Ethernet Ring Protection Switching Application Note

Alcatel·Lucent
Enterprise

Table of Contents

List of figures

About This Document

Purpose

The purpose of this document is to serve as a reference guide to the solutions that Alcatel-Lucent Enterprise can provide to it's mission-critical network customers. This solution is based on Recommendation ITU-T G.8032/Y.1344, which defines the protection switching protocol and mechanisms for Ethernet ring network topologies which will be described in detail in this document.

Audience

This guide is intended for Alcatel-Lucent Enterprise's Business Partner sales and pre-sales staff and customers.

Scope

This document focuses on the Ethernet Ring Protection Switching (ERPS) mechanism version 2, which describes in detials the steps and actions performed by the ring nodes during normal and failure conditions in the network. It also provides configuration commands related to the ERPS ring network.

Acronyms

AOS	Alcatel Operating System
DHL	Dual Home Link
DNF	Do Not Flush
ERP	Ethernet Ring Protection
ERP v2	Ethernet Ring Protection version 2
ERPS	Ethernet Ring Protection Switching
FDB	Forwarding Database
FRR	Fast Reroute
IoT	Internet of Things
ITU-T	International Telecommunication Union, Telecommunication Standardization Sector
LACP	Link Aggregation Control Protocol
LSP	Label Switched Path
MAC	Media Acces Control
MPLS	Multi-protocol label switching
MSTP	Multiple Spanning Tree Protocol
NNI	Network-to-Network Interface
NR	No Request
RB	RPL Blocked
R-APS	Ring Automatic Protection Switching
RPL	Ring Protection Link
RSTP	Rapid Spanning Tree Protocol
SF	Signal Failure
STP	Spanning Tree Protocol
SVLAN	Service VLAN
VC	Virtual Chassis
VLAN	Virtual Local Area Network
VRRP	Virtual Router Redundancy Protocol

1. Abstract

Ethernet ring protection (ERP) switching, is a widely-used technology in the core, distribution and access network layers of telecommunication networks providing fast sub-second convergence and low resource requirements for its control plane operation with simple configuration. It is designed for networks with physical ring or interconnected ring topologies and provides a protected, single-path communication between any two nodes in the network.

The ERP technology and its protection mechanism are specified by the Telecommunication Standardization Sector of the International Telecommunication Union ((ITU-T). At the moment of this writing, the latest version of the "Recommendation ITU-T G.8032/Y.1344 (2020) – Corrigendum 1" document was released in February 2022.

Alcatel-Lucent Enterprise provides support for ERP across its wired, hardened and non-hardened Alcatel-Lucent OmniSwitch® portfolio, both as single nodes and in a Virtual Chassis (VC) configuration. Moreover, ALE enhances mission-critical networks running the ERP protocol with numerous additional features such as advanced secure onboarding of users, devices and Internet of Things (IoTs) using the Access Guarding framework.

2. Introduction

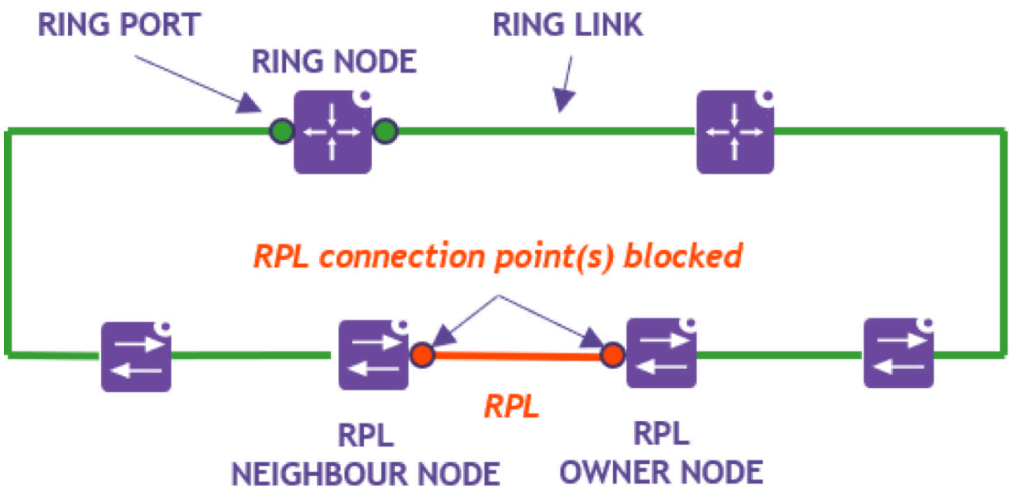
The concept of the Ethernet ring protection architecture is based on traffic flow on all links except for one, called the Ring Protection Link (RPL). During normal operation, the RPL is blocked to prevent network loops. In the event of a ring failure (either link or node), the RPL is unblocked to allow traffic to flow over it and maintain network connectivity. Coordination of ring protection actions at all nodes is handled by the Ring Automatic Protection Switching (R-APS) protocol.

3. Basic concept

An Ethernet ring is a collection of Ethernet ring nodes which are network elements with (at least) two ring ports, connected to two adjacent Ethernet ring nodes forming a closed physical loop. All Ethernet ring nodes must be able to support the ERP control plane functionality and control the blocking and unblocking of traffic through the ring ports.

