

BCMSN Study Notes

S.J. Iveson

Preface

This document is only relevant:

- **Up to version 12.05 of the Cisco IOS.**
- **To the CCNP v2.0 track.**
- **To the material taught on the BCMSN course.**

This document should not be used as a cheat to be memorised or learnt by rote in order to pass the relevant CCNP exam without attending a course or having any practical experience. This is generally not possible with Cisco exams anyway, due to their structure and depth.

This guide is intended for reference or last minute revision by candidates who are due to take the exam after having taken a course and/or having had a fair amount of relevant practical experience.

These notes were created by the author while studying for the exam on a Global Knowledge/Geotrain course. The author passed first time with a score of 825.

If you have any comments regarding this document please e-mail: sjiveson@routerzone.com or visit my website at www.routerzone.com.

Steven Iveson 2000

This document is not sponsored by, endorsed by or affiliated with Cisco Systems, Inc.

CONTENTS

CONTENTS	3
SECTION ONE: Traditional Campus Networks	6
Flow Control & Ethernet	7
A Traditional Campus Network:.....	9
ISSUE: Excessive Collisions - SOLUTION: Bridging	10
ISSUE: Broadcast Storms - SOLUTION: Routing.....	10
ISSUE: Router Latency/Cost - SOLUTION: Bridges & Routers.....	11
SOLUTION: Multilayer Switching	11
SECTION TWO: Emerging Campus Networks	13
Today's requirements include:	14
Services, Traffic Flows:.....	14
Switching, Layers.....	15
Switching, Layer 2 (Data Link)	15
Switching, Layer 3 (Network).....	15
a.k.a Hardware Based Routing.....	15
Switching, Layer 4.....	16
Multilayer Switching - MLS.....	16
Catalyst Multilayer Switching Caching.....	16
SECTION THREE: The Hierarchical Model	17
Campus Network Building Blocks.....	19
Switch Blocks;	20
Core Block;.....	20
Collapsed Core;.....	20
Dual Core;.....	20
Core Sizing:	21
Layer 2 Core; Backbone Scaling (Layer 2 – Layer 3 – Layer 2)	21
Layer 3 Core; Backbone Scaling (Layer 2 – Layer 3 – Layer 3)	21
SECTION FOUR: Configuring a Switch Block	22
Basic Device & Port Configuration	23
VLAN Challenges, Characteristics and Problems	25
VLAN Boundaries:	27
VLAN Membership:.....	29
Configuring VLANs	29
Link Types:.....	30
Trunk Negotiation	31
Configuring Trunk Links.....	32
VTP – VLAN Trunk Protocol	33
VTP Domains	34
VTP Modes	34
VTP Advertisements	35
VTP Configuration Tasks.....	36
VTP Pruning	37
SECTION FIVE: Spanning Tree & Redundant Links	38
Spanning Tree Protocol	39
Bridge Protocol Data Unit (BPDU).....	39
Spanning Tree & VLANs.....	42
Scaling STP	43
Fast EtherChannel.....	45
PortFast	46
UplinkFast	46
BackboneFast	47
APPENDIX	64

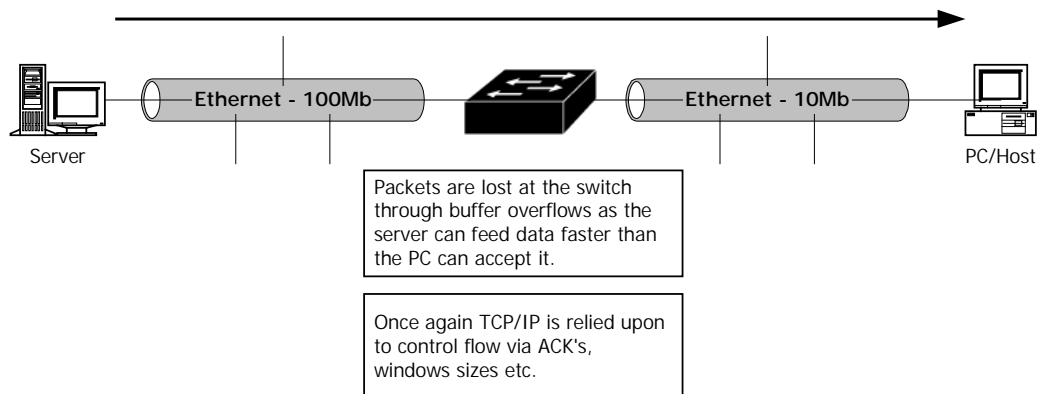
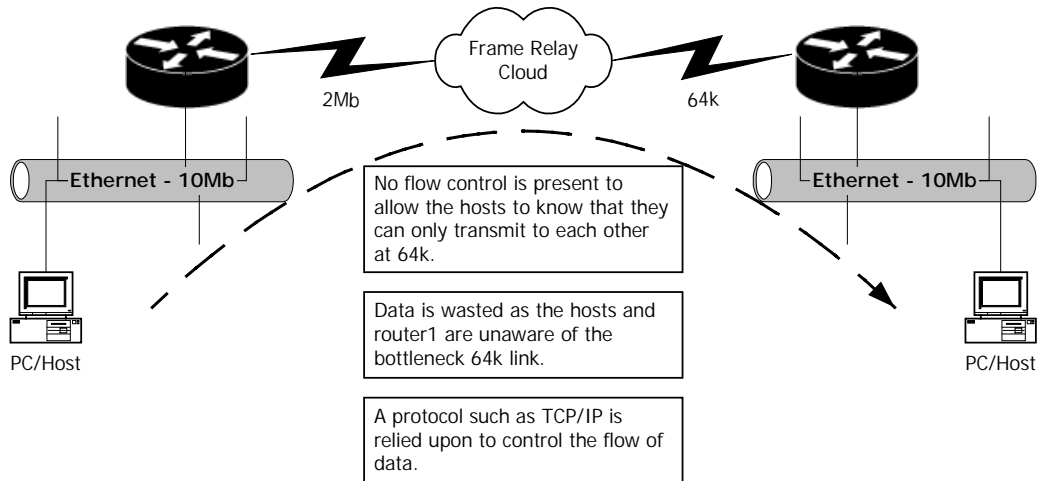
Switching Types.....	65
Ping	65
XCast Transmissions	65
Acronyms & Terms.....	66
Standards.....	66
Lineage	67

SECTION ONE: Traditional Campus Networks

Flow Control & Ethernet

There is no 'built in' flow control, or logic within the design of an Ethernet network.

For example:



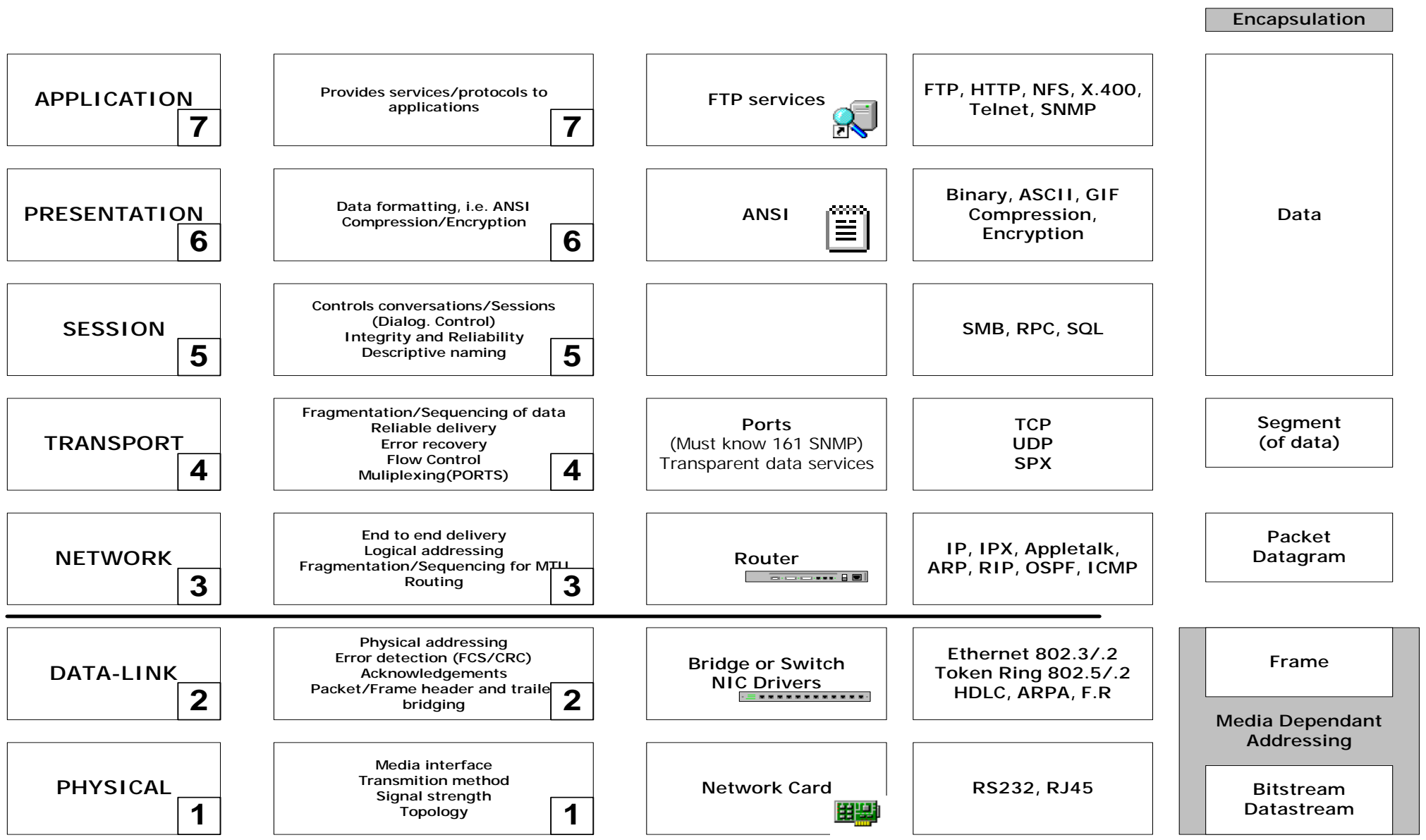
An error free network is impossible. Bottlenecks should be avoided to prevent data loss and re-transmission.

- Broadcasts, which all machines receive, travel all the way up the 7-layer model whether intended for the receiving host or not.
- Multicasts, which all machines also receive use a phsdo MAC address only travel to layer 2 if not intended for the receiving host.

Collisions were traditionally required to control media access. Using a switch mostly removes this problem by creating network segments with only 2 hosts attached at a time. I.e. every port is a network segment/collision domain.

Reducing the number of machines in a collision domain allows you to increase the number of broadcasts without affecting network performance. WE NEED BROADCASTS...

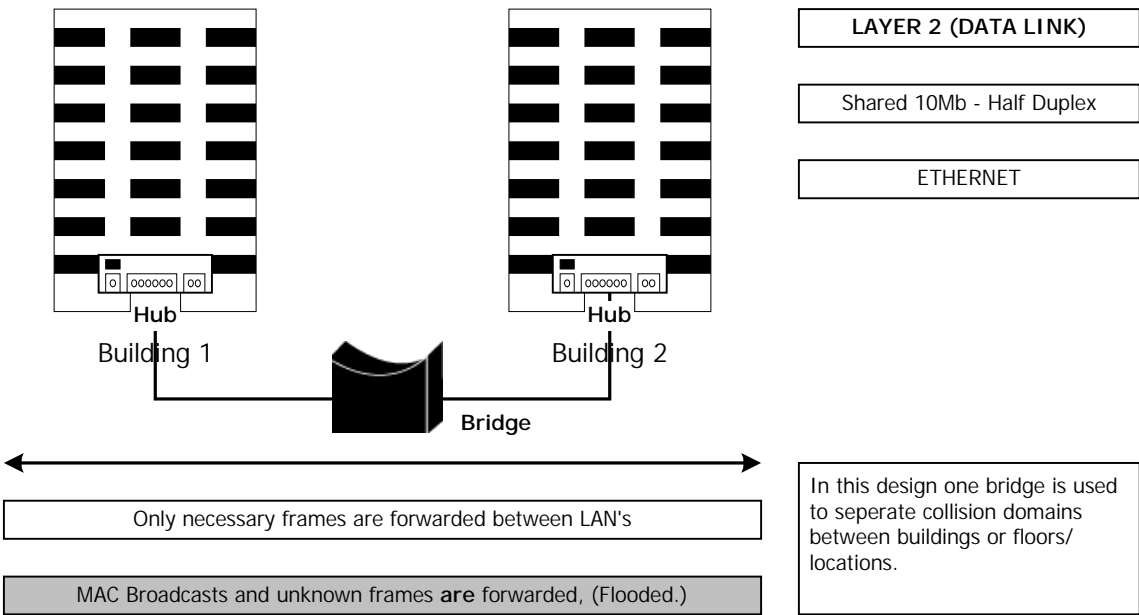
The default Ethernet frame size is 15kb.



