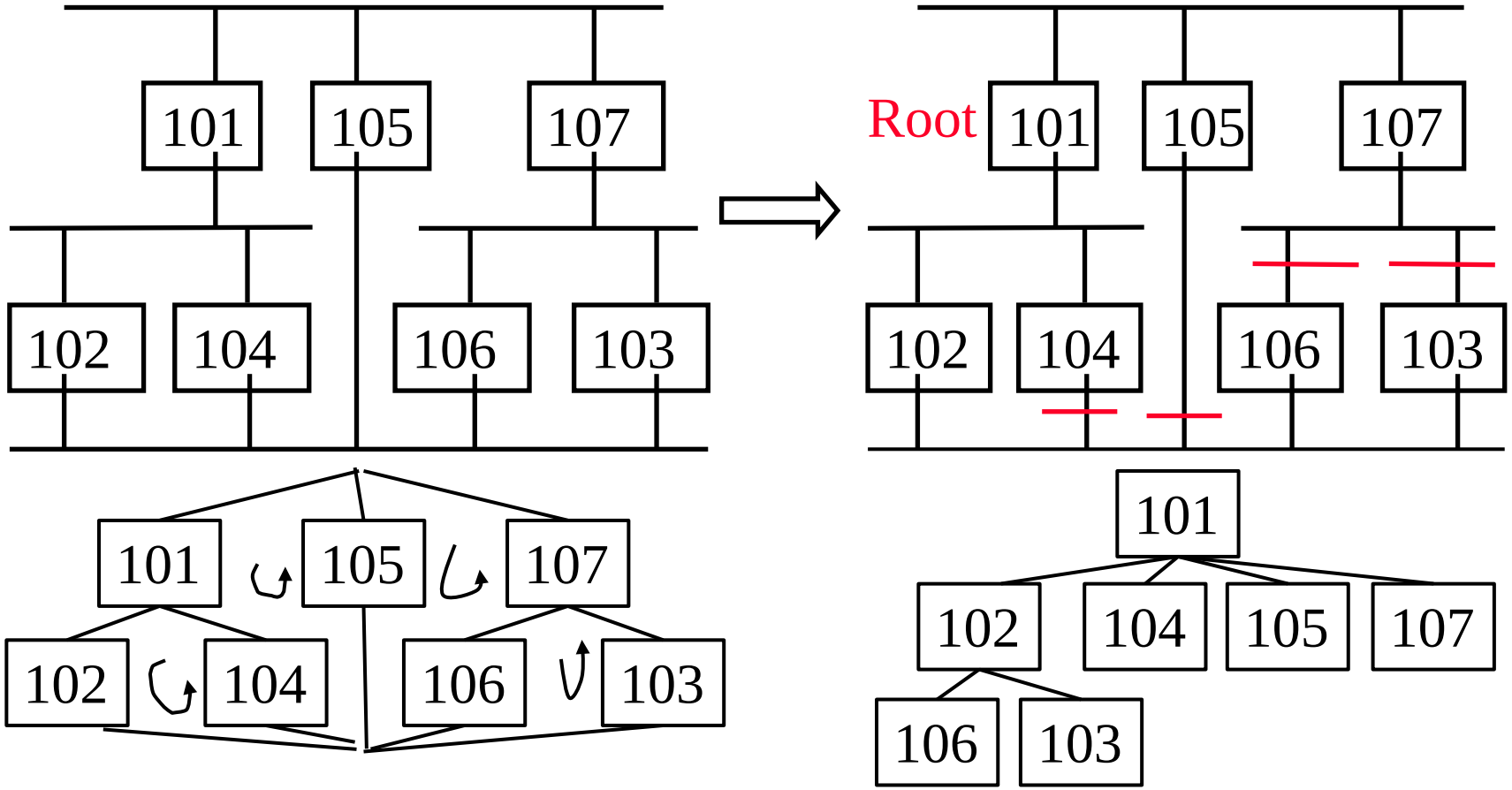
- In a spanning tree, the bridge with lowest Bridge ID will become root, which means it will have the highest priority (lowest first 16-bit priority)? Is there a particular reason for doing this? What does higher priority mean to the bridge (e.g. traffic through this bridge will have less chance to be blocked)?

Root does not have any special privilege. It simply has to to forward more traffic. Generally you give highest priority to your most powerful bridge. The priority overrides the address.

- Why form the original topology to a spanning tree? It seems some links are wasted and the distance between some nodes is farther.

Original topology is fixed and is done based on location and convenience. Spanning tree ensures that there are no loops.

Spanning Tree Example



Student Questions

- Should we kill the links between 102 and 104, 104 and 106, 106 and 103? Can 107 be the root?

??

- Could you explain more about this example?

Sure.

- When there is an equal cost path to the root, how STP decides which path to block?

Randomly. Or by Equal cost multipath (hashing).

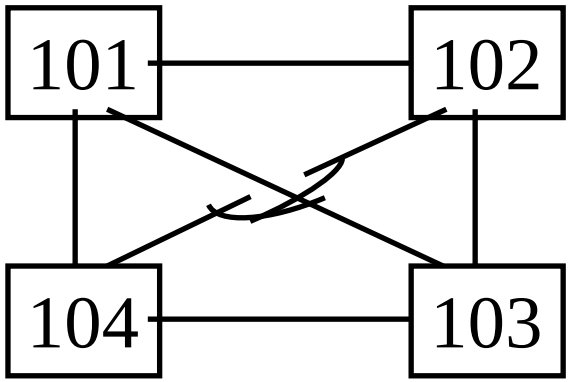
- *Could you please draw it as a tree and explain.*

Done.

Ref: Cisco, "Understanding Spanning-Tree Protocol Topology Changes," http://www.cisco.com/en/US/tech/tk389/tk621/technologies_tech_note09186a0080094797.shtml

Homework 4

- Which links in the following diagram will be blocked by spanning tree? Justify your answer.



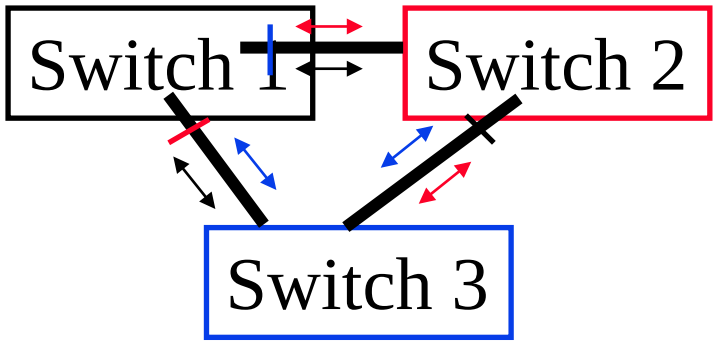
Student Questions

Enhancements to STP

- A topology change can result in 1 minute of traffic loss with STP ⇒ All TCP connections break
- Rapid Spanning Tree Protocol (RSTP)
  IEEE 802.1w-2001 incorporated in IEEE 802.1D-2004
- One tree for all VLANs ⇒ Common spanning tree
- Many trees ⇒ Multiple spanning tree (MST) protocol
  IEEE 802.1s-2002 incorporated in IEEE 802.1Q-2005
- One or more VLANs per tree.