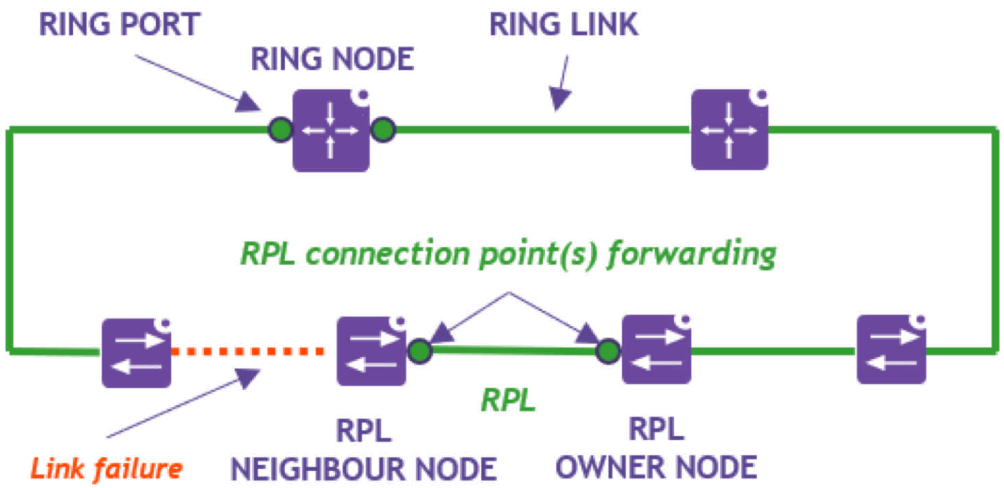
In the Ethernet ring, the RPL is blocked for traffic to prevent loops. The node that is adjacent to the RPL and responsible for blocking traffic on its end of the RPL is called the RPL Owner Node. Optionally, if configured, the RPL Neighbour Node is responsible for blocking traffic on the other side of the RPL. By blocking traffic at the RPL, loops can be avoided in the idle network, for example where no failure has been detected.

Figure 1: Network under normal condition



In the event of link or node failure on the ring network for example a network protection state, the RPL owner node is responsible for unblocking its end of the RPL, as well as the RPL Neighbour Node, if configured, to allow for communication between all the nodes in the network. This is done so that the network can continue to function even in the event of a failure.

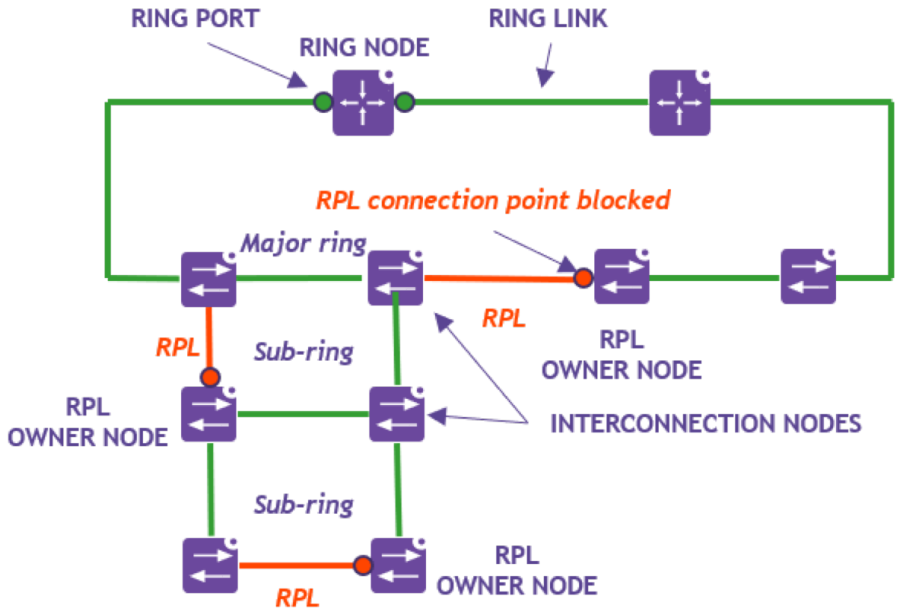
Figure 2: Network under ring failure condition



It is up to the network administrator to decide whether the network will automatically revert to a state where the RPL owner is responsible for blocking the traffic once the network failure has been repaired, or, reverting to the normal condition will be postponed and initiated in a controlled environment during a scheduled maintenance window.

The network presented in **Figures 1 and 2** consists of a single Ethernet ring. A multi-ring network, such as the one shown in Figure 3, uses the same protection switching mechanism and protocol as a single-ring network. Multi-ring networks consist of a major ring and one or more sub-rings. Sub-rings do not form a closed loop themselves, but traffic is looped through sub-ring links and the link between interconnection nodes. Interconnection nodes would have no more than one ring accessed by two ring ports and one or more sub-rings accessed through a single port.

Figure 3: Example of multi-ring network topology



4. Principle of operation

Ethernet ring protection switching, a mechanism used to ensure high availability in Ethernet ring networks, is based on the following two principles:

- **Loop avoidance**, this principle guarantees that at any time traffic can flow through all but one ring link. In case of normal condition mode, when no failure is detected, this is achieved through the use of the RPL node blocking traffic on the RPL on its end. In the event of ring failure, the RPL node is responsible for unblocking its end of the RPL to allow for traffic forwarding. One exception to this mechanism is a failure of the RPL itself.
- **Ethernet flow forwarding process, which includes learning, forwarding and filtering database mechanisms**. In the event of a ring failure event this process initiates protection switching of the traffic on all ring nodes.