A Traditional Campus Network:

- A building or collection of buildings connected within one network that consists of multiple LANs, contained within a fixed geographical area.
- The primary technology used is LAN based. The company normally owns the wires used within the campus. All links are the property of and controlled by the organisation in question

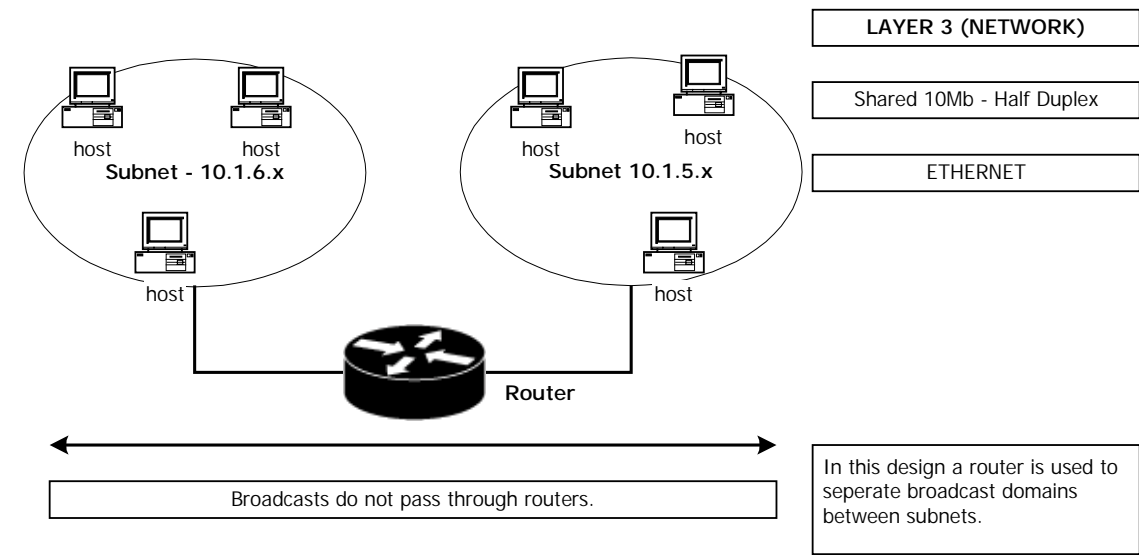
ISSUE: Excessive Collisions - SOLUTION: Bridging



Continued Problems:

- Broadcasts can still flood the network. (i.e. WINS, DHCP requests and also SAP and RIP.)
- Loops can occur causing broadcast storms.
- All workstations have to process broadcasts all the way up to the application layer.

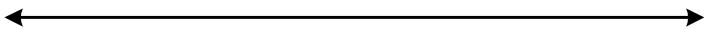
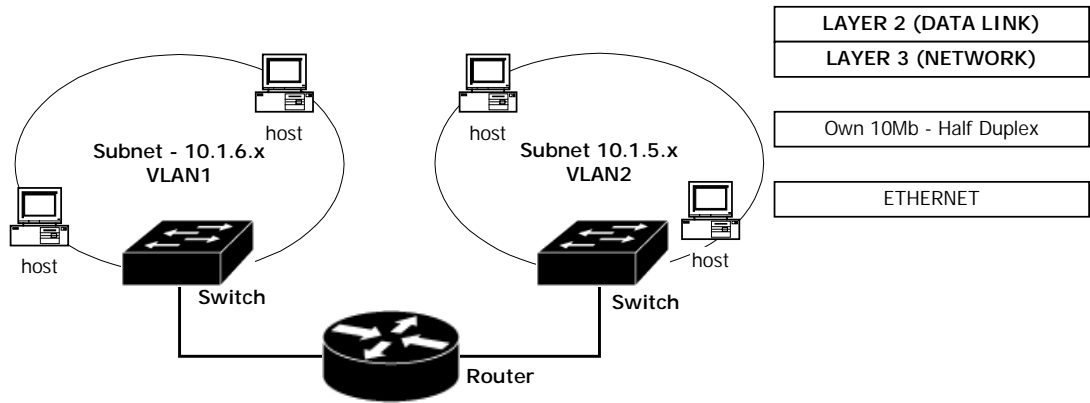
ISSUE: Broadcast Storms - SOLUTION: Routing



Continued Problems:

- Required servers etc. much be located on the same segment.
- Processing packets can create a bottleneck in the network, high latency.
- Routers are expensive and not suitable for separating large numbers of hosts.

ISSUE: Router Latency/Cost - SOLUTION: Bridges & Routers



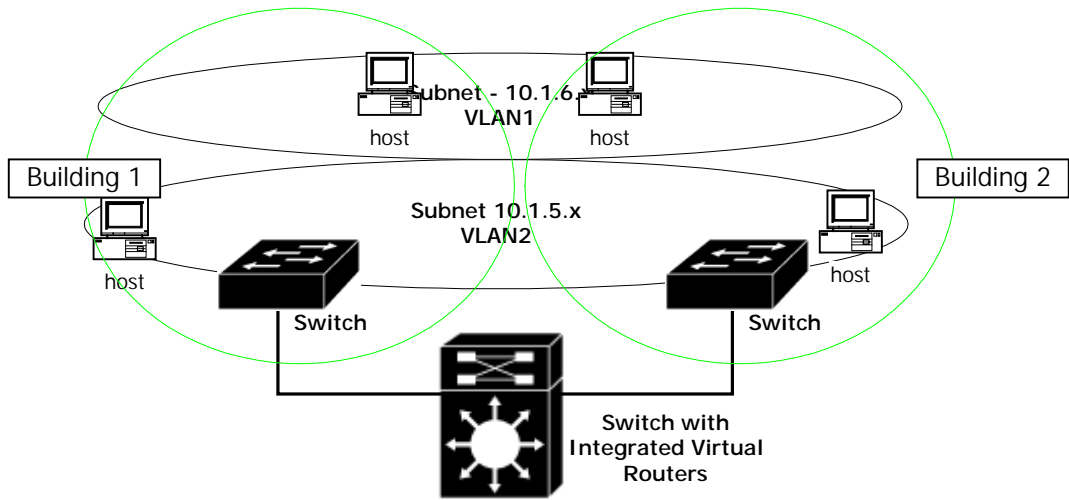
Broadcasts do not pass through routers. Switches prevent collisions occurring.

VLAN's allow hosts in differing physical locations to be grouped as if on a single shared media.

In this design a router is used to separate broadcast domains between subnets and a switch/VLAN is used to minimize collisions.

- Continued Problems:**
- Required servers etc. much be located on the same segment.
 - Processing packets at the router can create a bottleneck in the network and cause high latency.
 - Routers are expensive and not suitable for separating large numbers of hosts.

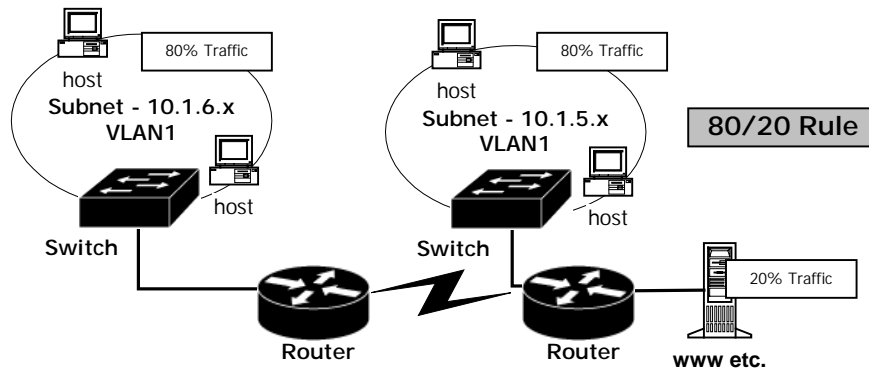
SOLUTION: Multilayer Switching



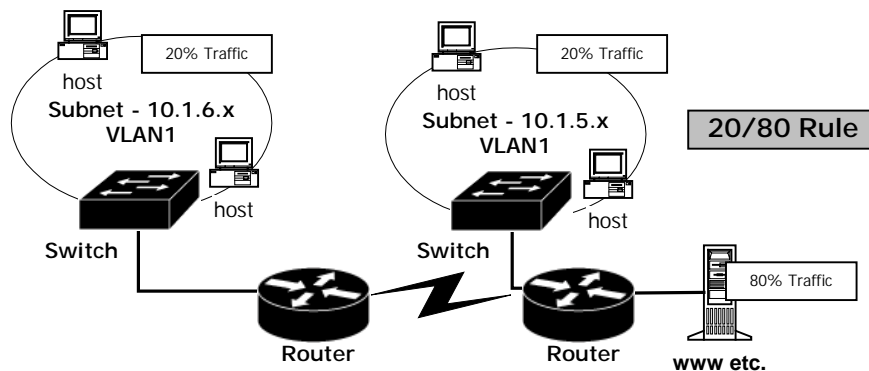
Traffic Flow

Users with common interests and work patterns are generally placed in the same logical network as the servers and services they access most frequently. This confines most network traffic to the LAN and minimizes the impact on the backbone.

Traditionally, in a well-designed network, 80% of traffic is local and 20% crosses the network backbone.



Now the situation is reversed due to new applications and services, which are hard to predict, such as web access and also because of server consolidation. This is known as the 20/80 rule.



Because of this, the layer three, (routing,) performance of the network needs to match the speed of layer two, (bridging/switching,) processing.




VLANS assume most traffic is local, now most traffic is remote and crosses subnets and VLANS so routing, (layer 3,) technology is required at switching speeds.

SECTION TWO: Emerging Campus Networks

Today's requirements include:

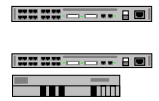
- Fast convergence (network adapts to change quickly)
- Control of data paths and flows (for users and/or applications)
- Control of data paths and flows for fail over (network always available)
- Scalable size and bandwidth
- Support for the new 20/80 rule
- Support for multicasting (new applications)

Services, Traffic Flows:

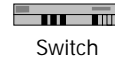
- **Local** services. Connected with switches, traffic remains within the subnet/VLAN
 - No routing, traffic does not use the network backbone 
 - Layer 2
- **Remote** services. Are not on the same subnet but are geographically close to the user, connected via switches and routers, traffic crosses subnets/VLANs.
 - Routed, traffic **may** use the network backbone 
 - Layer 2 and Layer 3
- **Enterprise** services. Are common to all users, traffic will cross subnets and the network backbone.
 - Switched and routed, traffic **will** use the network backbone
 - Placed close to the network backbone in own subnet
 - Layer 2 and Layer 3 (and then layer 2) 

Switching, Layers

Model Layer	PDU Type	Device Type	Based Upon
4 – Transport	TCP Segments	Routers/Access Lists	TCP/Logical Ports/Application
3 – Network	Packets	Routers	TCP/IP Address
2 – Data Link	Frames	Switches/Bridges	MAC Address
1 – Physical		NICs	



Switching, Layer 2 (Data Link)



Switch

- Hardware bridging
 - Low latency
 - Uses ASICs (Application Specific Integrated Circuits)
- Wire speed performance
- Increased network bandwidth (Full duplex, 100Mb)
- Low cost
- NO MODIFICATION OF PACKET REQUIRED

- Flatter network design, fewer hosts per physical subnet, more segments
- Has same problems as bridging regarding broadcasts and STP
- STILL A NEED FOR LAYER 3 FUNCTIONALITY

Switching, Layer 3 (Network)



- Enables the use of the following router features;
 - Security
 - Multiple/Optimal paths
 - Traffic management
 - Broadcast and multicast control
 - Logical addressing
- Hardware based packet forwarding
- High speed packet switching
- High speed scalability
- Low latency, lower cost
- Flow accounting, security and QoS
- A layer three switch can be used in the network to replace routers.

The first packet is passed to the router; the rest of the flow is switched based upon the routing decision.

a.k.a Hardware Based Routing

How it works:

- The first packet is sent to a router switch module, (card in the switch/RSM.)
- Normal router processing is performed on the packet
 - Path determination
 - Security via access lists etc.
 - Packet expiration
- The rest of the flow is directly switched, (or not,) based upon the routing decision made on the first packet, bypassing the RSM.

Switching, Layer 4

Works in a similar way to layer 3 switching with the addition of considering applications. I.e. TCP or UDP port numbers.

Layer 4 switching allows:

- The use of extended access lists
- Accounting via Netflow switching
- Prioritisation by application
- QoS (i.e. granting one application more bandwidth than another.)

Layer 4 switching requires larger forwarding table capacity, particularly within the core of a network. With layer 2 and 3 switching the table is based upon the number of network devices, with layer 4 the table is based upon the number of devices * the number of applications and conversations in use. (This is a memory, not processor, overhead.)