Student Questions

MSTP (Multiple Spanning Tree)



- MSTP (Multiple STP)
  IEEE 802.1s-2002 incorporated in IEEE 802.1Q-2005
- Each tree serves a group of VLANs.
- A bridge port could be in forwarding state for some VLANs and blocked state for others.

Student Questions

IS-IS Protocol

- Intermediate System to Intermediate System (IS-IS) is a protocol to build routing tables. Link-State routing protocol ⇒ Each nodes sends its connectivity (link state) information to all nodes in the network
- Dijkstra's algorithm is then used by each node to build its routing table.
- Similar to OSPF (Open Shortest Path First).
- OSPF is designed for IPv4 and then extended for IPv6.
  IS-IS is general enough to be used with any type of addresses
- OSPF is designed to run on the top of IP
  IS-IS is general enough to be used on any transport
  ⇒ Adopted by Ethernet

Ref: http://en.wikipedia.org/wiki/IS-IS

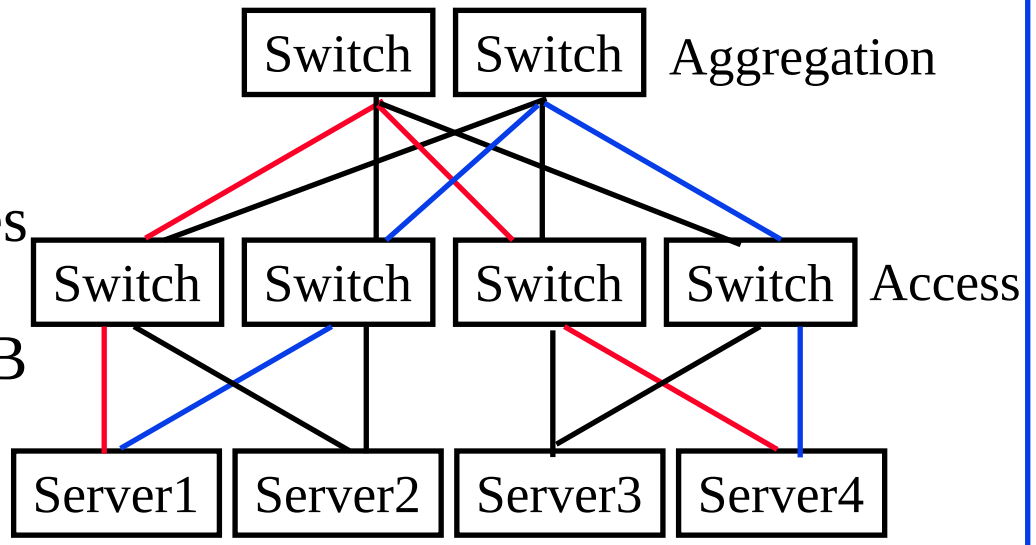
Student Questions

- Can IS-IS be used as a routing protocol for layer 3?

Yes. And, it is used in some networks.

Shortest Path Bridging

- IEEE 802.1aq-2012 (later incorporated in 802.1Q-2014)
- Allows all links to be used ⇒ Better CapEx
- IS-IS link state protocol (similar to OSPF) is used to build shortest path trees for each node to every other node within the SPB domain
- Equal-cost multi-path (ECMP) used to distribute load



Ref: http://en.wikipedia.org/wiki/Shortest_Path_Bridging

Student Questions

- How does OSPF prevent Loops?
- *Using Dijkstra's algorithm.*
- Do we use the shortest path bridging in data centers

Yes.

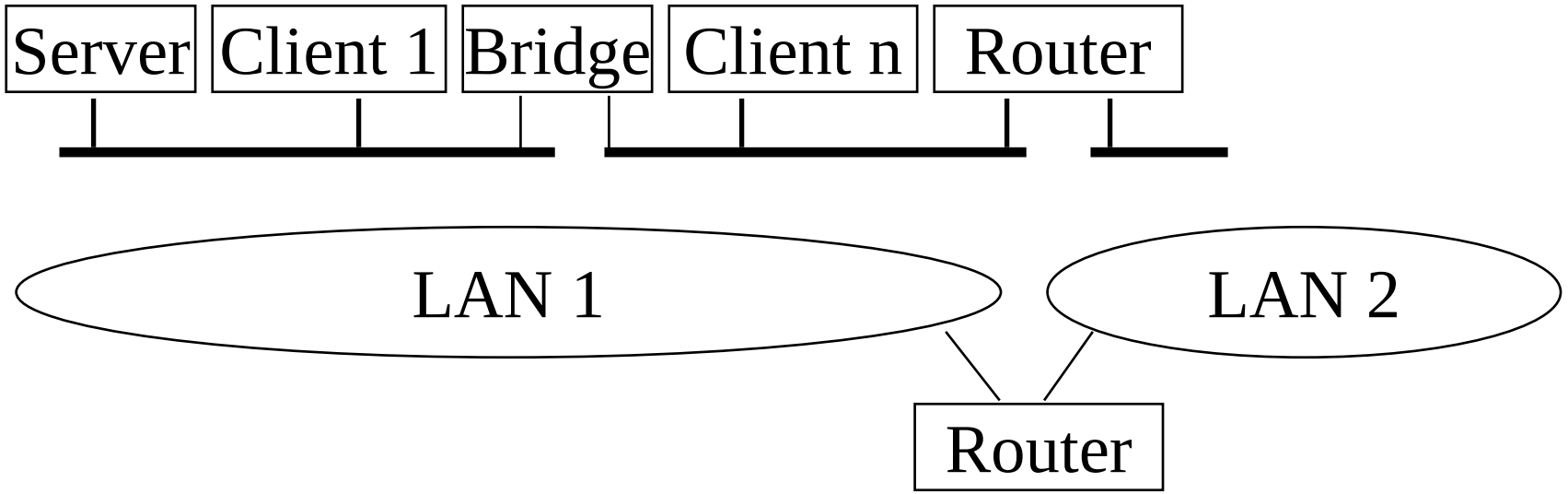
- Is the shortest path always the path with the minimum number of hops or can we consider other cost functions like congestion?

