Algorhyme

I think that I shall never see
a graph more lovely than a tree.
A tree whose crucial property
is loop-free connectivity.
A tree that must be sure to span
so packet can reach every LAN.
First, the root must be selected.
By ID, it is elected.
Least-cost paths from root are traced.
In the tree, these paths are placed.
A mesh is made by folks like me,
then bridges find a spanning tree.
–
Radia Perlman

Network Technology 2 –

Lennart Franked
email:lennart.franked@miun.se
Tel:060-148683

Informationsteknologi och medier / Informations- och Kommunikationssystem (ITM/IKS)
Mittuniversitetet

19/03-2013

Redundancy issues

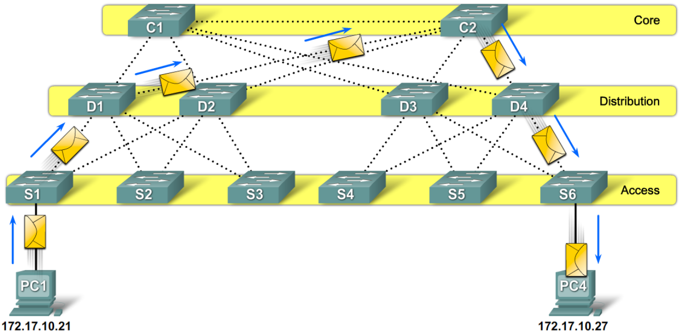

Figure 1 : Redundant network[2]

Redundancy issues

Issues

- Redundancy causes layer 2 loops
- Ethernet frames do *not* have a TTL field.

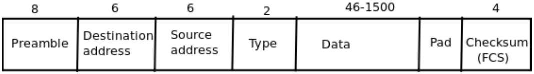

Figure 2 : DIX Ethernet header

Broadcast loop

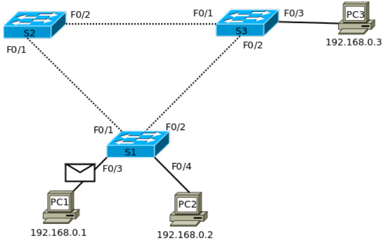

Figure 3 : Broadcast loop[2]

Table 1 : S1 MAC table

Destination	Port
PC1	F0/3

Table 2 : S2 MAC table

Destination	Port

Table 3 : S3 MAC table

Destination	Port

Broadcast loop

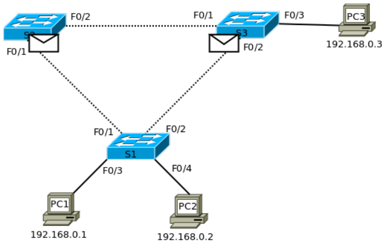

Figure 4 : Broadcast loop[2]

Table 4 : S1 MAC table

Destination	Port
PC1	F0/3

Table 5 : S2 MAC table

Destination	Port
PC1	F0/1

Table 6 : S3 MAC table

Destination	Port
PC1	F0/2

Broadcast loop

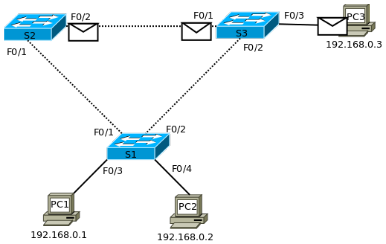

Figure 5 : Broadcast loop[2]

Table 7 : S1 MAC table

Destination	Port
PC1	F0/3

Table 8 : S2 MAC table

Destination	Port
PC1	F0/1

Table 9 : S3 MAC table

Destination	Port
PC1	F0/2

Broadcast loop

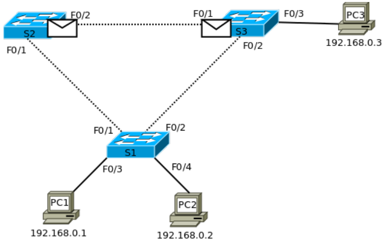

Figure 6 : Broadcast loop[2]