Multilayer Switching - MLS

- A combination of layer 2 and 3 switching in general and layer 2, 3 and 4 switching on Catalyst switches
- High-speed scalability
- Low latency
- ROUTE ONCE, SWITCH MANY

Catalyst Multilayer Switching Caching

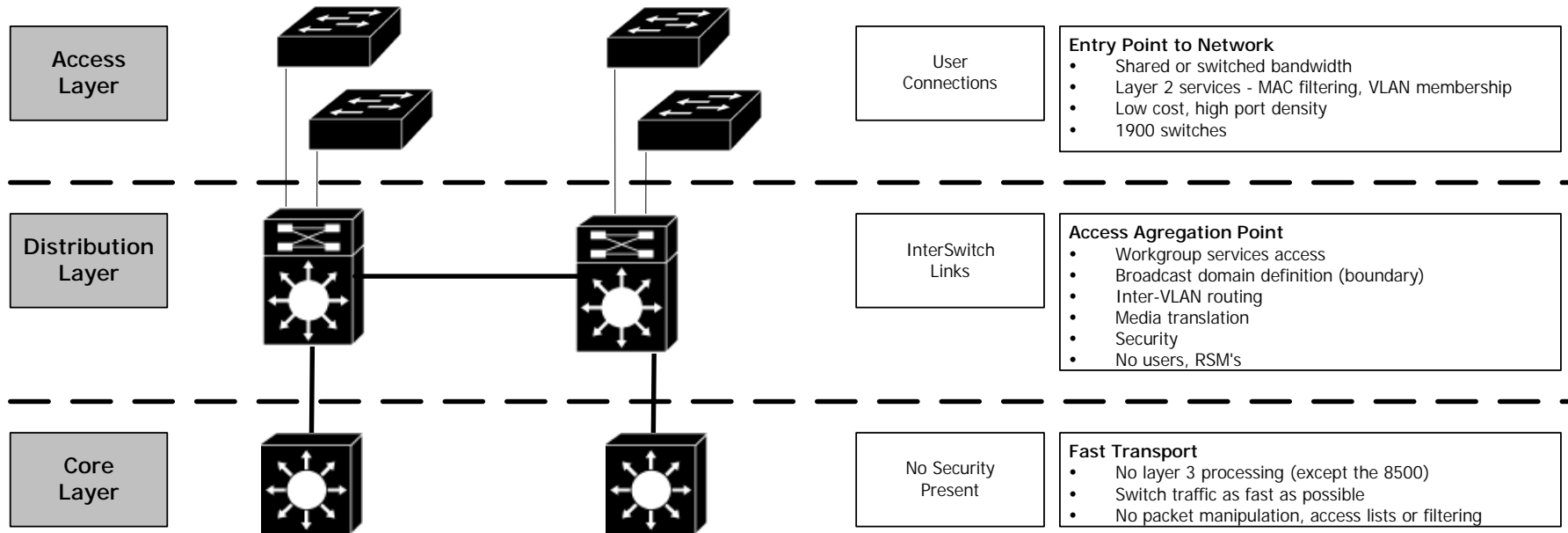
When operating as a layer 3 switch, the switch caches flows based upon IP Address.

When operating as a layer 4 switch, the switch caches flows based upon source and destination addresses and ports.

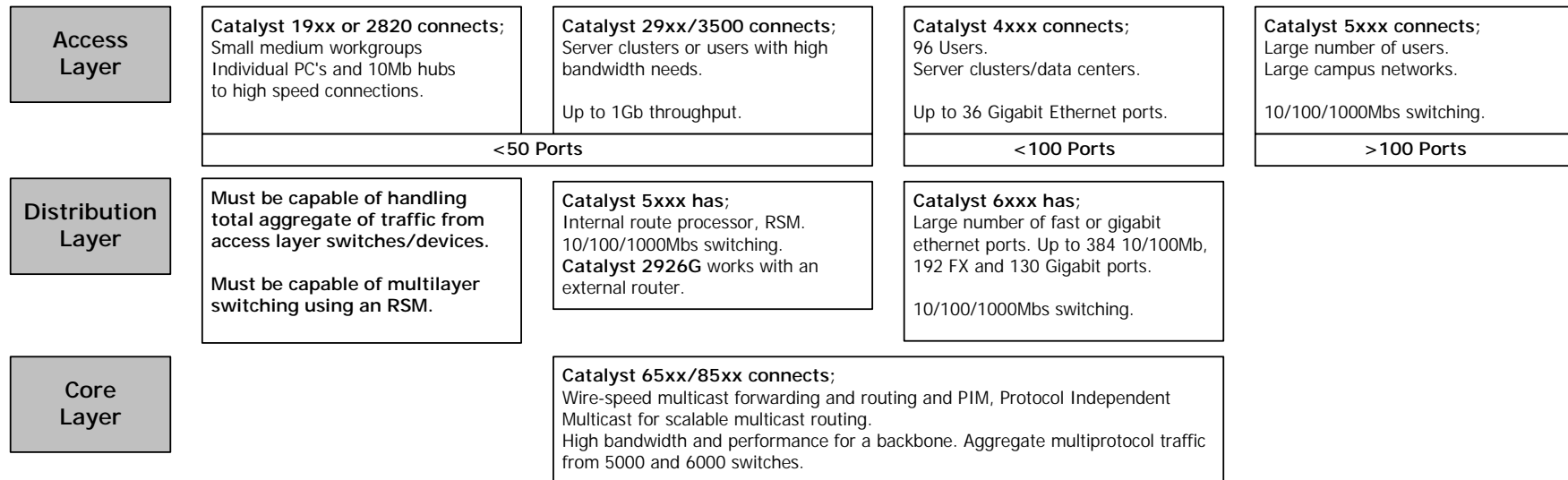
There is no performance difference between these modes as all switching is performed in hardware.

SECTION THREE: The Hierarchical Model

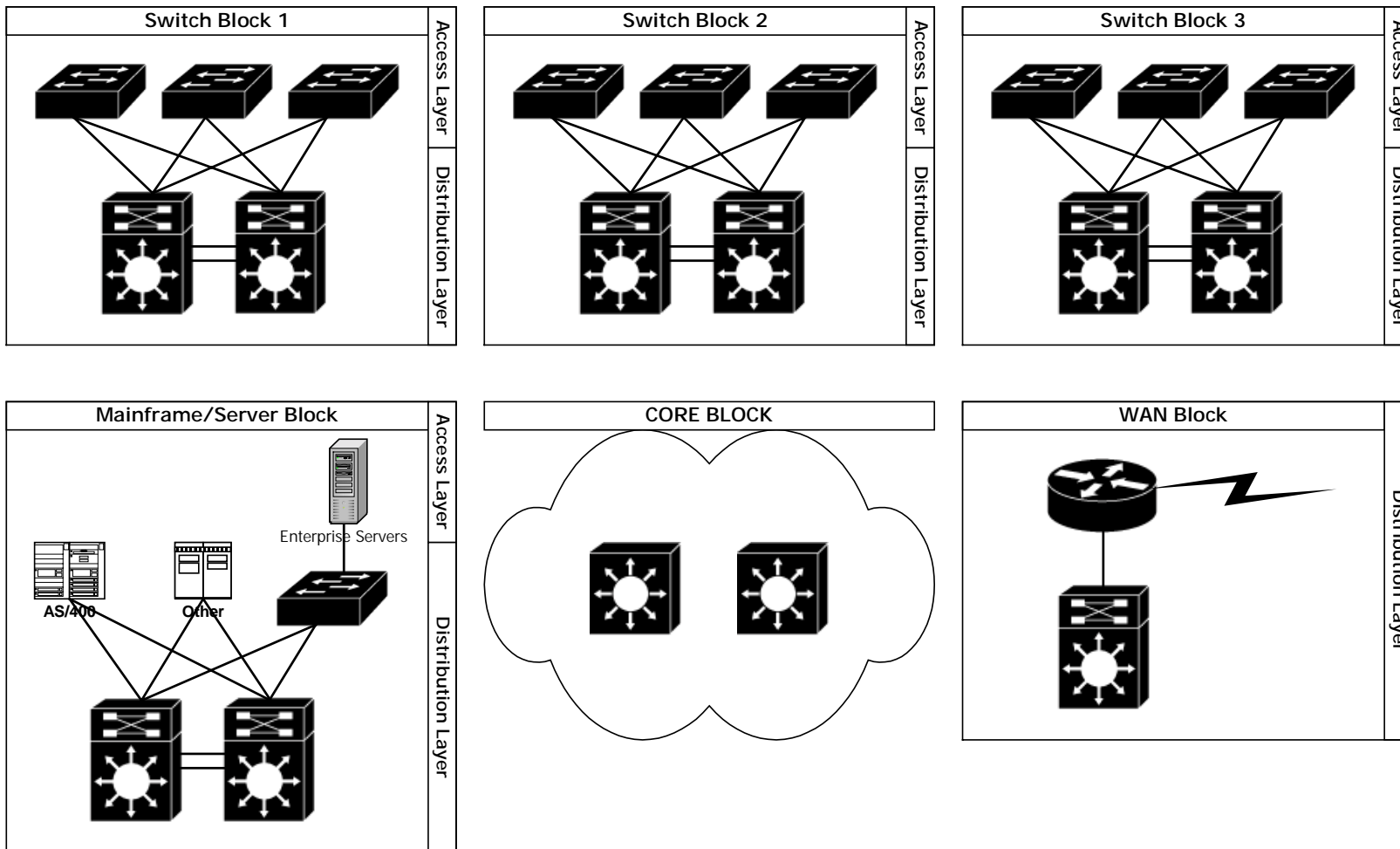
The Three Layer Model



Suitable Cisco Products



Campus Network Building Blocks



General campus network design is Layer 2 devices, Layer 3 devices and then Layer 2 devices again, (can be Layer 3.)

Switch Blocks;

Broadcast storms are contained within the block
Layer 2 and 3 switching (i.e. router functionality)
Support one or more subnets/VLANS

Access devices provide dedicated bandwidth to each port
All subnets have own VLAN or all subnets are in one VLAN
Access devices have redundant connections to distribution devices (using STP)
Catalyst 2900, 2820 and 1900 series

Distribution devices are central connection point
Provide layer 3 functions
Shields the block against failures in the rest of the network

Max 2000 users per switch block (in general, could be lower depending on traffic etc.)

Core Block;

Required if there are more than one switch block
No processor intensive work, such as routing, security etc.
Must be very fast

Individual subnets connect all distribution and core devices
At least 2 subnets
No trunk links
Switches, minimum of 2 recommended
Catalyst 55xx or 65xx for layer 2 or 85xx for layer 3

Links between distribution and core devices should at a minimum be able to support the load handled by the distribution device.
Links between core switches in the same core subnet should be able to handle the aggregate traffic of the distribution devices.

Collapsed Core;

When the same devices handle both distribution and core layers, i.e. the ML switches in two switch blocks.

Each device has a redundant link to a device in the other block. STP and HSRP are used.

Dual Core;

Two layer 2 switches connect 2 or more switch blocks
Redundant links are used between all core and distribution devices
Core devices are not linked to each other

Provides 2 equal cost paths and twice the bandwidth, uses HSRP, STP is not required at the core.

Core Sizing:

A routing protocol is used to maintain the current state of the network
The more routers on the network the longer it takes for updates and changes to propagate through the network

Maximum No. of Supported Blocks by Routing Protocol

Routing Protocol	Max. No. of Routing Peers	No. of Subnet Links to the Core	Max. No. of Supported Blocks
OSPF	50	2	25
EIGRP	50	2	25
RIP	30	2	15

Core layer switches must be able to scale to:

n^* (load per link at 100% capacity) = total core switch capacity

n = the number of distribution links

Layer 2 Core; Backbone Scaling (Layer 2 – Layer 3 – Layer 2)

Spanning tree prohibits core interconnections

As the number of core devices increases you need to increase the number of links from distribution switches to maintain redundancy

Routing protocols dictate the max. number of equal cost paths and limit the number of core switches

Core switches should not be interconnected as this requires the use of Spanning Tree which compromises the high performance of the core switches

Layer 3 Core; Backbone Scaling (Layer 2 – Layer 3 – Layer 3)

Used for:

- Fast convergence
 - Removes the need for Spanning Tree at the core with large numbers of core devices because of a large number of switch blocks. (Redundancy and speed are maintained without using slow ST.)
- Automatic load balancing
 - Best utilises the multiple redundant links without a need to manually configure them
- Eliminating peering problems
 - Removes router peering issues, each router does not need to maintain information about all other peer routers as in a layer 2 core as a hierarchy is created and distribution devices are not considered peers of all other distribution devices
 - Generally very expensive and only used in large campus networks with more than 100 switch blocks.

SECTION FOUR: Configuring a Switch Block

Basic Device & Port Configuration

Set based switch	Command (IOS) based switch
2926, 2926G, 1948G, 4000, 5000 and 6000	1900/2800/2900xl

Setting user passwords: (minimum 4 character)

switch (enable) set password (password is encrypted) `Switch(config)#enable password level 1 password`

Setting enable passwords: (minimum 4 character)

switch (enable) set enablepass (encrypted) `Switch(config)#enable password level 15 password`

switch (enable) clear enablepass `Switch(config)#no enable password level no.`

Setting a system prompt/hostname:

switch (enable) set prompt *name* `Switch(config)#hostname name`

Setting an IP address:

switch (enable) set interface sc0 *address netmask ...* `Switch(config)#ip address address netmask`
 [SC0 is an in-band logical interface.]