You can specify any cost function, such as data rate, dollar cost, or hops.

- I assume flow matching is expensive because we need TCAMs. Particularly, if we want to use shortest path bridging in data centers with a lot of short flows, isn't it too expensive?

Flow matching is done by hashing the source and destination L2 or L3 addresses. In higher-layer switches, it can include higher layer info, such as, TCP port #. TCAMs are not used for this.

What is a LAN?

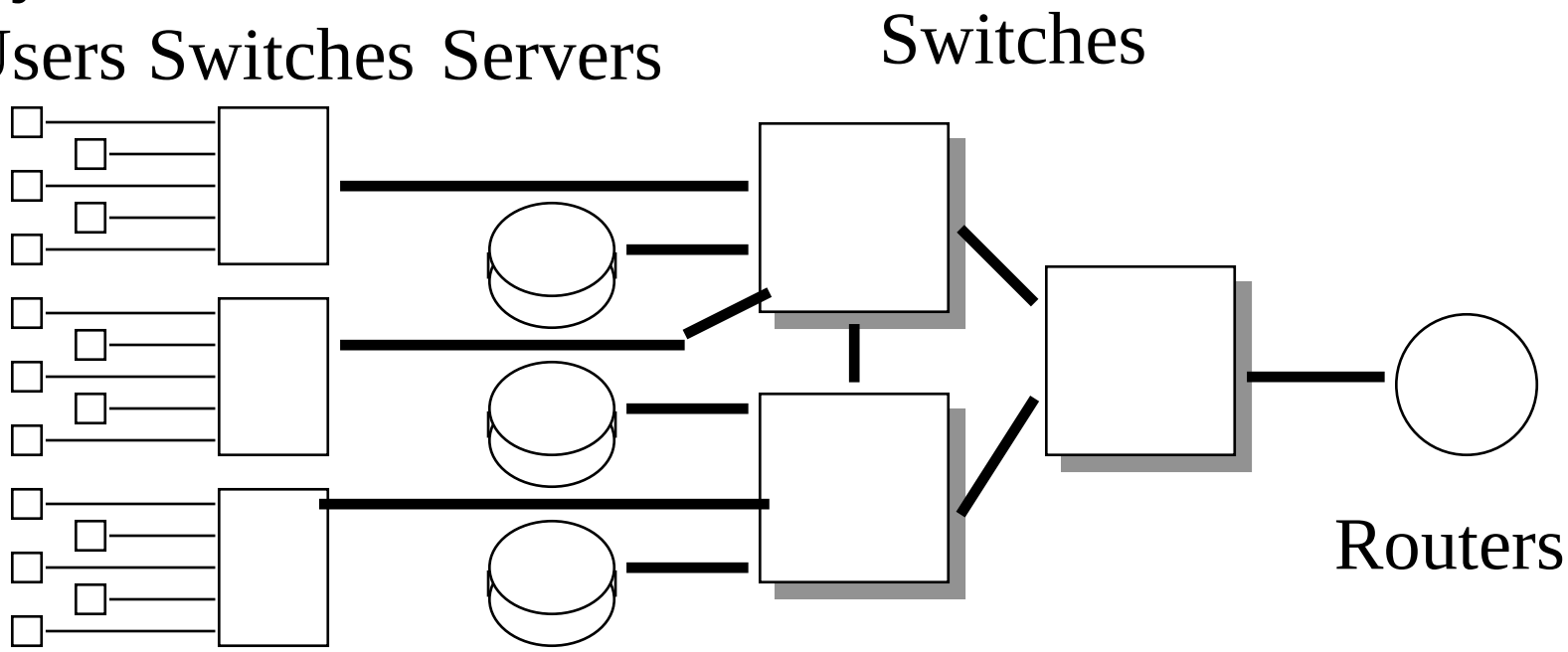


- LAN = Single broadcast domain = Subnet
- No routing between members of a LAN
- Routing required between LANs

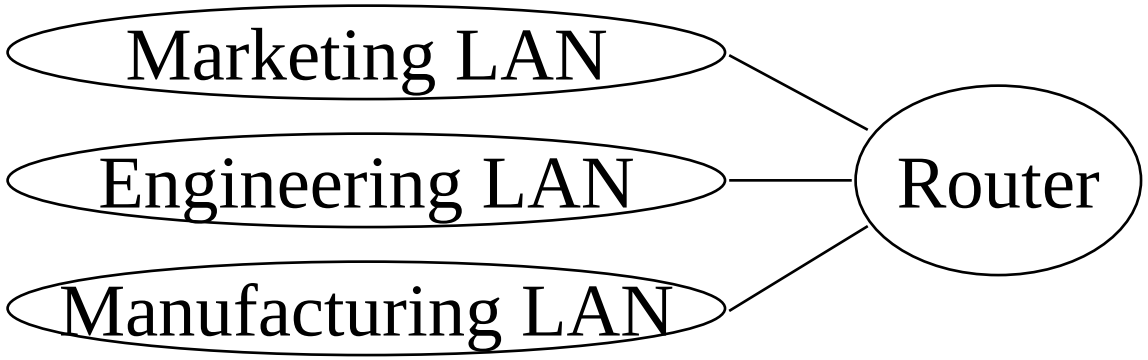
Student Questions

What is a Virtual LAN?