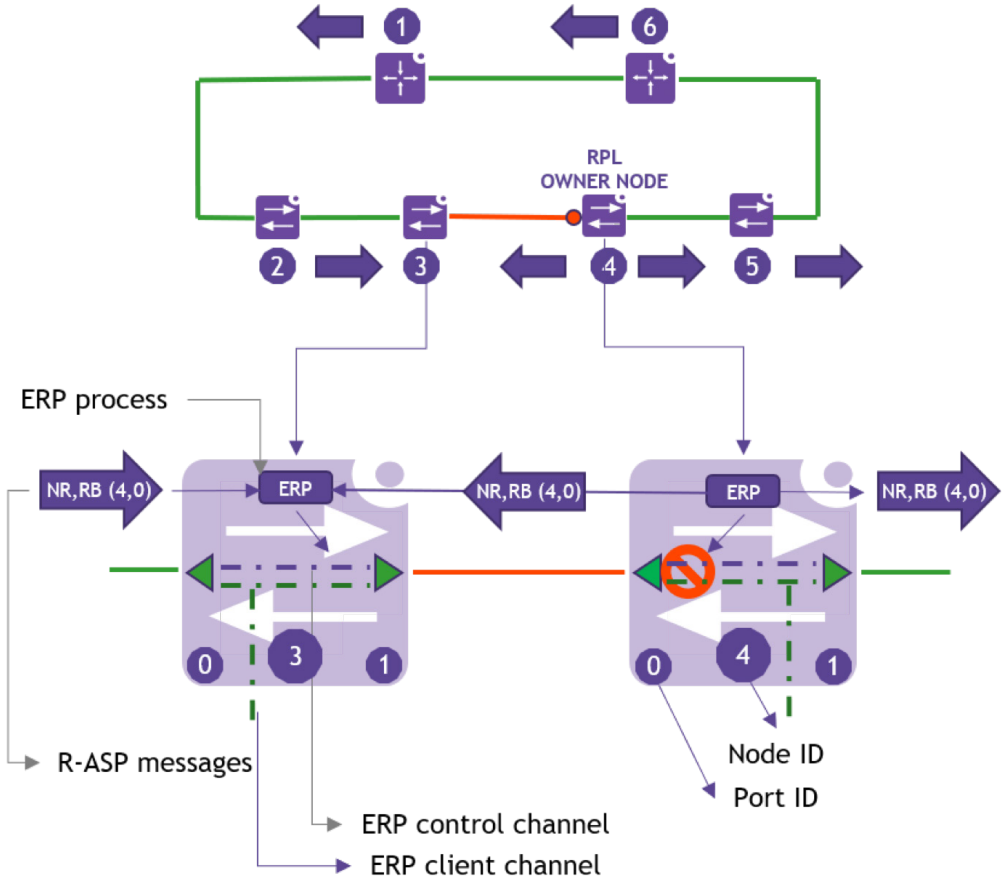
The Ring Automatic Protection Switching (R-APS) protocol is responsible for coordinating the above-mentioned processes among all nodes in the ring.

4.1 Normal condition

During the network normal condition, when all physical ring links are operational, the RPL owner node is responsible for blocking the RPL port and unblocking the non-RPL port for traffic. All other ring nodes unblock both ring ports, thus allowing for traffic to be forwarded in both directions over both ring ports. Traffic is also forwarded through user ports on all ring nodes where the VLAN domain is extended.

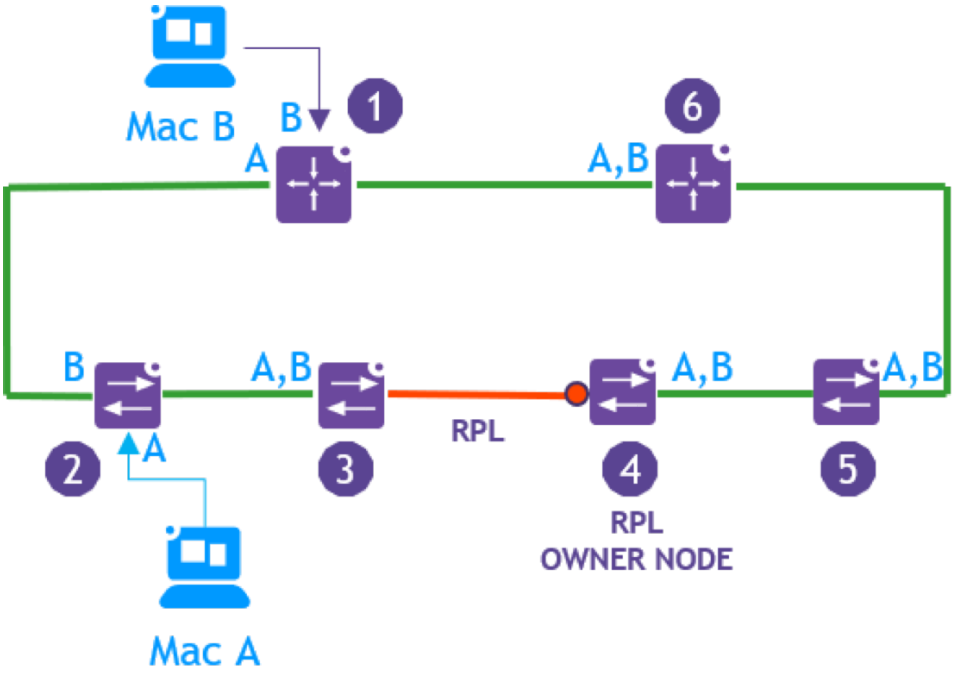
All ring nodes are in Idle state with only the RPL node sending R-APS messages, using a dedicated VLAN for it, informing other nodes in the ring that the RPL is blocked (RB) and there is no request (NR) for any action on their side. These messages contain the RPL node ID and the RPL port ID.

Figure 4: Normal condition of the network with nodes in idle state



During the network normal condition, communication between users A and B will be performed over the single available path which is over the direct ring link between ring nodes 1 and 2. **Figure 5** also presents the forwarding database for each node with learned MAC addresses of users A and B on the respected ports.