Table 10 : S1 MAC table

Destination	Port
PC1	F0/3

Table 11 : S2 MAC table

Destination	Port
PC1	F0/2

Table 12 : S3 MAC table

Destination	Port
PC1	F0/1

Broadcast loop

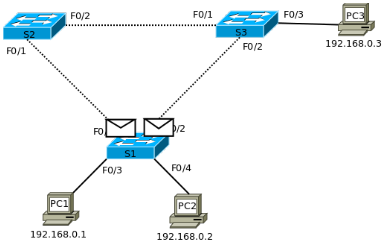

Figure 7 : Broadcast loop[2]

Table 13 : S1 MAC table

Destination	Port
PC1	F0/1

Table 14 : S2 MAC table

Destination	Port
PC1	F0/2

Table 15 : S3 MAC table

Destination	Port
PC1	F0/1

- Loop will continue until one of the links is broken.

Consequences of layer 2 loops

Layer 2 loop will result in:

- Broadcast storms
- Duplicate unicast
- Congestion
  - Packet Loss
  - increased delay

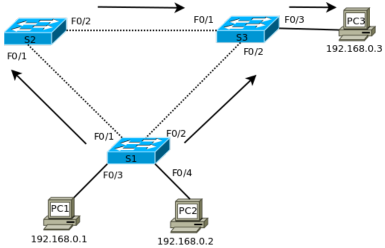

Figure 8 : Duplicate unicasts

Mittuniversitetet MID SWEDEN UNIVERSITY

STP

- Developed by Radia Perlman early 1980
- Published as IEEE 802.1D 1990
- Ensures that only one logical path exist between all destinations on the network.
- Blocking redundant paths

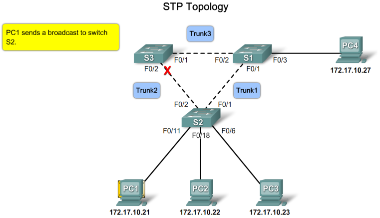

Figure 9 : Spanning Tree Protocol[2]

Spanning-tree Algorithm

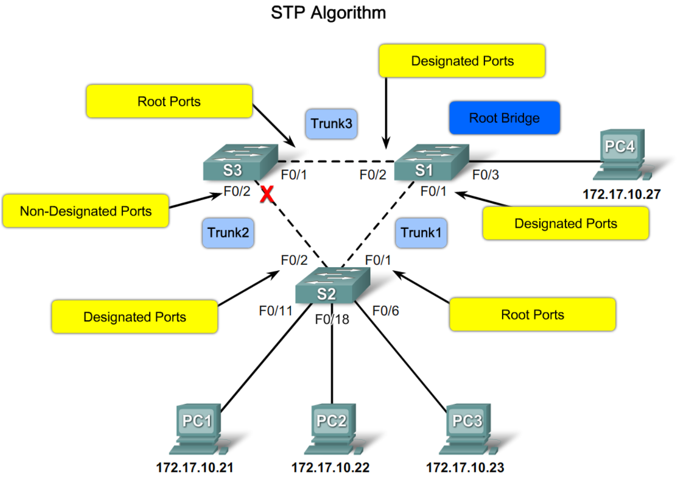

Figure 10 : Overview – Spanning-tree Algorithm[2]

STP components

Main components in STP

- Root bridge
- Spanning-tree Algorithm
- Port Roles
- Bridge Protocol Data Unit (BPDU)

Root bridge

Purpose of a root bridge

- Serves as a reference point to all spanning-tree calculations
- The root bridge is determined by an election process.
- The switch with the lowest bridge ID will be elected as root bridge.
- Bridge ID is exchanged using BPDU frames.

Port roles

- STA determines which ports need to be blocked to prevent a loop
- Three types of ports
  - Root port
    - Closest to the root bridge
  - Designated port
    - Non root port, used to forward frames.
  - Nondesignated port
    - Port blocked to prevent routing loops.
  - Disabled port
    - Administrativly shut down.