- **Physical View**



- **Logical View**

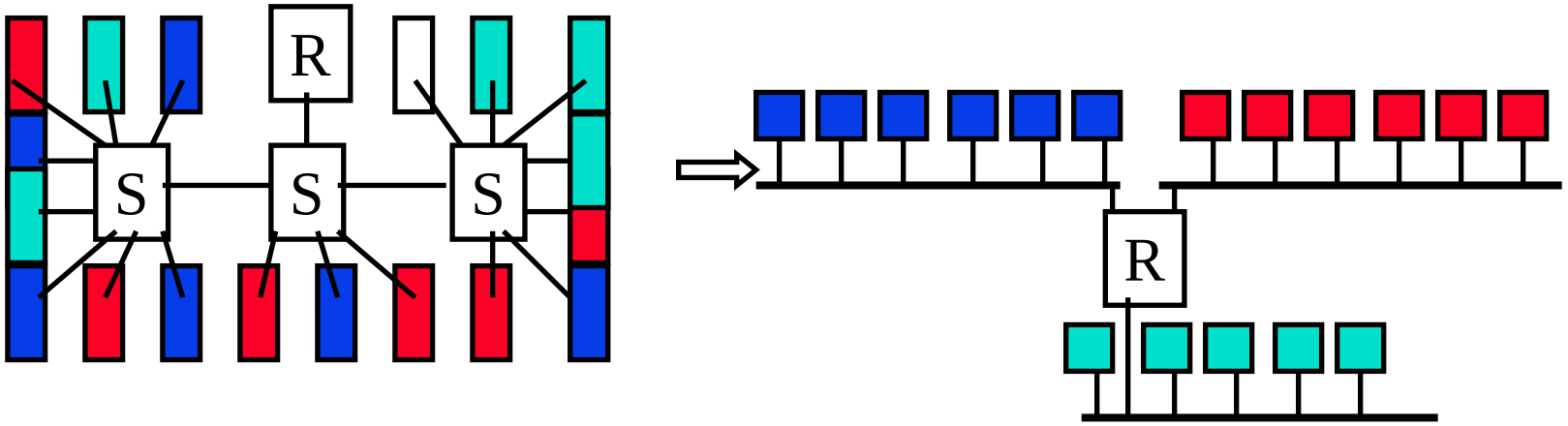


Student Questions

- What are the disadvantages of VLANs?

Complexity in switches.
⇒ Many cheap switches may not have this feature.

Virtual LAN



- Virtual LAN = Broadcasts and multicast goes only to the nodes in the virtual LAN
- LAN membership defined by the network manager ⇒ Virtual

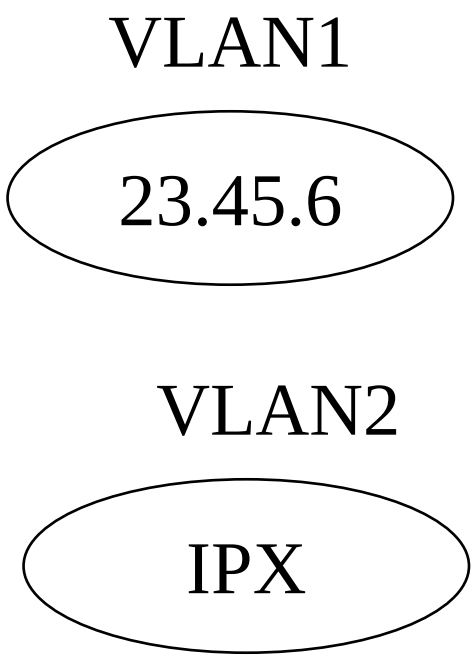
Student Questions

Types of Virtual LANs

- Layer-1 VLAN = Group of Physical ports
- Layer-2 VLAN = Group of MAC addresses
- Layer-3 VLAN = IP subnet

Switch Port	VLAN 1	VLAN 2
A1	√	
A2		√
A3	√	
B1		√
B2	√	

VLAN1	VLAN2
A1B234565600	21B234565600
D34578923434	634578923434
1345678903333	8345678903333
3438473450555	9438473450555
4387434304343	5387434304343
4780357056135	6780357056135
4153953470641	9153953470641
3473436374133	0473436374133
3403847333412	8403847333412
3483434343143	8483434343143
4343134134234	0343134134234



Student Questions

- I imagine different types of VLANs are managed by the device at their corresponding layers? For example a router takes care of layer-3 VLANs.

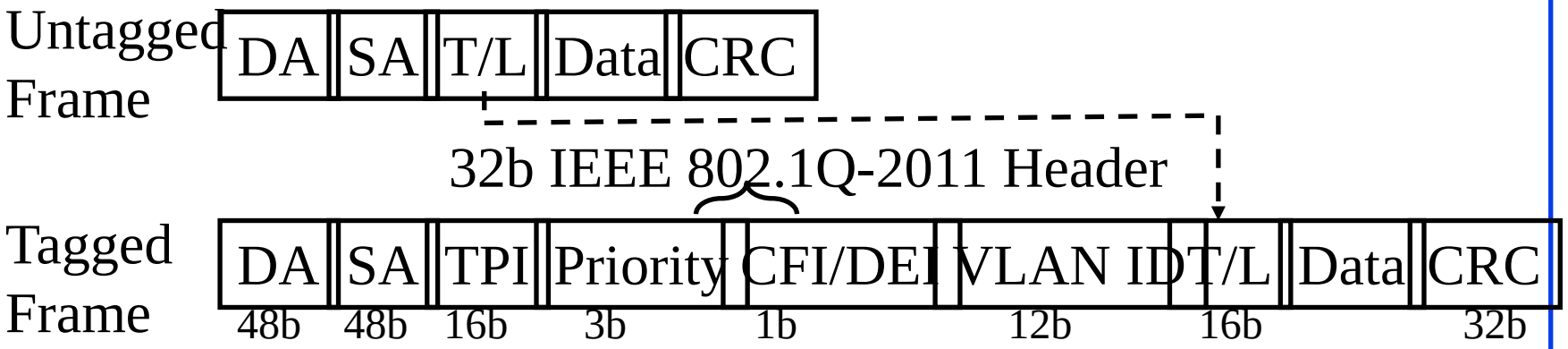
For L3-VLANs, you need L3 information and so only L3 device can manage it. DHCP is an L3 device. It can take care of L2, L3 VLANs. L1 VLANs are done manually or by L1 wiring automation.

- After configuring VLAN, when a port belonging to a VLAN of the switch receives a broadcast frame, will all hosts belonging to the same VLAN receive the broadcast frame?

Yes. VLAN==Shared broadcast

IEEE 802.1Q-2011 Tag

- Tag Protocol Identifier (TPI)
- Priority Code Point (PCP): 3 bits = 8 priorities 0..7 (High)
- Canonical Format Indicator (CFI): 0 ⇒ Standard Ethernet, 1 ⇒ IBM Token Ring format (non-canonical or non-standard)
- CFI now replaced by Drop Eligibility Indicator (DEI)
- VLAN Identifier (12 bits ⇒ 4095 VLANs)
- Switches forward based on MAC address + VLAN ID
  Unknown addresses are flooded.