Figure 5: Status of FDB on each node during normal conditions



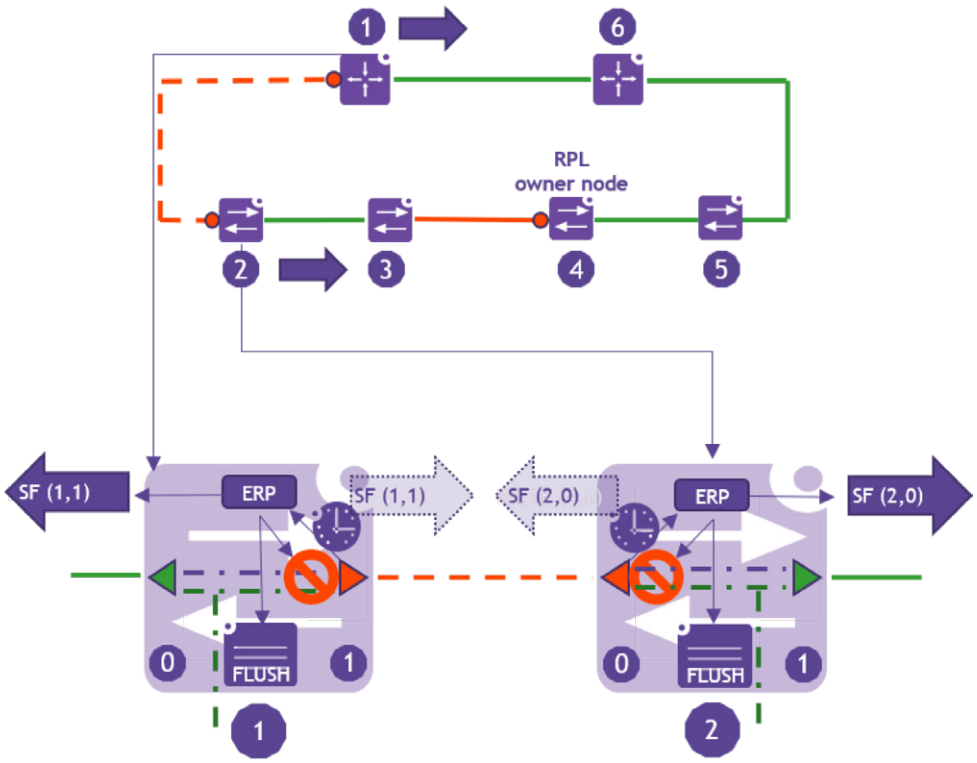
4.2 Single link failure condition

Detecting ring failure

In the event of a failure that occurred on the ring link, the associated ring nodes (nodes 1 and 2 in **Figure 6**) will detect the link failure on their local ports. This detection will initiate their hold-off timer before taking any protection switching action to allow automatic link recovery from intermittent link faults within this period that could be configured between 0 and 10 seconds. After the expiration of the hold-off time these ring nodes will:

- Internally block for traffic on the failed ring port
- Perform flushing of their FDB
- Start sending R-APS messages with Signal Failure (SF) code and affected node and port ID on both ring ports. Nodes will continue to send these messages for as long as the link failure persists.

Figure 6: Protection state activities of nodes with local link failure



Flushing FDB of ring nodes

All other ring nodes in the network will start receiving two different R-APS messages with a Signal Failure code, one sent by each ring node that has detected the local link failure. On receiving each R-APS message for the first time, the nodes will flush their forwarding database (twice in the case of a ring failure). The ring nodes with a local link failure will flush their FDB for the second time when they receive a R-APS message with a Signal Failure code from the node on the opposite side of the failed link.

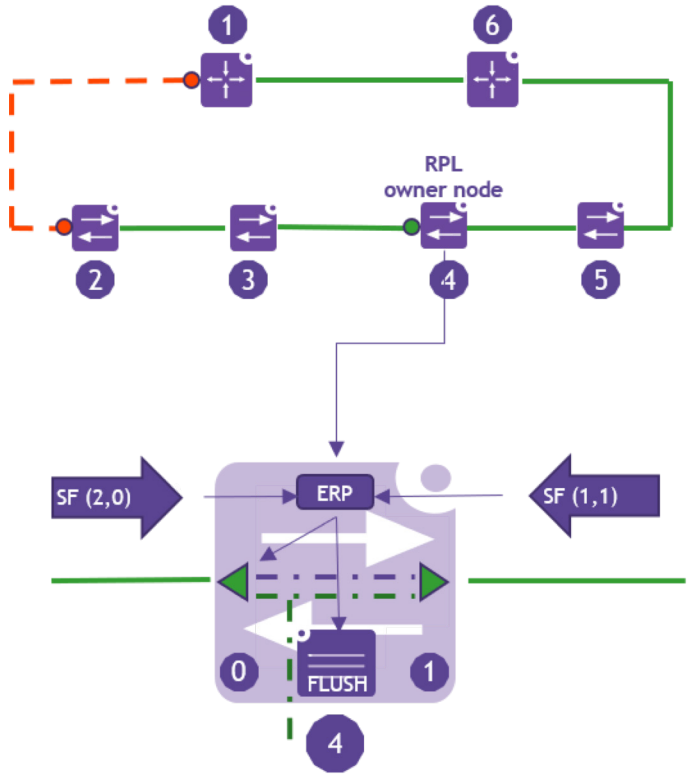
As previously mentioned, nodes with a local link failure will continue sending R-APS messages with the SF code for as long as the failure condition exists. However, reception of subsequent messages will not cause the FDB to be flushed again. To avoid repeated flushing of the FDB, the ring nodes will check the pair (Node ID, Blocked Port ID) within the R-APS SF message and will not respond to the receipt of messages with already known and stored pair of IDs.

This mechanism helps avoid continuous flushes of the FDB and ensures that the network can quickly recover from a failure and redirect traffic to an alternate path.