- By default the assigned ip address is placed into VLAN1, normally the management VLAN.
- In a large switched network the management VLAN, and therefore all the switches, should be assigned an address in the same ip network, i.e. 16.10.1.xxx.

switch (enable) ?? `Switch(config)#no ip address`

Viewing IP information:

switch (enable) show interface `Switch#show ip`

Giving an interface a description:

switch (enable) set port name mod/no. *description* `Switch(config-if)description "description"`
 [Maximum 20 characters] [Quotation marks are required if spaces are used]

switch (enable) set port name mod/no. *blank* `Switch(config-if)#description blank`

Defining Port Speed:

switch (enable) set port speed mod/no. 10|100|auto `Port speed is fixed in IOS based switches`
 [Using auto is not recommended]

Displaying Port Speed:

switch (enable) show port mod/no. `Switch#show int eth x/x`

Defining Port Line Mode (Duplex):

Using full duplex mode eliminates collisions and is recommended for server to server, server to switch and switch to switch connections.

switch (enable) set port duplex *mod/no. full/half*
 [full is the default for 100Tx ports]
 [half is the default for 10BT ports]

```
Switch(config-if)duplex auto|full|full-flow-control|half
[auto is the default for 100Tx ports]
[half is the default for 10BT ports]
```

If a ports speed is set to **auto** you will not be able to change the ports line/duplex setting.

Displaying Port Line Mode (Duplex):

switch (enable) show port mod/no.

```
Switch#show int eth x/x
```

Enabling a port:

switch (enable) set port enable *mod/no.*

```
Switch(config-if)#no shutdown
```

Dis-abling a port:

switch (enable) ??

```
Switch(config-if)#shutdown
```

*the new command; **switch (enable) set port host** automatically enables PortFast and sets trunking mode and EtherChannel mode to off.

VLAN Challenges, Characteristics and Problems

Layer2 Network Challenges:

1. Flat network structure
Every device sees every packet that is transmitted. Each port is in its own collision domain and therefore the distance limitation rules of Ethernet no longer apply. This leads to Layer 2 networks that are larger than traditional Ethernet networks. As the network grows the number of packets each station must process increases, (broadcasts etc,) every host has to process every broadcast as if it was intended for itself.
2. Security
There is no simple way to provide for security. Users may access all devices on the network.
3. Managing multiple paths
Layer 2 switches do not support redundant paths to a destination and are not capable of intelligent load balancing of traffic.
4. Problem Containment
Layer 2 networks are unable to contain problems caused by faulty devices, loops or excessive broadcasts, potentially allowing the network to fail.

VLAN Characteristics:

All hosts in a VLAN are members of the same broadcast domain; members of other VLANs do not receive broadcasts.

A VLAN is a *logical* subnet or segment, a physical subnet is made up of devices sharing a physical cable segment or hub.

Members of a logical subnet can exist anywhere in the switch block.

A router is required to communicate between logical subnets/VLANs as would be required to communicate between physical subnets.

VLAN membership is based on switch port number, (static,) but may be based on MAC address, (dynamic.)

Local VLANs are geographical, defined in a specific geographical area, usually the wiring closet.

End to end VLANs are defined throughout the entire switch fabric and can span several wiring closets or buildings.

VLAN Solutions for Layer 2 Network Problems:

1. Efficient bandwidth utilization

VLANs solve scalability problems found in large, flat networks by dividing the network into smaller broadcast domains or subnets. For information to pass to another VLAN it must be routed by a layer 3 process.

2. Security

By forcing inter-VLAN traffic to be routed by a layer 3 process all usual layer 3 security can be applied, i.e. access lists.

3. Load-balancing multiple paths:

Layer 3 routing protocols can intelligently determine the best path to a destination and can load balance when there are multiple paths to a destination.

4. Isolation of problem components:

In a flat layer 2 network, a faulty device, a loop or excessive broadcasts could impact the entire network and cause it to fail. Layer 3 routing prevents problems from spreading outside of a VLAN and limits its effects to the source VLAN.

Scaling the Switch Blocks with VLANs:

The number of VLANs in a switch block depends on

- Traffic patterns, (80/20 or 20/80.)
- The types of applications in use.
- Network management needs.
- Group common interests.
- The IP addressing scheme used.

Cisco recommends that IP subnets should correspond to VLANs, i.e. hosts in a VLAN should all be using the same IP subnet/network, hosts in the same IP subnet/network should all be in the same VLAN.

Cisco recommends that VLANs should not extend outside of the layer 2 domain of the distribution switch, i.e. should not cross the core layer.

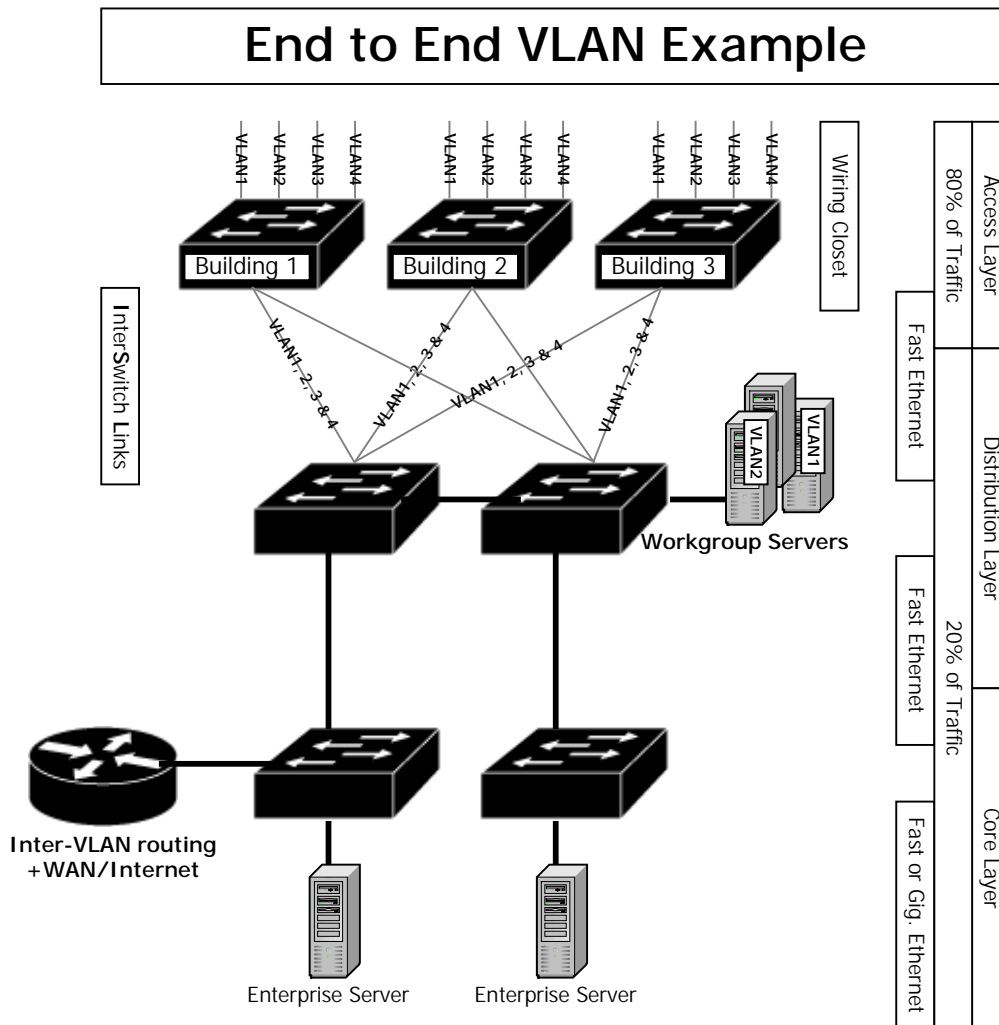
Example: 2000 users reside in the same building. You are using a class C IP subnets.

Only 254 users are allowed in each class C subnet, $2000/254 = 7.87$ therefore you will require a minimum of 8 VLANs in the switch block.

VLAN Boundaries:

End to End VLANs:

Users are grouped independent of physical location and according to common usage of server or by project team/department.
 80% of traffic is local, 20% is remote.
 As a user moves around the campus VLAN membership remains the same, (eventually all switches become members of all VLANs.)
 VLANs have a common set of security requirements.
 Users are generally in the same VLAN as their workgroup server, if possible.



Local VLANs

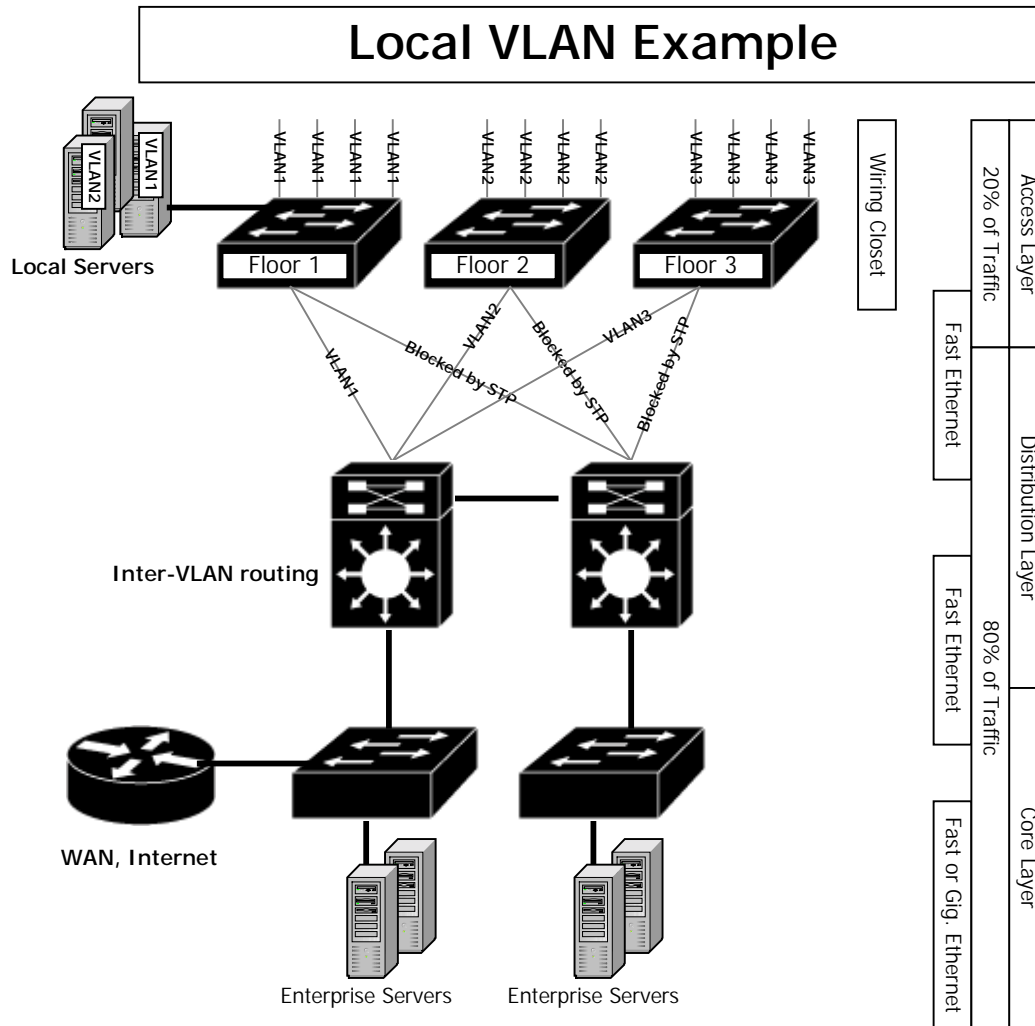
Improvement over end to end VLANs, accommodating new traffic patterns and the new 20/80 rule.

20% of traffic is local, 80% is remote.

Created according to geographical boundaries. Building/single switch.

Deterministic, consistent method of getting to resources.

Easier to manage and conceptualise than VLANs which span different geographical areas.



VLAN Membership:

Dynamic

- VLAN membership is assigned dynamically, based on the MAC address of the connected device. As a device enters the network the device queries a database for VLAN membership.
- Requires additional Cisco software.
- Has the overhead of requiring end-station MAC address lists and custom filtering tables.

Static

- VLAN membership is manually assigned to a port, by an administrator, regardless of what is attached.
- The attaching device is unaware that the VLAN exists.
- Good when moves are controlled and managed.
- Secure, easy to configure and straightforward to monitor.

The switch is responsible for identifying which VLAN information comes from and ensuring all other members of the VLAN receive the information but members of other VLANs do not.

Configuring VLANs

Set based switch	Command (IOS) based switch
2926, 2926G, 1948G, 4000, 5000 and 6000	1900/2800/2900xl

Creating a VLAN:

switch (enable) set vlan *no.* mod/*no.*

Switch(config)#

Naming a VLAN:

switch (enable) set vlan *no.* name *name*

Switch(config)#

Deleting a VLAN:

switch (enable) clear vlan *no.*

Switc(config)#

Viewing VLAN information:

switch (enable) show vlan

Switch(config)#

Setting an IP address:

switch (enable) set interface sc0 *address netmask* ...
[SC0 is an in-band logical interface.]

Switch(config)#

Link Types:

Access Links

- Are a member of only one VLAN, known as the *native* VLAN of the port.
- A switch removes any VLAN information from frames being sent to devices on these ports.
- The end device is unaware it is a member of a VLAN.
- The port cannot send or receive information from other VLANs unless routed.

Trunk Links

- Are capable of carrying multiple VLAN traffic.
- Typically used to connect switches to other switches or switches to routers.
- Cisco only supports trunk links on fast and gigabit Ethernet ports.
- A trunk link does not belong to a specific VLAN; it just transports VLAN data between devices.
- A trunk link can carry data for all or only some VLANs.
- A trunk link may have a native VLAN, which is used if the trunk link fails for any reason.

