Chapter 9: Rapid Spanning Tree Protocol (RSTP)

9.1 Introduction

In practice, STP can actually take anything from 30 to 50 s to re-converge following a topology change. In order to provide a faster recovery, the IEEE introduced Rapid Spanning Tree Protocol (RSTP) (IEEE 802.1w) – having backwards compatibility with STP.

Probably the most important feature introduced by 802.1w is rapid transition. RSTP provides significantly faster spanning tree convergence – responding to a topology change within 3 Hello times (3 2 s (default|)) or a few milliseconds of a physical link failure.

Whilst the legacy STP waited passively for the network to converge before it turned a port into the forwarding state, RSTP is able to actively confirm that a port can safely transition to the forwarding state without having to rely on any timer configuration.

As opposed to STP, RSTP clearly distinguishes between the **state** of a port and the **role** of a port.

9.2 RSTP port states

Whilst, as previously described, STP defines five port states, RSTP defines only three: **Forwarding**; **Learning**; and a new state, **Discarding**.

Discarding describes a port where all the received frames are discarded and no learning takes place. There are thus no entries in the filtering database that point to this port and no traffic being forwarded across it.

9.3 RSTP port roles

Whereas STP defines the port roles of designated and root, RSTP adds three new port roles: **Disabled** (formally defined as a port state in STP); **Backup** port; and **Alternate** port. In STP these latter two new roles would have been called blocked.

A backup port is paralleled to a designated port such that if a switch has a designated port to a particular collision domain, and has a second port connecting to that same collision domain, the port is declared a backup port. This is detected by the switch as it senses its own Configuration BPDU arriving on that port.

An alternate port is essentially a backup to the root port: if the root port is lost, the alternate port is quickly used as the new root port

Figure 9.1 illustrates the concept of port roles. Since Switch 14 is the root bridge switch in this network, both of its ports are designated ports. Switches 82, 65 and 47 have selected Ports 82/1, 65/1 and 47/1, respectively, as their root ports because they are the ports with the lowest root path cost on each switch.

Since only one of the links on a segment should forward traffic to and from that segment, the port connecting this bridge to the network segment is referred to as the **Designated Port**. On this basis, Ports 82/2 and 47/2 on Switches 82 and 47, respectively, are declared designated ports. This means that, assuming the ports have cleared the learning state, all ports on Switches 14, 82, and 47 are in the forwarding state.

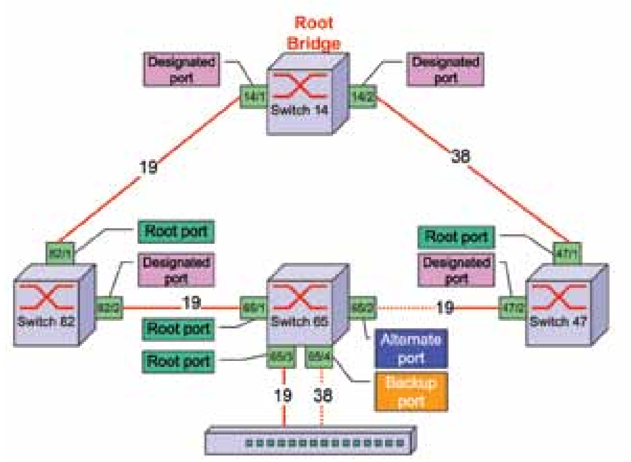

Figure 9.1: Illustration of RST port roles.

Switch 65 has selected Port 3 as the designated port for that collision domain because it has detected that there is no other switch in that domain. It has, however, determined that Ports 3 and 4 are both connected to the same domain, and has declared Port 4, which has a high root port cost, as a backup port. Since Port 2 is neither a root port (Port 1 has a lower root path cost); a designated port (Switch 47 has a lower root path cost and is connected to that collision domain); nor a backup port (since it is not connected to the same collision domain as any other Switch 65 port); it is declared an alternate port. This means that Ports 4 and 2 are both in the discarding state.

In STP, the Configuration BPDUs are initiated by the root and are passed down through the spanning tree. If there is a failure, the Configuration BPDUs cannot get past this point – thus indicating to all the devices below that point that something is wrong. Unfortunately, whilst they know something is wrong they do not know 'what' or 'where' the problem lies.

In RSTP, each switch is responsible for initiating its own Configuration BPDUs locally – based on a timer. Thus, if a switch stops seeing Configuration BPDUs on its root port, it knows the failure is with the designated switch for that collision domain.

A second feature of RSTP is that if a switch receives inferior information about the root bridge or root path cost on a designated port then, instead of discarding it (as in STP), it stores that information (rather than discarding it) and sends its own Configuration BPDU. This means that, in the event the switch loses its root port, it has up-to-date information on all of its ports and can immediately select a new root port.