Unblocking RPL

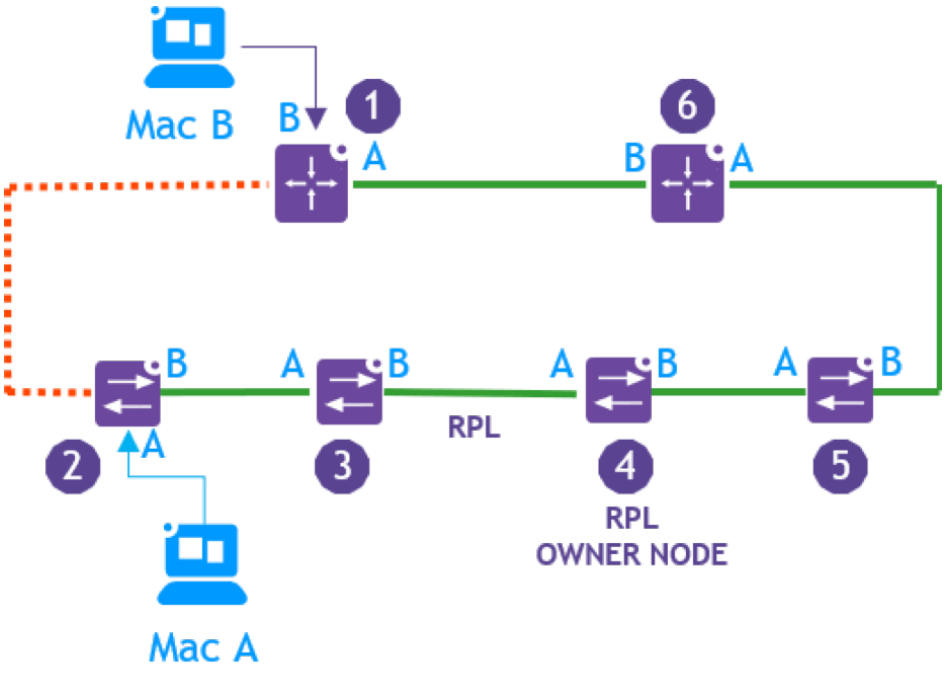
On receipt of the first R-APS message with the SF code, the RPL owner will also unblock its end of the RPL to allow communication over the RPL, thus providing reachability between all the nodes in the ring. This way, the network will remain connected even in the event of a failure.

Figure 7: RPL owner node upon receiving R-APS SF messages



Due to the FDB flushing, initial communication between users A and B will be flooded. However, once two-way communication is established, the MAC addresses of the users will be learned on the respective ports. This enables further traffic forwarding based on new FDB table entries.

Figure 8: FDB entries during failure conditions of the link between nodes 1 and 2



Exception in case of RPL failure

In this case where the failed link is the RPL itself, the process remains the same for the nodes at both ends of the RPL, except that a Do-Not-Flush (DNF) flag is included in the R-APS messages along with the SF condition sent to other ring nodes. The DNF flag indicates to the receiving ring nodes that there is no need to flush their FDBs as the topology is not changed. The same link is unavailable for traffic forwarding in the event of an RPL failure as it was when the failure didn't exist. There is no need to expect any changes in the local FDBs in order to flush them. If the node on the other side of the RPL is configured as an RPL neighbour, its R-APS message will also include the DNF flag.

4.3 Single link failure recovery (revertive and non-revertive modes)

Upon detection of a failed link recovery, nodes at both ends of the link will start a guard timer during which both nodes will ignore newly received R-APS messages. This is to prevent the nodes from receiving and acting on outdated R-APS messages which could create a loop in the network.

Recovery of one end of the recovered link

Both nodes also change their R-APS messages and initiate sending R-APS No Request (NR) messages instead of R-APS SF messages effectively informing other nodes in the network that the link failure no longer exists. Both nodes continue to ignore R-APS NR messages received from the other node on the recovered link until the guard time expires meaning they continue to block traffic on the port of the recovered link. Once the guard time has expired, both nodes start processing the R-APS NR messages received from the node on the other side of the recovered link and compare their own ID with the ID received from the other node. The node with the lower priority ID will unblock its port of the recovered link for traffic while the node with the higher ID will continue to block traffic on the port of the recovered link resulting in a link blocked on one end. At this point the RPL is still open for traffic.

Revertive mode - Recovery of RPL link

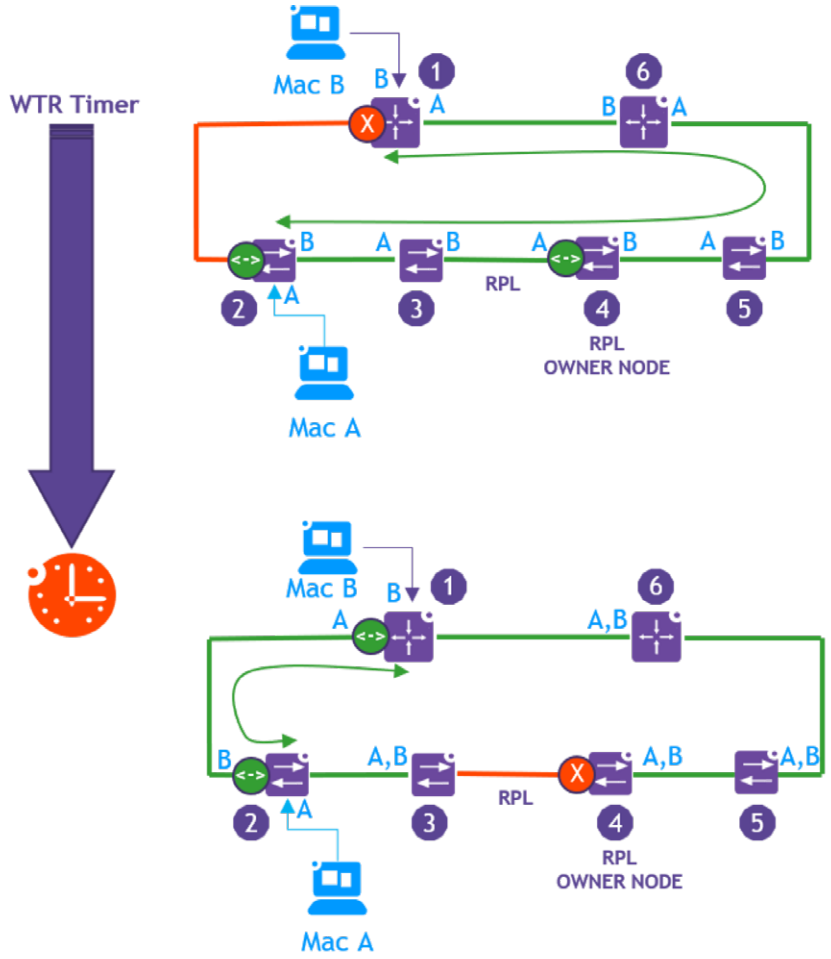
Upon receiving of R-APS NR messages as described above, the RPL owner node will start the Wait-to-Restore timer (WTR), which is configured to be longer than the guard timer, allowing the network to stabilise and preventing the operation of the protection switching due to intermittent link failure defects. When this timer expires, the RPL owner node:

