



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Network Security

Security in Network Layer

1

Security in Datalink layer

- Attack Switch
 - MAC flooding
 - Fake MAC address
- Attack ARP protocol
 - ARP Spoofing
 - DoS
- Attack DHCP protocol
 - DHCP Spoofing
 - DHCP Starvation
- Attack VLAN



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

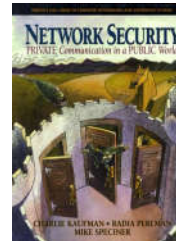
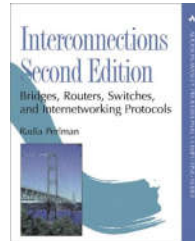
2

2

Attack STP



Radia Perlman,
networking hero!



- Ethernet bridges and switches can implement the **IEEE 802.1D Spanning-Tree Protocol** and use the spanning-tree algorithm to **construct a loop free shortest path network**.
- Radia Perlman “is the inventor of the spanning tree algorithm used by bridges (switches), and the mechanisms that make link state routing protocols such as IS-IS (which she designed) and OSPF (which adopted many of the ideas) stable and efficient. Her thesis on sabotage-proof networks is well-known in the security community.”

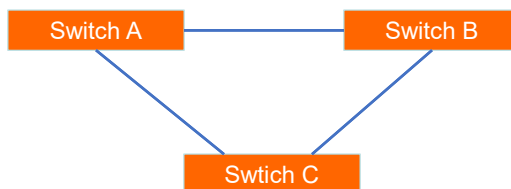


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

<http://www.equipecom.com/radia.html>

3

Switching Loop



- When there is more than one path between two switches
- What are the potential problems?



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

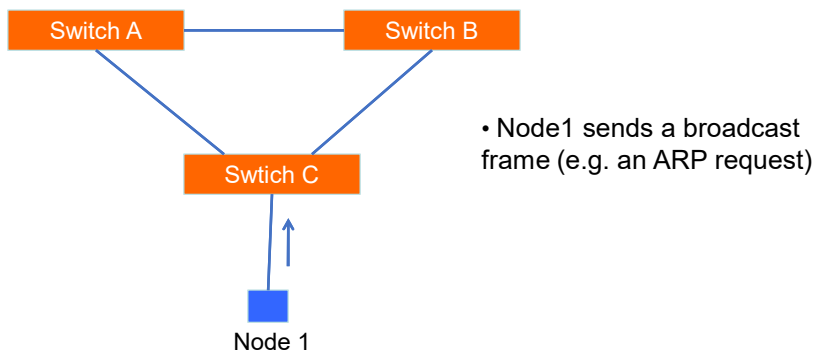
4

Switching Loop

- If there is more than one path between two switches:
 - Forwarding tables become unstable
 - Source MAC addresses are repeatedly seen coming from different ports
 - Switches will broadcast each other's broadcasts
 - All available bandwidth is utilized
 - Switch processors cannot handle the load

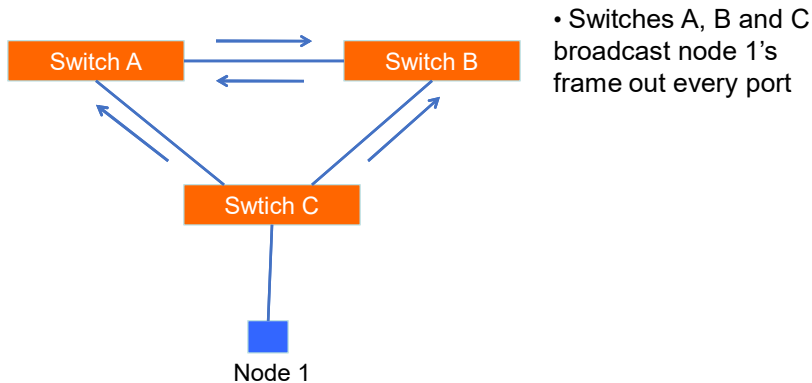
5

Switching Loop



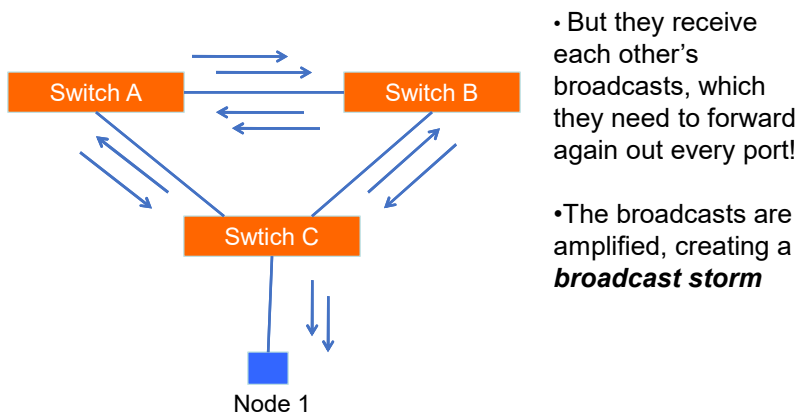
6

Switching Loop



7

Switching Loop



8

Good Switching Loops

- But you can take advantage of loops!
 - Redundant paths improve resilience when:
 - A switch fails
 - Wiring breaks
- How to achieve redundancy without creating dangerous traffic loops?



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

9

Redundancy

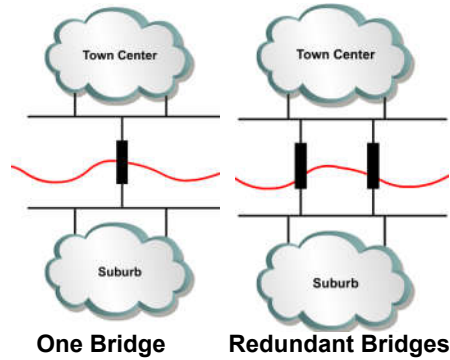
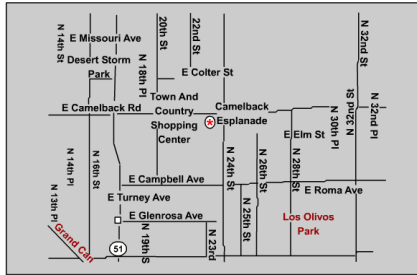
- Achieving such a goal requires extremely reliable networks.
- Reliability in networks is achieved by reliable equipment and by designing networks that are tolerant to failures and faults.
- The network is designed to reconverge rapidly so that the fault is bypassed.
- Fault tolerance is achieved by redundancy.
- Redundancy means to be in excess or exceeding what is usual and natural.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

10

Redundant topologies

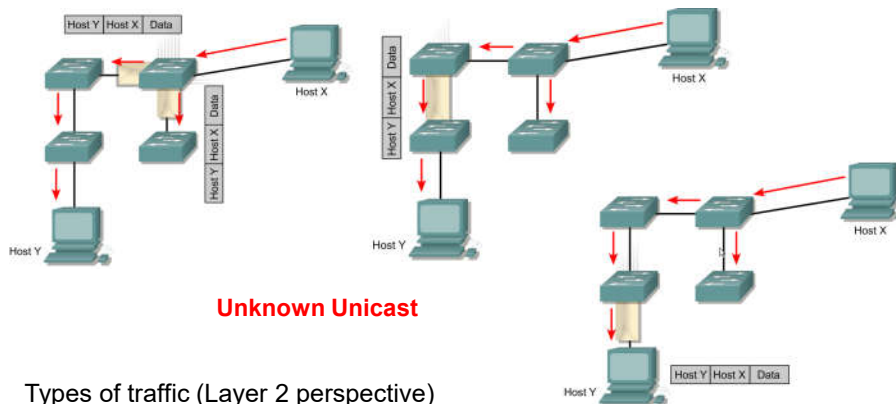


- A network of roads is a global example of a redundant topology.
- If one road is closed for repair there is likely an alternate route to the destination



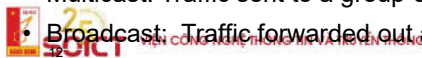
11

Types of Traffic



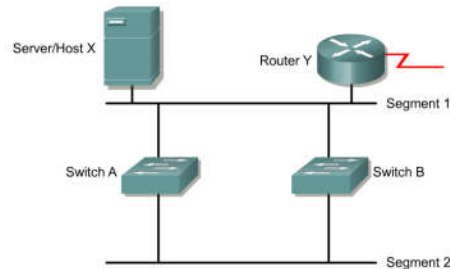
Types of traffic (Layer 2 perspective)

- Known Unicast: Destination addresses are in Switch Tables
- Unknown Unicast: Destination addresses are not in Switch Tables
- Multicast: Traffic sent to a group of addresses
- Broadcast: Traffic forwarded out all interfaces except incoming interface.



12

Redundant switched topologies



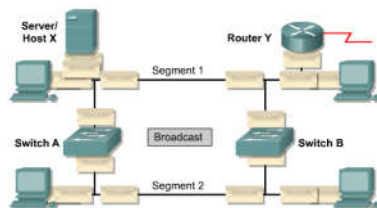
- Switches learn the MAC addresses of devices on their ports so that data can be properly forwarded to the destination.
- Switches will flood frames for unknown destinations until they learn the MAC addresses of the devices.
- Broadcasts and multicasts are also flooded. (Unless switch is doing Multicast Snooping or IGMP)
- A redundant switched topology **may** (STP disabled) cause broadcast storms, multiple frame copies, and MAC address table instability problems.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

13

Broadcast Storm



A state in which a message that has been broadcast across a network results in even more responses, and each response results in still more responses in a snowball effect.

www.webopedia.com

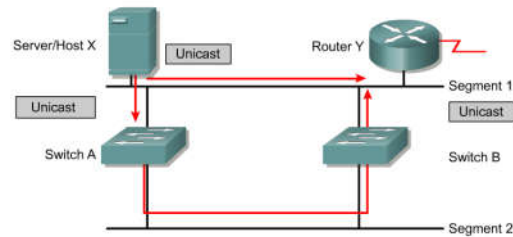
- Broadcasts and multicasts can cause problems in a switched network.
- If Host X sends a broadcast, like an ARP request for the Layer 2 address of the router, then Switch A will forward the broadcast out all ports.
- Switch B, being on the same segment, also forwards all broadcasts.
- Switch B sees all the broadcasts that Switch A forwarded and Switch A sees all the broadcasts that Switch B forwarded.
- Switch A sees the broadcasts and forwards them.
- Switch B sees the broadcasts and forwards them.
- The switches continue to propagate broadcast traffic over and over.
- This is called a broadcast storm.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

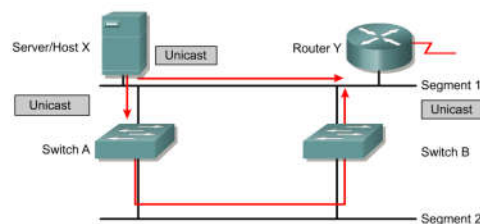
14

Multiple frame transmissions



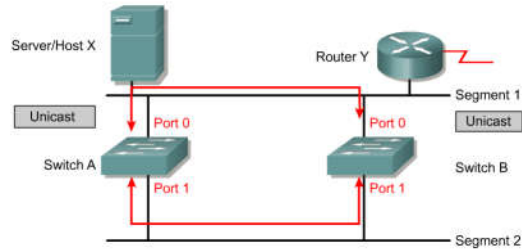
- In a redundant switched network it is possible for an end device to receive multiple frames.
- Assume that the MAC address of Router Y has been timed out by both switches.
- Also assume that Host X still has the MAC address of Router Y in its ARP cache and sends a unicast frame to Router Y.

Multiple frame transmissions

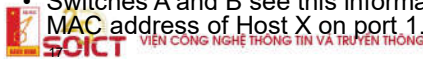


- The router receives the frame because it is on the same segment as Host X.
- Switch A does not have the MAC address of the Router Y and will therefore flood the frame out its ports. (Segment 2)
- Switch B also does not know which port Router Y is on.
- Note: Switch B will forward the the unicast onto Segment 2, creating multiple frames on that segment.
- After Switch B receives the frame from Switch A , it then floods the frame it received causing Router Y to receive multiple copies of the same frame.

Media access control database instability

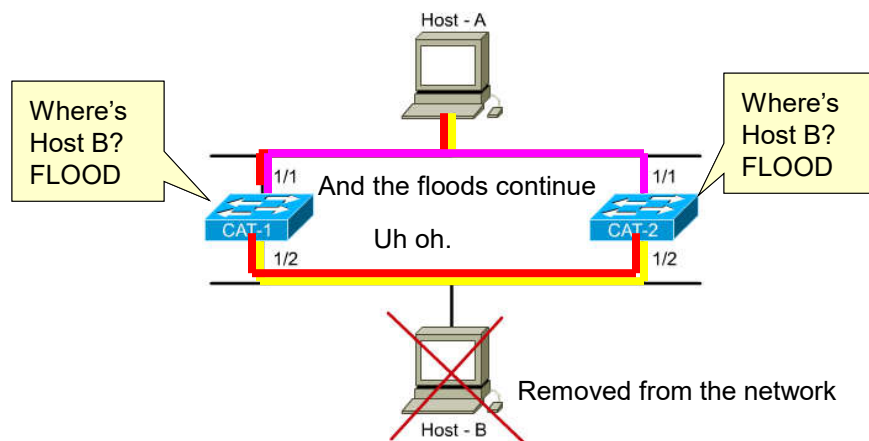


- In a redundant switched network it is possible for switches to learn the wrong information.
- A switch can incorrectly learn that a MAC address is on one port, when it is actually on a different port.
- Host X sends a frame directed to Router Y.
- Switches A and B learn the MAC address of Host X on port 0.
- The frame to Router Y is flooded on port 1 of both switches.
- Switches A and B see this information on port 1 and incorrectly learn the MAC address of Host X on port 1.



17

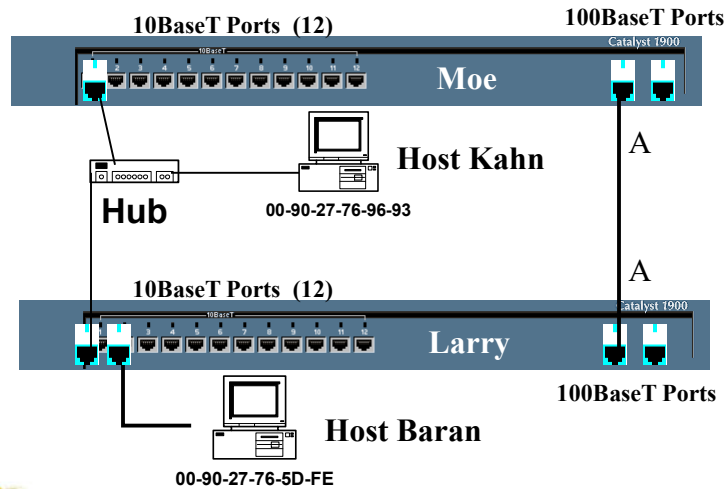
Layer 2 Loops - Flooded unicast frames



18

Redundant Paths and No Spanning Tree

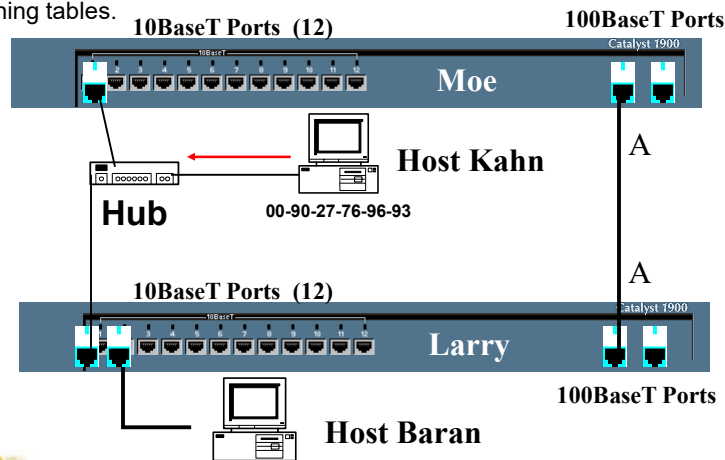
Another problem, incorrect MAC Address Tables



19

Redundant Paths and No Spanning Tree

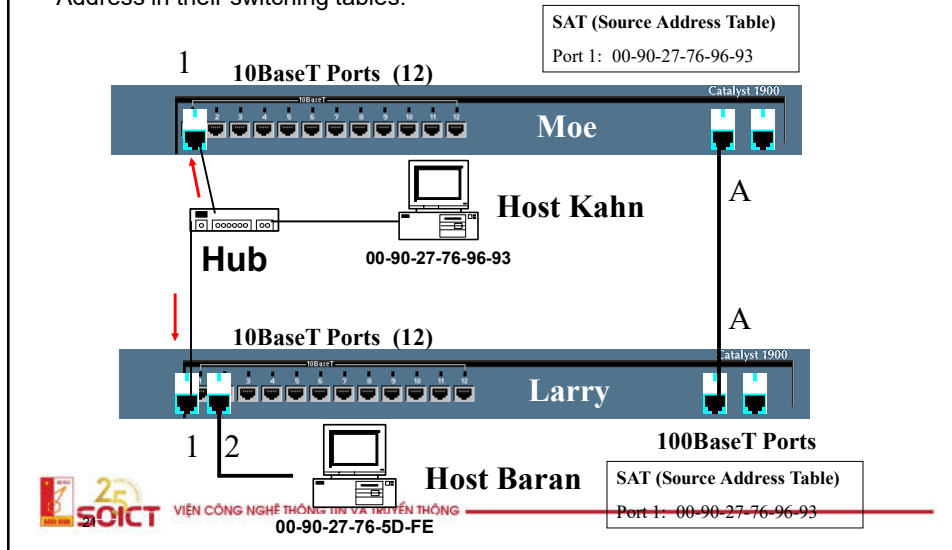
Host Kahn sends an Ethernet frame to Host Baran. Both Switch Moe and Switch Larry see the frame and record Host Kahn's Mac Address in their switching tables.