To achieve fast convergence on a port, RSTP relies on two new variables: **Edge** Ports and **Link** Ports.

Edge ports

An edge port is a designated port that operates in a half-duplex mode and connects to a collision domain where there are no other switches present. This includes any port that directly attaches to an end-station, a router, a server or a hub (see Figure 9.2). These ports are automatically made part of the spanning tree and transition directly to the forwarding state – skipping the listening and learning stages, regardless of what happens on the other ports. A switch can automatically detect an edge port by noting the absence of Configuration BPDUs from any attached system. Alternatively, it can be manually configured by the network administrator.

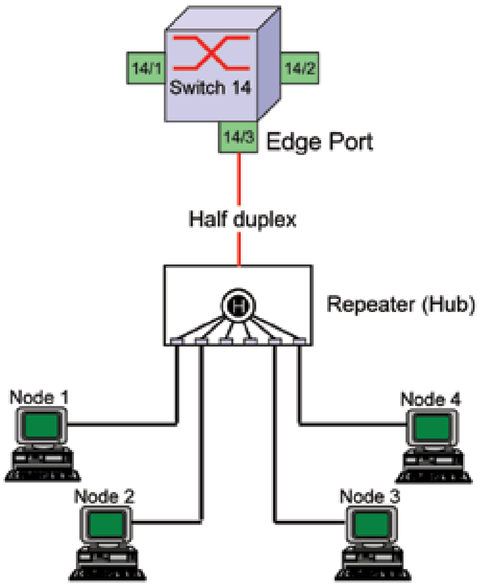

Figure 9.2: Example of edge port operating in a half-duplex mode and connecting to a hub, a collision domain where no other switches are present.

Link ports

Ports that are not edge ports are referred to as link ports and participate in the spanning tree process.

In switched networks, most links operate in full-duplex mode and are thus treated as point-to-point links by RSTP, which can only achieve rapid transition to the forwarding state on edge ports and on point-to-point links. The link port is automatically derived from the duplex mode of a port. A port that operates in full-duplex is assumed to be point-to-point, while a half-duplex port is considered as a shared port by default.

9.4 Fast recovery

Fast network recovery on RSTP encompasses a wide range of mechanisms.

If a designated port fails on a switch having a backup port, it engages immediately without fear of forming a loop since the switch still has the lowest root path cost for that collision domain, regardless of the number of switches attached to it. If either a root port or a designated port fails, and it is in half duplex mode (e.g., attached to a hub), then the original STP is engaged.

For point-to-point link ports: if failure occurs on a designated port, the responsibility for recovery lies with the switches in the tree below the failure; with the switch actually experiencing the failure no longer providing the designated port for that collision domain.

In the event of a root port failure (e.g., the port physically goes offline or it stops seeing Configuration BPDUs from the up-tree switch), a switch immediately places all the designated ports into the discarding state to prevent loops from forming. It selects its next best port to be the new root port – moving it into the learning state on its way to forwarding. It then sends a Configuration BPDU proposal on all of its other link ports.

This configuration BPDU notifies the surrounding switches and gives them an opportunity to reconfigure. However, switches receiving the Configuration BDPU on a backup, alternate, or designated port will not change their configuration since, if they were not using the switch that detected the loss to get to the root bridge previously, they will certainly not use it when it changes to a worse path.

The only switches of concern are those receiving the Configuration BDPU on their root port – who immediately place their designated ports into discarding state; select a new root port; and respond to the original switch with a Configuration BPDU response.

The original switch can thus reconfigure and determine which roles to assign to which ports, on a port-by-port basis, with no fear of forming loops.

This essentially creates a cascading effect, away from the root bridge, down through the spanning tree, which is redeployed with no need for any switch to wait in a listening state. In effect, no switch activates a port until it is certain that that port cannot participate in a loop. This is a major aspect that allows RSTP to achieve faster convergence times than STP.

RSTP maintains backup details regarding the discarding status of ports – thus avoiding timeouts if the current forwarding ports were to fail or BPDUs were not received on the root port within a certain interval.

9.5 Summary

In summary, RSTP provides significantly faster spanning tree convergence than STP. As distinct from STP, RSTP assumes that the three STP port states (listening, blocking and disabled) are all the same and do not forward frames or learn MAC addresses. Hence, RSTP places them all into a new state called discarding state. Learning and forwarding ports remain more or less the same.

Furthermore, unlike STP, where bridges only send out a BDPU when they see one on their root port, RSTP-enabled switches send out BPDUs every Hello time, containing current information.

And finally, whilst STP defines only two port types (root and designated), RSTP includes two additional port types – alternate ports and backup ports. An alternate port has an alternative path, or parts, to the root but is currently in a discarding state (and may be considered as an additional unused root port). A backup port can be considered as an additional unused designated port.