IBM

@server

xSeries Education

CIGESM & Spanning Tree Protocol

Spanning Tree Overview for CIGESM

What is STP?

- STP = Spanning Tree Protocol

- Flavors of STP
  - IEEE 802.1d – Basic STP.
  - IEEE 802.1w – RSTP.
  - IEEE 802.1s – MSTP.

- Cisco has Per VLAN Spanning Tree called PVST (ie Basic STP), and Rapid PVST (ie RSTP)

Need For Spanning Tree Protocol

- Loops in Switches are bad, worse than in routers.

    - In routers, additional copies are not spawned.

- Full Connectivity and no loops naturally lead to a "tree" topology

- No guarantee of shortest path forwarding between nodes in network

Host2

Switch 2
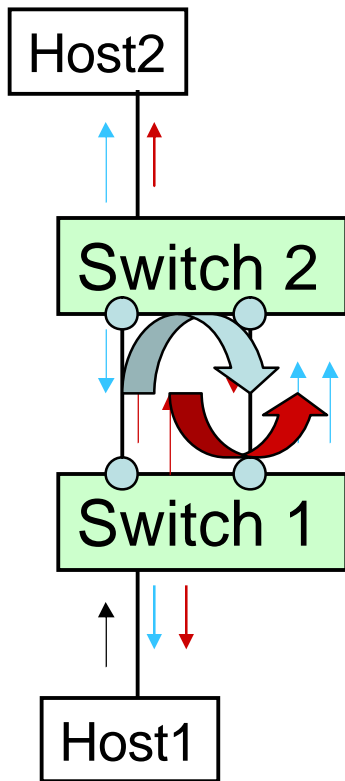
Switch 1

Host1

- Host1 wants to talk to Host2

- Host1 broadcasts ARP request

- Switch1 floods broadcasts

- Broadcasts loop forever

Basic STP terminology definition

- Bridge ID
  - an 8-byte value consists of a 2-byte priority followed by a 6-byte MAC address
  - the 6 byte MAC address guarantees uniqueness
  - the most significant 2 bytes can be changed by user to control the total 8 byte value
- Port ID
  - a 2 byte value consists a 1 byte priority followed by a 1 byte port identifier value
- Root Bridge
  - one for the whole network
  - chosen based on lowest bridge ID
  - good to place in the center of the network to optimize for shortest path between any 2 nodes in network
- Designated Bridge
  - one per LAN
  - provides lowest path cost to root

Basic STP terminology definition cont'd

- Root Port
  - one in each non-root bridge
  - provides best path to the root bridge from this bridge
- Designated Ports
  - 0,or more on a bridge
  - all ports which connect the bridge to LANs on which it is the designated bridge (i.e. traffic from the LAN to the root bridge goes through these ports)
- Alternate Ports
  - 0 or more per bridge
  - currently in non forwarding state, could forward sometime in the future
- BPDU – Bridge Protocol Data Units
  - Exchanged between adjacent network elements

STP Port States

- Blocking - Traffic not forwarded; No address learning
- Listening - Traffic not forwarded; No address learning
- Learning - Traffic not forwarded; address learning enabled
- Forwarding - traffic forwarded; address learning enabled
- State Transition : blocking -> listening -> learning -> forwarding

STP Algorithm in a nutshell

- Elect a single bridge, among all the bridges on all the LANs, to be root bridge

- Calculate the distance of the shortest path from themselves to the root bridge - root port

- For each LAN, elect a designated bridge from among the bridges residing on the LAN - designate ports

- Make root port and designated ports forwarding. They are part of the Spanning Tree.

- Make all other ports blocking

Cisco STP Enhancements

- Standard STP converges very slowly on change; so Cisco came up with fast convergence enhancements:
  - Port Fast
  - Uplink Fast and CrossStack UplinkFast
  - Backbone Fast

- Designed to be used by access switches

- A link between a bridge and a worskstation can not introduce any loop and therefore can go to forwarding without passing through the listening and learning states

- Designed to be used between access and distribution layer

- Used to rapidly recover a root port that fails

- It needs at least one redundant link to be useful

Cisco's Enhancement - BackboneFast

- Designed to be used in the backbone layer
- Two new BPDU types are needed
  – RLQ Request (Root Link Query Request)
  – RLQ Reply (Root Link Query Reply)
- Must be enabled in all the switches to work

Cisco's Enhancement – BackboneFast (continued)

Main goal: detect indirect link failures and maxage sooner if needed

Main idea: when a port receives an inferior BPDU, it could mean that the designated bridge has lost connectivity with the root bridge. There are two cases:

- Inferior BPDU received on the root port à sure
- Inferior BPDU received on a blocked port à possible

STP Options

- BPDU Filter
  - Associated with PortFast ports. Filters BPDUs on send and receive from an interface.
  - Enabled on 14 Blade server interfaces by default
- BPDU Guard
  - Associated with PortFast ports. Puts interface in error-disabled state when BPDU is received.
- Root Guard
  - Blocks switch connected to interface from becoming STP Root of the L2 network. Interface will go into Blocked state. Reason is "root inconsistency".