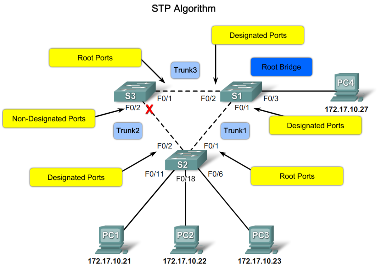

Figure 11 : Spanning Tree Protocol[2]

Spanning-tree Algorithm

- Calculates the best path to a root bridge
- Each link between two switches are given a cost value
- Path to root bridge is the sum of all egress port costs
- Best path to a root bridge is the path with the lowest cost

Table 16 : Path cost

Link speed	Cost (revised IEEE)	cost (old IEEE)
10Gbit/s	2	1
1Gbit/s	4	1
100Mbit/s	19	10
10Mbit/s	100	100

Bridge Protocol Data Unit

- First 4 bytes identify frame type
- Next four fields identify the root bridge and path cost to root bridge
- Last four fields contains update information.
- BPDU frames are sent every other second on all trunk interfaces.

The BPDU Fields

Field #	Bytes	Field
4	2	Protocol ID
	1	Version
	1	Message type
	1	Flags
8	8	Root ID
	4	Cost of path
	8	Bridge ID
	2	Port ID
12	2	Message age
	2	Max age
	2	Hello time
	2	Forward delay

Figure 12 : A BPDU frame[2]

Bridge ID

- Bridge priority
  - Modifiable value
- Extended system ID
  - Contains VLAN ID for the VLAN associated with this STP tree.
- MAC address

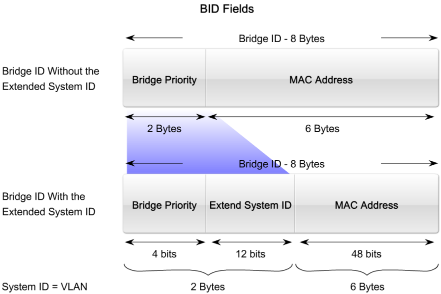

Figure 13 : Bridge ID[2]

BPDU timers

- BPD Timers.
  - Hello time - BPDU send interval
  - Forward delay - Time spent in listening and learning mode
  - Maximum age - How long a switch port saves BPDU information.

BPDU Timers

Hello time	The hello time is the time between each BPDU frame that is sent on a port. This is equal to 2 seconds by default, but can be tuned to be between 1 and 10 seconds.
Forward delay	The forward delay is the time spent in the listening and learning state. This is by default equal to 15 seconds for each state, but can be tuned to be between 4 and 30 seconds.
Maximum age	The max age timer controls the maximum length of time a switch port saves configuration BPDU information. This is 20 seconds by default, but can be tuned to be between 6 and 40 seconds.

Figure 14 : BPDU Timers[2]

Port states

- Each port involved in STP transitions through five possible states.
  - Blocking - Does not participate in the frame forwarding.
  - Listening - Elects root bridge, root ports, designated port and non-designated port.
  - Learning - Prepares to participate in frame forwarding.
  - Forwarding - Participates in the active STP topology
  - Disabled - Administratively down.

Port States

Processes	Blocking	Listening	Learning	Forwarding	Disable
Receives and process BPDUs	✓	✓[1]	✓	✓	✗
Forward data frames received on interface	✗	✗	✗	✓	✗
Forward data frames switched from another interface	✗	✗	✗	✓	✗
Learn MAC addresses	✗	✗	✓	✓	✗

[1]Return to blocking if not lowest cost path to root bridge

Figure 15 : physically separated networks[2]

Transitions

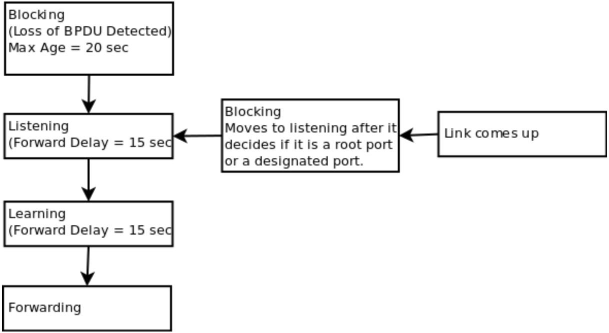

Figure 16 : STP port state transitions[2]

PortFast

Cisco PortFast

- Transitions from blocked to forwarding immediately.
- Minimize the time that access ports must wait.
- Used on access ports connected to an end device.