Ref: Canonical vs. MSB Addresses, http://support.lexmark.com/index?page=content&id=HO1299
Ref: G. Santana, "Data Center Virtualization Fundamentals," Cisco Press, 2014, ISBN:1587143240

Student Questions

Link Layer Discovery Protocol (LLDP)

- IEEE 802.1AB-2009
- Neighbor discovery by periodic advertisements
- Every minute a LLC frame is sent on every port to neighbors
- LLDP frame contains information in the form of Type-Length-Value (TLV)
- Types: My Chassis ID, My Port ID, Time-to-live, Port description (Manufacturer, product name, version), Administratively assigned system name, capabilities, MAC address, IP Address, Power-via-MDI, Link aggregation, maximum frame size, …

Ref: M. Srinivasan, "Tutorial on LLDP," http://www.eetimes.com/document.asp?doc_id=1272069
Ref: http://en.wikipedia.org/wiki/Link_Layer_Discovery_Protocol

Student Questions

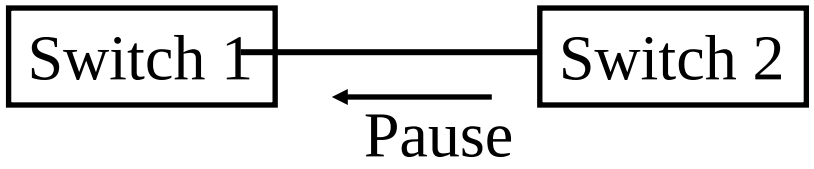
Data Center Bridging

- Goal: To enable storage traffic over Ethernet
- Four Standards:
  - Priority-based Flow Control (IEEE 802.1Qbb-2011)
  - Enhanced Transmission Selection (IEEE 802.1Qaz-2011)
  - Congestion Control (IEEE 802.1Qau-2010)
  - Data Center Bridging Exchange (IEEE 802.1Qaz-2011)
- All of these are now incorporated in IEEE 802.1Q-2014

Ref: M. Hagen, "Data Center Bridging Tutorial," http://www.iol.unh.edu/services/testing/dcb/training/DCB-Tutorial.pdf

Student Questions

Ethernet Flow Control: Pause Frame

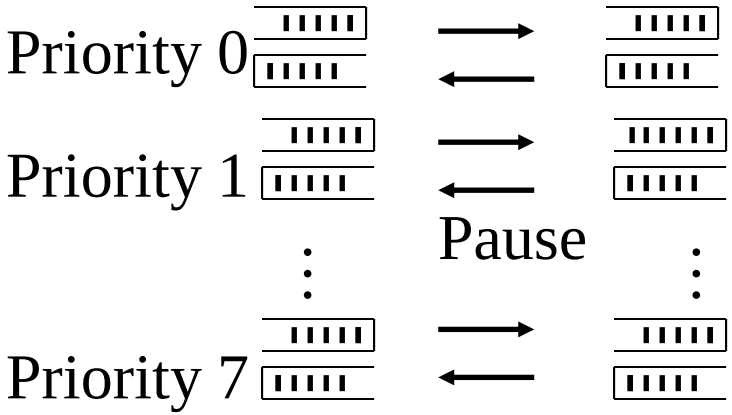


- Defined in IEEE 802.3x-1997. A form of on-off flow control.
- A receiving switch can stop the adjoining sending switch by sending a "Pause" frame.
  Stops the sender from sending any further information for a time specified in the pause frame.
- The frame is addressed to a standard (well-known) multicast address. This address is acted upon but not forwarded.
- Stops all traffic. Causes congestion backup.

Ref: http://en.wikipedia.org/wiki/Ethernet_flow_control

Student Questions

Priority-based Flow Control (PFC)



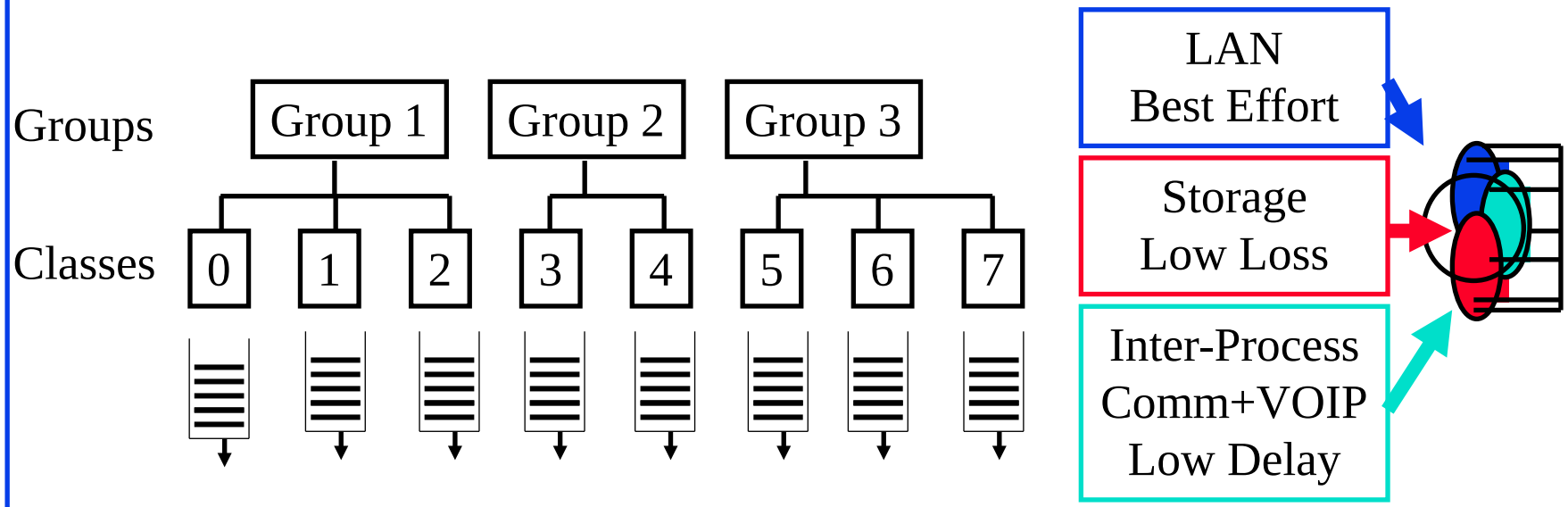
- IEEE 802.1Qbb-2011
- IEEE 802.1Qbb-2011 allows any single priority to be stopped. Others keep sending

Ref: J. L. White, "Technical Overview of Data Center Networks," SNIA, 2013, http://www.snia.org/sites/default/education/tutorials/2012/fall/networking/JosephWhite_Technical%20Overview%20of%20Data%20Center%20Networks.pdf

Student Questions

Enhanced Transmission Selection

- IEEE 802.1Qaz-2011
- Goal: Guarantee bandwidth for applications sharing a link
- Traffic is divided in to 8 classes (not priorities)
- The classes are grouped.
- Standard requires min 3 groups: 1 with PFC (Storage with low loss), 1 W/O PFC (LAN), 1 Strict Priority (Inter-process communication and VOIP with low latency)



Student Questions

- Is the reservation end-to-end? or, we have to manually reserve and define each group on each switch along the path? Can it be used with RSVP?

Here all reservations are L2. So on one extended LAN only. That extended LAN can have many L2 switches in between.