Loop Example - Yes STP

Host2

Switch 2

F   F

F   B

Switch 1

Host1

○ **Forwarding**
● **Blocking**
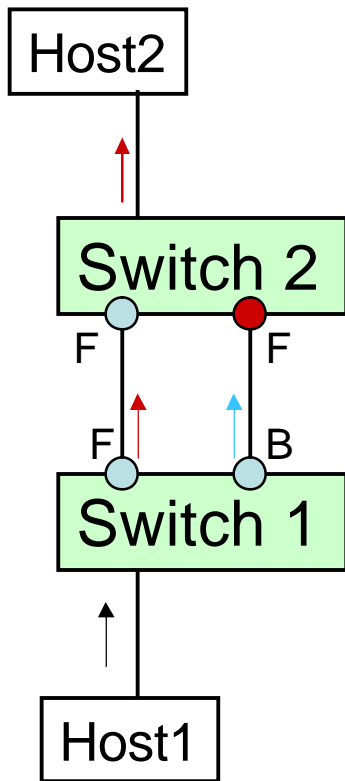
- Host1 wants to talk to Host2
- Host1 broadcasts ARP request
- Switch1 floods the broadcast
- broadcast goes around the loop once and is dropped

STP Fast Convergence Enhancements Cisco's solution

- Standard STP converges very slowly on change; so Cisco came up with fast convergence enhancements:
  - Port Fast
  - Uplink Fast and CrossStack UplinkFast
  - Backbone Fast

» 802.1w RSTP

Rapid Spanning Tree Protocol

- 802.1w is not a new protocol, but an enhancement of 802.1D
- 802.1W(RSTP) is 802.1D(STP) compatible
- The main enhancement provided by 802.1W is the significant reduction in the time taken to reconfigure the active topology in the face of changes to the physical topology or its configuration parameters

RSTP (continued)

- Aims to transition
  - Root & Designated ports to forwarding ASAP
  - Alternate & Backup ports to blocking.
- Basic functionality works on a point-to-point link.
- Requires explicit handshaking between bridges (proposal-agreement)
- Provides for Protocol Migration.
- 802.1D timers are used only as last resort (eg. fwdelay)