- Starts blocking traffic over the RPL - flushes its FDB and
- Starts sending R-APS messages with the code No_Request, RPL-_Blocked (NR, RB) informing all ring nodes that the RPL is blocked

Upon receiving R-APS (NR, RB) messages, all ring nodes should unblock any blocked link ports. Therefore, the remaining blocked port on the node with the higher ID at one end of the recovered link will now be unblocked for traffic. As the RPL is already blocking at this point there will be no loop in the network. In addition, all ring nodes should, upon receiving R-APS (NR, RB) message, also flush their own FDB to allow for faster convergence. The FDB flush is only performed upon receiving the first R-APS (NR, RB) message.

Exceptionally, the above-mentioned process may be stopped if a new R-APS SF message is received by the RPL owner during the WTR time countdown.

Figure 9: Revertive mode of single link failure recovery



RPL failure recovery

In the event of the RPL failure, upon recovery, the same process described above is initiated in the network, except that the R_APS messages sent from the RPL owner have the Do-Not-Flush bit set. The Do-Not-Flush bit is set in the R-APS message so that other nodes in the network know that they don't need to flush their own FDBs, which would increase the convergence time.

Non-revertive mode recovery

In non-revertive mode, neither the RPL owner nor any other ring node will act upon receiving of R-APS NR messages from the nodes on either side of the recovered link. Only upon a "clear" command initiated by a network administrator on the RPL owner node, non-revertive operation is cleared and the recovery process will continue as described above in Revertive mode – recovery of RPL link.

The network may be deployed in non-revertive mode to minimise the potential impact of automatic convergence following link recovery, and scheduled maintenance windows are preferred. In this mode, the network administrator has full control over the recovery process, allowing them to perform network recovery without causing unintended service disruption.

5. Interconnected rings

5.1 Connecting distribution/access layer

Modern mission-critical networks use a hierarchical design in which the core layer is built on ring topology to which distribution and/or access layer devices are connected in a redundant way to prevent a single point of node or link failure. Redundantly connected devices require technology that can prevent forming Layer 2 loops by automatically detecting and disabling traffic forwarding across all the links.

Any variant of Spanning Tree, such as Rapid Spanning Tree Protocol (RSTP) and Multiple Spanning Tree Protocol (MSTP), could be used to create a loop-free logical topology. Although the ERP ports do not participate in the STP process, BPDUs could still be flooded over the ring ports and links. However, any form of Spanning Tree Protocol is not recommended since they increase convergence time and could potentially cause network instability.

For single distribution/access deployed nodes, it is possible to use Dual Home Link (DHL) technology. The forwarding status of each VLAN is modified by DHL to prevent network loops and maintain connectivity to the core ring network when one of the links fails. However, this technology is not a solution where multiple nodes are required in the access layer with a limited fibre optics availability.

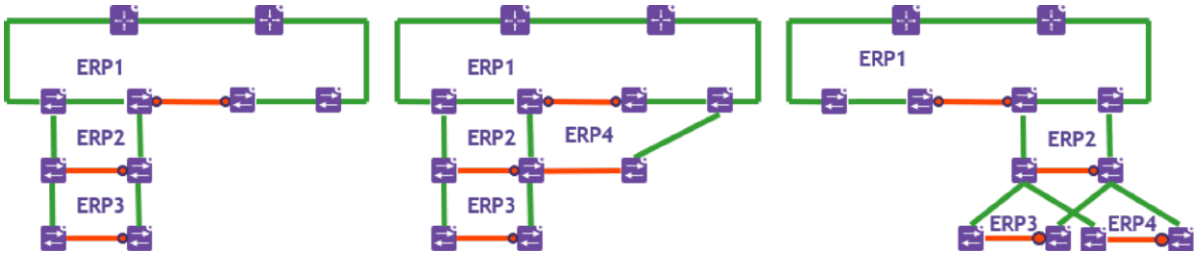
The recommended solution to connect one or more access devices to the core ring network is to use a multi-ring topology which is supported with ERP v2.

5.2 Multi-ring designs

The standard hierarchical network design often uses a single major ring in the core layer of the network with multiple sub-rings deployed in a distribution/access layer of the network. A sub-ring is also an Ethernet ring that is connected to a major ring through the pair of interconnected ring nodes. The sub-ring itself, does not form a physical closed loop. A closed loop is created together with one or more links between the interconnection nodes controlled by the major ring.

The sub-ring links are controlled by the ERP sub-ring control process while the links between interconnection nodes are controlled by the ERP control process of the major ring.