20

Redundant Paths and No Spanning Tree

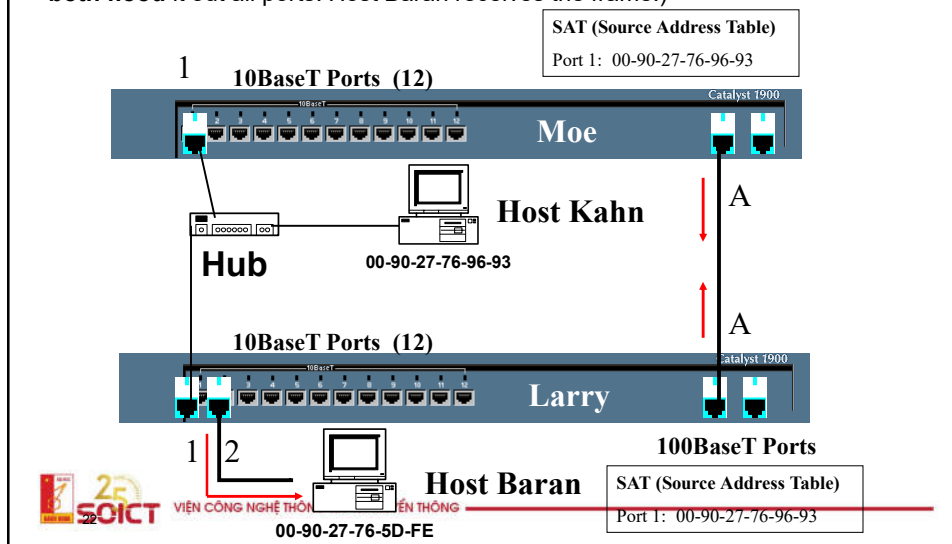
Both Switch Moe and Switch Larry see the frame and record Host Kahn's Mac Address in their switching tables.



21

Redundant Paths and No Spanning Tree

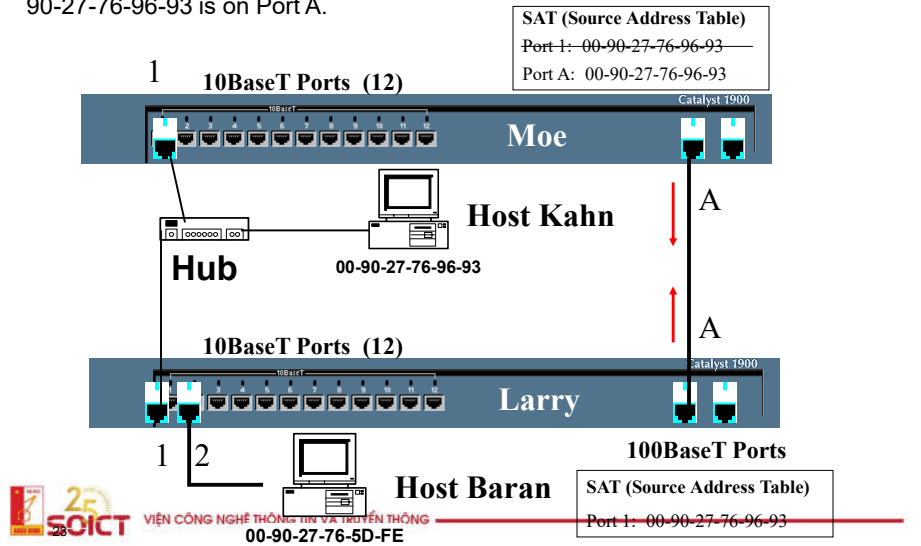
Both Switches do not have the **destination MAC address** in their table so they **both flood** it out all ports. Host Baran receives the frame.)



22

Redundant Paths and No Spanning Tree

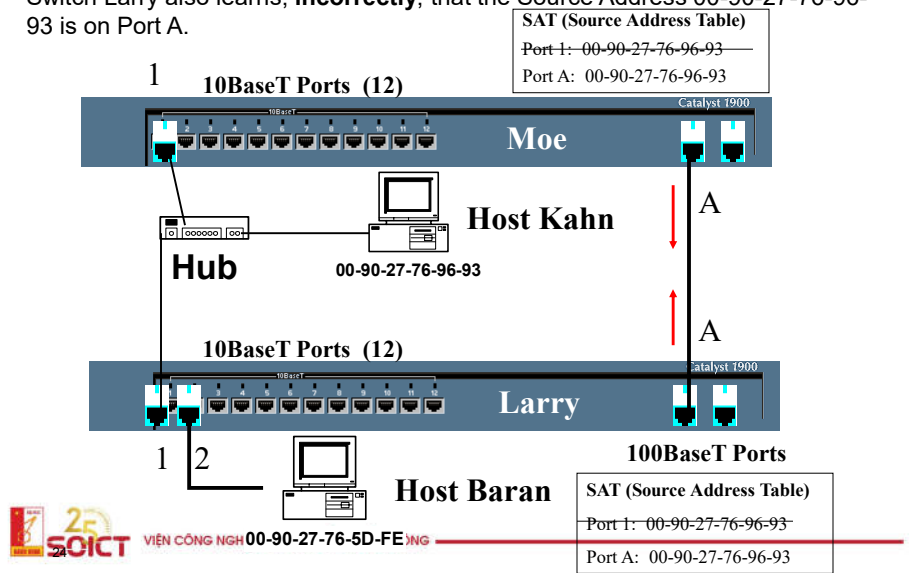
Switch Moe now learns, **incorrectly**, that the Source Address 00-90-27-76-96-93 is on Port A.



23

Redundant Paths and No Spanning Tree

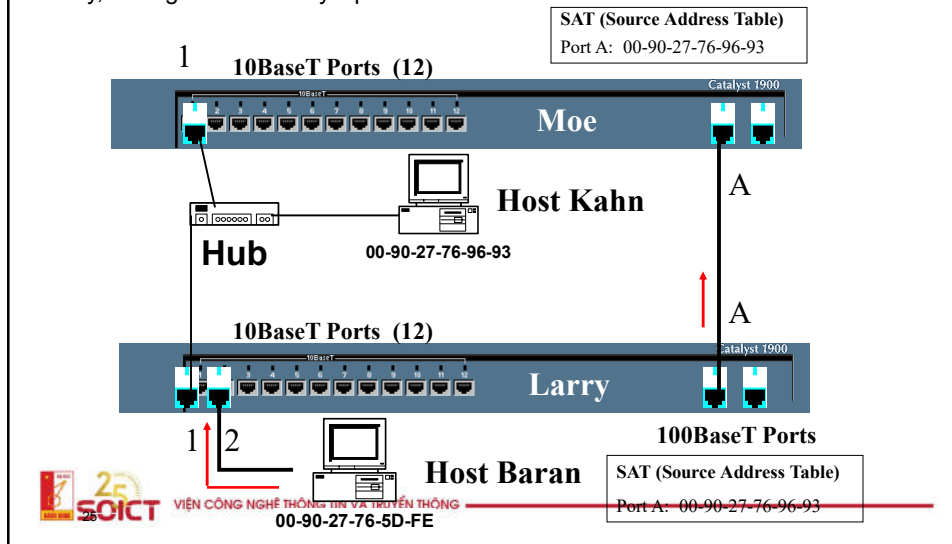
Switch Larry also learns, **incorrectly**, that the Source Address 00-90-27-76-96-93 is on Port A.



24

Redundant Paths and No Spanning Tree

Now, when Host Baran sends a frame to Host Kahn, it will be sent the longer way, through Switch Larry's port A.



25

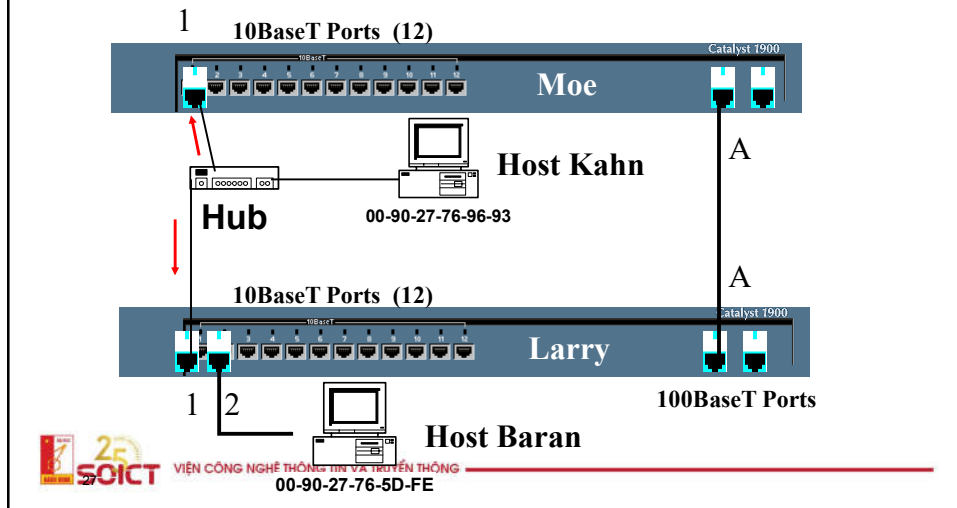
Redundant Paths and No Spanning Tree

- Then, the same confusion happens, but this time with Host Baran.
- Okay, maybe not the end of the world.
- Frames will just take a longer path and you may also see other “unexpected results.”
- But what about broadcast frames, like ARP Requests?

26

Broadcasts and No Spanning Tree

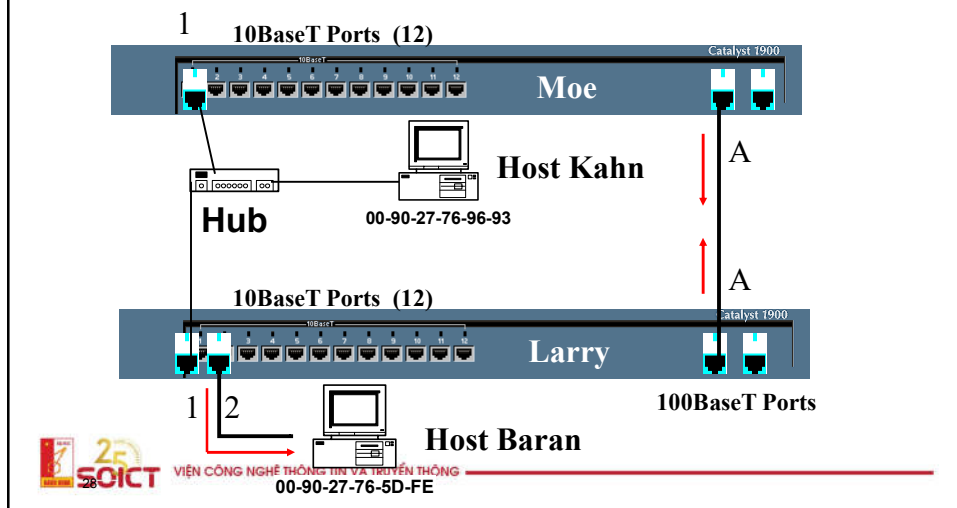
Lets, leave the switching tables alone and just look at what happens with the frames. Host Kahn sends out a layer 2 broadcast frame, like an ARP Request.



27

Broadcasts and No Spanning Tree

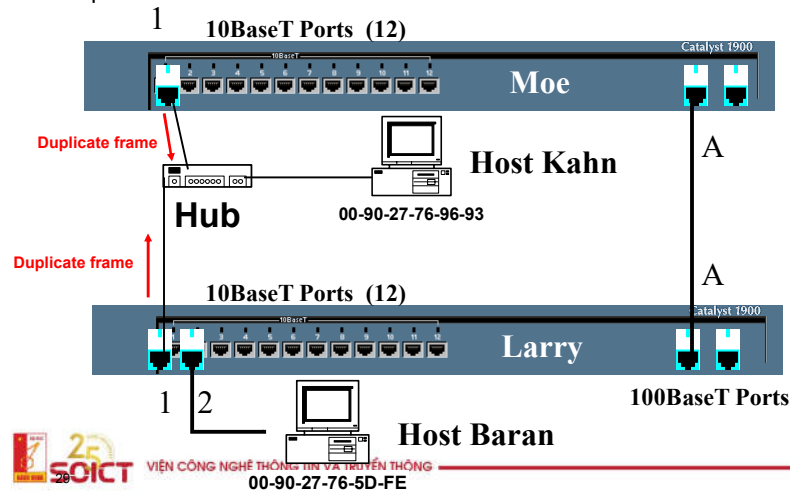
Because it is a layer 2 broadcast frame, both switches, Moe and Larry, **flood the frame out all ports**, including their port A's.



28

Broadcasts and No Spanning Tree

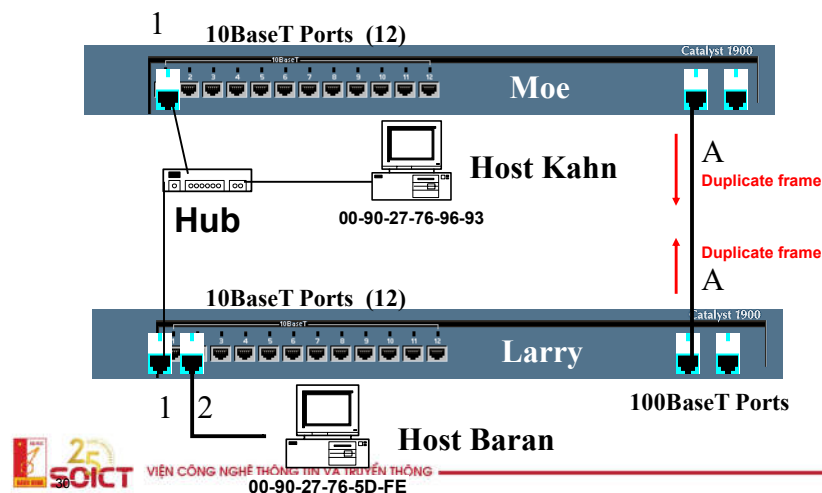
Both switches receive the same broadcast, but on a different port. Doing what switches do, **both switches flood the duplicate broadcast frame** out their other ports.