Spanning Tree Flavors

- DEC

- IBM

- IEEE - 802.1D

- VLAN- BRIDGE

- CISCO - PVST

- CISCO – PVST +

- CISCO - MISTP

- IEEE - 802.1W / RSTP

- IEEE - 802.1s / MSTP

- 2 ESMs in chassis

- 2 Upstream switches linked to both ESMs

Core or Gateway router
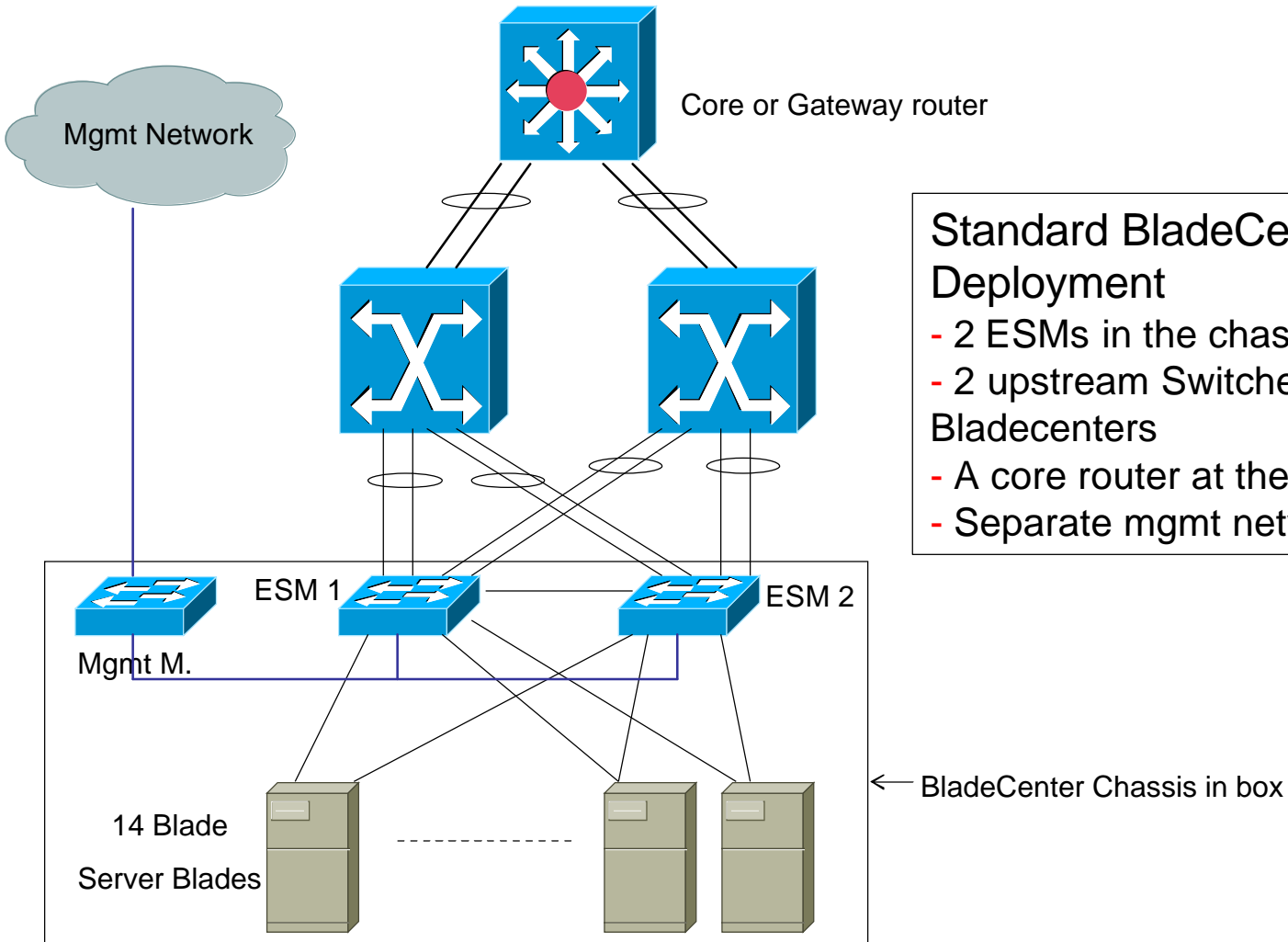
Mgmt Network

Standard BladeCenter Redundancy
Deployment
- 2 ESMs in the chassis
- 2 upstream Switches aggregating the
Bladecenters
- A core router at the top
- Separate mgmt network in RED

ESM 1

ESM 2

Mgmt M.

← BladeCenter Chassis in box

14 Blade

Server Blades

Internal ESM L2 Traffic Flow