Figure 10: Examples of multi-ring ERP networks

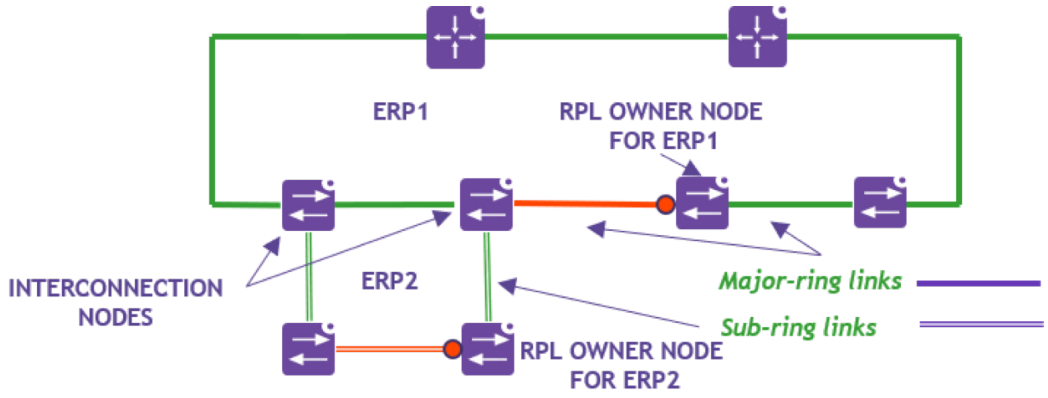


In order for the R-APS protocol and related protection switching mechanisms to be applicable to multi-ring networks, the following principles must be respected:

- The R-APS protocol is not shared across Ethernet ring interconnections
- On each ring port, each R-APS control protocol and protected VLANs are controlled by only one Ethernet ring
- Each major or sub-ring must have its own RPL

Following these principles, the network in **Figure 11** is designed to have two ERP instances, one running within the major ring and one within the sub-ring. The ring link between the two interconnection nodes is controlled only by the ERP1 instance and the ring ports on both sides of this link are controlled by ERP1 instance and related R-APS protocol of ERP1 instance. On the interconnection nodes, there are two ERP instances running. The ERP2 instance is configured on a single port, neither of the two ring ports are controlled by the ERP1 instance. The RPL owner functionality is configured on two different nodes but, more importantly, two different links are used as RPL for each ERP instance. The interconnection node can provide the functionality, or the RPL owner as well, as long as different links are used as RPL for different ERP instances.

Figure 11: Multi-ring network and related configuration



5.3 R-APS Virtual Channel

The R-APS protocol uses a dedicated R-APS channel VLAN for forwarding R-APS messages. By default, this R-APS channel is blocked on the same ring port where the traffic channels for protected VLANs are blocked. At any given time in an idle network, one ring port is blocked while others are forwarding.

Since the link(s) between interconnection nodes are controlled by the ERP1 instance, the ERP2 instance at the interconnection nodes must use the R-APS virtual channel for its R-APS protocol messages. When there are multiple sub-rings using the same shared links between interconnecting nodes, different VLANs will be used for the different R-APS virtual channels to clearly distinguish them from the R-APS messages of the ERP1 instance.

While this is the default behaviour in the AOS ERP network, it is possible to use sub-rings without virtual channels. In this case, R-APS messages are terminated at the interconnection nodes, but not blocked at the RPL of the sub-ring.

6. Convergence of ERP networks

6.1 ERP convergence

The ERP protocol is a technology widely used in mission-critical networks because it can complete the protection switching operations within a 50 ms interval. However, this time interval of 50 ms, as specified in the standard, is only respected under the following conditions:

- There is no congestion in an Ethernet ring network
- All nodes are in the idle state
- The number of nodes in the ring is less than 16
- The ring fibre length is less than 1200km

If any of these conditions are not met, the protection switching mechanism may take longer than 50ms.

When calculating the overall protection switching time for an ERP network, several components must be considered, including:

- Time to detect a link failure by directly connected ring nodes
- Time to generate and transmit R-APS SF message by the directly connected nodes on each side of the failed link
- Ring propagation delay of the R-APS SF message to reach each node in the network
- R-APS SF message processing and FDB flushing by receiving nodes

It's worth noting that the FDB flush mechanism is not part of the R-APS protocol and ERP control plane, but it is an important step in protection switching and can have a significant impact on the overall convergence time. The same applies for ring nodes running virtual chassis (VC) technology by which multiple physical nodes operate as a single logical node. ALE's support of ERP and VC technologies combined, increases resilience in the network in several ways:

- By deploying ring links consisting of multiple fibre links that are, using the LACP protocol, connected to different physical nodes of a ring node running VC
- By dual-homing connectivity of the access layer devices to different physical nodes running VC

However, combining these two technologies in the network may increase its convergence time due to the added complexity of the control plane and the synchronisation of FDBs flushes.

6.2 End-to-end convergence

The 50 ms period for the protection switching mechanism should not be interpreted as an end-to end convergence time. In addition to the conditions mentioned for the 50ms protection switching mechanism, there are several other factors that can influence end-to-end network convergence. Some of these factors are:

- Convergence of routing protocols including Virtual Router Redundancy Protocol (VRRP) when present in the network
- Layer 2 loop prevention protocols used between access and core layers

These factors can significantly increase the overall network convergence time.

Convergence in IP/MPLS networks, which also occurs in less than 50ms, is often used as an example of fast convergence in telecommunications networks. However, it is important to note that this is (again) only link protection for a Label Switched Path (LSP), a path through an MPLS network, and not end-to-end network convergence. The Fast Reroute (FRR) feature in MPLS networks can provide fast protection for LSPs in the event of link (or node) failure by pre-computing and installing backup next-hop paths before the failure occurs. However, FRR does not provide end-to-end network convergence and service recovery within 50 ms for every event of network failure.