Example

```
Switch(config-if)#spanning-tree portfast
```

STP convergence

STP convergence process can be seen as three steps

- Elect the root bridge
- Elect the root ports
- Elect the designated and nondesignated ports

Elect the root bridge

- Based on the lowest Bridge ID
- If priority are the same, lowest MAC-address will be used

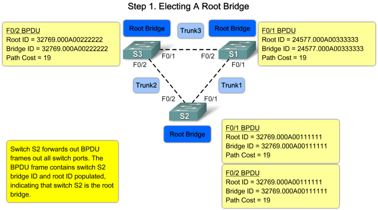

Figure 17 : Electing root[2]

Root port

- Compares the path cost to the root bridge.
- Switch port with the lowest cost to root will be set as root port.
- If equal cost, port priority will be used.
- If equal port priority, lowest port number will be used.

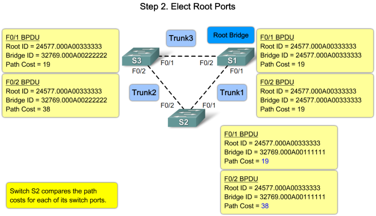

Figure 18 : Elect root port[2]

Designated and non designated ports

- Port with the lowest root path cost becomes designated port.
- Port with the highest root path cost becomes non designated port.

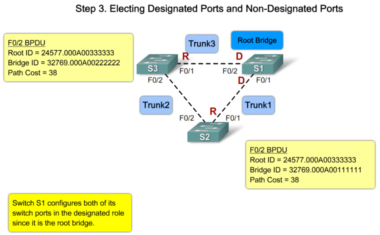

Figure 19 : Elect designated port[2]

Topology changes

- If a topology change occur, the first affected switch will inform the root bridge using a TCN (Topology Change Notification)
- Each switch receiving a TCN BPDU will respond with a TCA (Topology Change Acknowledgment)
- These frames will be sent until root bridge responds.

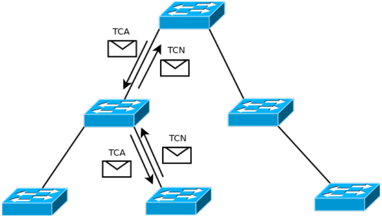

Figure 20 : Response to topology change[2]

Topology changes

- Root bridge responds with a Topology Change frame
- Each receiving switch will change the aging time to the forwarding delay (15 sec default)
- This allows the switch to relearn the MAC addresses available on the network.

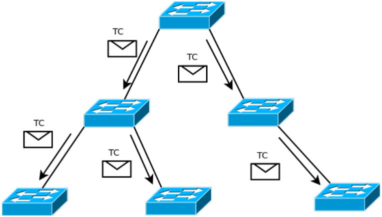

Figure 21 : Response to topology change[2]

STP variants

Cisco and STP Variants

Cisco Proprietary	PVST • Uses the Cisco proprietary ISL trunking protocol • Each VLAN has an instance of spanning tree • Ability to load balance traffic at layer-2 • Includes extensions BackboneFast, UplinkFast, and PortFast
	PVST+ • Supports ISL and IEEE 802.1Q trunking • Supports Cisco proprietary STP extensions • Adds BPDU guard and Root guard enhancements
	rapid-PVST+ • Based on IEEE802.1w standard • Has faster convergence than 802.1D
IEEE Standard	RSTP • Introduced in 1982 provides faster convergence than 802.1D • Implements generic versions of the Cisco proprietary STP extensions • IEEE has incorporated RSTP into 802.1D, identifying the specification as IEEE 802.1D-2004
	MSTP • Multiple VLANs can be mapped to the same spanning-tree instance • Inspired by the Cisco Multiple Instances Spanning Tree Protocol (MISTP), • IEEE 802.1Q-2003 now includes MSTP

Figure 22 : Variants of STP[2]

Per VLAN Spanning Tree +

PVST+

- Supports multiple VLAN.
- One tree per VLAN.
- Allows us to set a different root for each VLAN.
- Default STP protocol on a Cisco Catalyst switch.