29

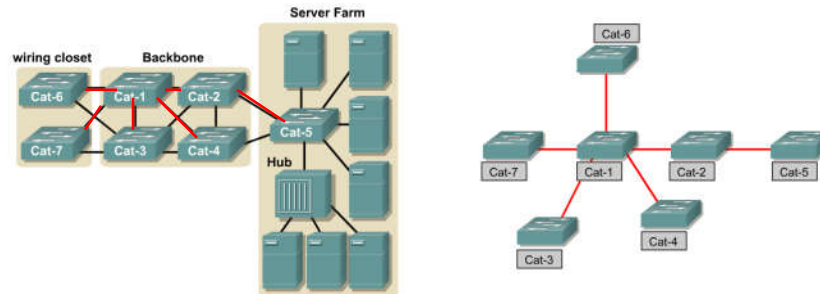
Broadcasts and No Spanning Tree

Here we go again, with the switches flooding the same broadcast again out its other ports. This results in duplicate frames, known as a **broadcast storm**!



30

Redundant topology and spanning tree



- It is a spanning tree because all devices in the network are reachable or spanned.
- The algorithm used to create this loop free logical topology is the **spanning-tree algorithm**.
- This algorithm can take a relatively long time to converge.
- A new algorithm called the **rapid spanning-tree algorithm** is being introduced to reduce the time for a network to compute a loop free logical topology.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

33

Traditional Spanning Tree (802.1d)

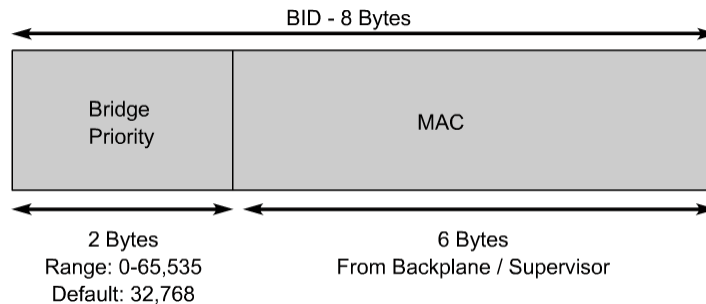
- First Step:
 - Decide on a point of reference: the **Root Bridge**
 - The election process is based on the Bridge ID, which is composed of:
 - The Bridge Priority: A two-byte value that is configurable
 - The MAC address: A unique, hardcoded address that cannot be changed.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

34

Bridge ID (BID)



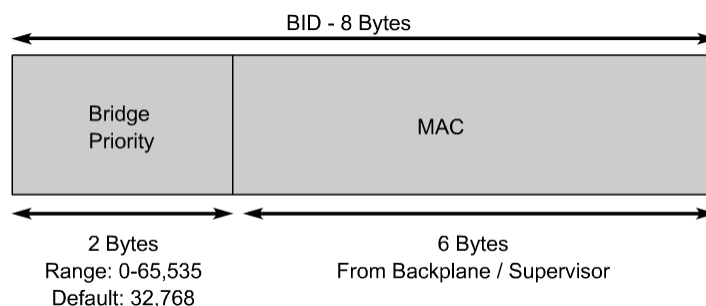
- **Bridge ID (BID)** is used to identify each bridge/switch.
- The BID is used in determining the center of the network, in respect to STP, known as the root bridge.
- Consists of two components:
 - **A 2-byte Bridge Priority:** Cisco switch defaults to **32,768** or 0x8000.
 - **A 6-byte MAC address**



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

35

Bridge ID (BID)



- **Bridge Priority** is usually expressed in **decimal format** and the **MAC address** in the BID is usually expressed in **hexadecimal format**.
- BID is used to elect a root bridge (coming)
- **Lowest Bridge ID is the root.**
- If all devices have the same priority, the bridge with the lowest MAC address becomes the root bridge. (Yikes!)



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

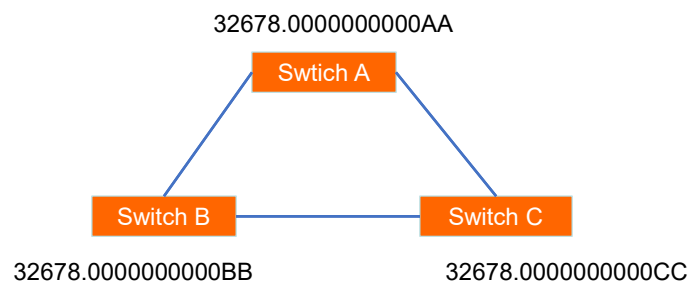
36

Root Bridge Selection (802.1d)

- Each switch starts by sending out BPDUs with a Root Bridge ID equal to its own Bridge ID
 - *I am the root!*
- Received BPDUs are analyzed to see if a lower Root Bridge ID is being announced
 - If so, each switch replaces the value of the advertised Root Bridge ID with this new lower ID
- Eventually, they all agree on who the Root Bridge is

37

Root Bridge Selection (802.1d)



- All switches have the same priority.
- Who is the elected root bridge?

38

Root Port Selection (802.1d)

- Now each switch needs to figure out where it is in relation to the Root Bridge
 - Each switch needs to determine its **Root Port**
 - The key is to find the port with the lowest **Root Path Cost**
 - The cumulative cost of all the links leading to the Root Bridge

Root Port Selection (802.1d)

- Each link on a switch has a **Path Cost**
 - Inversely proportional to the link speed
 - e.g. The faster the link, the lower the cost

Link Speed	STP Cost
10 Mbps	100
100 Mbps	19
1 Gbps	4
10 Gbps	2

Root Port Selection (802.1d)

- **Root Path Cost** is the accumulation of a link's Path Cost and the Path Costs learned from neighboring Switches.
 - It answers the question: *How much does it cost to reach the Root Bridge through this port?*



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

41

Root Port Selection (802.1d)

1. Root Bridge sends out BPDUs with a Root Path Cost value of 0
2. Neighbor receives BPDU and adds port's Path Cost to Root Path Cost received
3. Neighbor sends out BPDUs with new cumulative value as Root Path Cost
4. Other neighbor's down the line keep adding in the same fashion



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

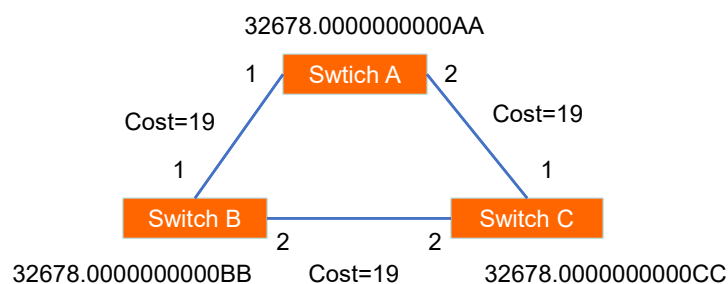
42

Root Port Selection (802.1d)

- On each switch, the port where the lowest Root Path Cost was received becomes the **Root Port**
 - This is the port with the best path to the Root Bridge

43

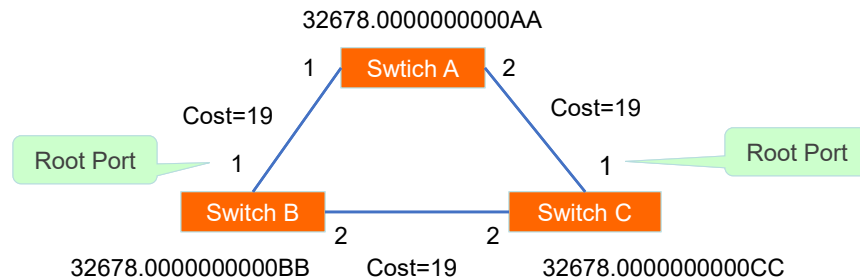
Root Port Selection (802.1d)



- What is the Path Cost on each Port?
- What is the Root Port on each switch?

44

Root Port Selection (802.1d)



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

45

Electing Designated Ports (802.1d)

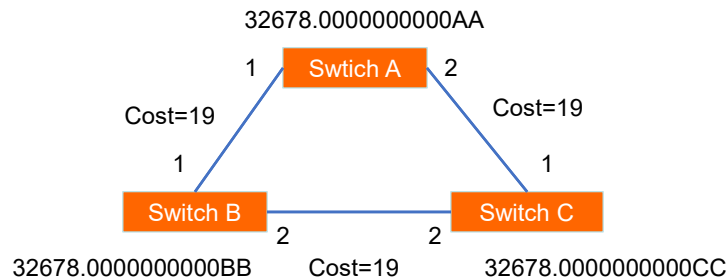
- OK, we now have selected root ports but we haven't solved the loop problem yet, have we
 - The links are still active!
- Each network segment needs to have only one switch forwarding traffic to and from that segment
- Switches then need to identify one **Designated Port** per link
 - The one with the lowest cumulative Root Path Cost to the Root Bridge



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

46

Electing Designated Ports(802.1d)



- Which port should be the Designated Port on each segment?



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

47

Electing Designated Ports (802.1d)

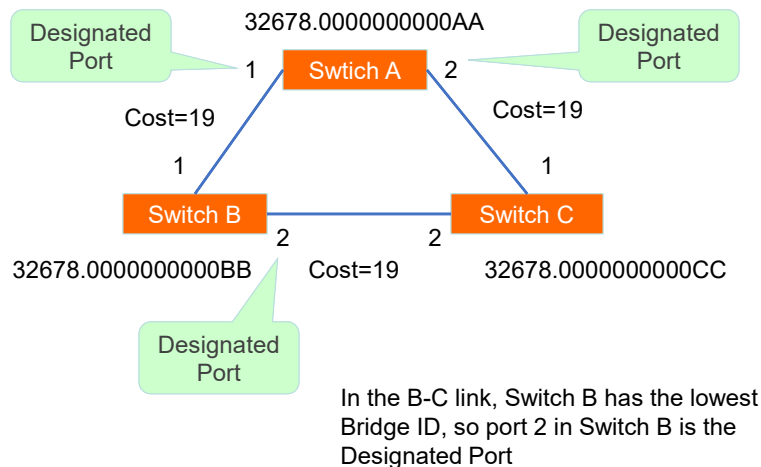
- Two or more ports in a segment having identical Root Path Costs is possible, which results in a tie condition
- All STP decisions are based on the following sequence of conditions:
 - Lowest Root Bridge ID
 - Lowest Root Path Cost to Root Bridge
 - Lowest Sender Bridge ID
 - Lowest Sender Port ID



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

48

Electing Designated Ports(802.1d)



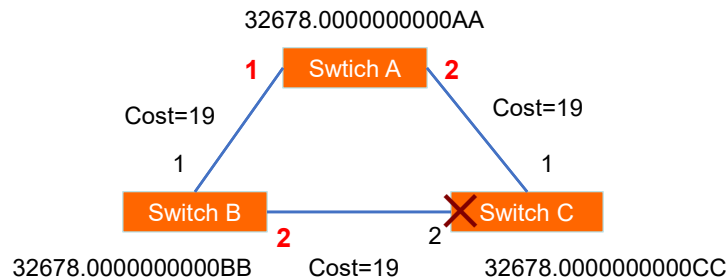
49

Blocking a port

- Any port that is not elected as either a Root Port, nor a Designated Port is put into the **Blocking State**.
- This step effectively breaks the loop and completes the Spanning Tree.

50

Designated Ports on each segment (802.1d)



- Port 2 in Switch C is then put into the **Blocking State** because it is **neither a Root Port nor a Designated Port**



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

51

Four-Step STP Decision Sequence

- When creating a loop-free topology, STP always uses the same four-step decision sequence:

Four-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

- Bridges use Configuration BPDUs during this four-step process.
 - There is another type of BPDU known as Topology Change Notification (TCN) BPDU.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

52

Four-Step STP Decision Sequence

BPDU key concepts:

- Bridges save a copy of only the best BPDU seen on every port.
- When making this evaluation, it considers all of the BPDUs received on the port, as well as the BPDU that would be sent on that port.
- As every BPDU arrives, it is checked against this four-step sequence to see if it is more attractive (lower in value) than the existing BPDU saved for that port.
- Only the lowest value BPDU is saved.
- Bridges send configuration BPDUs until a more attractive BPDU is received.
- Okay, lets see how this is used...

53

Three Steps of Initial STP Convergence

- The STP algorithm uses three simple steps to converge on a loop-free topology.
- Switches go through three steps for their initial convergence:

STP Convergence

Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

- All STP decisions are based on a the following predetermined sequence:

Four-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

54

Three Steps of Initial STP Convergence

STP Convergence

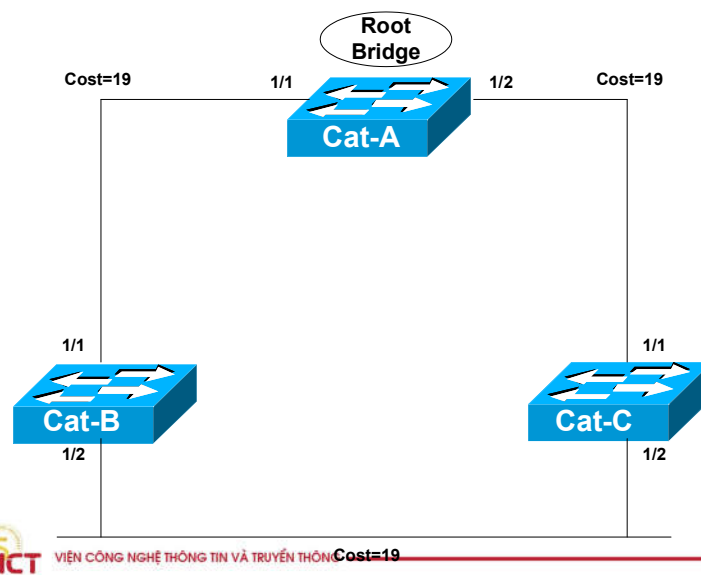
Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

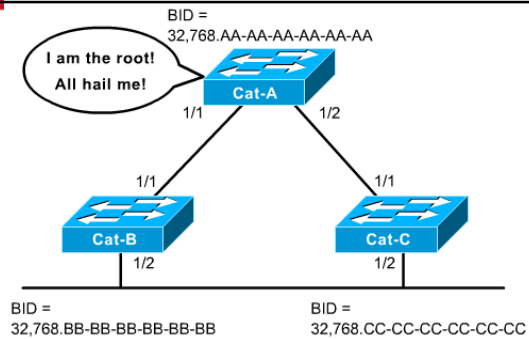
55

Step 1 Elect one Root Bridge



56

Step 1 Elect one Root Bridge



- When the network first starts, all bridges are announcing a chaotic mix of BPDUs.
- All bridges immediately begin applying the four-step sequence decision process.
- Switches need to elect a single Root Bridge.
- Switch with the **lowest BID** wins!
- Note: Many texts refer to the term “highest priority” which is the “lowest” BID value.
- This is known as the “Root War.”

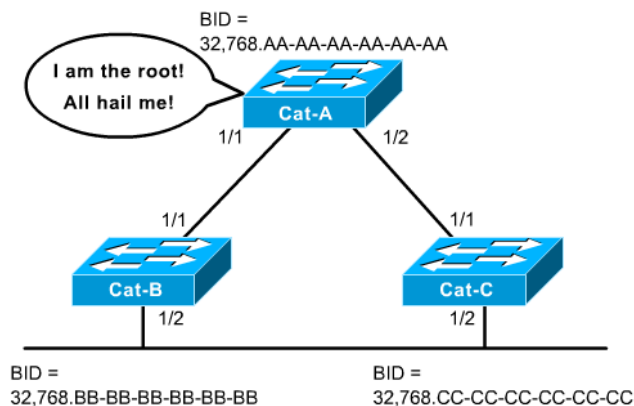


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

57

Step 1 Elect one Root Bridge

Cat-A has the lowest Bridge MAC Address, so it wins the Root War!



All 3 switches have the same default Bridge Priority value of 32,768



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

58

Step 1 Elect one Root Bridge

BPDUs

802.3 Header

Destination: 01:80:C2:00:00:00 Mcast 802.1d Bridge group
Source: 00:D0:C0:F5:18:D1
LLC Length: 38

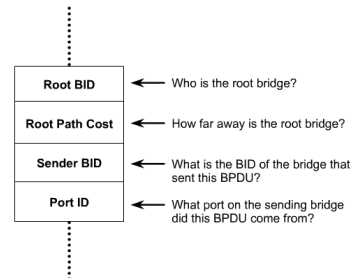
802.2 Logical Link Control (LLC) Header

Dest. SAP: 0x42 802.1 Bridge Spanning Tree
Source SAP: 0x42 802.1 Bridge Spanning Tree
Command: 0x03 Unnumbered Information

802.1 - Bridge Spanning Tree

Protocol Identifier: 0
Protocol Version ID: 0
Message Type: 0 Configuration Message
Flags: %00000000
Root Priority/ID: 0x8000/ 00:D0:C0:F5:18:C0
Cost of Path To Root: 0x00000000 (0)
Bridge Priority/ID: 0x8000/ 00:D0:C0:F5:18:C0
Port Priority/ID: 0x80/ 0x1D
Message Age: 0/256 seconds (exactly 0 seconds)
Maximum Age: 5120/256 seconds (exactly 20 seconds)
Hello Time: 512/256 seconds (exactly 2 seconds)
Forward Delay: 3840/256 seconds (exactly 15 seconds)

Its all done with BPDUs!