Hard Filter that prevents traffic flow between the External ports and Mgmt Module ports

ESM

4 External Ports

2 Mgmt Module Ports

CPU / IOS

14 Internal Blade ports

Internal ESM Layer 2 Traffic cont.
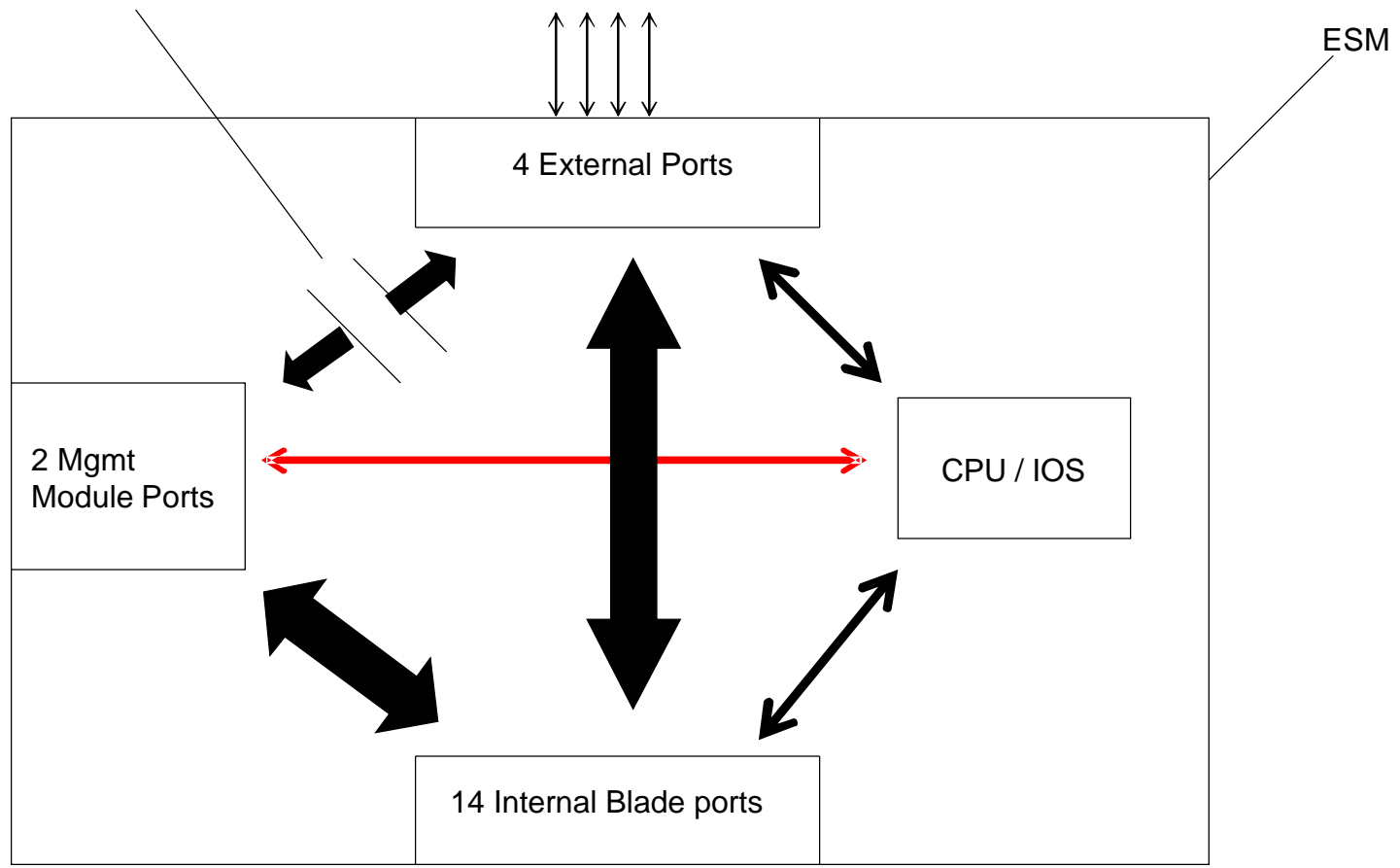
- STP doesn't know about hard filter to block layer 2 traffic from mgmt module to external ports.
  - hence the port cost.
- Two ESMs in same chassis exchange BPDUs across the Mgmt Module ports on all VLANs configured on these Ethernet interfaces.
- Layer 2 traffic can be switched across the Mgmt Module ports from BladeServer Ports.