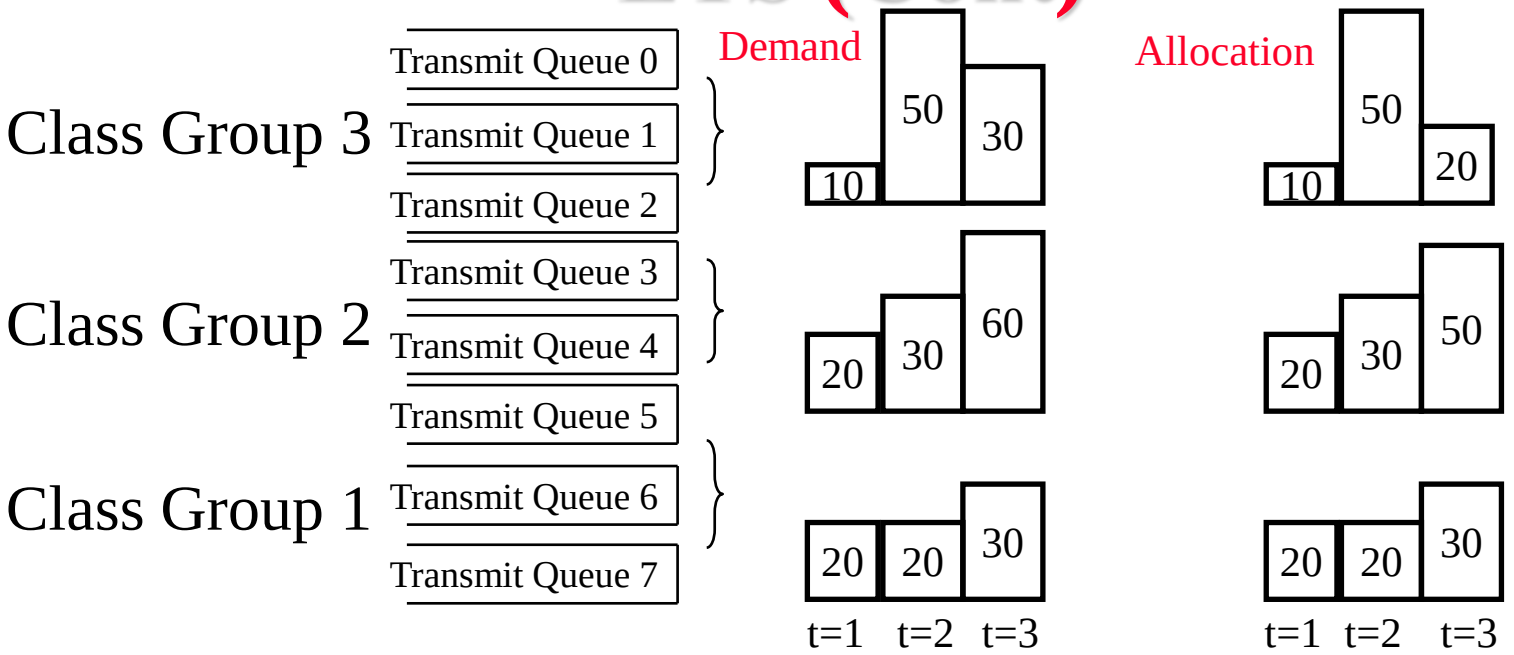
- Priority flow control is referred to as Class-based Flow Control, how classes are related to priorities? what is the main difference between PFC and strict priority?

Priority creates an orderly queue.
Classes do not necessarilty have an order.

- So in a word, ETS was invented to guarantee the minimum bandwidth required to ensure the normal operation of each network?

For real time video services

ETS (Cont)



- Bandwidth allocated per class group in 1% increment but 10% precision (±10% error).
- Max 75% allocated ⇒ Min 25% best effort
- Fairness within a group
- All unused bandwidth is available to all classes wanting more bandwidth. Allocation algorithm **not** defined.
- Example: Group 1=20%, Group 2=30%

Student Questions

- Is the left demand and the right allocation? Is it supposed to coordinate with the numbers on slide 31?

Yes, Yes!

- What do the six different figures on the top right represent?

Top=Group 2, Middle-Group 2, Bottom=Group 1
Left=Demand, Right=Allocation

- In this slide, we defined max-min fairness as an allocation algorithm for ETS, but in slide 29 we are saying "allocation algorithm is not defined" can you please elaborate?

The standard does not require any particular algorithm. Companies can choose their own. Most like max-min fairness.

- Does 10% precision imply minimum 10% allocation? Otherwise, allocation would be negative.

Precision is related to how accurately you measure it.
1 Mbps ≈ 1.2 Mbps, 1.1 Mbps, 1.001 Mpbs

A ETS Fairness Example

- **Max-Min Fairness**: Giving more to any one should not require decreasing to someone with less allocation (Help the poorest first)
- **Example**: In a 3-class group bridge, Groups 1 and 2 have a minimum guaranteed bandwidth of 20% and 30%, respectively.
  In a particular time slot, the traffic demands for group 1, 2, and 3 are 30%, 60%, 30%, respectively. How much should each group get?
- Iteration 1: Group 1 = 20, Group 2= 30,
  Unallocated = 50, Unsatisfied groups = 3
  Fair allocation of unallocated bandwidth = 50/3 per group
- Iteration 2: Group 1 = 20+10 (can't use more), Group 2=30+50/3, Group 3=50/3
  Total Used = 280/3, Unallocated = 20/3, Unsatisfied groups =2,
  Fair share of unallocated bandwidth = 10/3 per group
- Iteration 3: Group 1 = 30, Group 2= 30+50/3+10/3,
  Group 3 = 50/3+10/3
  Total Used = 100, Unallocated = 0 ⇒ Done.

Student Questions

- Could you please explain the example again?

Sure.

- Does the max-min fairness some extent avoid network congestion? For example, an ill-behaved flow consisting of large data packets will only affect itself rather than other flows.

Congestion: Load > Capacity
Fairness: Equal distribution of capacity

- If the guaranteed bandwidth is more than expected demand for one group, should we reallocate the extra bandwidth to the rest groups or just keep it?

Yes, if actual demand < Guaranteed.

- How expensive is it to perform bandwidth allocation adjustment for the groups? In general how often is it performed on a network (e.g. checking the demanded bandwidth and do the fairness calculation)?

Vendor specific. Not sure. But once per second would be reasonable on a Gbps link. It would depend upon the cost of the link.

- In iteration 2, why the total used is 280/3, and unallocated is 20/3?

Used=20+10+30+(50/3)+(50/3) = 280/3
Unallocated = Left over = 100-(280/3)=20/3

- Is it possible the total demand on bandwidth excess 100 percent?