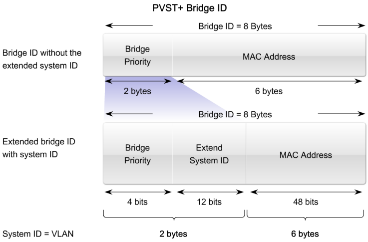

Figure 23 : BPDU Bridge ID[2]

Per VLAN Spanning Tree +

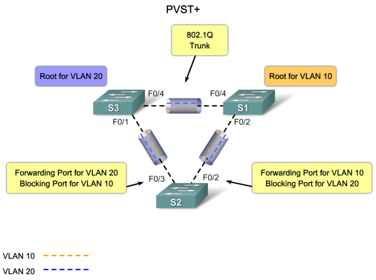

Figure 24 : PVST+[2]

Example (Macro)

```
Switch1(config)#spanning-tree vlan
10 root primary
Switch1(config)#spanning-tree vlan
20 root secondary
```

Example (Static)

```
Switch1(config)#spanning-tree vlan
10 priority 0
Switch1(config)#spanning-tree vlan
20 priority 4096
```

RSTP

IEEE 802.1w

- Final version was published 2002-02-07
- Preferred protocol for preventing layer 2 loops.
- Based on many Cisco enhancements of STP.
- Backwards compatible with original STP.
- Election process is the same as in 802.1D
- Same BPDU format

Port States

Table 17 : Port states[1]

STP (802.1D) Port State	RSTP (802.1W) Port State	Port included in active topology	Is Port learning MAC Addresses?
Disabled	Discarding	No	No
Blocking	Discarding	No	No
Listening	Discarding	Yes	No
Learning	Learning	Yes	Yes
Forwarding	Forwarding	Yes	Yes

Port Roles

RSTP Port Roles

- Root Port
  - The port that receives the best BPDU.
  - Closest to the root bridge in terms of path cost.
- Designated Port
  - The port that sends the best BPDU on the segment that it is connected to.
- Alternate
  - Receives a BPDU that is more useful than its own, and less useful than the BPDU received from the root-port.
  - Originates from another bridge than the designated bridge.
  - Alternate path to root.
  - Blocked port.
- Backup
  - Receives a more useful BPDU but less useful than the BPDU received from the root-port.
  - Both BPDUs originate from the same designated bridge.
  - Blocked port.
  - Redundant link to root.

Access port

- Edge Port
  - Same as Cisco's PortFast port
  - Stops being an Edge port if a BPDU is received on that port.

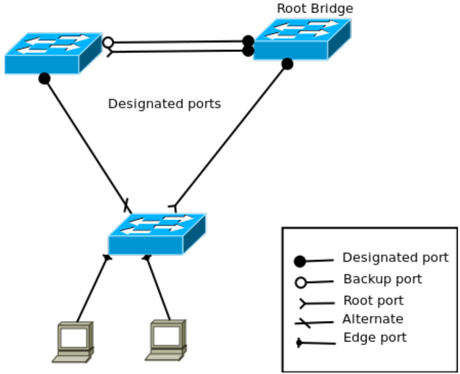

Figure 25 : RSTP link types

BPDU

BPDU in RSTP

- In 802.1D bridges only relayed BPDU.
- Made it impossible to know if the root is down, or a designated switch.
- In RSTP each bridge sends its own BPDU every hello-time.
- Works as a keep-alive between bridges.
- Allows for quick failure detection.

The Rapid in RSTP

Rapid Transition

- In 802.1D a designated port must wait two times the forwarding delay before it transitions to the forwarding state.
- RSTP keeps track of port and link type to know if it can make a rapid transition to forwarding state.
- Two variables are used to decide if it can make a rapid transition or not.
  - Edge port
  - Link type
    - point-to-point (Full-duplex)
    - shared (Half-duplex)

Link Types

- Rapid transition is only possible on Edge-ports and point-to-point links.
- On shared-links, timers will be used.

Ports that are in sync

A port is in sync if:

- It is in blocking state (Discarding).
- The port is an edge-port.

If a port is in sync it will not have to participate in a proposal/agreement process.