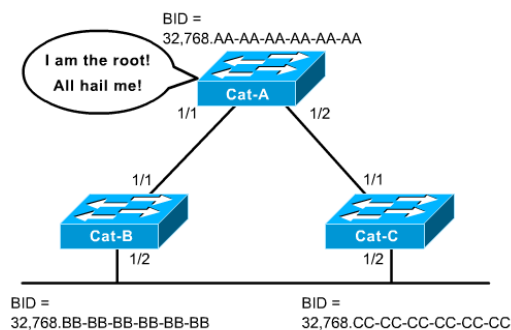


Configuration BPDUs are sent every 2 seconds by default.

SOICT VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

59

Step 1 Elect one Root Bridge



- At the beginning, all bridges assume they are the center of the universe and declare themselves as the Root Bridge, by placing its own BID in the Root BID field of the BPDU.
- Once all of the switches see that Cat-A has the lowest BID, they are all in agreement that Cat-A is the Root Bridge.

SOICT VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

60

Three Steps of Initial STP Convergence

STP Convergence

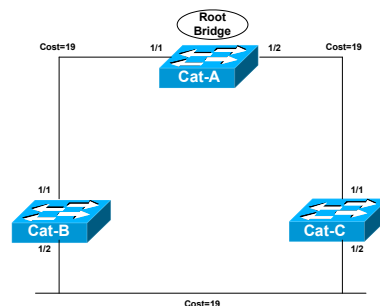
Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

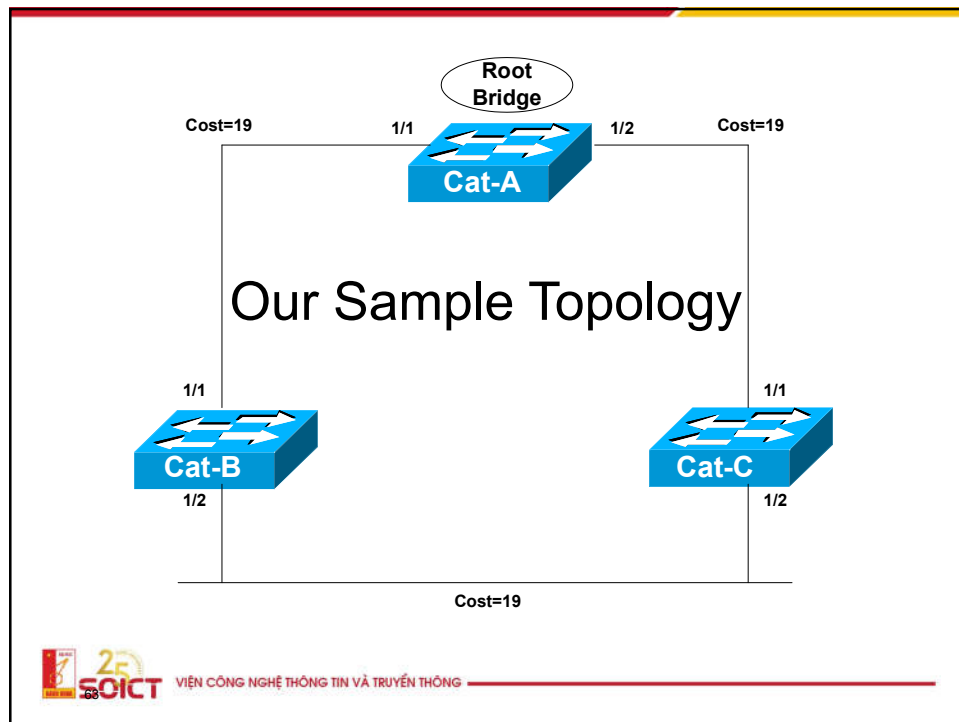
61

Step 2 Elect Root Ports

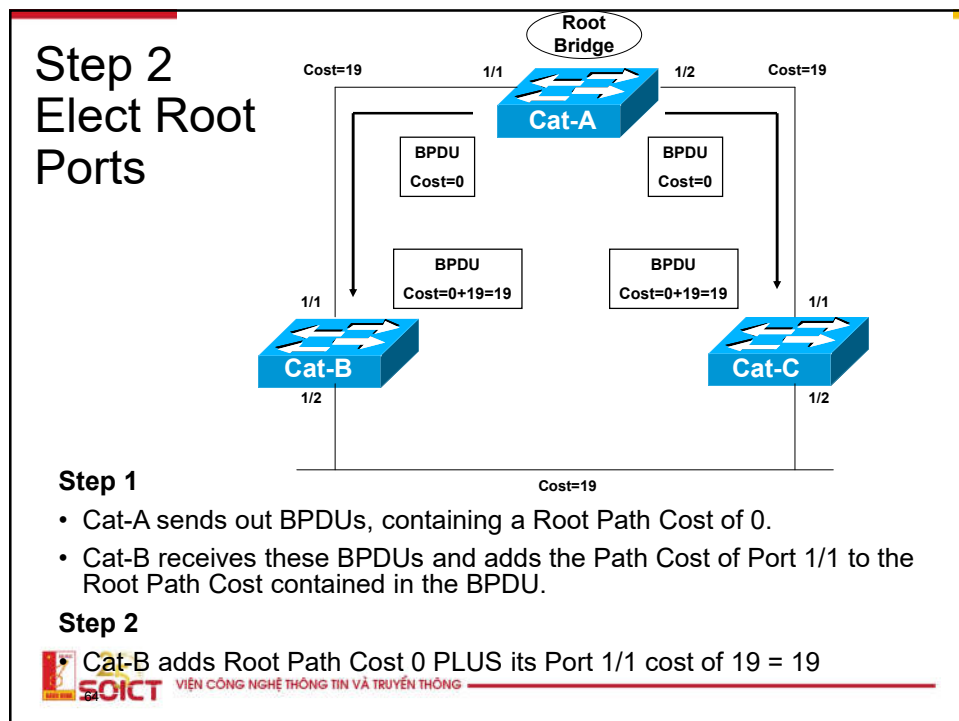


- Now that the Root War has been won, switches move on to selecting **Root Ports**.
- A bridge's **Root Port** is the *port closest to the Root Bridge*.
- Bridges use the **cost** to determine closeness.
- **Every non-Root Bridge will select one Root Port!**
- Specifically, bridges track the **Root Path Cost**, the cumulative cost of all links to the Root Bridge.

62

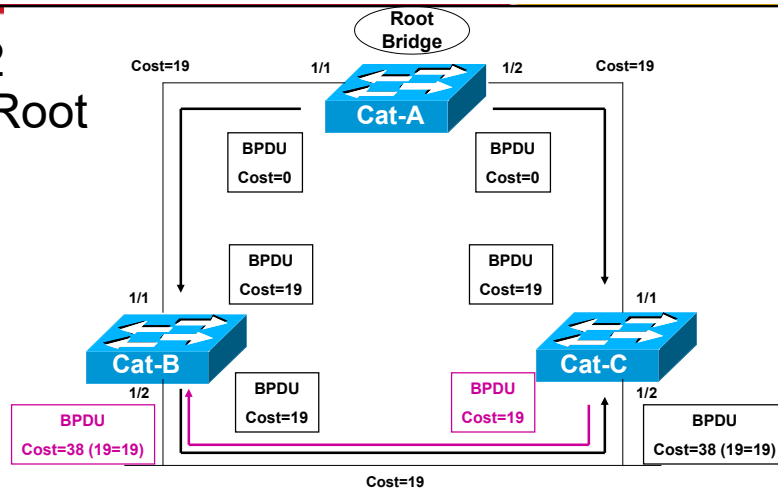


63



64

Step 2 Elect Root Ports



Step 3

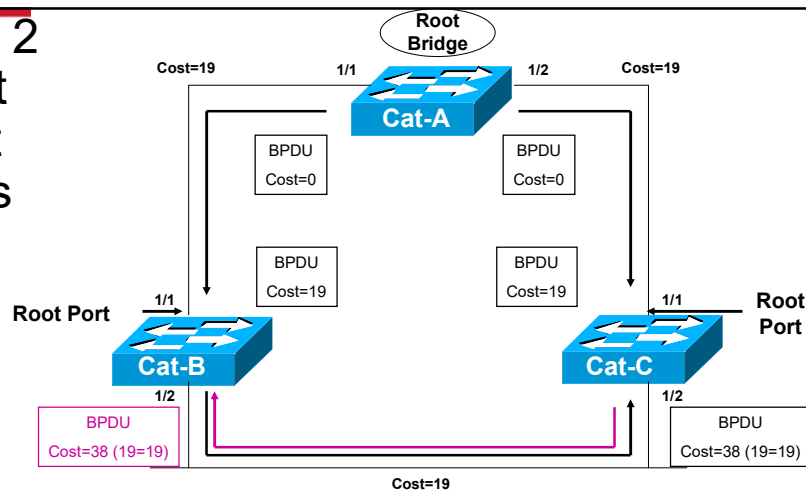
- Cat-B uses this value of 19 internally and sends BPDUs with a Root Path Cost of 19 out Port 1/2.

Step 4

- Cat-C receives the BPDU from Cat-B, and increased the Root Path Cost to 38 (19+19). (Same with Cat-C sending to Cat-B.)

65

Step 2 Elect Root Ports



Step 5

- Cat-B calculates that it can reach the Root Bridge at a cost of 19 via Port 1/1 as opposed to a cost of 38 via Port 1/2.
- Port 1/1 becomes the Root Port for Cat-B, the port closest to the Root Bridge.
- Cat-C goes through a similar calculation. Note: Both Cat-B:1/2 and Cat-C:1/2 save the best BPDU of 19 (its own).

66

Three Steps of Initial STP Convergence

STP Convergence

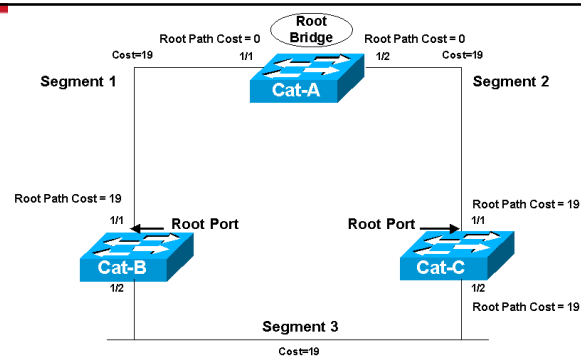
Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

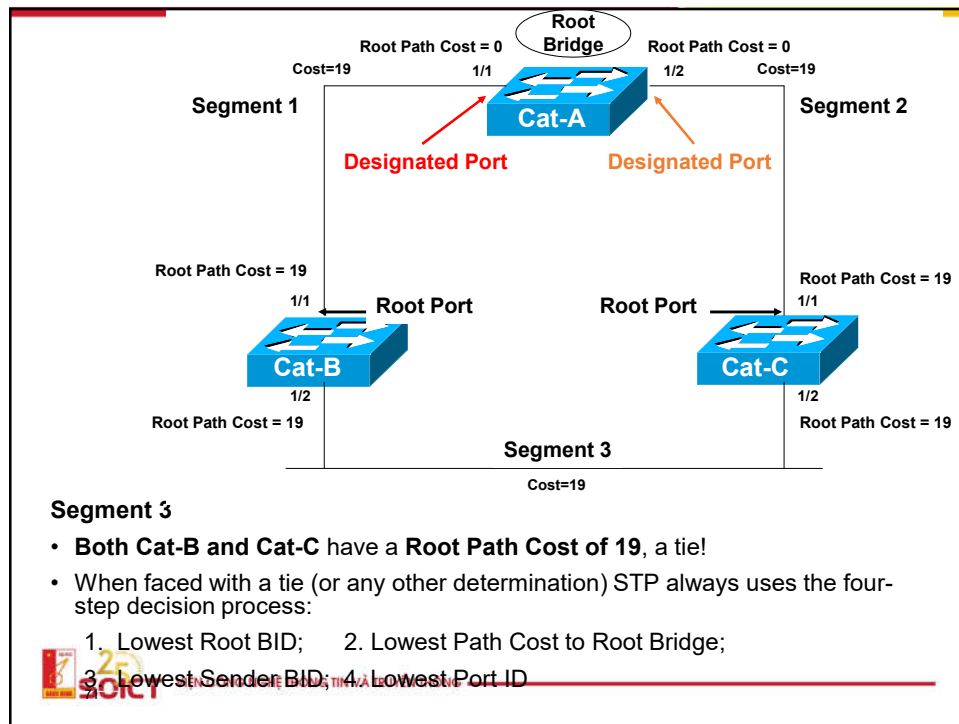
67

Step 3 Elect Designated Ports

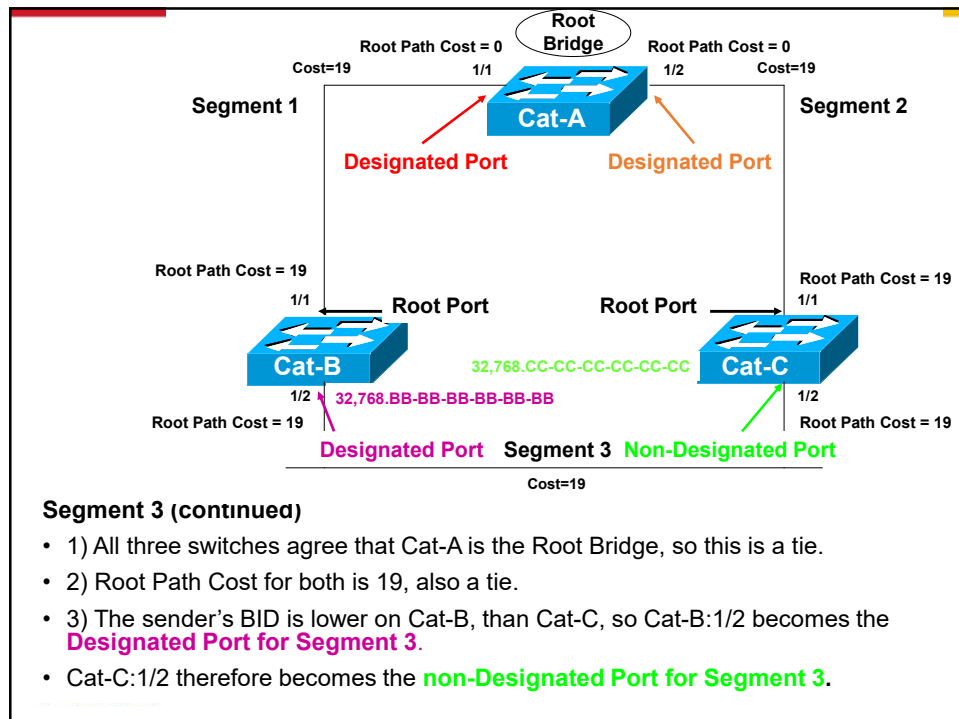


- The loop prevention part of STP becomes evident during this step, electing designated ports.
- A **Designated Port** functions as *the single bridge port that both sends and receives traffic to and from that segment and the Root Bridge*.
- **Each segment in a bridged network has one Designated Port, chosen based on cumulative Root Path Cost to the Root Bridge.**
- The switch containing the Designated Port is referred to as the **Designated Bridge** for that segment.
- To locate Designated Ports, let's take a look at each segment.