Yes. Two users each could send 100% → Load=200%

Tabular Method for Max-Min Fairness

Iteration		1	2	3	Total	Unused	# Unsatisfied
	Demand	30	60	30	120		
1	Guaranteed Allocation	20	30	0	50	50	
	Total Used	20	30	0	50	50	3
2	Additional Allocation	16.7	16.7	16.7			
	Total Used	30	46.7	16.7	93.3	6.7	2
3	Additional Allocation	0	3.3	3.3			
	Total Used	30	50	20	100	0	2

- Iterations end when either unused capacity or # of unsatisfied groups is zero.

Student Questions

- So the satisfaction is based on the demand not based on providing the minimum guaranteed BW?

Satisfaction = Allocated/Demand

- In the test do we have to use fractions instead of floating points?

You can use either.

- How do you define "poorest" ? In this example (30,60,30) why not to work on satisfying the largest number of groups first? ignoring 60 and satisfying 30 and 30 first?

Poor = Low demand

- If the demand is less than the guaranteed are we allowed to reallocate the unused BW?

Yes, particularly if others need it. Also, there are "best effort" customers whose guarantee is zero.

- Is the demand for each group dynamic? Like if after iteration3, the demand of group 1, 2, 3 change to 10, 10, 10. Will the bandwidth assigned to each group be decreased?

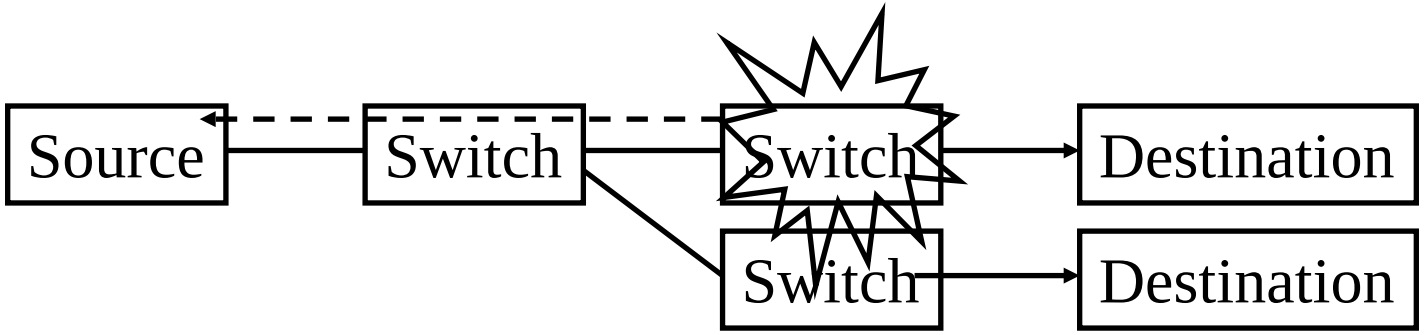
Yes. It is dynamic. We have to reallocate every few seconds.

Homework 4B

- What would be max-min allocation for a 4 group system in which group 1 through 3 are guaranteed 10%, 20%, and 30% respectively. The demands on a 100 Gbps system are 1 Gbps, 4Gbps, and 35 Gbps, and 70 Gbps.

Student Questions

Quantized Congestion Notification (QCN)



- IEEE 802.1Qau-2010 Dynamic Congestion Notification
- A source quench message is sent by the congested switch direct to the source. The source reduces its rate for that flow.
- Sources need to keep per-flow states and control mechanisms
- Easy for switch manufacturers but complex for hosts. Implemented in switches but not in hosts ⇒ Not effective.
- The source may be a router in a subnet and not the real source ⇒ Router will drop the traffic. QCN does not help in this case.

Ref: *I. Pepelnjak,* "DCB Congestion Notification (802.1Qau)," http://blog.ipspace.net/2010/11/data-center-bridging-dcb-congestion.html

Student Questions

- So we use the AIMD instead of QCN?

AIMD is an algorithm for systems with no explicit rate feedback. QCN simply provides a source-quench feedback, so it can be used with AIMD.

- In the video, we were unable to see the slide for QCN. Could you explain it again?

DCBX

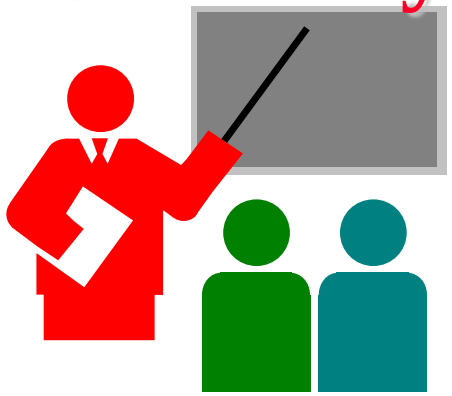
- Data Center Bridging eXchange, IEEE 802.1Qaz-2011
- Uses LLDP to negotiate quality metrics and capabilities for Priority-based Flow Control, Enhanced Transmission Selection, and Quantized Congestion Notification
- New TLV's
  - Priority group definition
  - Group bandwidth allocation
  - PFC enablement per priority
  - QCN enablement
  - DCB protocol profiles
  - FCoE and iSCSI profiles

Student Questions

- What aspects does DCBX extend on the basis of DCB?

It defines a protocol for exchange of parameters.

Summary



1. Ethernet's use of IDs as addresses makes it very easy to move systems in the data center ⇒ Keep traffic on the same Ethernet
2. Spanning tree is wasteful of resources and slow.
   Ethernet now uses shortest path bridging (similar to OSPF)
3. VLANs allow different non-trusting entities to share an Ethernet network
4. Data center bridging extensions reduce the packet loss by enhanced transmission selection and Priority-based flow control

Student Questions

List of Acronyms

- BER Bit Error Rate
- BPDU Bridge Protocol Data Unit
- CD Collision Detection
- CFI Canonical Format Indicator
- CRC Cyclic Redundancy Check
- CSMA Carrier Sense Multiple Access with Collision Detection
- DA Destination Address
- DCB Data Center Bridging
- DCBX Data Center Bridging eXtension
- DEI Drop Eligibility Indicator
- DNS Domain Name System
- ECMP Equal-cost multi-path
- ETS Enhanced Transmission Selection
- GB Giga Byte

Student Questions

List of Acronyms (Cont)

ID	Identifier
IP	Internet Protocol
IEEE	Institution of Electrical and Electronics Engineers
IS-IS	Intermediate System to Intermediate System
iSCSI	Internet Small Computer System Interface
LACP	Link Aggregation Control Protocol
LAN	Local Area Network
LLC	Logical Link Control
LLDP	Link Layer Discovery Protocol
MAC	Media Access Control
MDI	Medium Dependent Interface
MSB	Most significant byte first
MST	Multiple Spanning Tree
MSTP	Multiple Spanning Tree Protocol
OAM	Operations, Administration, and Management