Proposal and agreement process in RSTP

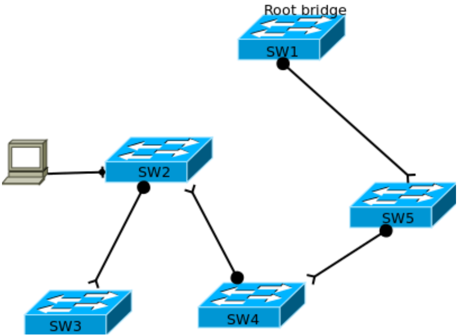

Figure 26 : RSTP topology

- A link will be added between SW1 and SW2.

Proposal and agreement process in RSTP – SW1 <-> SW2

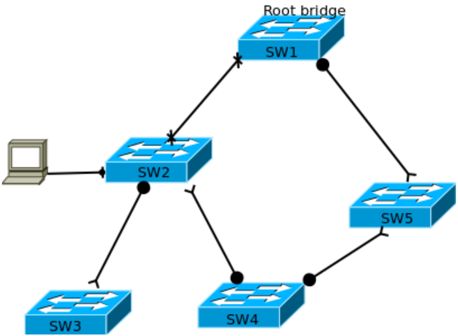

Figure 27 : RSTP topology

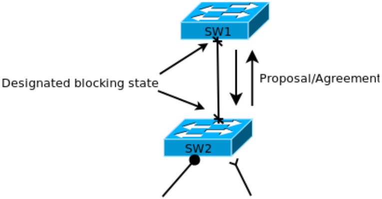

Figure 28 : Negotiation between SW1 and SW2

Proposal and agreement process in RSTP – SW2 <-> SW3,SW4

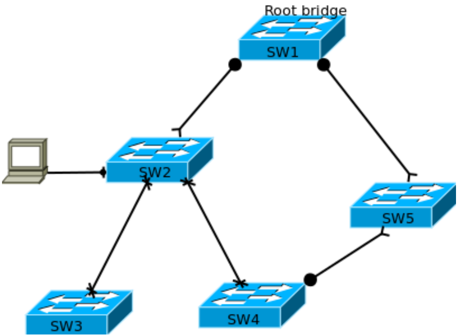

Figure 29 : RSTP topology

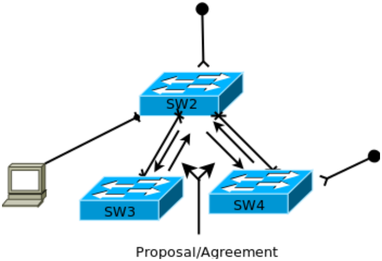

Figure 30 : Negotiation between SW2 - SW3 and SW2 - SW4

Proposal and agreement process in RSTP – SW4 <-> SW5

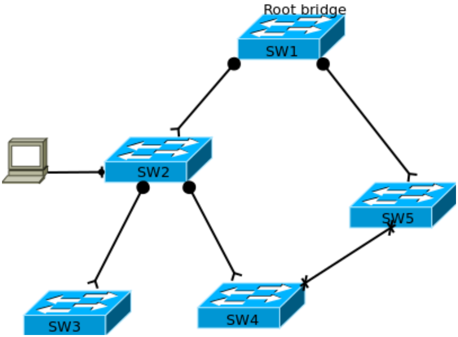

Figure 31 : RSTP topology

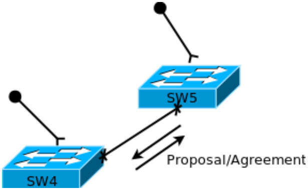

Figure 32 : Negotiation between SW4 and SW5

Proposal and agreement process in RSTP – SW5 <-> SW1

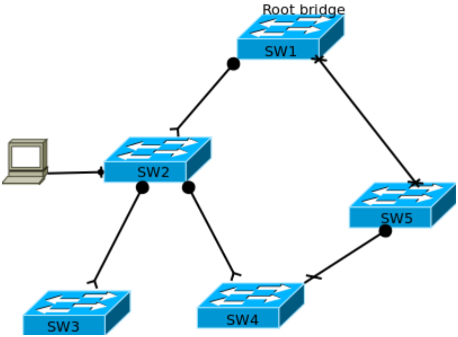

Figure 33 : RSTP topology

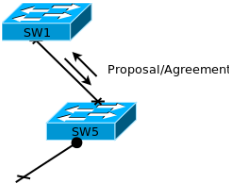

Figure 34 : Negotiation between SW5 and SW1

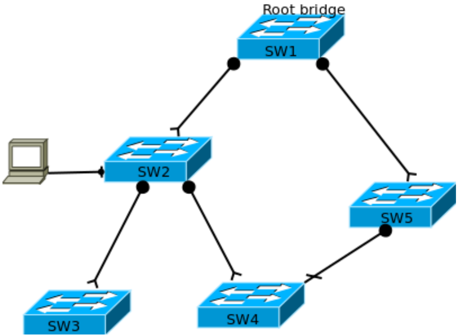

Figure 35 : RSTP topology

Inter-VLAN-routing

Chapter 6 - Inter-VLAN Routing

Traditional Inter-VLAN Routing

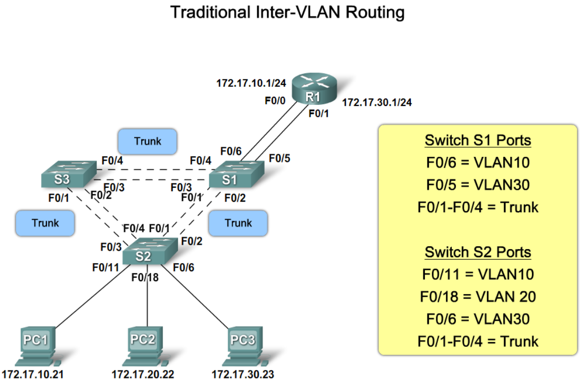

Figure 36 : Separate physical interface per VLAN[2]

Traditional Inter-VLAN Routing

```
R1#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)#interface f0/0
R1(config-if)#ip address 172.17.10.1 255.255.255.0
R1(config-if)#no shutdown
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0,
changed state to up