STP Changes for ESM

Changes to CIGESM that effect STP.

Changes are from standard Cisco switch behavior.

1. **Portfast and BPDU Filter enabled by Default on BladeServer ports (1-14).**

   Can be disabled by users if they choose.

2. **Path Cost on Mgmt Module ports (15&16) changed to 100 from 19 (which is default for 100Mbps).**

   Can be modified by users if they choose.

   Allows STP to block these interfaces if loop is detected

3. **ESM Will not allow the current Mgmt VLAN to block on the Mgmt Module Ports. Always in forwarding state.**

   This is not modifiable by users.

4. **Port to redundant Mgmt Module is disabled.**

   Done on purpose to prevent STP loops. Will become active when ESM receives event that new Mgmt Module is active.

5. **Rapid PVSTP enabled by default, not regular PVSTP.**

```
Switch_ESM_1 # show spanning-tree
VLAN0002
  Spanning tree enabled protocol rstp
  Root ID    Priority    32768
             Address     00b0.8e7f.88f8
             Cost        24
             Port        17 (GigabitEthernet0/17)
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     0003.fd62.c240
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300


Interface        Role Sts Cost       Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Gi0/10           Desg FWD 4           128.10   Edge P2p
Gi0/17           Root FWD 10          128.10   P2p


VLAN0010
 Spanning tree enabled protocol rstp
  Root ID    Priority    32778
             Address     0003.fd62.c240
             This bridge is the root
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32778  (priority 32768 sys-id-ext 10)
             Address     0003.fd62.c240
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300

Interface        Role Sts Cost       Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Gi0/16           Desg FWD 100         128.16   P2p
```

- Use another network element to be STP Root. ESM should not be root.
  - most customers will do this anyway.
  - There is no known problem is this occurs, just not a good idea
- Don't un-configure port cost on Mgmt Module ports.
- Keep the ESM Mgmt Vlan separate from user traffic.
  - default Mgmt vlan is 1.

IEEE Documents

- IEEE 802.1D - Media Access Control (MAC) Bridges

- IEEE 802.1Q - Virtual Bridged Local Area Networks

- IEEE 802.1w - Rapid Reconfiguration (Supp. to 802.1D) (D9)

- IEEE 802.1s - Multiple Spanning Tree (Supp. to 802.1Q) (D9)

- IEEE 802.1t   - Local and Metropolotan area Networks: Common Specifications

- UDLD – UniDirectional Link Detect
    - **Layer 2 Protocol between end points**
- When Unidirectional is Link detected, port is administratively shut down
    - **If other end is not running UDLD, port is shut down.**
- A Unidirectional link can cause a Layer 2 loop.
- Typically used on fiber optic connections, but can be used on copper twisted pair as well.

IBM
@server™
xSeries Education
Feature Configuration

Switch Features and Configurations

- Spanning Tree Protocol and Enhancements
- Interface Types
- VTP and VLAN Configuration
- Trunk Configuration
- EtherChannel
- SPAN
- Access Control List

Interface Types

- Access Port – belong to one VLAN, packets not tagged
- Trunk Port – carry traffic of multiple VLANs
- Port-channel – logical interface of EtherChannel Port Groups
- Switched Virtual Interface (SVI) – Layer 3 representation of the (management) VLAN

VLAN Configuration

- Normal range VLAN
  - VLAN ID 1 to 1005
  - Can only add/remove/change when in VTP server or transparent mode
  - saved in VLAN database file vlan.dat
  - VLAN 1, and 1002 to 1005 are specially created and cannot be removed
- Extended range VLAN
  - VLAN ID 1006 to 4094
  - must be in VTP transparent mode when create, not recognized by VTP
  - not saved in VLAN database file vlan.dat
  - must use config-vlan mode
  - - cannot be included in the prunning eligible range
  - cannot be configured by VMPS

VLAN Configuration

- Create a new VLAN
  - VLAN database mode

  ```
  switch# vlan database
  ```

  - config-vlan mode

  ```
  switch(config)# vlan VLAN_#
  ```

- Statically assign a port to the VLAN

  ```
  switch(config-if)# switchport mode access

  switch(config-if)# switchport access vlan VLAN_#
  ```

- Dynamically assign a port to the VLAN

  ```
  switch(config-if)# switchport mode access

  switch(config-if)# switchport access vlan dynamic
  ```

Configure Trunk

- External ports are in dynamic desirable DTP mode
    - enables DTP (Cisco's dynamic trunking protocol)
    - actively negotiates trunking state (access or trunk) and encapsulation (802.1Q)
    - for non-Cisco switches, set trunk mode manually
- Other DTP modes:
    - access: a user port and cannot be a trunk
    - trunk: a trunk and negotiates trunking with other port
    - dynamic auto: passively waits to be contacted.