Student Questions

List of Acronyms (Cont)

- OSPF Open Shortest Path First
- OUI Organizationally Unique Identifier
- PCP Priority Code Point
- PFC Priority-based Flow Control
- PHY Physical layer
- QCN Quantized Congestion Notification
- QoS Quality of Service
- RSTP Rapid Spanning Tree Protocol
- SA Source Address
- SNIA Storage Networking Industries Association
- SPB Shortest Path Bridging
- STP Spanning Tree Protocol
- TCP Transmission Control Protocol
- TLV Type-Length-Value
- TPI Tag Protocol Identifier
- VLAN Virtual Local Area Network
- VM Virtual machine

Student Questions

List of Acronyms (Cont)

- VOIP Voice over IP
- WAN Wide Area Network
- WiFi Wireless Fidelity
- WiMAX Wireless Interoperability for Microwave Access

Student Questions

Reading List

- G. Santana, "Data Center Virtualization Fundamentals," Cisco Press, 2014, ISBN:1587143240
- Enterasys, "Enterasys Design Center Networking - Connectivity and Topology Design Guide," 2013, http://www.enterasys.com/company/literature/datacenter-design-guide-wp.pdf
- Cisco, "Understanding Spanning-Tree Protocol Topology Changes," http://www.cisco.com/en/US/tech/tk389/tk621/technologies_tech_note09186a0080094797.shtml
- Cisco, Understanding Rapid Spanning Tree Protocol (802.1w), http://www.cisco.com/en/US/tech/tk389/tk621/technologies_white_paper09186a0080094cfa.shtml
- Canonical vs. MSB Addresses, http://support.lexmark.com/index?page=3Dcontent&id=3DHO1299

Student Questions

Reading List (Cont)

- M. Hagen, "Data Center Bridging Tutorial," http://www.iol.unh.edu/services/testing/dcb/training/DCB-Tutorial.pdf
- J. L. White, "Technical Overview of Data Center Networks," SNIA, 2013, http://www.snia.org/sites/default/education/tutorials/2012/fall/networking/JosephWhite_Technical%20Overview%20of%20Data%20Center%20Networks.pdf
- I. Pepelnjak, "DCB Congestion Notification (802.1Qau)," http://blog.ipspace.net/2010/11/data-center-bridging-dcb-congestion.html

Student Questions

Wikipedia Links

- http://en.wikipedia.org/wiki/10-gigabit_Ethernet
- http://en.wikipedia.org/wiki/100_Gigabit_Ethernet
- http://en.wikipedia.org/wiki/Data_center
- http://en.wikipedia.org/wiki/Data_center_bridging
- http://en.wikipedia.org/wiki/Data_link_layer
- http://en.wikipedia.org/wiki/EtherChannel
- http://en.wikipedia.org/wiki/Ethernet
- http://en.wikipedia.org/wiki/Ethernet_flow_control
- http://en.wikipedia.org/wiki/Ethernet_frame
- http://en.wikipedia.org/wiki/Ethernet_physical_layer
- http://en.wikipedia.org/wiki/EtherType
- http://en.wikipedia.org/wiki/Fast_Ethernet
- http://en.wikipedia.org/wiki/Gigabit_Ethernet

Student Questions

Wikipedia Links (Cont)

- http://en.wikipedia.org/wiki/IEEE_802.1aq
- http://en.wikipedia.org/wiki/IEEE_802.1D
- http://en.wikipedia.org/wiki/IEEE_802.1Q
- http://en.wikipedia.org/wiki/IEEE_802.3
- http://en.wikipedia.org/wiki/IEEE_P802.1p
- http://en.wikipedia.org/wiki/IS-IS
- http://en.wikipedia.org/wiki/Link_Aggregation
- http://en.wikipedia.org/wiki/Link_Aggregation_Control_Protocol
- http://en.wikipedia.org/wiki/Link_layer
- http://en.wikipedia.org/wiki/Link_Layer_Discovery_Protocol
- http://en.wikipedia.org/wiki/Logical_link_control
- http://en.wikipedia.org/wiki/MAC_address
- http://en.wikipedia.org/wiki/MC-LAG

Student Questions

Wikipedia Links (Cont)

- http://en.wikipedia.org/wiki/Media_Independent_Interface
- http://en.wikipedia.org/wiki/Minimum_spanning_tree
- http://en.wikipedia.org/wiki/Network_switch
- http://en.wikipedia.org/wiki/Organizationally_unique_identifier
- http://en.wikipedia.org/wiki/Port_Aggregation_Protocol
- http://en.wikipedia.org/wiki/Priority-based_flow_control
- http://en.wikipedia.org/wiki/RSTP
- http://en.wikipedia.org/wiki/Shortest_Path_Bridging
- http://en.wikipedia.org/wiki/Spanning_tree
- http://en.wikipedia.org/wiki/Spanning_Tree_Protocol
- http://en.wikipedia.org/wiki/Subnetwork_Access_Protocol
- http://en.wikipedia.org/wiki/Virtual_LAN

Student Questions

Scan This to Download These Slides

Raj Jain
http://rajjain.com

http://www.cse.wustl.edu/~jain/cse570-21/m_04dce.htm

Student Questions

- Slides on min-max fairness tabular method and QCN are missing.

If a slide takes less than 5 seconds, it does not appear in the video.

- When it comes to explaining calculation, it's not very clear in the video, because we cannot see which part of the ppt is being pointed while speaking. Is it possible to do an illustration from a blank paper?

Good feedback for future.

Related Modules

CSE567M: Computer Systems Analysis (Spring 2013),
https://www.youtube.com/playlist?list=PLjGG94etKypJEKjNAa1n_1X0bWWNyZcof

CSE473S: Introduction to Computer Networks (Fall 2011),

https://www.youtube.com/playlist?list=PLjGG94etKypJWOSPMh8Azcgy5e_10TiDw

Wireless and Mobile Networking (Spring 2016),
https://www.youtube.com/playlist?list=PLjGG94etKypKeb0nzyN9tSs_HCd5c4wXF

CSE571S: Network Security (Fall 2011),

https://www.youtube.com/playlist?list=PLjGG94etKypKvzfVtutHcPFJXumyyg93u

Video Podcasts of Prof. Raj Jain's Lectures,
https://www.youtube.com/channel/UCN4-5wzNP9-ruOzQMs-8NUw

Student Questions