R1(config-if)#interface f0/1
R1(config-if)#ip address 172.17.30.1 255.255.255.0
R1(config-if)#no shutdown
%LINK-5-CHANGED: Interface FastEthernet0/1, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/1,
changed state to up
R1(config-if)#end
R1#copy running-config startup-config
```

Figure 37 : Separate physical interface per VLAN[2]

Router-on-a-stick

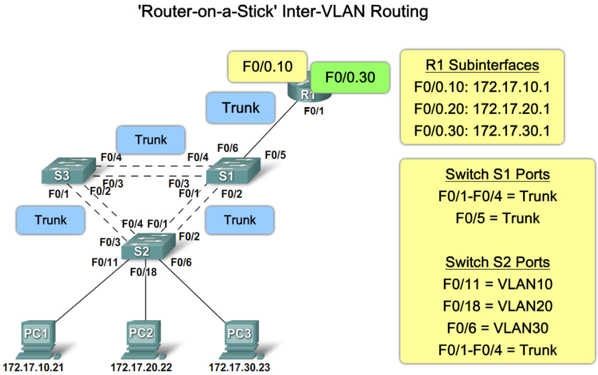

Figure 38 : Logical interface per VLAN[2]

Router-on-a-stick

```
R1#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
R1(config)#interface f0/0.10
R1(config-subif)#encapsulation dot1q 10
R1(config-subif)#ip address 172.17.10.1 255.255.255.0
R1(config-subif)#interface f0/0.30
R1(config-subif)#encapsulation dot1q 30
R1(config-subif)#ip address 172.17.30.1 255.255.255.0
R1(config-subif)#interface f0/0
R1(config-if)#no shutdown
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0,
changed state to up
%LINK-5-CHANGED: Interface FastEthernet0/0.10, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0.10,
changed state to up
%LINK-5-CHANGED: Interface FastEthernet0/0.30, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0.30,
```

Figure 39 : Logical interface per VLAN[2]

Questions?

Referenser

[1] Cisco. Understanding rapid spanning tree protocol (802.1w), 2006. URL http://www.cisco.com/en/US/tech/tk389/tk621/technologies_white_paper09186a0080094cfa.shtml.

[2] Wayne Lewis. *LAN switching and wireless : CCNA exploration companion guide*. Cisco, Indianapolis, Ind., 2008. ISBN 1-58713-207-9 (hardcover w/cd).