VLAN Frame Identification for Trunk Links

The techniques used to identify which VLAN a frame belongs to are:

ID Method	Encapsulation?	Tagging	Media	Proprietary
Cisco ISL	YES	NO	Ethernet	YES
802.1q	NO	YES	Ethernet	NO
LANE	NO	YES	ATM	NO
802.10	NO	YES	FDDI	YES

*Trunking capabilities are hardware-dependent. The **show port capabilities** command can be used to find out which trunking techniques are supported.

Frame Tagging (IEEE 802.1q, LANE, 802.10):

Frame identification a.k.a Frame Tagging, assigns a user-defined ID to each frame. This is known as a VLAN ID or colour.

The ID is placed in the *header* of each frame as it is forwarded on a trunk link, each switch examines the ID to determine the VLAN of the frame and transmits it to any relevant ports, if any. Any frame identification is removed when the frame is transmitted out of an access link/port.

If a trunk link is configured to transport frames for the VLAN in question it will be forwarded out of the trunk link port.

Supports up to $10^{12}-1$ VLANs

ISL:

External Frame Tagging, can only be used on Trunk links, can only be recognised by ISL aware devices.

Used between switches and routers.

Ethernet frames are encapsulated with a header that contains the VLAN ID.

The header is 26-bytes long, containing 10-bit VLAN ID, a 4-bit FCS tail is used for CRC. The ID is only added if the frame is forwarded out of a trunk link, all ISL information is removed if the frame is forwarded out of an access link.

The frame can exceed the Ethernet maximum transmission unit size of 1518-bytes.

802.1q:

Internal Frame tagging, appears to be a normal Ethernet frame and can be used on both Trunk and Access links.

A 4-byte tag header contains a tag protocol identifier (TPID) and tag control information (TCI) with the following parts:

Initial MAC Address	2-byte TPID 2-byte TCI	Initial Type/Data	New CRC
---------------------	---------------------------	-------------------	---------

TPID Contains:

A fixed value of 0x8100, indicates that the frame carries the 802.1q/802.1p tag information.

TCI Contains:

A 3-bit user priority
1-bit canonical format (CFI indicator)
12-bit VLAN Identifier (VID) which uniquely identifies the VLAN to which the frame belongs

The frame can exceed the Ethernet maximum transmission unit size of 1518-bytes.

Trunk Negotiation

A trunk is a point to point link between two Catalyst switch ports, or between a Catalyst switch and a router. Trunks allow you to extend VLANs from one Catalyst switch to another.

The Dynamic Trunking Protocol, DTP, manages trunk negotiation in Catalyst supervisor engine software release 4.2 and later.

DTP supports auto-negotiation of both ISL and 802.1q trunks.

In prior releases trunk negotiation was managed by the Dynamic Inter-Switch Link, DISL, protocol. DSL supports auto-negotiation of ISL trunks only. With supervisor engine software v4.1 you must manually configure 802.1q trunks on both ends of the links.

Prior to v4.1 802.1q trunk links are not supported.

During trunk negotiation ports will not participate in Spanning-Tree Protocol.

Configuring Trunk Links

Set based switch	Command (IOS) based switch
2926, 2926G, 1948G, 4000, 5000 and 6000	1900/2800/2900xl

Configuring a Port as a Trunk Link:

switch (enable) set trunk *mod/port* on type *type*
Type could be: isl|dot1q|dot10|lane

Switch(config-if)#trunk on

By default all VLANs are transported across a trunk link, 1-1000

*The neighbouring port is the port on the other side of the link, (on another switch/router.)

On

Places the port into permanent trunking mode.
Becomes a trunk link even if the neighbouring port does not agree to the change.
Does not allow the negotiation of an encapsulation type, which must be specified.
Always use this mode if you want a port to be a trunk link.

Off

Places the port in non-trunking mode.
Negotiates to convert the link into a non-trunk link.
Becomes a non-trunk port even if the neighbouring port does not agree to the change.

Desirable

The port actively attempts to convert the link to a trunk link.
The port becomes a trunk port if the neighbouring port is set to on, desirable or auto.

Auto

Makes the port 'willing' to convert the link to a trunk link.
The port becomes a trunk port if the neighbouring port is set to on or desirable.
This is the default mode for Fast and Gigabit Ethernet ports.
If both sides of the trunk link are left in this state the link will never become a trunk as neither side will be the first to ask to convert to a trunk.

Nonegotiate

Places the port into permanent trunking mode but prevents the port from generating DTP frames.
The neighbouring port must be manually configured as a trunk port to establish a trunk link.

Clearing VLANs from Trunk Links

By default all VLANs are transported across a trunk link. In certain circumstances a trunk link should not carry all VLANs, for example:

Broadcast suppression

All broadcasts are sent to every port in a VLAN. A trunk link is a member of the VLAN in question and must therefore pass all of the broadcasts. Resources,

(bandwidth/processing,) are wasted if there the VLAN is not used at the other end of the trunk link.

Topology Change

Changes that occur in the topology must also be propagated across the trunk link. If the VLAN is not used on the other end of the trunk link there is no need for the overhead of the topology change.

Set based switch	Command (IOS) based switch
2926, 2926G, 1948G, 4000, 5000 and 6000	1900/2800/2900xl

Clearing VLANs from Trunk Links:

switch (enable) clear trunk *mod/port vlan-range*

Switch(config-if)#?

If removing many VLANs from the link it may be easier to clear all VLANs from the link and then add the relevant VLANs. Remember: All VLANs are added by default, 1-1000.

Viewing Trunk Link Configuration:

switch (enable) show trunk *mod/port*

Switch#?

VTP – VLAN Trunk Protocol

VTP Benefits

VTP maintains VLAN configuration consistency throughout a network, managing the addition, deletion and renaming of VLANs network wide, using layer 2 trunk frames.

VTP allows you to make centralised changes that are communicated to all switches in the network.

VTP prevents inconsistencies, such as:

- Security violations caused by cross-connected VLANs with duplicate names
- Internal disconnections caused by VLANs mapped between LAN media types

VTP has the following benefits:

- Configuration consistency
- Mapping scheme for VLANs crossing mixed media types, allows VLANs to be trunked over mixed media.
- Accurate tracking and monitoring
- Dynamic reporting of added VLANs across a network

VTP Domains

A VTP domain is made up of one or more connected devices that share the same VTP domain name. Trunk ports are used to propagate VLAN information.

Switches can only participate in one VTP management domain.
Switches in different domains do not share VTP information.
Routers do not propagate VTP information.

Catalyst family switches advertise the following on trunk ports:

- Management Domain
- Configuration Revision Number
- Known VLANs and specific parameters

Transparent mode switches do not accept VTP information but will forward VTP information on trunk ports.

By default, management domains are set to non secure mode. Adding a password, **on every switch in the domain**, sets the domain to secure mode. (The password must be the same on all switches.)

When a switch receives an update with a higher configuration revision than that stored in its configuration database it updates the database accordingly and is prepared to receive traffic on trunk ports with the newly identified VLAN Ids, emulated LAN names and 802.10 SAIDs.

(VTP has its own NVRAM, a clear config all command does not clear the entire configuration, only a hard reboot will do so.)

VTP Modes

Catalyst family switches can operate in any of the following modes:

Server

- Default mode.
- You can create, modify and delete VLANs and specify VTP configuration parameters for the entire domain.
- Advertises VLAN configuration to other switches in the same domain.
- Synchronises VLAN configuration with other switches based on advertisements received over trunk links.
- Preserves global VLAN information on reboot.

Client

- Same as server but you are unable to modify any VLAN information.
- Does not preserve global VLAN information on reboot.

Transparent

- Does not participate in VTP.
- You are able to create, modify and delete local VLANs only.
- VTP v2 switches do forward VTP advertisements they receive.

Adding a Switch to an Existing VTP Domain:

- Enter the command `clear config all` to remove the existing configuration. This will not clear the VTP configuration revision number.
- Cold-boot the switch to clear the VTP NVRAM. This will reset the VTP configuration revision number to 0, ensuring that the new switch will not propagate any incorrect information.
- Determine the mode of the switch and include the mode when setting VTP domain information on the switch.
 - Server mode is the most common.
 - If using server mode verify that the VTP configuration revision is 0 before adding to the domain.
 - Cisco recommends several servers with all other switches in client mode to control VTP information.
 - Using secure mode by assigning a password to the domain is recommended.

VTP Advertisements

Every switch advertises its management domain, configuration revision number and the VLANs it knows about, including configuration parameters on its trunk ports.

These frames are sent to a multicast address so all neighbouring routers receive the information.

All advertisements start as configuration revision 0, when changes are made the revision number increments by 1. Advertisements are sent downstream only, never to the root bridge.

There are two types of advertisements:

- Requests from clients wanting to learn at bootup
- Response from servers

There are three types of messages:

- Advertisements requests from clients, server responds with summary and subset advertisements.
- Summary advertisements, every 300seconds on VLAN1 and every time a topology change occurs.
- Subset advertisements, detailed information about VLANs

Advertisements may contain:

- Management domain name – ads with a different domain name are ignored
- Configuration revision number – a higher number indicates the latest configuration
- MD5 Digest – a key that is sent when a password has been assigned, if the key doesn't match the update is ignored
- Updater identity – the identity of the switch sending the summary advertisement

When a client or server switch receives an advertisement it must be prepared to receive traffic from a new VLAN. A server can also instruct a client to delete a VLAN and disable all ports assigned to it.

Subset advertisements contain:

- VLAN Ids (ISL and 802.1Q)
- Emulated LAN names (for ATM LANE)

- 802.10 SAID values (FDDI)
- VTP domain name
- VTP configuration revision number
- VLAN configuration, including MTU size for each VLAN
- Frame format

VTP Configuration Tasks

Determine the version of VTP that you will use

VTP v1 is the default

VTP versions are now interoperable

VTP v2 should only be used if you need it's specific enhancements, such as support for Token Ring.

Switch (enable) **set vtp v2 enable**

If all switches in a domain support version 2, it only needs to be enable on one switch, the version change will propagate to all other switches.

Decide if the switch is to join an existing domain or if a new domain is to be created

To create a management domain, or add a switch to one;

Switch (enable) **set vtp domain *domain_name* password *password***

Domain names can be up to 32 characters long

Passwords must be between 8 and 64 characters long

Choose a VTP mode for the switch

Choose server mode if this is the first switch in a domain

Choose client mode if there are many other switches in the domain

Choose client mode if there are other switches in the domain and you would like this switch to be a server, switch to server mode once the latest VLAN configuration has been learnt

Choose transparent mode if this switch is not going to share VLAN information with other switches in the network. No VLAN information will be propagated. This can be a problem if many administrators are creating VLANs in your network as overlaps may occur.

Switch (enable) **set vtp domain *domainname* mode *server/client/transparent***

Viewing VTP Domains

Switch (enable) **show vtp domain**

Verifying VTP Traffic/Operation

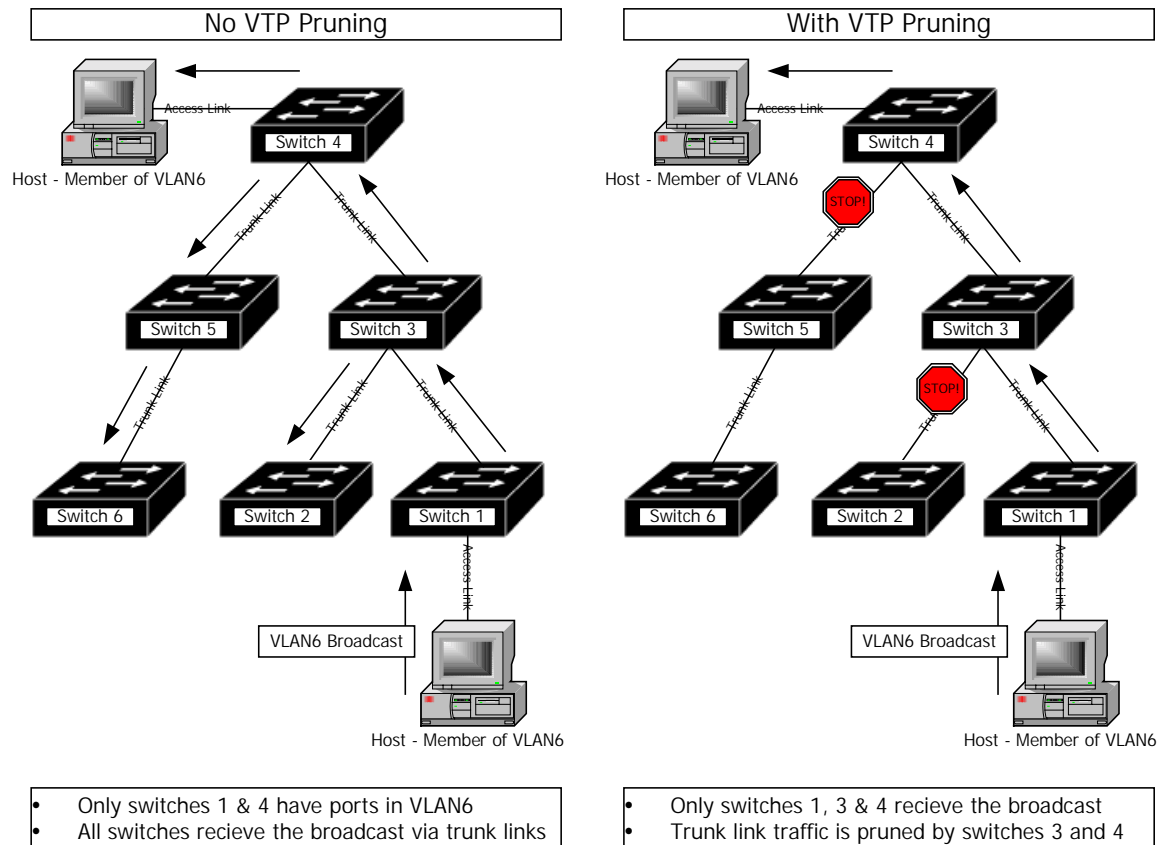
Switch (enable) **show vtp statistics**

Switch (enable) **clear vtp statistics**

VTP Pruning

Disabled by default

Reduces unnecessary bandwidth usage, i.e. broadcast, multicast packets, by restricting that traffic to only the trunk links that the traffic needs to use to reach network devices.