68



71



72

Spanning Tree Protocol States

- Disabled
 - Port is shut down
- Blocking
 - Not forwarding frames
 - Receiving BPDUs
- Listening
 - Not forwarding frames
 - Sending and receiving BPDUs



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

73

Spanning Tree Protocol States

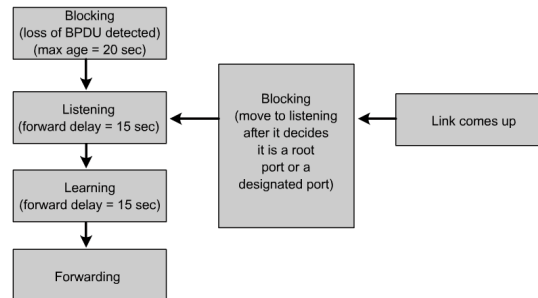
- Learning
 - Not forwarding frames
 - Sending and receiving BPDUs
 - Learning new MAC addresses
- Forwarding
 - Forwarding frames
 - Sending and receiving BPDUs
 - Learning new MAC addresses



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

74

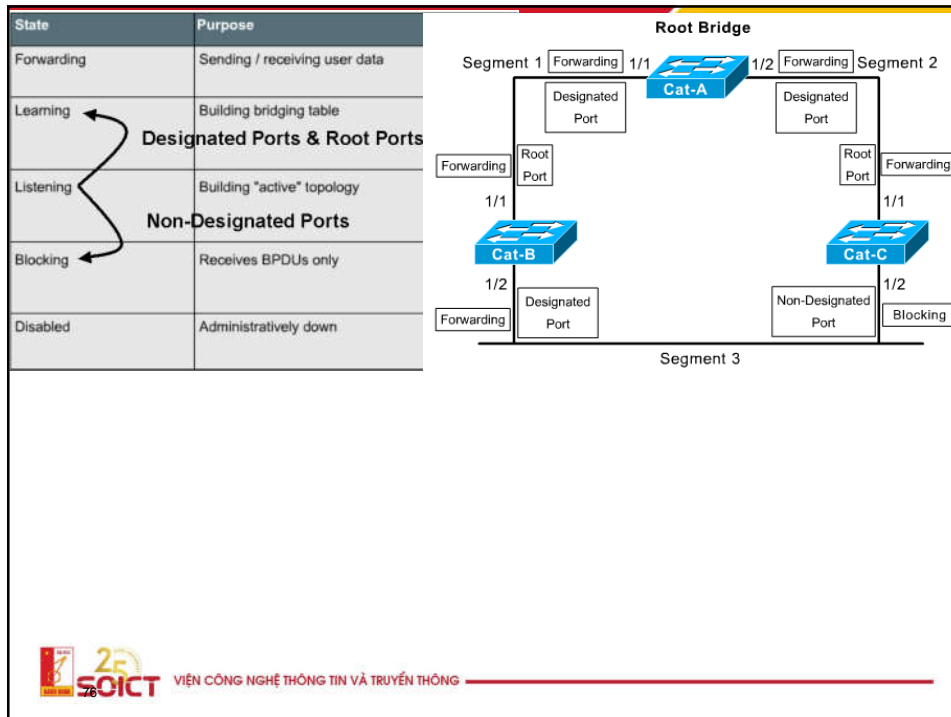
Stages of spanning-tree port states



- Time is required for (BPDU) protocol information to propagate throughout a switched network.
- Topology changes in one part of a network are not instantly known in other parts of the network.
- There is propagation delay.
- A switch should not change a port state from inactive (Blocking) to active (Forwarding) immediately, as this may cause data loops.
- Each port on a switch that is using the Spanning-Tree Protocol has one of five states.

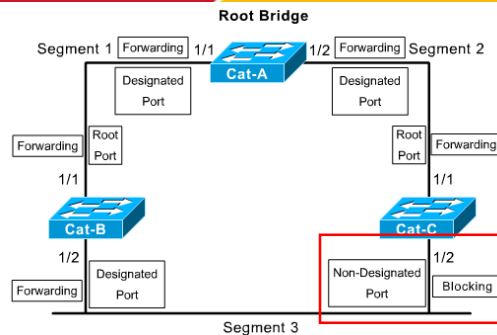
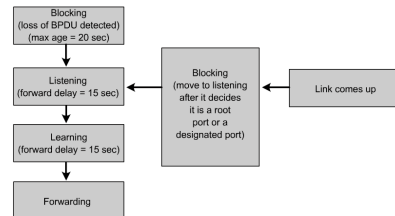


75



76

STP Port States



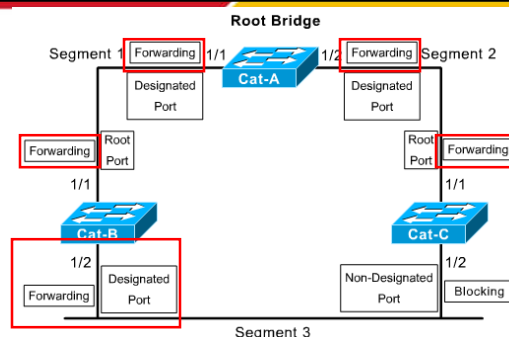
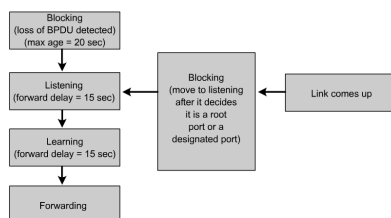
- In the **blocking state**, ports can only receive BPDUs.
 - Data frames are discarded and no addresses can be learned.
 - It may take up to 20 seconds to change from this state.
- Ports go from the blocked state to the **listening state**.
 - Switch **determines if there are any other paths to the root bridge**.
 - The **path that is not the least cost path to the root bridge goes back to the blocked state**.
 - The listening period is called the forward delay and lasts for 15 seconds.
 - In the listening state, user data is not being forwarded and MAC addresses are not being learned.
 - BPDUs are still processed.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

77

STP Port States



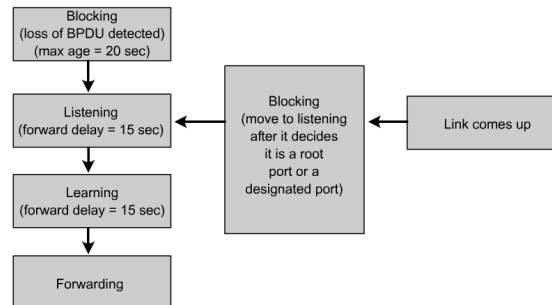
- Ports transition from the listening to the **learning state**.
 - In this state **user data is not forwarded, but MAC addresses are learned** from any traffic that is seen.
 - The learning state lasts for 15 seconds and is also called the forward delay.
 - BPDUs are still processed.
- A port goes from the learning state to the **forwarding state**.
 - In this state **user data is forwarded and MAC addresses continue to be learned**.
 - BPDUs are still processed.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

78

STP Timers

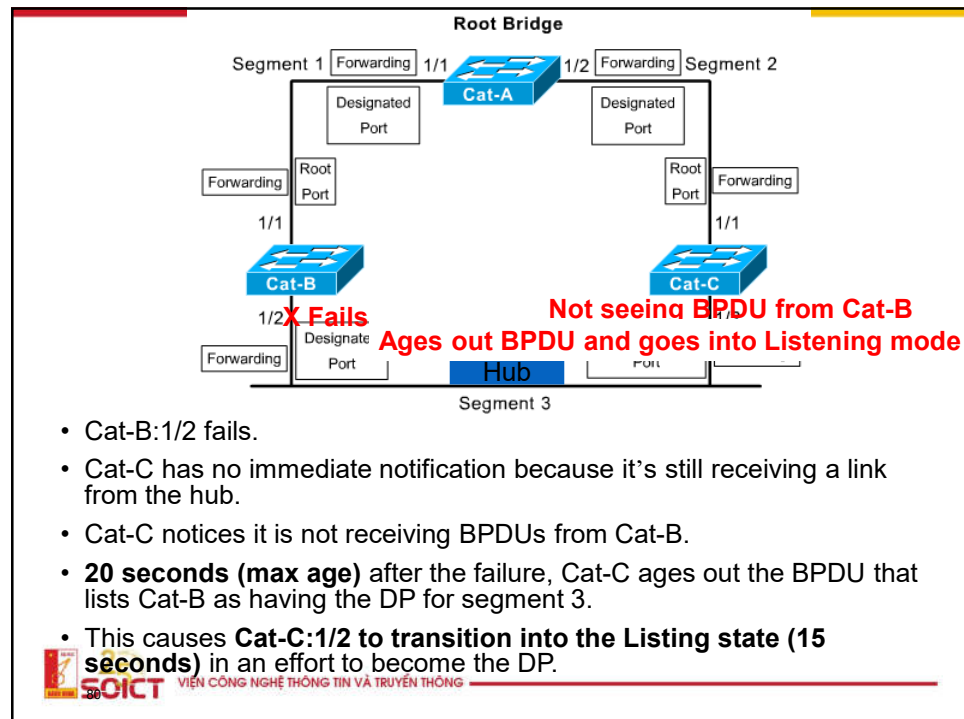


- Some details have been left out, such as timers, STP FSM, etc.
- The time values given for each state are the default values.
- These values have been calculated on an assumption that there will be a maximum of seven switches in any branch of the spanning tree from the root bridge.
- These are discussed in CCNP 3 Multilayer Switching.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

79

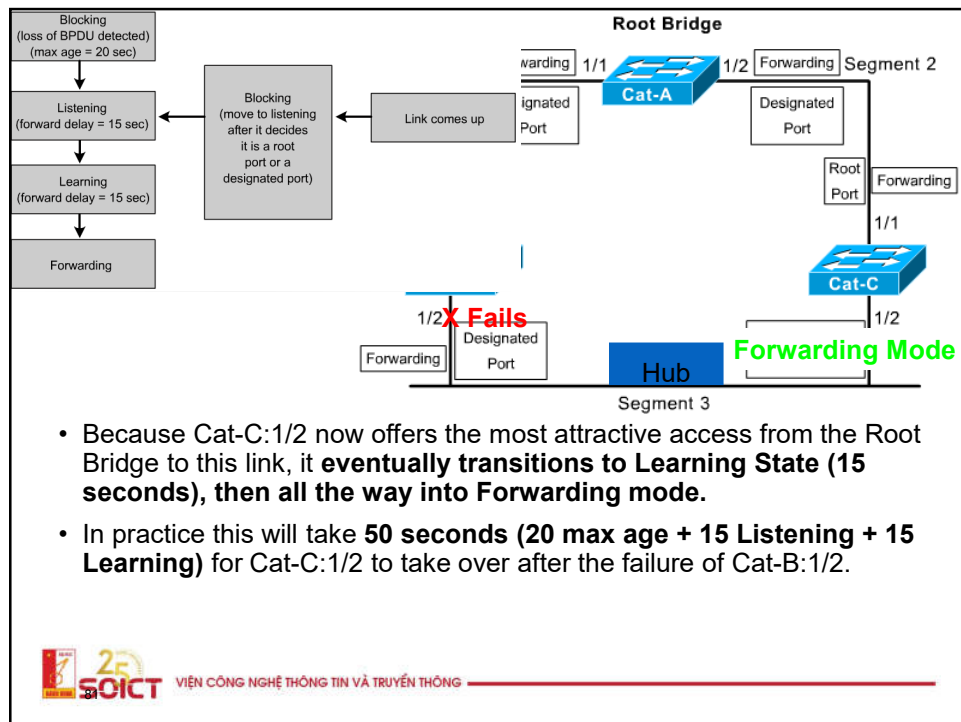


- Cat-B:1/2 fails.
- Cat-C has no immediate notification because it's still receiving a link from the hub.
- Cat-C notices it is not receiving BPDUs from Cat-B.
- **20 seconds (max age)** after the failure, Cat-C ages out the BPDU that lists Cat-B as having the DP for segment 3.
- This causes **Cat-C:1/2 to transition into the Listening state (15 seconds)** in an effort to become the DP.



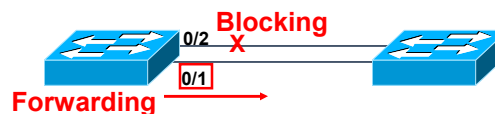
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

80



81

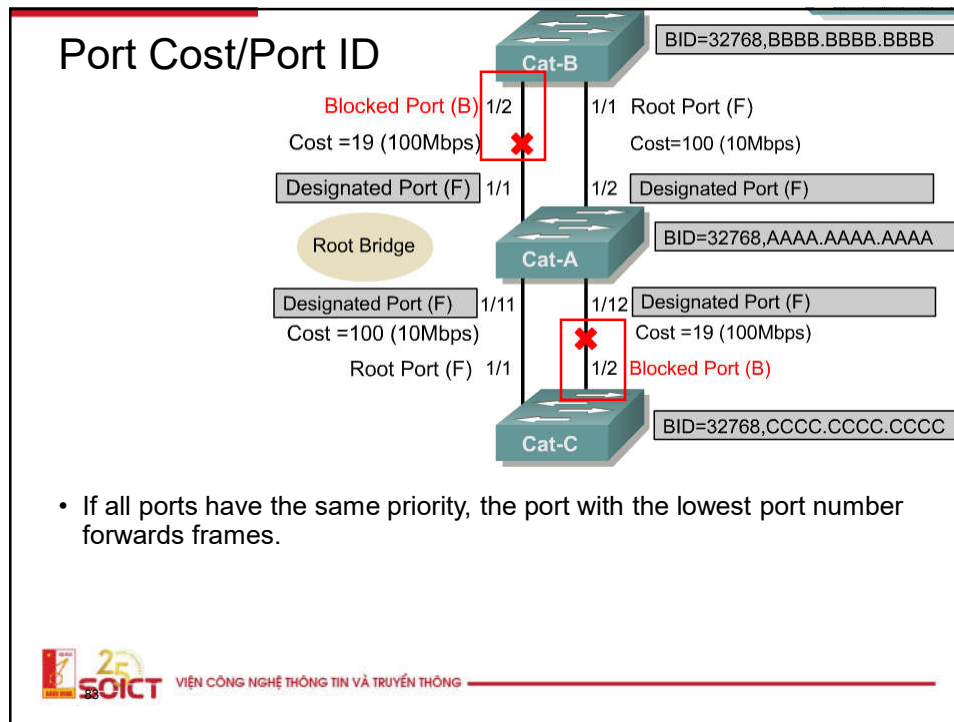
Port Cost/Port ID



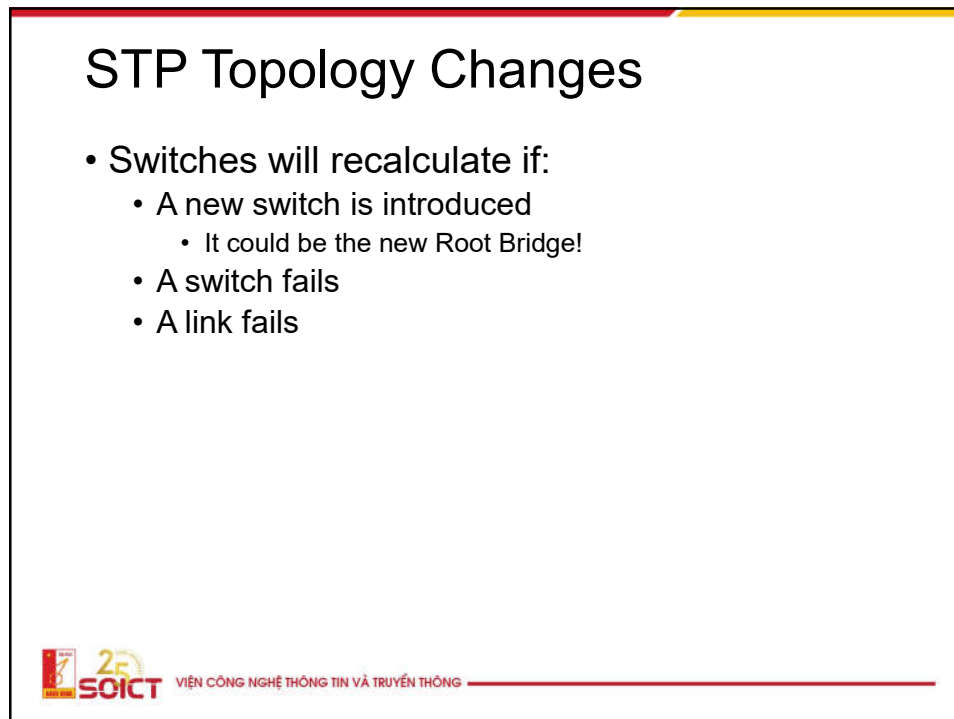
Assume path cost and port priorities are default (32). Port ID used in this case. Port 0/1 would forward because it's the lower than Port 0/2.