- ## Set trunk mode manually

switch(config-if)# **switchport mode trunk**

- ## Set trunk native VLAN

switch(config-if)# **switchport trunk native vlan** *VLAN_#*

- ## Configure allowed VLANs

switch(config-if)# **switchport trunk allowed vlan add** *VLAN_#*

Troubleshooting Trunk

- Check interface type and link status

  switch# **show interface state**

- Check the spanning tree state of the port

  switch# **show spanning-tree interface gig0/17**

- Check native VLAN on both sides to match

  switch# **show interface gig0/17 switchport**

Configure VTP

- VLAN Trunking Protocol (VTP)
  - runs over trunk links
  - synchronize the VLAN databases of all switches in the VTP domain
  - a VTP domain is an administrative group
  - All switches within the group must have the same VTP domain name
- VTP switch roles
  - Server: can create/delete/rename VLANs (Cisco default)
  - Client: cannot make VLAN changes
  - Transparent: can create/delete/rename VLANs; local database only; forward VTP advertisements (CIGESM default)
- VTP version
  - version 1
  - version 2: support Token Ring and others
  - not compatible (either all running version 1 or all running version 2)

- ## Set up VTP mode

  switch(config)# **vtp mode { server | client | transparent }**

- ## Set up VTP domain name

  switch(config)# **vtp domain** *NAME*

- ## Set up VTP version

  switch(config)# **vtp version { 1 | 2 }**

Configure EtherChannel

- A way of combining several physical links between switches into one logical connection to aggregate bandwidth

- Treated EtherChannel as one path by Spanning Tree protocol

- Load balancing on source MAC address or destination MAC address

- Achieves redundancy

Configure EtherChannel

- All ports must be the same speed and duplex

- All ports in the bundle should be enabled

- None of them can be a SPAN destination port

- Two channel protocols

- IEEE 802.3ad Link Aggregation Control Protocol (LACP)

  - active mode: port actively negotiates channeling

  - passive mode: wait to be contacted

- Cisco proprietary Port Aggregation Protocol (PAgP)

  - On mode: forms channel without PAgP

  - Auto mode: responds, do not initiate

  - Desirable mode: actively negotiates

Configure EtherChannel

- Bundle several interfaces

> switch(config)# **interface {** *interface* **| range** *interface – interface* **}**
>
> switch(config-if-range)# **channel-group** *NUMBER* **mode { active | passive | auto | on | desirable }**

- Configure channel protocol

> switch(config-if-range)# **channel-protocol { lacp | pagp }**

- Configure load-balancing

> switch(config# **port-channel load-balance { dst-mac | src-mac }**

- Verify EtherChannel

> switch# **show etherchannel summary**
>
> switch# **show etherchannel** *NUMBER* **port-channel**

- SPAN – Switched Port Analyzer (Port mirroring)
- Can specify traffic directions
    - Receive (Rx) – ingress of source port
    - Transmit (Tx) – egress of source port
    - Both
- Source port (monitored port) can be any type, in any VLAN
- Destination port (monitoring port) can be any physical port, but not port-channel, not SVI

- Create monitor session with source port

  switch(config)# **monitor session** *NUMBER* **source interface** *interface* **{ both | rx | tx }**

- Specify destination port

  switch(config)# **monitor session** *NUMBER* **destination interface** *interface*

- Verify SPAN

  switch# **show monitor [session** *NUMBER* **]**

Configure ACL

- ACL – Access Control List

- An ACL contains an ordered list of access control entries (ACE)

- Each ACE specifies permit or deny conditions the packets must satisfy to match the ACE

- CIGESM supports the following types of ACL

  - IP ACL filters IP, TCP, and UDP traffic

  - Ethernet or MAC ACL filters Layer 2 traffic

  - MAC extended access lists use source and destination MAC addresses and optional protocol type information for matching operations

  - Standard IP access lists use source addresses for matching operations

  - Extended IP access lists use source and destination addresses and optional protocol type information for matching operations

Configure ACL

- Numbered IP standard ACL

  switch (config)# **access-list 2 deny host 171.69.198.102**

  switch (config)# **access-list 2 permit any**

  switch# **show access-lists**

- Numbered IP extended ACL

  switch(config)# **access-list 102 deny tcp 171.69.198.0 0.0.0.255 172.20.52.0 0.0.0.255 eq telnet**