Enabling pruning on a server enables pruning for the entire management domain.
 VLAN1 is always prune ineligible
 VLANs 2-1000 are prune eligible by default (VTP will prune traffic for these VLANs)

To make a specific VLAN prune ineligible:

Switch (enable) **clear vtp pruneeligible *vlan_range***

To make a specific VLAN prune eligible:

Switch (enable) **set vtp pruneeligible *vlan_range***

To verify VTP pruning:

Switch (enable) **show trunk *mod/port***

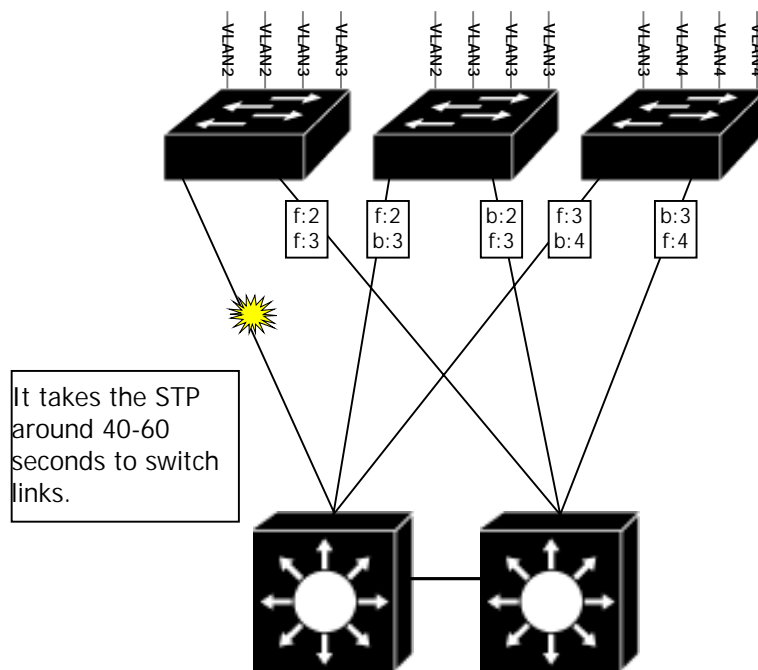
SECTION FIVE: Spanning Tree & Redundant Links

Spanning Tree Protocol

Prevents/Eliminates bridging loops – ensures that there is only one path to every destination.

Allows the use of redundant links in layer 2 devices.
STP disables redundant paths, (blocks data only.)
STP is enabled by default on Cisco switches.

In the case of link failure STP would bring up the link that was previously disabled:



Bridge Protocol Data Unit (BPDU)

All switches in a LAN participating in Spanning Tree Protocol gather information on other switches in the network through an exchange of data messages. These messages are called Bridge Protocol Data Units, BPDU.

The exchange of BPDUs results in:

- The election of a root switch for the stable spanning-tree network topology
- The election of a designated switch for every switched segment
- The removal of loops in the switched network by placing redundant switch ports in a backup state

BPDUs are transmitted every 2 seconds on every port to ensure a stable, loop free topology. The speed of links is the metric used.

The root bridge is generally the switch with the fastest links.???

Root Bridge Selection (Election)

At start-up a switch assumes that it is the root bridge and sets its root ID to the bridge ID.

The bridge ID is made up of a 2-byte priority, 0x8000 by default on Cisco switches, and a 6-byte MAC address. The priority field is configurable.

The lower the bridge ID the more likely a switch will become the root.

If all devices have the same priority, the bridge with the lowest MAC address becomes the root bridge.

Designated Bridge Selection

For Ethernet segments with more than 1 switch a designated bridge is elected. This switch is normally the one with the fewest number of hops to the root bridge.

Root Port

Every switch defines one port as the root port, this is the port with the fewest number of hops to the root bridge.

Root Association

Once a root bridge is elected, every switch must form an association with the root bridge.

Switches do this by listening for BPDUs on all ports, if a switch receives BPDU's on multiple ports then the switch has a redundant path to the root bridge.

One path will have to be disabled. To decide which port will forward and which will block data 2 BPDU fields are look at:

Path cost
Port ID

Path cost is analysed first. The port with the lowest path cost wins. Path cost is based on the link speed and the number of links the BPDU traversed.

If path cost is equal for both ports the port ID is used, the port with the lowest ID wins.

STP Port States

Blocked

- All ports start in this state to prevent loops
- A port stays in this state if spanning tree determines a lower cost path to the root

Listen

- Ports use this time to attempt to learn if there are any other paths to the root bridge
- Port can listen to frames but not send or receive data
- Port can NOT add information to it's address table
- Used to ensure the port does not create a loop
- Ports listen for a time called the fwd delay

Learn

- Port can listen to frames but not send or receive data
- Similar to the listen state except the port can add information it has learnt to its address table
- Port learns for a time called the fwd delay

Forward

- Port can send and receive data
- A port is only put in this state if there are no redundant links or the port determines it has the best path to the root

Default BPDU Timers

Default timers are used to ensure network topology convergence after a change and take into account bandwidth, delays and latency.

Based on a diameter of 7 switches, (can be modified,) by default.

Hello Time = 2 seconds
 Maximum Time or Max Age = 20 seconds
 Forward (Fwd) Delay = 15 seconds

Default STP Timers

Default timers can be modified, but this should only be done at times of great instability and timers should be increased only.

From Blocking to Listening = 20 seconds (Max Age)
 From Listening to Learning = 15 seconds (Fwd Delay)
 From Learning to Forwarding = 15 seconds (Fwd Delay)

STP Commands

Set based switch	Command (IOS) based switch
2926, 2926G, 1948G, 4000, 5000 and 6000	1900/2800/2900xl

Disabling Spanning Tree:

switch (enable) set spantree disable
switch (enable) set spantree disable *mod/port*

```
Switch(config)#no spantree vlan-list
Switch(config)#
```

Enabling Spanning Tree:

switch (enable) set spantree enable (all)
switch (enable) set spantree enable *mod/port*

```
Switch(config-if)#spantree vlan-list
```

Verifying Spanning Tree:

switch (enable) show spantree *vlan*

```
Switch#show spantree
```

Spanning Tree & VLANs

Per-VLAN Spanning Tree (PVST)

- Solution to the scaling and stability of large spanning tree networks, implements a Spanning Tree for every VLAN.
- One instance of spanning tree maintains a loop-free topology in each VLAN, separate timers etc.
- For 1900 and 2820 switches a maximum of 64 VLANs is supported, other must do without STP. STP is enabled by default on VLANs 1-64.
- Trunk ports block only required VLAN traffic, not whole port.

Having separate spanning trees has the following benefits:

- Spanning Tree topology is reduced (per VLAN)
- Scalability is improved and convergence times reduced
- Faster recovery and better reliability

Having separate spanning trees has the following disadvantages:

- Greater utilization of switch resources/processing for every VLAN
- Greater utilization of bandwidth to support BPDUs for every VLAN

Configure separate root bridges for each VLAN for optimum redundancy/performance.

Common/Mono Spanning Tree (CST)

- STP runs on VLAN1, has a single root bridge
- IEEE 802.1Q standard

Having a common spanning tree has the following benefits:

- Less processing overhead on the switch
- Fewer BPDUs consuming bandwidth

Having a common spanning tree has the following disadvantages:

- Single root bridge which may mean less than optimal paths for some devices
- Large spanning tree topology resulting in increased convergence times and frequent reconfiguration.

Per-VLAN Spanning Tree PLUS (PVST+)

- Interoperates with CST (802.1Q)
- Tunnels PVST BPDUs across the 802.1Q VLAN as multicast data
- Backward compatible with PVST through ISL trunking
- Blocks ports that receive inconsistent BPDUs
- No user intervention is required

Scaling STP

Root Bridge Placement:

- Proper placement of the root bridge is essential to ensure optimum spanning tree paths.
- This also provides deterministic paths for data.
- You can specify a root and a secondary root for every VLAN, in the case of root failure.
- The root is generally a distribution switch near the centre of the network.

Set based switch
2926, 2926G, 1948G, 4000, 5000 and 6000

Setting a Switch as the Root:

switch (enable) set spantree root vlans (dia *no*) (hello *hello-time*)

Dia = network diameter, i.e 3 switches from centre to edge of network (this adjusts the default BPDU timers)

Setting a Switch as the Secondary Root:

switch (enable) set spantree root secondary vlans (dia *no*) (hello *hello-time*)

Dia = network diameter, i.e 3 switches from centre to edge of network (this adjusts the default BPDU timers)

Verifying Spanning Tree:

switch (enable) show spantree *vlan* (or *mod/port*) (active)

active = ???

Port & Path Costs and Priority

- Port cost influences how Spanning Tree chooses between two paths to the root bridge.
- The lower the port cost the more likely a port will be used to forward (depending on the total path cost.)
- Path costs are the sum of the cost of all ports passed between this switch and the root.
- Port priority defaults to 32, the lowest value port is used, if all ports have the same value the lowest port number is used.

Setting Port Cost:

switch (enable) set spantree portcost 4/8 *cost*
Range is 1-65535, cost is normally 1000/LAN speed

Switch(config-if)#spantree cost *cost*

Verifying Port Cost:

switch (enable) show spantree *mod/port*

Switch(config)#

Setting Port Priority:

switch (enable) set spantree portpri *mod/port priority*
(*vlans*)

Switch(config)#

Priority range is 0-63

PER VLAN:

switch (enable) set spantree portvlanpri *mod/port*
priority (vlans)

Priority range is 0-63

This is good for splitting VLANs over 2 ports

```
Switch(config-if)#spantree priority value
```

Priority range is 0-255

Verifying Port Priority:

switch (enable) show spantree *mod/port*

```
Switch(config)#
```

Fast EtherChannel

- Allows for the load balancing and redundancy of trunk links between switches.
- Fast or Gigabit Ethernet only. 200Mb to 800Mb or 1 to 4 full duplex connections on Fast Ethernet.
- If a link is lost traffic is rerouted to the remaining links transparently.
- Both ends of an EtherChannel should be identically configured, speed, duplex, VLANs etc.

Redundancy

- Data is distributed by 'flow', i.e. data from host A to host C uses channel 1, data from host B to host D uses channel 2 and so on.
- When a link/channel fails the next channel takes over once it has re-learnt the address, (this takes milliseconds.)

PAGP – Port Aggregation Protocol

- Additional feature of EtherChannel, aids in the automatic creation of Fast EtherChannel links.
- PagP packets are sent between all Fast EtherChannel ports and if neighbouring ports can be paired this is done automatically. The links become a single port in Spanning Tree.
- Ports must be in the same VLAN or trunk ports, dynamic VLANs cannot be used.

ENSURE ALL PORTS ARE IDENTICALLY CONFIGURED BEFORE CONTINUING:

Determining if a Line Card is Capable of EtherChannel:

switch (enable) show port capabilities *mod/port*
port is not required

Switch(config)#

Creating an EtherChannel:

switch (enable) set port channel *mod/ports*
on/off/auto/desirable

Switch(config)#port-channel mode
on/off/auto/desirable

Verifying an EtherChannel:

switch (enable) show port channel 1
port is not required

Switch#show interface

PortFast

- Spanning Tree runs on all ports of a switch and may make a port wait up to 50 seconds before data may be sent through it.
- This is unnecessary if a workstation or server is attached.
- PortFast allows a port to enter the forwarding state almost immediately by reducing the listening and learning states.

Enabling PortFast on a Port:

switch (enable) set spantree portfast *mod/port* enable `Switch(config-if)#spantree start-forwarding`

Disabling PortFast on a Port:

switch (enable) set spantree portfast *mod/port* disable `Switch(config-if)#no spantree start-forwarding`

*The 'uplink fast' command should be used on the Catalyst 2900XL switch.

UplinkFast

- Speeds Spanning Tree convergence.
- Allows a blocked port to begin forwarding almost immediately when the switch detects the failure of the forwarding link.
- Designed specifically for the access layer, i.e. end leaf nodes of the Spanning Tree.
- Should only be used when conforming with the Cisco 3 layer model, i.e. the access switch has a single uplink to two distribution layer switches.
- Root bridges should not use UplinkFast.
- Not available on the 8500 series.
- An uplink group is required, with a root port and a set of blocked ports. (This allows load balancing)
- When failure of the root port is detected the switch transfers a blocked port to the forwarding state within 3 to 4 seconds, bypassing the listening and learning states.
- UplinkFast applies to all VLANs and cannot be configured for individual VLANs.