- If the path cost and bridge IDs are equal (as in the case of parallel links), the switch goes to the port priority as a tiebreaker.
- Lowest port priority wins (all ports set to 32).
- You can set the priority from 0 – 63.
- If all ports have the same priority, the port with the lowest port number forwards frames.

82



83



84

STP Convergence Recap

- Recall that switches go through three steps for their initial convergence:

STP Convergence

Step 1 Elect one Root Bridge

Step 2 Elect Root Ports

Step 3 Elect Designated Ports

- Also, all STP decisions are based on a the following predetermined sequence:

Four-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

85

Attack STP

- Taking over the root bridge
 - The BPDU sent out announces that the attacker's system has a lower bridge priority.
- DoS using a flood of config BPDUs
 - The attacker broadcasts out an STP configuration/topology change BPDU in an attempt to force an STP recalculation
- Defence:
 - Root guard
 - BPDU guard
 - BPDU filtering



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

86

86



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Security in Network Layer

87

History of Network Protocols

- Infrastructure protocols were designed when security concerns were almost non-existing
- Trust was assumed
- Recall early history of Internet
- Connected major universities with government labs ... in fact, commercial use was at first prohibited
- Main goal for DARPA Internet Program
 - Share large service machines on ARPANET
- Many protocol specifications focused **only** on operational aspects ... overlooked security implications ...



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Hey, we're all friends!!

88

88

Vulnerabilities in Protocols

- During last twenty years, many vulnerabilities have been identified in TCP/IP stacks of most systems
- Protocol weaknesses due to:
 - Design of Protocol and
 - Daily operation and configuration



TCP/IP Suite Problems



• Problems

Can you think of some security problems with design of TCP/IP suite?

- IP addresses are not validated
- Hosts can not be authenticated
- Trivial to spoof packets as coming from a trusted host
- Remote utilities assumes trust between hosts
- Encryption not typically used, and not for headers



Protocol Attacks



- What type of network attacks are common in today's Internet?
 - Denial of Service (DoS) and Distributed Denial of Service (DdoS)
 - Man in the Middle Attack
 - Eavesdropping network traffic
 - Application Security Attacks
 - Web Based Attacks
 - SQL Injection
 - Crosssite Scripting
 - Driveby Malware



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

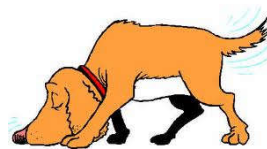
91

91

Protocol Attack Techniques

• Sniffing Traffic

- **Eavesdropping** on a network
- “Wiretap” programs ... [name one program](#)
- **Wireless networks**
 - Easier to see all the traffic, put NIC into Monitor mode
- **Wired networks**
 - NIC needs to be in promiscuous mode
 - Must do ARP spoofing or other attack to get all packets forwarded to you
- Can only see traffic from subnet you are tapped into



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

92

92

Which Protocols

- **TCP/IP Protocol Suite**
 - Application Layer - DNS
 - Transport Layer - UDP/TCP
 - **Network Layer - IP/ICMP/BGP**
 - Data Link Layer - ARP




VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

93

93

TCP/IP Problems

- **Steve Bellovin**


Who does this look like?

 - AT&T Bell labs researcher
 - One of the first to publicize problems in TCP/IP protocols
 - Wrote his original paper in 1989
 - Documented many problems
 - Some problems no longer relevant

Updated Paper - 2004

<https://www.cs.columbia.edu/~smb/papers/badesp.pdf>



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

94

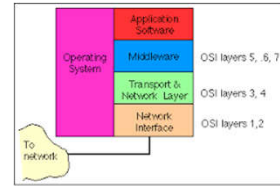
94

94

Problems Summary

• Network layer

- IP Spoofing attacks
 - TCP Sequence numbers not random → Can be predicted
- Trusted Hosts
 - Used remote Linux utilities to violate trust
 - Hardly ever used these days .. we won't cover it
- ICMP Messages
 - Used them to perform DoS, routing re-direction
- Routing Protocols
 - RIP, BGP have authentication problems



IP Source Routing Abuse

• Routing Information Protocol (RIP)

- Used to propagate routing information on local networks
- Routers need to exchange information using routing protocols
- Typically will exchange information every so many seconds
- IP Source routing feature
 - Allows source machine to specify path packet will take through network



Internet Protocol

- Connectionless
 - Unreliable
 - Best effort
- Specify Options
 - Source Route

Version	Header Length
Type of Service	
Total Length	
Identification	
Flags	Fragment Offset
Time to Live	
Protocol	
Header Checksum	
Source Address of Originating Host	
Destination Address of Target Host	
Options	
Padding	
IP Data	



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

97

97

IP Source Routing Abuse

- Example of MITM (Man-In-The-Middle) Attacks
 - Send bogus routing information trying to impersonate a particular host
 - Want packets to be sent to the attacker machine
 - Attacker can intercept packets and gain passwords, credit card numbers or other sensitive information



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

98

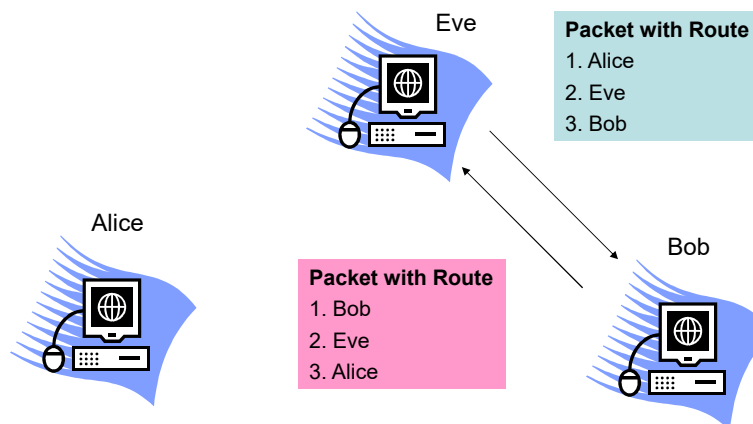
98

Steps in Source Route Attack

Attack Steps (three hosts, Eve, Alice and Bob)

1. Eve generates packets with fake source route
2. Packets claim to come from Alice
3. Source route includes Eve's IP
 Eve looks like a router between Alice and Bob
 Bob is the destination
4. Routers between Eve and Bob read source route and deliver packets to Bob via Eve

Steps in Source Route Attack



Steps in Source Routing Abuse

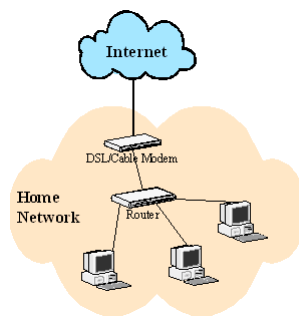
Attack Steps

1. Bob responds by sending packets through Eve to Alice
2. Eve never forwards packets to Alice, doesn't need to even do a DoS on Alice

Comment

This attack doesn't work across the Internet
 Most gateways block Source Routed packets
 Yet, not blocked on internal networks
 Insiders can get away with this type of attack

Other Routing Vulnerabilities



ICMP

- What is ICMP protocol used for?
 - Internet Control Message Protocol (ICMP)
 - Mostly ... Used to send error messages
 - Requested service is not available, or that host or router could not be reached

http://en.wikipedia.org/wiki/Internet_Control_Message_Protocol



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

103

103

ICMP Messages

- 0 Echo Reply
- 3 Destination Unreachable
- 4 Source Quench
- 5 Redirect
- 8 Echo Request
- 11 Time Exceeded
- 12 Parameter Problem
- 13 Timestamp
- 14 Timestamp Reply
- 15 Information Request
- 16 Information Reply



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

104

104

ICMP Messages



- **Destination Unreachable message**
 - ICMP message generated by host or its inbound gateway to inform client
 - Destination is unreachable for some reason
 - Destination Unreachable message may be generated as a result of
 - TCP, UDP or another ICMP transmission



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

105

105

ICMP Messages



- **The Source Quench,**
 - Message requests sender to decrease traffic rate of messages to a router or host
 - Message may be generated if router or host does not have sufficient buffer space to process the request, or
 - May occur if router or host's buffer is approaching its limit



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

106

106

ICMP Attacks

- Attacks Reported in Bellovin Paper

- ICMP Redirect message

- Used by gateways to advise hosts of better routes, Some limitations on how its used

Must be

- Tied to existing connection
- Must only be sent from first gateway to originating host



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

107

107

107

ICMP Attacks

- Attacks Reported in Bellovin Paper

- ICMP Redirect message

1. Host C sends a Syn packet to S via A, a router
2. Before packet can get there, Host X, our attacker, sends an ICMP redirect for Host X to C spoofing the address A
3. C now redirects packets to X
4. X forwards packets to S to avoid suspicion



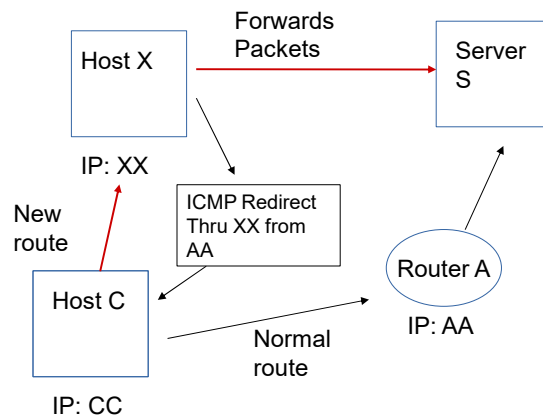
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

108

108

108

ICMP Redirect



ICMP Attacks

• ICMP Current Attacks

– ICMP Redirect

- Still a threat if not ignored
- Current recommendation is to turn off redirects on CISCO routers
- Routing protocol takes care of best paths, hosts should ignore ICMP redirect messages

ICMP Attacks



- **More Current Attacks**
- Other ways ICMP is used to compromise
 - **ICMP Source Quench**
 - Slows down transmission of traffic essentially performing a partial DoS on itself
 - **ICMP DoS**
 - Attacker could use either ICMP **Time exceeded** or **Destination unreachable** messages. Both messages can cause host to drop a connection
 - Attacker can simply forge one of these ICMP messages, and send it to one or both communicating hosts ... their connection will then be broken



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

111

111

ICMP Attacks



More Attacks

- **SMURF Attack**
- Generate ping stream (ICMP echo request) to
 - Network broadcast address
 - Spoofed source IP set to victim host
- Every host on ping target network will generate ping reply (ICMP echo reply)
- Amplified ping reply stream can easily overwhelm victim's network connection

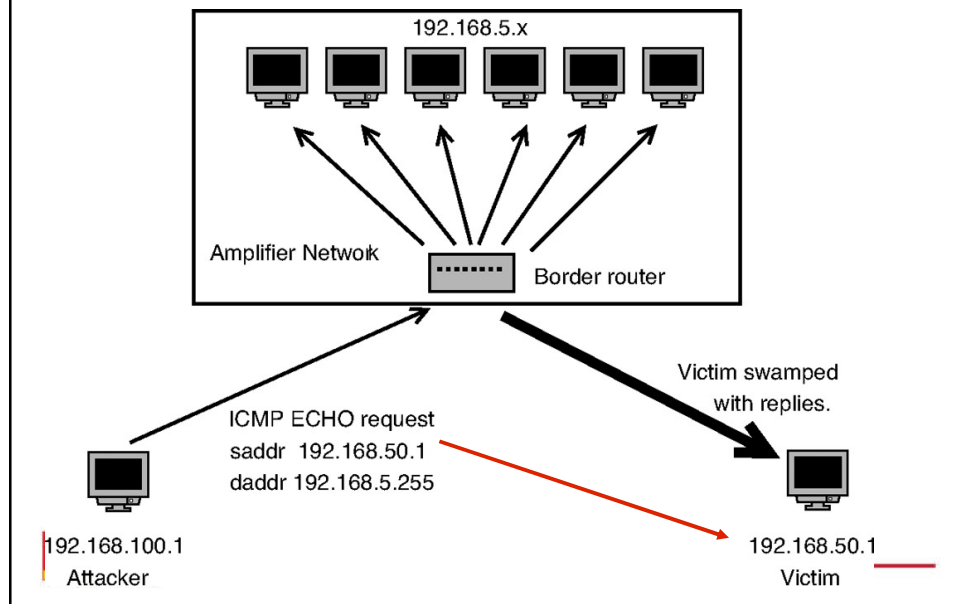


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

112

112

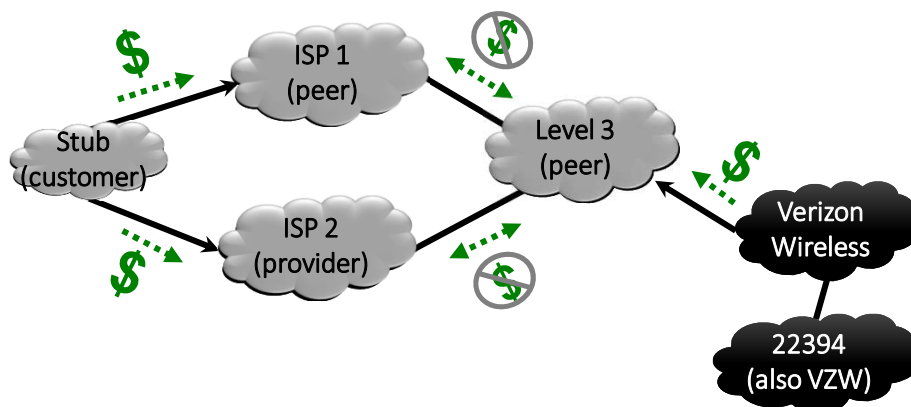
Smurf Attack



113

BGP: The Internet's Routing Protocol

A simple model of AS-level business relationships.



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

114

BGP-4

- **BGP** = Border Gateway Protocol
- Is a **Policy-Based** routing protocol
- Is the **EGP** of today's global Internet
- Relatively simple protocol, but configuration is complex and the entire world can see, and be impacted by, your mistakes.