  switch(config)# **access-list 102 permit tcp any any**

- Named MAC extended ACL

  switch(config)# **mac access-list extended my-mac-acl**

  switch(config-ext-macl)# **deny any any decnet-iv**

  switch(config-ext-macl)# **permit any any**

  switch# **show access-list**

- Apply IP ACL at physical interface

  switch(config-if)# **ip access-group 2 in**

  switch# **show ip interface** *interface*

- Apply MAC ACL at physical interface

  switch(config-if)# **mac access-group my-mac-acl in**

  switch# **show mac access-group interface** *interface*

IBM
@server
xSeries Education
Serial Over LAN

Server Serial Connections

- Initially used the serial interface in the servers

- Currently used in a terminal server environment

- Remotely manage the server

- Remotely load scripts for configuring the server

Stand-alone Environment

BladeCenter™ Technical Training

LAN Connection

Terminal Server

Serial Ports

IBM Server   IBM Server   IBM Server

Rack-optimized SOL

LAN Connection

Terminal Server

Network

Rack of 1U Servers

Remote IT Specialist
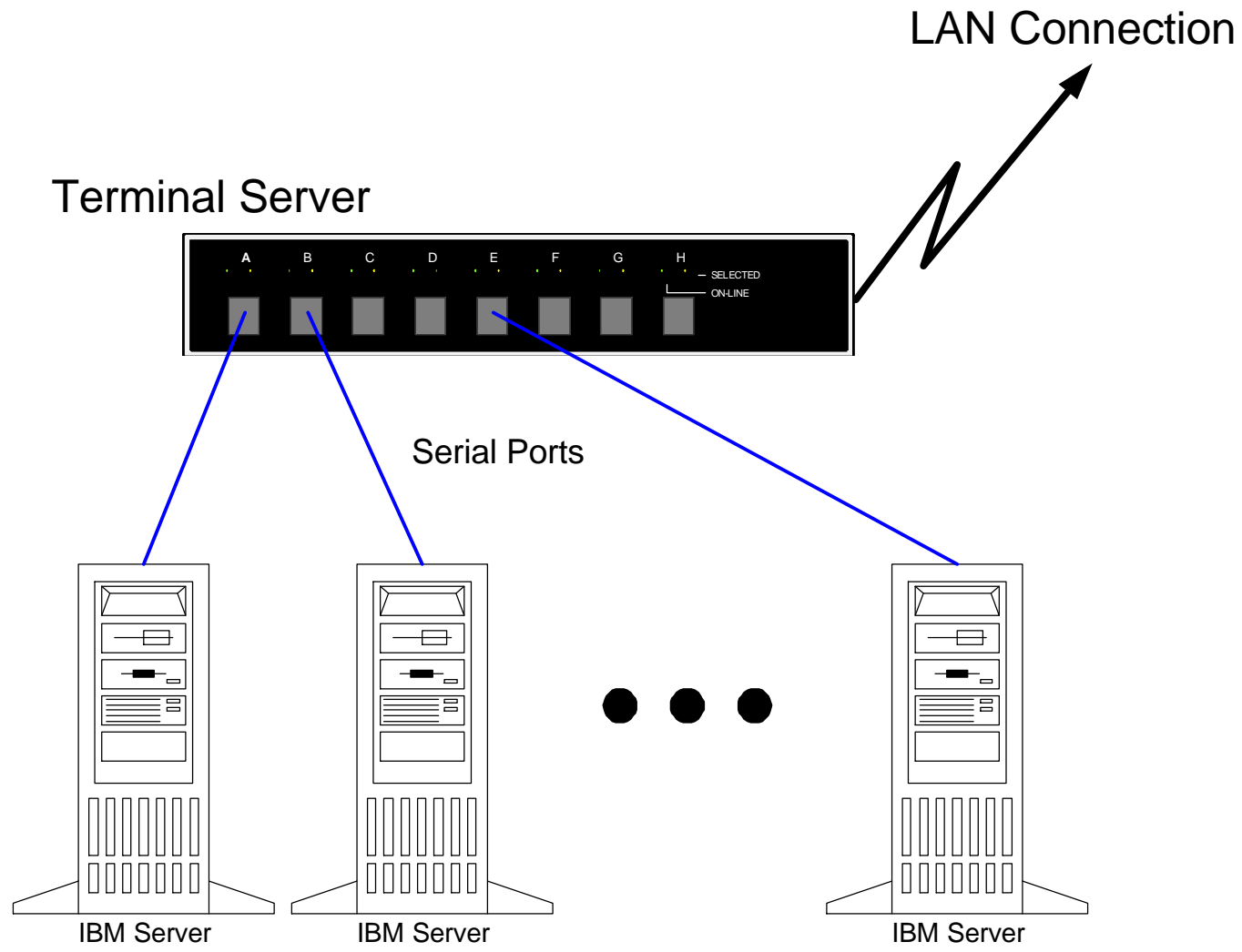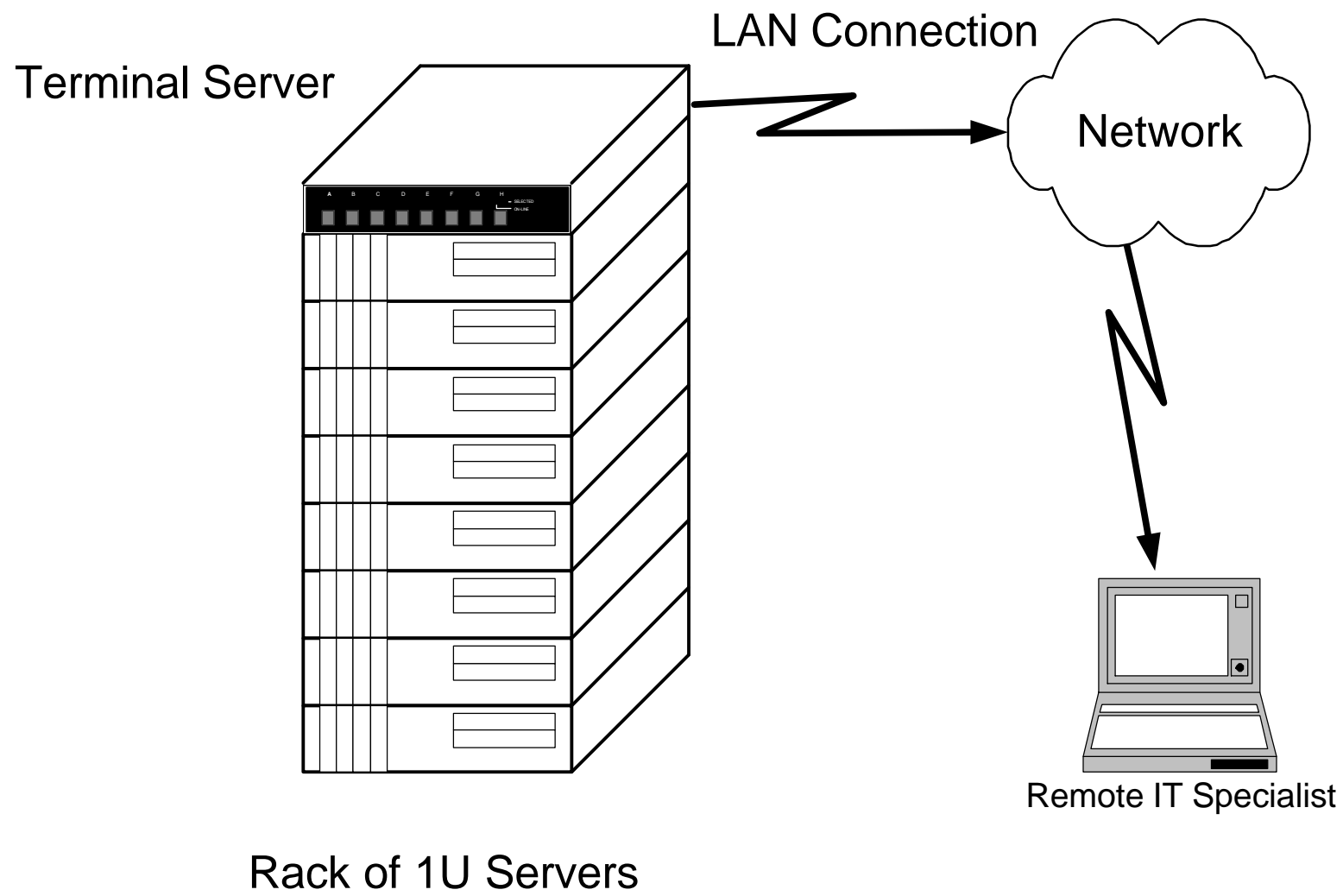
What is Serial Over LAN in a Blade?