Enabling UplinkFast on a Port:

switch (enable) set spantree uplinkfast enable *on/off* `Switch(config)#uplink-fast`

These commands increase the path cost of all ports making it unlikely that the switch will become the root bridge.

Verifying UplinkFast:

switch (enable) show spantree uplinkfast

`Switch(config)#show uplink-fast`
`Switch(config)#show uplink-fast statistics`

BackboneFast

- Initiated when a root or blocked port on a switch receives inferior BPDUs from its designated bridge.
- An inferior BPDU identifies one switch as both the Root Bridge and designated bridge.
- Must be enabled on all switches in order to work.
- Enables faster convergence in the event of a backbone link failure.

When a switch receives an inferior BPDU it indicates that an indirect link has failed - that the designated bridge has lost its connection to the root bridge.

Under normal STP operation the switch ignores inferior BPDUs for the Max Age time. (i.e. From the Blocking to Listening states.)

The switch then tries to determine if it has an alternative path to the root:

- If the inferior BPDU arrives on a blocked port, the root port and other blocked ports become alternative paths to the root bridge.
- If the inferior BPDU arrives on the root port, all blocked ports become alternative paths to the root bridge.
- If the inferior BPDU arrives on the root port and there are no blocked ports the switch assumes it has lost connectivity to the root bridge, expires the Max Age timer to expire and becomes the root bridge according to Spanning Tree rules.

With BackboneFast:

- The switch immediately uses the alternative (blocked) paths to transmit a BPDU called a Root Link Query PDU.
- If an alternative path to the root bridge is found the Max Age timer is expired and ports that can connect to the root bridge are moved out of the blocking state.
- If no alternative is found the Max Age timer is still expired???

Enabling BackboneFast on a Switch:

```
switch (enable) set spantree backbonefast
```

Verifying BackboneFast on a Switch:

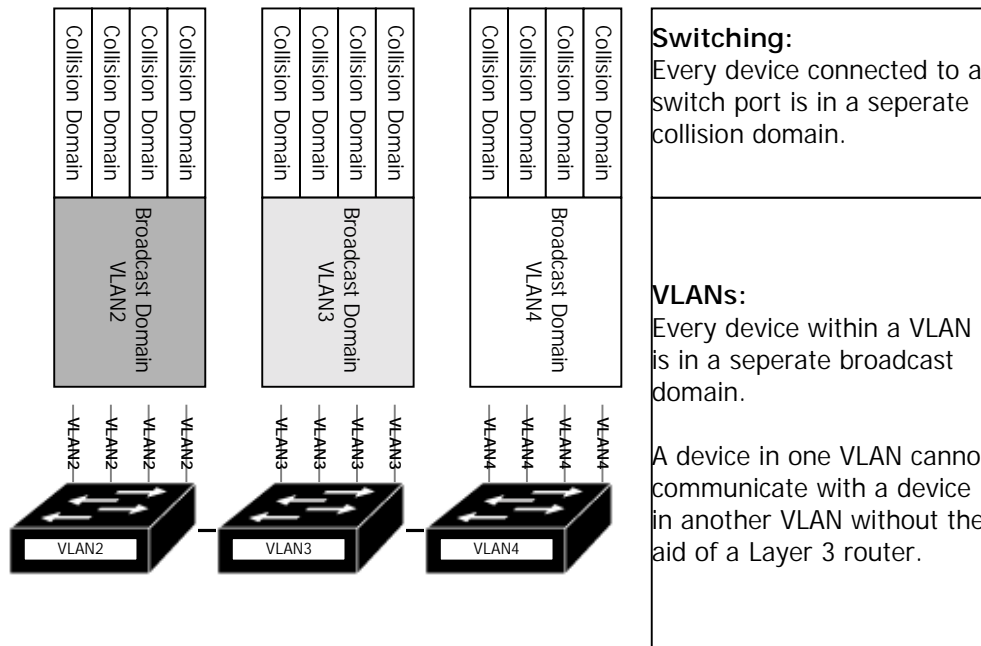
```
switch (enable) show spantree backbonefast
```

SECTION SIX: InterVLAN Routing

InterVLAN Communications

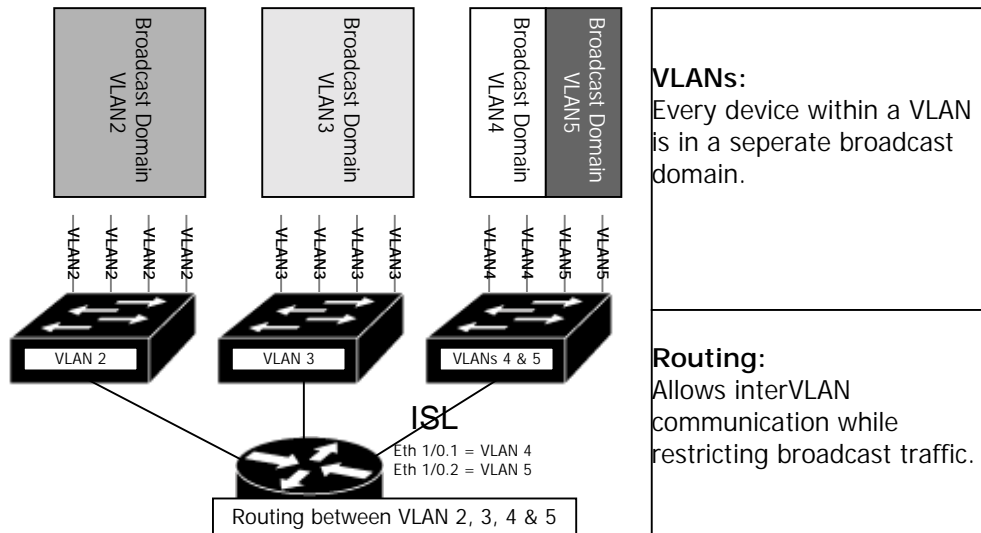
- Switches effectively bridge every port, making every port a separate collision domain, effectively removing collisions and contention of media access.
- VLANs control the size of a broadcast domain and keep traffic local, ensuring broadcasts do not flood and degrade the network. VLANs are logical network subnets.

However, devices in one VLAN do need to communicate with devices in another and this requires a routing device or process:



Traditional Routing

Adding a router between VLANs solves this problem and allows access to shared network resources, remote sites and other remote services.

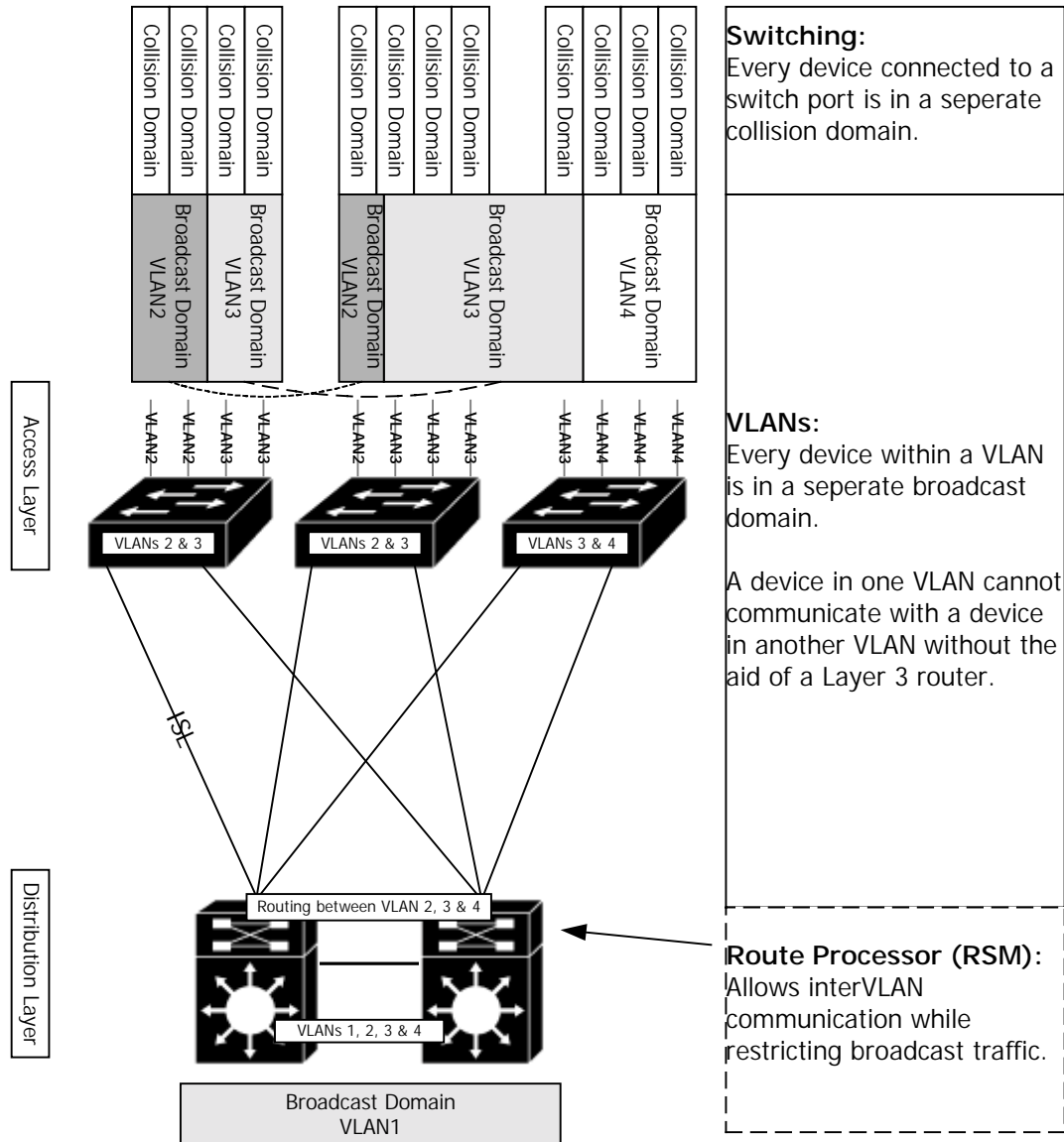


When using a physical router you can route multiple VLAN traffic in two ways.

- Use multiple links from a switch or group of switches each connecting to a separate interface on the route. This is a waste of resources.
- Use ISL to trunk the VLAN traffic to a single interface with multiple sub-interfaces.

Integrated Routing/Layer 3 Switching

A route processor within a switch can be used rather than a physical router:



When configuring routing between VLANs you should consider:

- The sharing of resources between VLANs
- Load balancing
- Redundant links
- Addressing

Because it is the only layer to employ routing the distribution layer is the demarcation between the access layer and core layer networks.

Distribution layer route processors therefore provide:

- Broadcast Control
- Segmentation
- Termination of collision domains (if hubs are used at the access layer)

Physical routers that support MLS in conjunction with a switch are:

- 7500, 7200, 4700 and 4500
- MLSP, Multi-Layer Switch Protocol software must be installed in order for these routers to provide layer 3 services to the switch

Catalyst 5500 switches use a Route Switch Module (RSM) for integrated Layer 3 switching and can route up to 256 VLANs.

Catalyst 5000 switches use a Route Switch Feature Card (RSFC) for integrated Layer 3 switching.

Catalyst 6000/6500 switches use a Multilayer Switch Module for integrated Layer 3 switching.

Configuring VLAN Interfaces on a Route Processor

Locating the Route Processor:

```
switch (enable) show module
```

Accessing the Route Processor:

```
switch (enable) session mod-number
```

mod-number is the module the RSM is installed to.

Naming the Route Processor:

```
router(config)# hostname name
```

Enabling a Routing Protocol:

```
router(config)# ip routing
router(config)#router routing-protocol
router(config-router)#network network-number – must be a network physically connected
```

Configuring a VLAN Interface:

This assumes you have already configured the relevant VLANs on the switch.
 The RSM automatically encapsulates packets using ISL.
 Initially all VLAN interfaces are shutdown, you need to enter the 'no shutdown' command to enable the interface.

```
router(config)# interface vlan-interface-number – same as VLAN number on switch
```

Assign an IP address.

```
router(config-if)#ip address address subnet-mask
```

An RSM has a global MAC address that applies to all interfaces on the device, specifying a unique MAC address per interface enhance the operation of the interface on some Catalyst switches.

Defining a Default Gateway (On the Switch):

The default gateway on the switch enables it to communicate with devices in other subnets. To configure a default gateway, a default route must be added that points to the gateway router in the same subnet as the internal sc0 interface on the switch.

```
Switch (enable) set ip route destination gateway metric  

Destination = default for default route or network number  

Gateway = address of the relevant router  

Metric = optional, 0 = local network, 1 = remote network.
```

Verifying the Default Gateway (On the Switch):

```
switch (enable) show ip route
```

```
Switch(config)#show ip
```

Configuring a VLAN Interface on an External Route Processor

Sub-interfaces should be used:

- Identify the interface to be used (will be a trunk link)
- Define the VLAN encapsulation to be used (generally ISL)
- Assign an IP address

Identify the Interface:

```
router(config)# interface Ethernet slot/port.subinterface-number
```

Define VLAN Encapsulation:

```
router(config-if)#encapsulation isl vlan-number  

Vlan-number is the number of the VLAN this sub-interface will carry traffic for
```

Assign an IP address:

```
router(config-if)#ip address address subnet-mask
```

Defining a Default Gateway:

```
router(config)#ip default-gateway address
```

SECTION SEVEN: Multilayer Switching

Fundamentals

Multilayer Switching, (MLS,) provides high performance hardware-based Layer 3 switching on Catalyst switches. MLS switches IP packet flows between subnets using advanced Application-Specific Integrated Circuit, (ASIC,) hardware, bypassing processor intensive packet routing from routers.