- **1989 : BGP-1 [RFC 1105]**
 - Replacement for EGP (1984, RFC 904)
- **1990 : BGP-2 [RFC 1163]**
- **1991 : BGP-3 [RFC 1267]**
- **1995 : BGP-4 [RFC 1771]**
 - Support for Classless Interdomain Routing (CIDR)

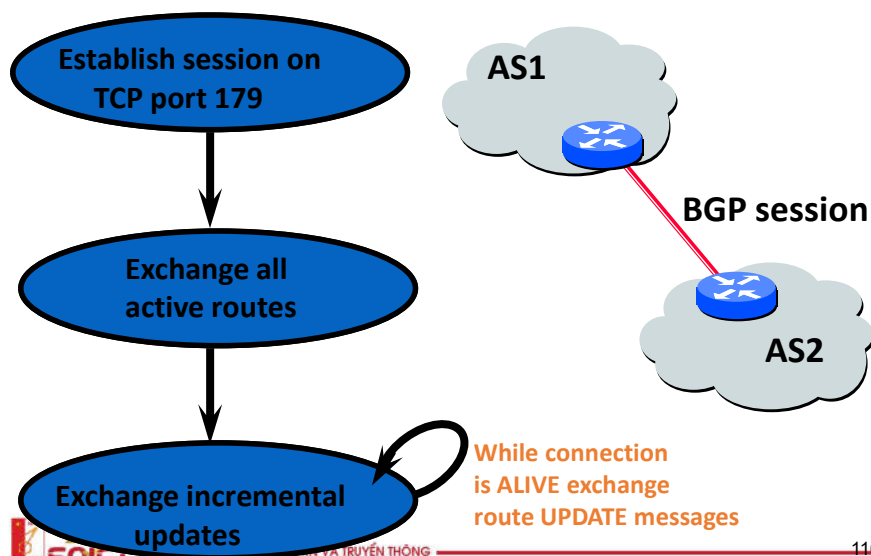


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

115

115

BGP Operations (Simplified)



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

116

116

Interconnecting BGP Peers

- BGP uses TCP to connect peers
- Advantages:
 - Simplifies BGP
 - No need for periodic refresh - routes are valid until withdrawn, or the connection is lost
 - Incremental updates
- Disadvantages
 - Congestion control on a routing protocol?
 - Inherits TCP vulnerabilities!
 - Poor interaction during high load



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

117

117

Four Types of BGP Messages

- **Open** : Establish a peering session.
- **Keep Alive** : Handshake at regular intervals.
- **Notification** : Shuts down a peering session.
- **Update** : Announcing new routes or withdrawing previously announced routes.

announcement
=
prefix + attributes values



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

118

118

Policy with BGP

- BGP provides capability for enforcing various policies
- Policies are **not** part of BGP: they are provided to BGP as configuration information
- BGP enforces policies by **choosing paths from multiple alternatives** and **controlling advertisement to other AS's**
- Import policy
 - What to do with routes learned from neighbors?
 - Selecting best path
- Export policy
 - What routes to announce to neighbors?
 - Depends on relationship with neighbor

119

Examples of BGP Policies

- A multi-homed AS refuses to act as transit
 - Limit path advertisement
- A multi-homed AS can become transit for some AS's
 - Only advertise paths to some AS's
 - Eg: A Tier-2 provider multi-homed to Tier-1 providers
- An AS can favor or disfavor certain AS's for traffic transit from itself

120

Export Policy

- An AS exports only best paths to its neighbors
 - Guarantees that once the route is announced the AS is willing to transit traffic on that route
- To Customers
 - Announce all routes learned from peers, providers and customers, and self-origin routes
- To Providers
 - Announce routes learned from customers and self-origin routes
- To Peers
 - Announce routes learned from customers and self-origin routes

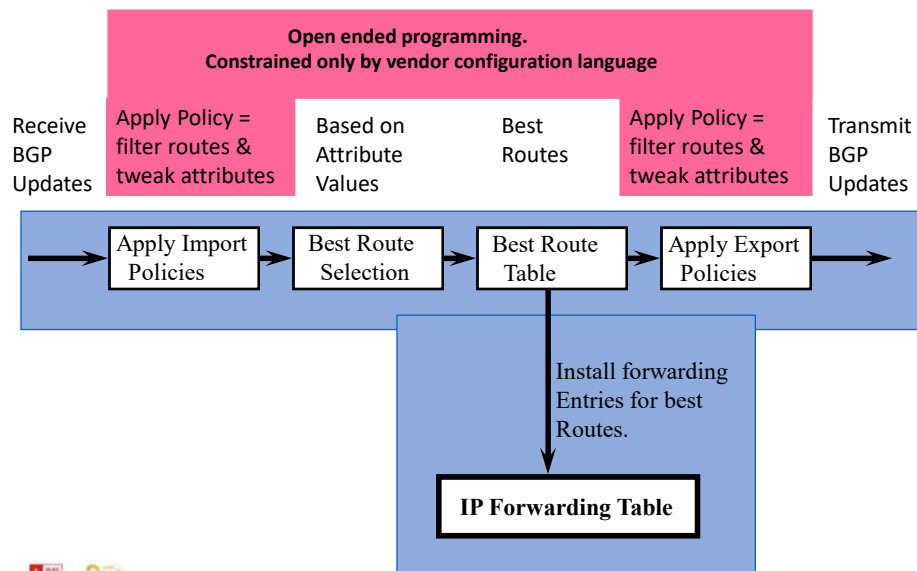


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

121

121

BGP Route Processing



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

122

122

BGP UPDATE Message

- List of withdrawn routes
- Network layer reachability information
 - List of reachable prefixes
- Path attributes
 - Origin
 - Path
 - Metrics
- All prefixes advertised in message have same path attributes



Path Selection Criteria

- Information based on path attributes
- Attributes + external (policy) information
- Examples:
 - Hop count
 - Policy considerations
 - Preference for AS
 - Presence or absence of certain AS
 - Path origin
 - Link dynamics



Important BGP Attributes

- Local Preference
- AS-Path
- MED
- Next hop



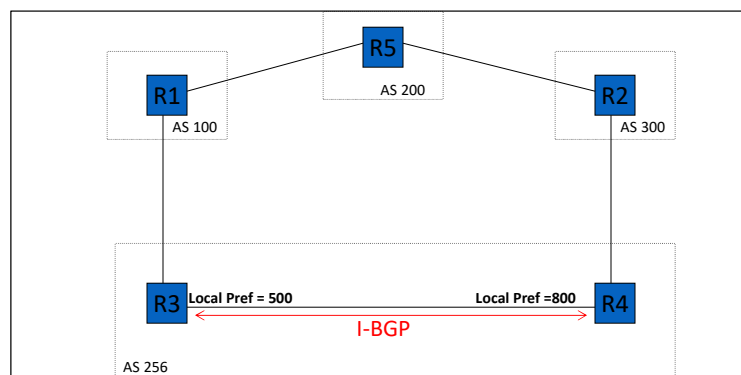
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TÀI CHÍNH - 15744 - Fall 2004 Lecture 3

125

125

LOCAL PREF

- Local (within an AS) mechanism to provide relative priority among BGP routers



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TÀI CHÍNH - 15744 - Fall 2004 Lecture 3

126

126

LOCAL PREF – Common Uses

- Handle routes advertised to multi-homed transit customers
 - Should use direct connection (multihoming typically has a primary/backup arrangement)
- Peering vs. transit
 - Prefer to use peering connection, why?
- In general, customer > peer > provider
 - Use LOCAL PREF to ensure this



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TÀI CHÍNH

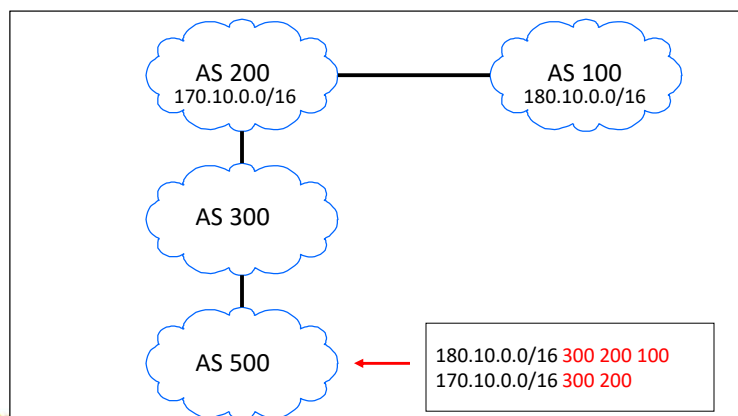
15744-Fall 2004 Lecture 3

127

127

AS_PATH

- List of traversed AS's
- Useful for loop checking and for path-based route selection (length, regexp)



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TÀI CHÍNH

15744-Fall 2004 Lecture 3

128

128

Multi-Exit Discriminator (MED)

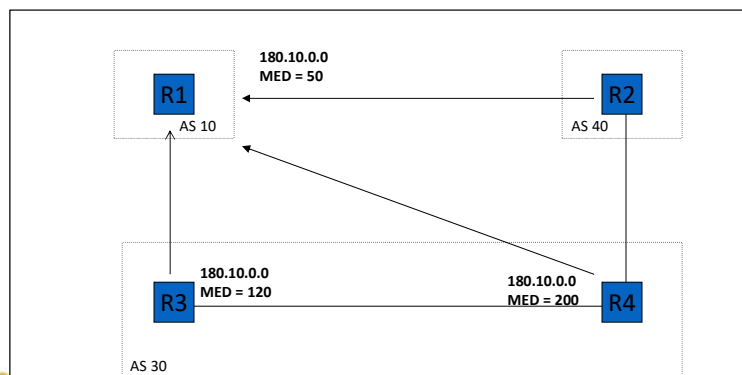
- Hint to external neighbors about the preferred path into an AS
 - Non-transitive attribute
 - Different AS choose different scales
- Used when two AS's connect to each other in more than one place



129

MED

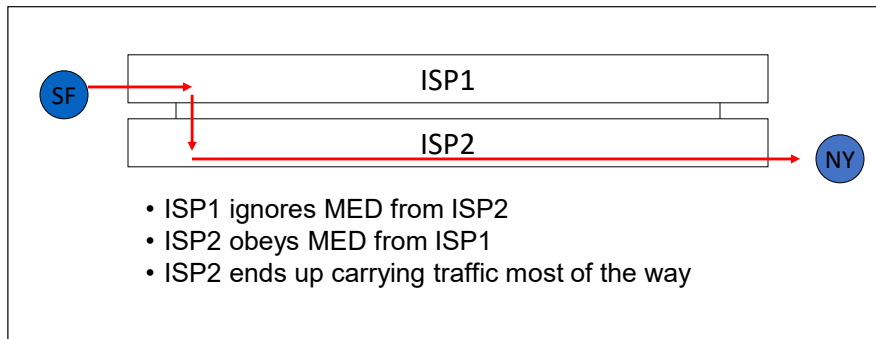
- Typically used when two ASes peer at multiple locations
- Hint to R1 to use R3 over R4 link
- Cannot compare AS40's values to AS30's



130

MED

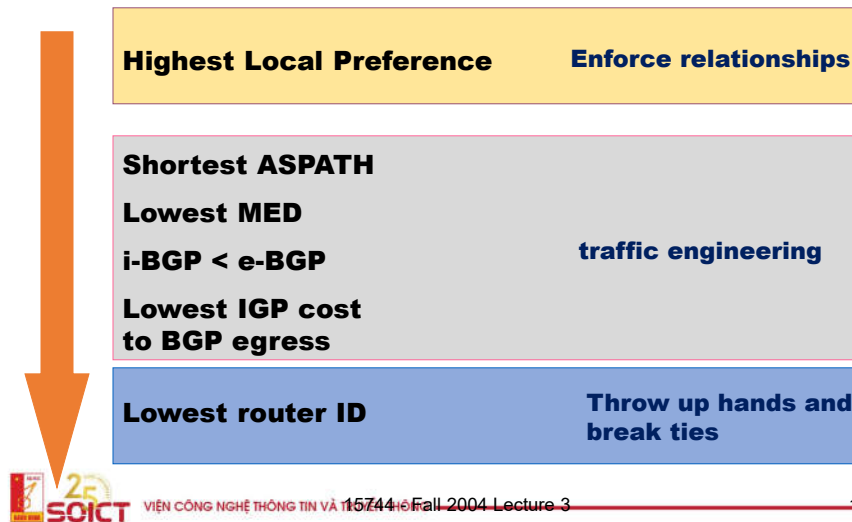
- MED is typically used in provider/subscriber scenarios
- It can lead to unfairness if used between ISP because it may force one ISP to carry more traffic:



Other Attributes

- ORIGIN
 - Source of route (IGP, EGP, other)
- NEXT_HOP
 - Address of next hop router to use
- Check out <http://www.cisco.com> for full explanation
- Question: Too many choices/ attributes how to select routes !

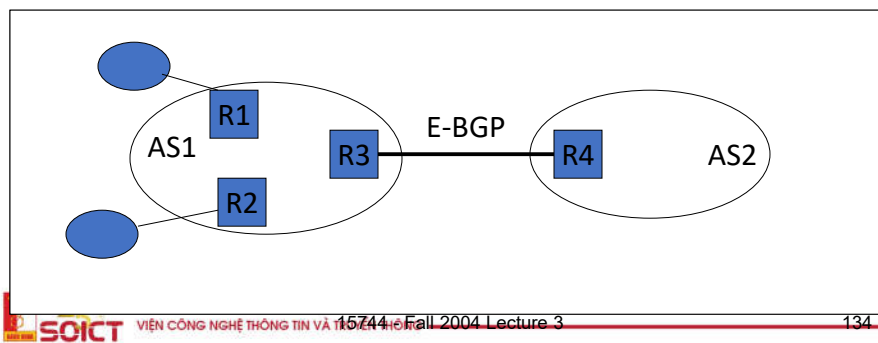
Route Selection Process



133

Internal vs. External BGP

- BGP can be used by R3 and R4 to learn routes
- How do R1 and R2 learn routes?
 - Option 1: Inject routes in IGP
 - Only works for small routing tables
 - Option 2: Use I-BGP



134

Internal BGP (I-BGP)

- Same messages as E-BGP
- Different rules about re-advertising prefixes:
 - Prefix learned from E-BGP can be advertised to I-BGP neighbor and vice-versa, but
 - Prefix learned from one I-BGP neighbor **cannot** be advertised to another I-BGP neighbor
 - Reason: no AS PATH within the same AS and thus danger of looping.



135

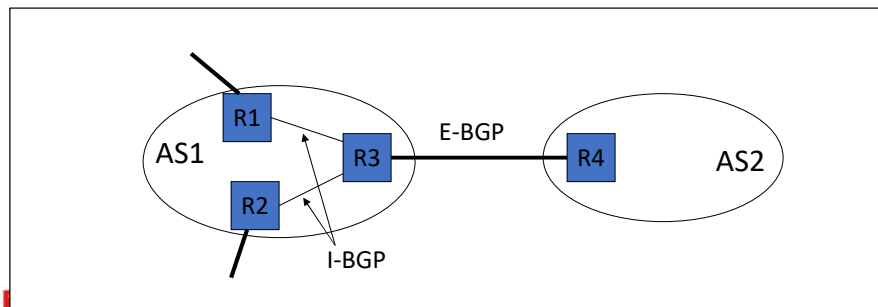
Internal BGP (I-BGP)

- R3 can tell R1 and R2 prefixes from R4
- R3 can tell R4 prefixes from R1 and R2
- R3 cannot tell R2 prefixes from R1

R2 can only find these prefixes through a *direct connection* to R1

Result: I-BGP routers must be fully connected (via TCP)!

- contrast with E-BGP sessions that map to physical links



136

BGP Session Security



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

137

137

TCP Connection Underlying BGP Session

- BGP session runs over TCP
 - TCP connection between neighboring routers
 - BGP messages sent over TCP connection
 - Makes BGP vulnerable to attacks on TCP
- Main kinds of attacks
 - Against confidentiality: eavesdropping
 - Against integrity: tampering
 - Against performance: denial-of-service
- Main defenses
 - Message authentication or encryption
 - Limiting access to physical path between routers
 - Defensive filtering to block unexpected packets



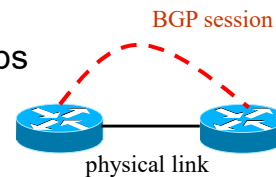
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

138

138

Attacks Against Confidentiality

- Eavesdropping
 - Monitoring the messages on the BGP session
 - ... by tapping the link(s) between the neighbors
- Reveals sensitive information
 - Inference of business relationships
 - Analysis of network stability
- Reasons why it may be hard
 - Challenging to tap the link
 - Often, eBGP session traverses just one link
 - ... and may be hard to get access to tap it
 - Encryption may obscure message contents
 - BGP neighbors may run BGP over IPSec

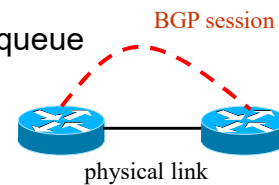


Attacking Message Integrity

- Tampering
 - Man-in-the-middle tampers with the messages
 - Insert, delete, modify, or replay messages
- Leads to incorrect BGP behavior
 - Delete: neighbor doesn't learn the new route
 - Insert/modify: neighbor learns bogus route
- Reasons why it may be hard
 - Getting in-between the two routers is hard
 - Use of authentication (signatures) or encryption
 - Spoofing TCP packets the right way is hard
 - Getting past source-address packet filters
 - Generating the right TCP sequence number

Denial-of-Service Attacks, Part 1

- Overload the link between the routers
 - To cause packet loss and delay
 - ... disrupting the performance of the BGP session
- Relatively easy to do
 - Can send traffic between end hosts as long as the packets traverse the link which you can figure out from traceroute
- Easy to defend
 - Give higher priority to BGP packets
 - E.g., by putting packets in separate queue



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

141

141

Denial-of-Service Attacks, Part 2

- Third party sends bogus TCP packets
 - FIN/RST to close the session
 - SYN flooding to overload the router
- Leads to disruptions in BGP
 - Session reset, causing transient routing changes
 - Route-flapping, which may trigger flap damping
- Reasons why it may be hard
 - Spoofing TCP packets the right way is hard
 - Difficult to send FIN/RST with the right TCP header
 - Packet filters may block the SYN flooding
 - Filter packets to BGP port from unexpected source
 - ... or destined to router from unexpected source



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

142

142

Exploiting the IP TTL Field

- BGP speakers are usually one hop apart
 - To thwart an attacker, can check that the packets carrying the BGP message have not traveled far
- IP Time-to-Live (TTL) field
 - Decrement once per hop
 - Avoids packets staying in network forever
- Generalized TTL Security Mechanism (RFC 3682)
 - Send BGP packets with initial TTL of 255
 - Receiving BGP speaker checks that TTL is 254
 - ... and flags and/or discards the packet others
- Hard for third-party to inject packets remotely



143

Validity of the routing information: Origin authentication



144

Prefix Hijacking

The diagram shows a network topology with seven Autonomous Systems (ASes) represented by clouds, numbered 1 through 7. AS 1 is at the bottom left, connected to AS 2. AS 2 is connected to AS 3. AS 3 is connected to AS 4 and AS 7. AS 4 is connected to AS 7 and AS 5. AS 7 is connected to AS 5. AS 5 is connected to AS 6. A red arrow points from AS 1 to the IP address 12.34.0.0/16, indicating a legitimate advertisement. A green arrow points from AS 6 to the same IP address, indicating a hijacked advertisement. A green arrow also points from AS 1 to AS 6, representing traffic being redirected to the hijacker.