BIOS

Serial traffic

H8

Ethernet chip

VLAN 3
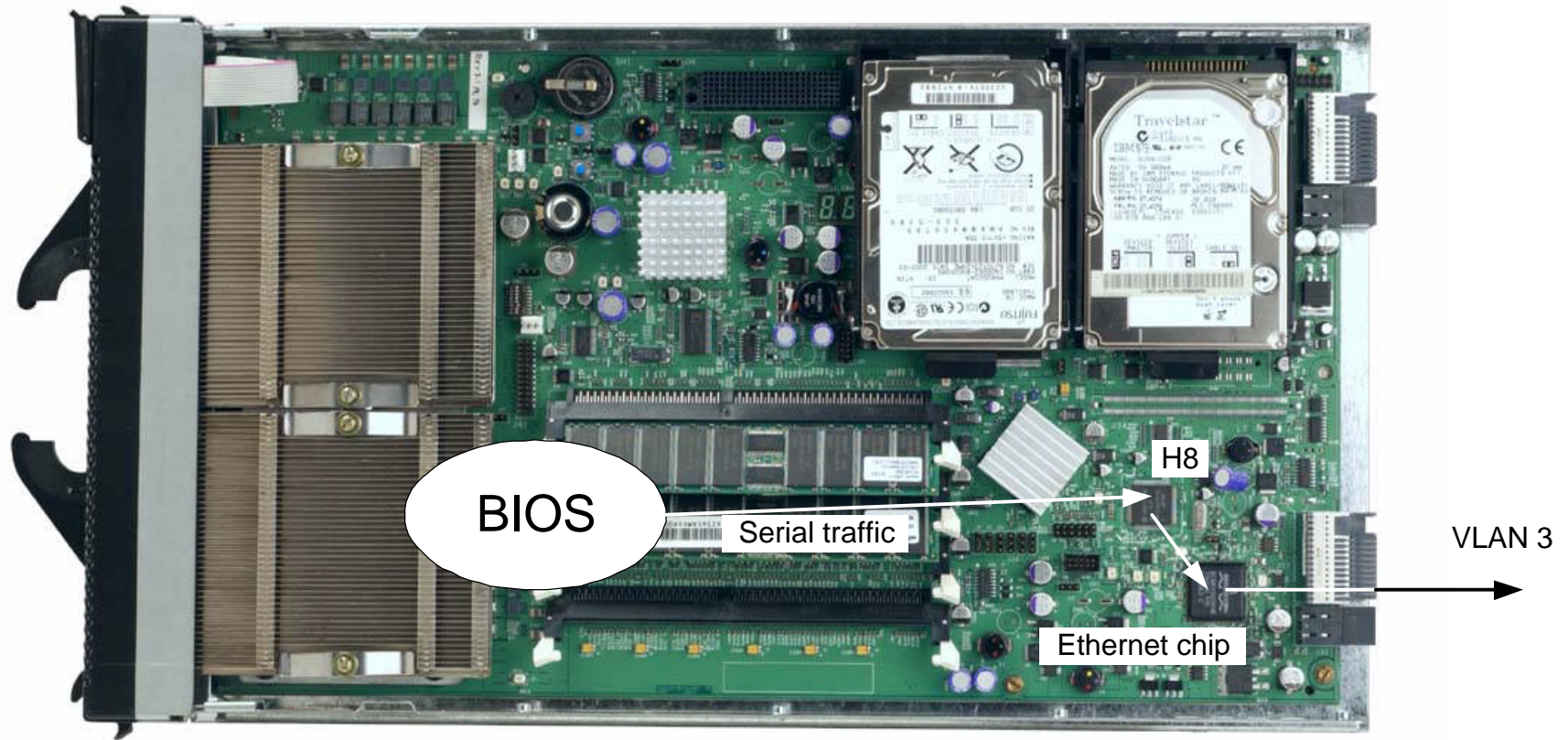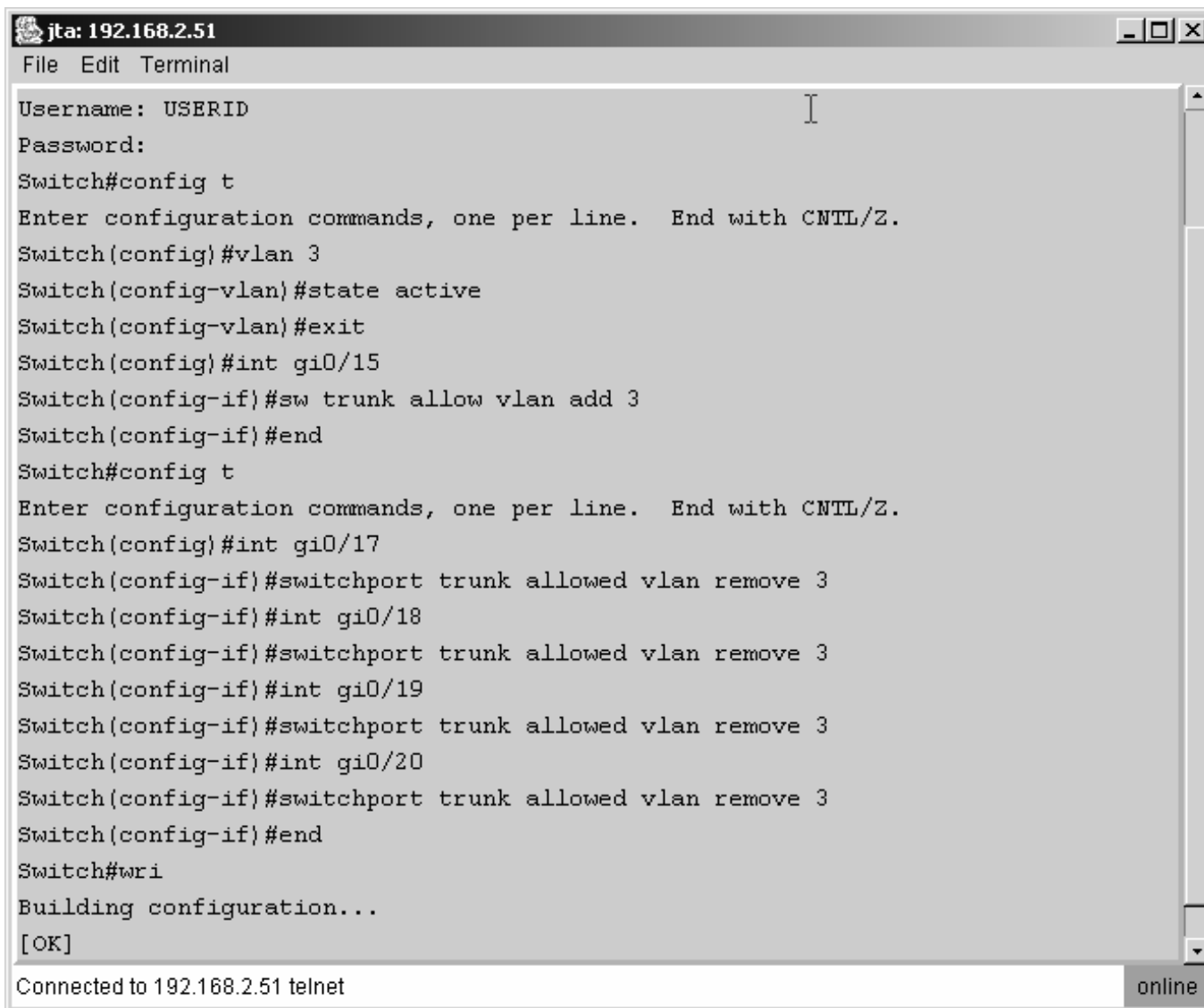
The H8 forms the SOL packet
The Ethernet chip adds the packet to its data stream

Switch Module - SOL

SWITCH MODULE
(1 and 2)

Ethernet Switch Hardware

4

14

(CPU)

to SW 1
to SW 2
to SW 3
to SW 4

MID PLANE

SERVER BLADE
(1 of 14)

Dual Ethernet "Adapter"

Optional Daughter Card

P C I

Blade Subsystems

Ethernet Switch

MGMT MODULE
(1 of 2)

Internal Mgmt
Network Connections:
2 100Mbps ports
(1 per mgmt module)

SOL Configuration - CIGESM

BladeCenter™ Technical Training

SOL Configuration Management Module

BladeCenter™ Technical Training

SOL Configuration Management Module

BladeCenter™ Technical Training

BladeCenter™ Technical Training