MLS is based upon flows. A flow is a conversation between two hosts within a specific timeframe, for example HTTP packets flowing between a source and destination. This would be a separate flow from that of FTP packets flowing between the same two hosts.

MLS is a technique used to increase IP routing performance by handling packet switching and re-write functions in hardware. It supports traditional routing functions and protocols but forwards packets normally handled by a router via Layer 3 switches where a relevant path exists.

Requirements:

- Catalyst 2926G, 5000 or 6000 switches with supervisor engine software release 4.1(1)
- Cisco IOS Release 11.3(2)WA4(4) or later
- Supervisor Engine III or III F with the NFFC II or Supervisor Engine IIG or IIIG
- Route Switch Feature Card (RSFC)

- Routers: 7500, 7200, 4700, 4500 or 8500

MLS-RP Advertisement:

When an MLS-RP, (Route Processor,) is activated in a campus network, the MLS-RP sends out multicast hello messages every 15 seconds. (The CGMP multicast address is used, with a unique protocol type.)

These messages are sent to all switches in the network and contain:

- The MAC addresses used by the MLS-RP on its interfaces that are participating in MLS.
- Access list information
- Additions and deletions of routes

Switches with Layer 3 capabilities process the hello message; those without simply pass the hello messages on.

When a switch receives the frame, the device extracts all MAC addresses and the associated VLAN ID for that address. The MLS-SE records the addresses of the MLS-RPs in the MLS-SE Content-Addressable Memory, (CAM,) table.

MLS Cache Entries

MLS maintains a cache of flows and stores statistics for each. All packets in a flow are compared to a cache.

Cache entries are one-way, communication between host A and host B is a separate flow to communication between host B and host A.

If the MLS cache contains an entry that matches the packet in a flow, the MLS-SE, (switch engine,) switches the flow, bypassing the router or route processor.

Initial Frame Flow

If the MLS cache does not contain an entry that matches the packet the following occurs:

- The switch receives an incoming frame and looks at the destination MAC address.
- The switch recognises the destination MAC address of the frame as the address of the MLS-RP because the switch initially received this destination address in a layer 3 hello message and stored that address in the CAM table.
- The MLS-SE then checks the MLS cache to determine if a MLS flow is already established for this flow.
 - If the frame is the first in a flow there will not be an entry in the cache. Because the frame contained a RP destination address the switch recognises the potential for Layer 3 switching of the frame and flow.
 - On the initial packet the switch does not have all the information for a Layer 3 switch for the frame. Therefore, the switch forwards the frame to the addressed RP.
 - This process creates a candidate entry in the MLS cache.
 - The RP receives the frame and consults the routing table to determine if it has knowledge of a route to the destination address.
 - If the RP finds the destination address a new Layer 2 header is constructed which contains the RP's MAC address as the source MAC address.
 - The RP also enters the MAC address of the destination host or next hop router as the destination MAC address.
 - The frame is then forwarded back to the MLS-SE.
- When the switch receives the frame it now knows out of which port to forward the frame, based on the CAM table. The switch also recognises the MAC address in the source field as that of the MLS-RP.
- This recognition triggers the checking of the MLS cache to see if there is an entry for this particular route processor.
- The switch compares the XTAGs for both the candidate entry in the MLS cache and the returned frame. (SE to RP, RP to SE.)
- If the XTAGs match the frame came from the same RP for the same flow. The switch records the information from the returned frame in the MLS cache.
- The switch forwards the frame out of the appropriate port using the destination MAC address.
- The second frame becomes the 'enable' entry in the MLS cache and the partial entry for that flow is completed.

Note: The MLS-SE must see both sides of the flow going from source to destination in order to perform Layer 3 switching.

Subsequent Flows

- When the switch receives subsequent packets in the flow the switch recognises that the frames contain the MAC address of the route processor.
- The switch checks the MLS cache and finds the entry matching the flow in question.
- The switch rewrites the Layer 2 frame header, changing the destination MAC address to the MAC address of the 'real' destination and the source MAC address to the MAC address of the MLS-RP.
- Layer 3 information is unchanged with the exception of the TTL field, in the IP header, which is decremented by 1. The checksum is also recomputed. The SE rewrites the switched Layer 3 packets so that they appear to have been routed by a route processor.

The state and identity of the flow are maintained while traffic is active. When traffic for a flow ceases the entry ages out.

- Partial or candidate entries will remain in the cache for five seconds with no enabled entry before timing out.
- Cached entries that are complete, where the switch has captured both the candidate and enabling packet, will remain in the cache as long as packets in that flow are detected.

Commands that Disable MLS

Any command that requires every packet in a flow to be manipulated a route processor precludes the MLS function.

The following commands are some of those that disable MLS on a given interface:

- no ip routing - Purges all MLS cache entries and disables MLS on the MLS-RP
- ip security - Disables MLS on the interface
- ip tcp compression-connections - Disables MLS on the interface
- ip tcp header-compression - Disables MLS on the interface
- clear ip-route – Removes the MLS cache entries in all switches performing Layer 3 switching for this MLS-RP

Enabling MLS on a RP

Enabling MLS on the RP:

`router(config)#mls rp ip` (As of IOS v12 MLS also routes IPX.)

Disabling MLS on the RP:

`router(config)#no mls rp ip`

Assigning a VLAN ID to an Interface on an External Router

Required only on an external router with a non-ISL interface.

RPs and ISL configured links automatically use VLAN Ids to identify interfaces.

External router interfaces are aware of subnets but not VLANs. To this end MLS requires each external router interface to have a VLAN ID assigned to it.

Assigning a VLAN ID:

```
router(config-if)#mls rp vlan-id vlan-id-num
```

Vlan-id-num represents the VLAN assigned to the interface, use the *no* keyword to remove an ID (this disables MLS for the interface.)

Assigning an MLS Interface to a VTP Domain??

An external route processor must be in the same VLAN Trunking Protocol domain as the switch.

If the switch is not in a VTP domain this is not necessary.

Adding the RP to a VTP Domain:

```
router(config)#interface vlan41
```

```
router(config-if)#mls rp vtp-domain domain-name
```

For an ISL interface only configure the primary interface, all sub-interfaces will inherit the domain membership.

Verifying the MLS VTP Domain:

```
router(config)#show mls rp
```

```
router(config)#show mls rp vtp-domain vtp-domain-name
```

Enabling MLS on an Interface

```
router(config-if)#mls rp ip
```

(Use the *no* keyword to disable MLS)

Assigning an MLS Management Interface

When an RSM or router is configured to participate in MLS the device uses MLSP to send hello messages etc. One interface on the MLS-RP must be identified as the management interface through which MLSP packets are sent and received. This can be any MLS interface connected to the switch.

If no management interface is configured no MLSP messages will be sent.

```
router(config-if)#mls rp management-interface      (use the no keyword to disable.)
```

Verifying MLS-RP Configuration

```
router#show mls rp
```

Verifying MLSP-RP Interface Configuration

```
router#show mls rp interface interface-number      (interface-number is the vlan number)
```

MLS Flow Masks

The MLS-SE uses flow masks to determine how packets are compared to MLS entries in the MLS cache. The flow mask mode is based on the access lists configured on the MLS router interfaces.

The MLS-SE supports three flow mask modes:

- Destination-IP
- Source-Destination-IP
- IP-Flow

When the flow mask changes the entire MLS cache is purged. Standard access lists look at the source address only.

Destination-IP Flow Masks

- Default flow mask, least specific
- The MLS-SE maintains one MLS entry for each dest. IP address
- All flows to a given dest. Address use this MLS entry
- The mode is used if there are no access lists configured on any MLS router interfaces

Source-Destination-IP Flow Masks

- The MLS-SE maintains one MLS entry for each source and destination IP address pair
- Used regardless of IP protocol port
- Used if there are standard access lists on any MLS interfaces

IP-Flow Flow Masks

- The MLS-SE maintains a MLS cache entry for every IP flow
- Includes source and destination IP addresses, protocol and protocol ports
- Used if there are extended access lists on any MLS interfaces

Access Lists:

Applying an access list to an interface purges all cached MLS flows

When applying an output access list normal MLS events occur, if the first packet passed the restrictions then the flow is switched.

When applying an input access list all packets are routed, as all incoming packets need to be examined.

For this reason routers with IOS v11.3 or later do not automatically support input access lists on interfaces configured for MLS. (To use access lists enter the following command in global configuration mode: mls rp ip input-acl.)

IP Routing Performance with MLS

MLS is enabled by default on capable switches.

If MLS has been disabled enter the following to enable it

Switch (enable) set mls enable

To disable MLS:

Switch (enable) set mls disable

Aging Out Cache Entries

Cache entries are removed for the following reasons:

- Candidate entries are aged in 5 seconds if no enable entry occurs
- An entry has not been detected for the specified aging time of 256 seconds, (default)
- The whole cache is purged when an access list is applied, a routing change occurs or MLS is disabled.

To Modify the Aging Time:

Switch (enable) set mls agingtime *time* (8 to 2032 seconds in 8 second increments)

Short Lived Flows

Short lived flows such as DNS or DHCP conversations take up valuable space in the cache.

To prevent this use Aging Fast, which removes entries from the cache if the MLS-SE does not detect a specified number of packets in a certain time period.

To Enable Aging Fast:

Switch (enable) set mls agingtime fast *fastagingtime packet-threshold*

Fastagingtime = Amount of time an entry remains in the cache, can be 32, 64, 96 or 128 seconds, default is 0

Packet-threshold = The number of packets that must be detected within the specified amount of time, default is 0

To Verify Aging Fast:

Switch (enable) show mls

Displaying MLS Cache Entries

To display all entries in the cache:

Switch (enable) show mls entry

To display all entries for a specific destination address:

Switch (enable) show mls entry destination *ip-address*

To display all entries for a specific source address:

Switch (enable) show mls entry source *ip-address*

Removing MLS Cache Entries

To remove all entries in the cache:

Switch (enable) clear mls entry

To remove all entries for a specific destination address:

Switch (enable) clear mls entry destination *ip-address*

To display all entries for a specific source address:

Switch (enable) clear mls entry source *ip-address*

APPENDIX

Switching Types

CUT THROUGH

The receiving port checks the MAC address table before forwarding the packet out of the correct port. This has a high impact when you have a large MAC address table.

ADD TRADITIONAL EXPLANATION

STORE AND FORWARD

The receiving port broadcasts the packet to all other ports
The supervisor card makes a forwarding decision and tells all the ports whether to send the packet or not.

Cisco use store and forward.

Ping

Destination Unreachable: The default gateway cannot reach the network.

Network/Host Unreachable: There is no entry in the route table.

XCast Transmissions

MULTICAST

All hosts receive a multicast at the same time, i.e. at a preset and fixed time.
Example: Television

(A compromise between UNI and BROADCAST.)

UNICAST

A single host requests data at any time, i.e. randomly.
Example: Watching a short presentation by clicking a link on a web browser.

Acronyms & Terms

Acronym	Stands For
ASIC	Application-Specific Integrated Circuit (Switch Hardware)
CAM	Content-Addressable Memory (MLS)
CST	Common Spanning Tree
DISL	Dynamic Inter-Switch Link
DTP	Dynamic Trunking Protocol
EARL	Enhanced Address Recognition Logic(EtherChannel)
EBC	Ethernet Bundle Controller (EtherChannel)
FCS	Frame Check Sequence
MLSP	Multi-Layer Switch Protocol (External Routers with Switches MLS)
MLS	Multi-Layer Switch/ing
MLS-RP	Multi-Layer Switch – Route Processor
MLS-SE	Multi-Layer Switch – Switch Engine
MSM	Multilayer Switch Module (Catalyst 6500/6000)
MST	Mono Spanning Tree
PagP	Port Aggregation Protocol
RLQ	Route Link Query (STP/BackboneFast)
RP	Route Processor
RSFC	Route Switch Feature Card (Catalyst 5000)
RSM	Route Switch Module (Catalyst 5500)
SAID	
SE	Switch Engine
STP	Spanning Tree Protocol
TPID	Tag Protocol Identifier
TCI	Tag Control Information
VTP	VLAN Trunk Protocol

Standards

Acronym	Stands For
802.1Q CST	Common Spanning Tree
802.1q	VLAN Trunking/Tagging
802.10	VLAN Trunking/Tagging over FDDI (Cisco Only)
802.3	Fast/Ethernet

Lineage

Hub/Bridge/Switch
Switching Layers, 2 or 3 or 4
VLAN
Switching/Routing
Switch Block/Hierarchical Model
VLAN again
VLAN boundary
Trunk Links and VTP
STP
STP and VLANs
STP Scaling
EtherChannel
PortFast
UplinkFast
BackboneFast
InterVLAN routing
Layer 3 switching
MLS