- Consequences for the affected ASes
 - Blackhole: data traffic is discarded
 - Snooping: data traffic is inspected, and then redirected
 - Impersonation: data traffic is sent to bogus destinations

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

446

73

Hijacking is Hard to Debug

- Real origin AS doesn't see the problem
 - Picks its own route
 - Might not even learn the bogus route
- May not cause loss of connectivity
 - E.g., if the bogus AS snoops and redirects
 - ... may only cause performance degradation
- Or, loss of connectivity is isolated
 - E.g., only for sources in parts of the Internet
- Diagnosing prefix hijacking
 - Analyzing updates from many vantage points
 - Launching traceroute from many vantage points

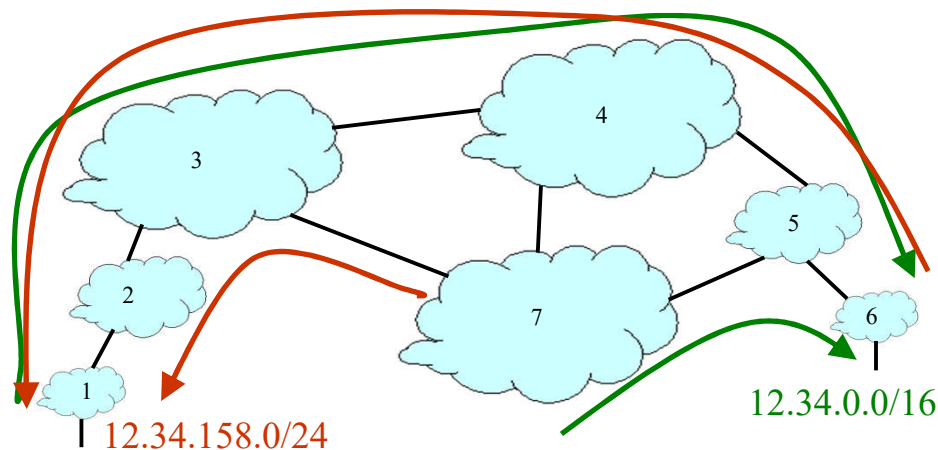


VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

147

147

Sub-Prefix Hijacking



- Originating a more-specific prefix
 - Every AS picks the bogus route for that prefix
- Traffic follows the longest matching prefix



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

148

148

How to Hijack a Prefix

- The hijacking AS has
 - Router with eBGP session(s)
 - Configured to originate the prefix
- Getting access to the router
 - Network operator makes configuration mistake
 - Disgruntled operator launches an attack
 - Outsider breaks into the router and reconfigures
- Getting other ASes to believe bogus route
 - Neighbor ASes not filtering the routes
 - ... e.g., by allowing only expected prefixes
 - But, specifying filters on *peering* links is hard



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

149

149

The February 24 YouTube Outage

- YouTube (AS 36561)
 - Web site www.youtube.com
 - Address block 208.65.152.0/22
- Pakistan Telecom (AS 17557)
 - Receives government order to block access to YouTube
 - Starts announcing 208.65.153.0/24 to PCCW (AS 3491)
 - All packets directed to YouTube get dropped on the floor
- Mistakes were made
 - AS 17557: announcing to everyone, not just customers
 - AS 3491: not filtering routes announced by AS 17557
- Lasted 100 minutes for some, 2 hours for others



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

150

150

Timeline (UTC Time)

- 18:47:45
 - First evidence of hijacked /24 route propagating in Asia
- 18:48:00
 - Several big trans-Pacific providers carrying the route
- 18:49:30
 - Bogus route fully propagated
- 20:07:25
 - YouTube starts advertising the /24 to attract traffic back
- 20:08:30
 - Many (but not all) providers are using the valid route



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml

151

151

Timeline (UTC Time)

- 20:18:43
 - YouTube starts announcing two more-specific /25 routes
- 20:19:37
 - Some more providers start using the /25 routes
- 20:50:59
 - AS 17557 starts prepending ("3491 17557 17557")
- 20:59:39
 - AS 3491 disconnects AS 17557
- 21:00:00
 - All is well, videos of cats flushing toilets are available



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml

152

152

Another Example: Spammers

- Spammers sending spam
 - Form a (bidirectional) TCP connection to a mail server
 - Send a bunch of spam e-mail
 - Disconnect and laugh all the way to the bank
- But, best not to use your real IP address
 - Relatively easy to trace back to you
- Could hijack someone's address space
 - But you might not receive all the (TCP) return traffic
 - And the legitimate owner of the address might notice
- How to evade detection
 - Hijack unused (i.e., unallocated) address block in BGP
 - Temporarily use the IP addresses to send your spam



153

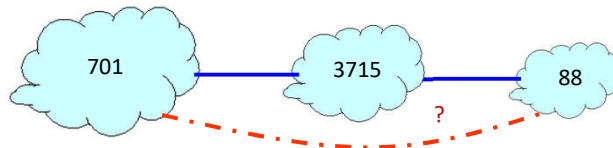
BGP AS Path



154

Bogus AS Paths

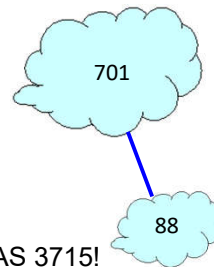
- Remove ASes from the AS path
 - E.g., turn “701 3715 88” into “701 88”
- Motivations
 - Make the AS path look shorter than it is
 - Attract sources that normally try to avoid AS 3715
 - Help AS 88 look like it is closer to the Internet’s core
- Who can tell that this AS path is a lie?
 - Maybe AS 88 *does* connect to AS 701 directly



155

Bogus AS Paths

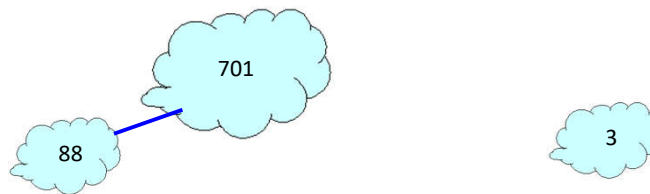
- Add ASes to the path
 - E.g., turn “701 88” into “701 3715 88”
- Motivations
 - Trigger loop detection in AS 3715
 - Denial-of-service attack on AS 3715
 - Or, blocking unwanted traffic coming from AS 3715!
 - Make your AS look like it has richer connectivity
- Who can tell the AS path is a lie?
 - AS 3715 could, if it could see the route
 - AS 88 could, but would it really care as long as it received data traffic meant for it?



156

Bogus AS Paths

- Adds AS hop(s) at the end of the path
 - E.g., turns “701 88” into “701 88 3”
- Motivations
 - Evade detection for a bogus route
 - E.g., by adding the legitimate AS to the end
- Hard to tell that the AS path is bogus...
 - Even if other ASes filter based on prefix ownership



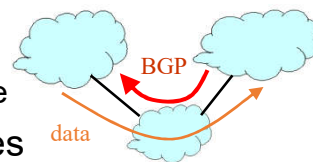
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

157

157

Invalid Paths

- AS exports a route it shouldn't
 - AS path is a valid sequence, but violated policy
- Example: customer misconfiguration
 - Exports routes from one provider to another
- ... interacts with provider policy
 - Provider prefers customer routes
 - ... so picks these as the best route
- ... leading the dire consequences
 - Directing all Internet traffic through customer
- Main defense
 - Filtering routes based on prefixes and AS path



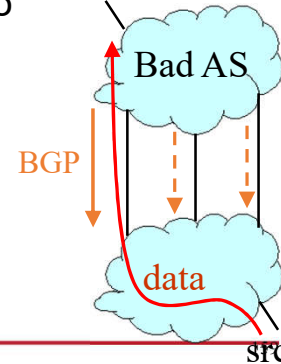
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

158

158

Missing/Inconsistent Routes

- Peers require consistent export
 - Prefix advertised at all peering points
 - Prefix advertised with same AS path length
- Reasons for violating the policy
 - Trick neighbor into “cold potato”^{dest}
 - Configuration mistake
- Main defense
 - Analyzing BGP updates
 - ... or data traffic
 - ... for signs of inconsistency



BGP Security Today

- Applying best common practices (BCPs)
 - Securing the session (authentication, encryption)
 - Filtering routes by prefix and AS path
 - Packet filters to block unexpected control traffic
- This is not good enough
 - Depends on vigilant application of BCPs
 - ... and not making configuration mistakes!
 - Doesn't address fundamental problems
 - Can't tell who owns the IP address block
 - Can't tell if the AS path is bogus or invalid
 - Can't be sure the data packets follow the chosen route

Proposed Enhancements to BGP



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

161

161

S-BGP Secure Version of BGP

- Address attestations
 - Claim the right to originate a prefix
 - Signed and distributed out-of-band
 - Checked through delegation chain from ICANN
- Route attestations
 - Distributed as an attribute in BGP update message
 - Signed by each AS as route traverses the network
 - Signature signs previously attached signatures
- S-BGP can validate
 - AS path indicates the order ASes were traversed
 - No intermediate ASes were added or removed



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

162

162

S-BGP Deployment Challenges

- Complete, accurate registries
 - E.g., of prefix ownership
- Public Key Infrastructure
 - To know the public key for any given AS
- Cryptographic operations
 - E.g., digital signatures on BGP messages
- Need to perform operations quickly
 - To avoid delaying response to routing changes
- Difficulty of incremental deployment
 - Hard to have a “flag day” to deploy S-BGP



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

163

163

Incrementally Deployable Schemes

- Monitoring BGP update messages
 - Use past history as an implicit registry
 - E.g., AS that announces each address block
 - E.g., AS-level edges and paths
- Out-of-band detection mechanism
 - Generate reports and alerts
 - Internet Alert Registry: <http://iar.cs.unm.edu/>
 - Prefix Hijack Alert System
 - Soft response to suspicious routes
 - Prefer routes that agree with the past
 - Delay adoption of unfamiliar routes when possible
 - Some (e.g., misconfiguration) will disappear on their own



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

164

164

What About Packet Forwarding?



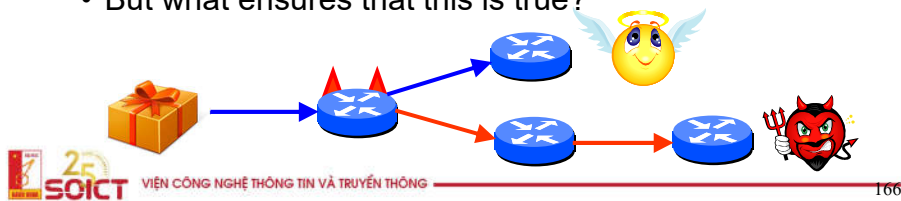
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

165

165

Control Plane Vs. Data Plane

- Control plane
 - BGP is a routing protocol
 - BGP security concerns validity of routing messages
 - I.e., did the BGP message follow the sequence of ASes listed in the AS-path attribute
- Data plane
 - Routers forward data packets
 - Supposedly along the path chosen in the control plane
 - But what ensures that this is true?



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

166

166

Data-Plane Attacks, Part 1

- Drop packets in the data plane
 - While still sending the routing announcements
- Easier to evade detection
 - Especially if you only drop some packets
 - Like, oh, say, BitTorrent or Skype traffic
- Even easier if you just slow down some traffic
 - How different are normal congestion and an attack?
 - Especially if you let ping/traceroute packets through?

167

Data-Plane Attacks, Part 2

- Send packets in a different direction
 - Disagreeing with the routing announcements
- Direct packets to a different destination
 - E.g., one the adversary controls
- What to do at that bogus destination?
 - Impersonate the legitimate destination (e.g., to perform identity theft, or promulgate false information)
 - Snoop on the traffic and forward along to real destination
- How to detect?
 - Traceroute? Longer than usual delays?
 - End-to-end checks, like site certificate or encryption?

168

Fortunately, Data-Plane Attacks are Harder

- Adversary must control a router along the path
 - So that the traffic flows through him
- How to get control a router
 - Buy access to a compromised router online
 - Guess the password
 - Exploit known router vulnerabilities
 - Insider attack (disgruntled network operator)
- Malice vs. greed
 - Malice: gain control of someone else's router
 - Greed: Verizon DSL blocks Skype to gently encourage me to pick up my landline phone to use Verizon long distance \$ervice ☺



What's the Internet to Do?



BGP is So Vulnerable

- Several high-profile outages
 - <http://merit.edu/mail.archives/nanog/1997-04/msg00380.html>
 - http://www.renesys.com/blog/2005/12/internetwide_nearcatastrophela.shtml
 - http://www.renesys.com/blog/2006/01/coned_steals_the_net.shtml
 - http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml
- Many smaller examples
 - Blackholing a single destination prefix
 - Hijacking unallocated addresses to send spam
- Why isn't it an even bigger deal?
 - Really, most big outages are configuration errors
 - Most bad guys want the Internet to stay up
 - ... so they can send unwanted traffic (e.g., spam, identity theft, denial-of-service attacks, port scans, ...)



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

171

171

BGP is So Hard to Fix

- Complex system
 - Large, with around 30,000 ASes
 - Decentralized control among competitive ASes
 - Core infrastructure that forms the Internet
- Hard to reach agreement on the right solution
 - S-BGP with public key infrastructure, registries, crypto?
 - Who should be in charge of running PKI and registries?
 - Worry about data-plane attacks or just control plane?
- Hard to deploy the solution once you pick it
 - Hard enough to get ASes to apply route filters
 - Now you want them to upgrade to a new protocol
 - ... all at the exact same moment?



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

172

172

Conclusions

- Internet protocols designed based on trust
 - The insiders are good guys
 - All bad guys are outside the network
- Border Gateway Protocol is very vulnerable
 - Glue that holds the Internet together
 - Hard for an AS to locally identify bogus routes
 - Attacks can have very serious global consequences
- Proposed solutions/approaches
 - Secure variants of the Border Gateway Protocol
 - Anomaly detection schemes, with automated response
 - Broader focus on data-plane availability



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

173

173

Encrypting and Decrypting With Keys

- Encrypt to hide message contents
 - Transforming message contents with a key
 - Message cannot be read without the right key
- Symmetric key cryptography
 - Same secret key for encrypting and decrypting
 - ... makes it hard to distribute the secret key
- Asymmetrical (or public key) cryptography
 - Sender uses public key to encrypt message
 - Can be distributed freely!
 - Receiver uses private key to decrypt message



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

174

174

Authenticating the Sender and Contents

- Digital signature for authentication
 - Data attached to the original message
 - ... to identify sender and detect tampering
 - Sender encrypts message digest with private key
 - Receiver decrypts message digest with public key
 - ... and compares with message digest it computes
- Certificate
 - Collection of information about a person or thing
 - ... with a digital signature attached
 - A trusted third party attaches the signature



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

175

175

Public Key Infrastructure (PKI)

- Problem: getting the right key
 - How do you find out someone's public key?
 - How do you know it isn't someone else's key?
- Certificate Authority (CA)
 - Bob takes public key and identifies himself to CA
 - CA signs Bob's public key with digital signature to create a certificate
 - Alice can get Bob's key and verify the certificate with the CA
- Register once, communicate everywhere
 - Each user only has the CA certify his key
 - Each user only needs to know the CA's public key



VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

176

176



177