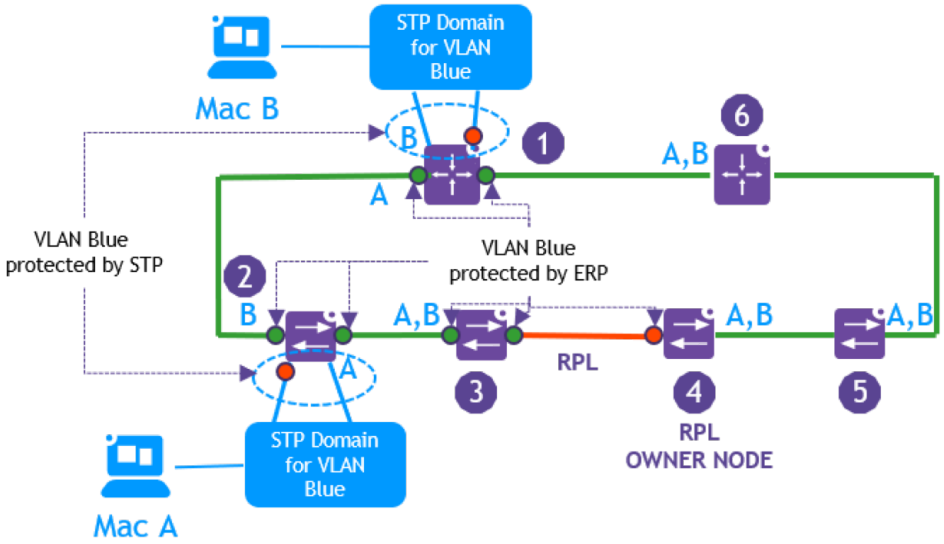
7. Interaction with other protocols in ERP networks

7.1 Spanning Tree

The Spanning Tree Protocol (STP) maintains a loop-free topology in the network providing data path redundancy. It ensures that there is only one data path between any two switches within the same Layer 2 domain regardless of the physical topology. By default, STP is active on all switches and a loop-free topology is automatically calculated based on STP parameters.

When configuring an ERP network by setting up a port as an ERP ring port, the Spanning Tree Protocol would be automatically disabled on that port since the port state would be controlled by the ERP protocol. However, the STP protocol will be active on all other switch ports and will determine the blocking or forwarding state of the VLANs configured on those ports. This way, the operation of the ERP protocol on the ring, together with the operation of STP on the connected network segments, provides for loop-free operation for all the VLANs configured throughout the larger network.

Figure 12: ERP and STP protected ports



With the default behaviour described above, one of the main drawbacks is the lack of support for network-wide redundancy. As shown in **Figure 12**, if network node 2 fails, the VLAN connectivity of the associated STP domain is completely disconnected from the network. To overcome this problem and provide network wide redundancy, an alternative architecture, such as an ERP sub-ring, should be considered.

7.2 VLAN stacking

The ALE implementation of ERP also supports ERP over an 802.1ad network, using VLAN stacking. The Network-to-Network Interface (NNI) ports can be configured as ERP ports to protect the Service VLANs (SVLANs) on the ring network. User-to-Network Interface (UNI) ports, on the other hand, cannot be used as ERP ring ports but are used to connect to customer networks.

8. ALE portfolio supporting ERP

ALE provides a broad range of products that support ERPv2 protocol:

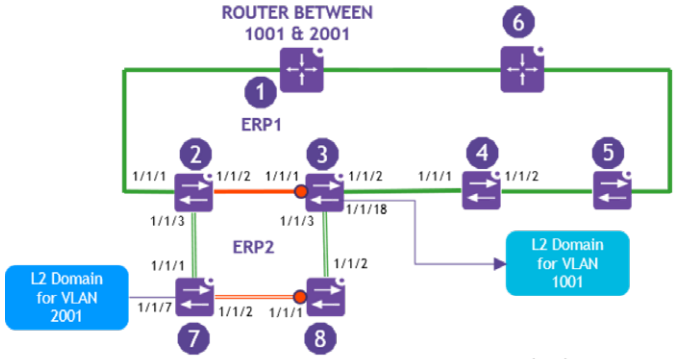
- Mission-critical enterprise core product lines OmniSwitch 9900 and OmniSwitch 6900
- Premium access OmniSwitch 6860E/N
- Advanced value access OmniSwitch 6560 and OmniSwitch 6570M
- Ruggedised and extended temperature OmniSwitch 6865 and OmniSwitch 6465/T

Such a wide range of nodes supporting the ERP protocol can fulfil various requirements for mission-critical networks with L3/L2 capabilities and fast convergence. These nodes can be deployed both indoors and outdoors, in extreme temperatures, and in harsh environments.

9. Appendix: ERP configuration

The network example in **Figure 13** is a ring topology with a major ring running ERP instance #1 and a sub-ring running ERP instance #2. The user VLAN domain 1001 is connected to node #3, which is also the RPL owner node for the major ring. The RPL owner for the sub-ring is node #8. The user VLAN domain 2001 is connected to node #7. Nodes 2 and 3 act as interconnection nodes, and routing between L2 domains is performed on node #1. The figure also shows configuration details related to the ERP configuration.

Figure 13: ERP configuration details of major ring and sub-ring



/ERP ring node #5
! ERP:
erp-ring 1 port1 1/1/1 port2 1/1/2 ~
~ service-vlan 1000 level 1
erp-ring 1 enable
! VLAN:
vlan 1000 members port 1/1/1-2 tagged
vlan 1001 members port 1/1/1-2 tagged
vlan 2000 members port 1/1/1-2 tagged
vlan 2001 members port 1/1/1-2 tagged

/ERP ring node #2
! ERP:
erp-ring 1 port1 1/1/1 port2 1/1/2 service-vlan 1000 level 1
erp-ring 2 sub-ring-port 1/1/3 service-vlan 2000 level 1
erp-ring 1 enable
erp-ring 2 enable
! VLAN:
vlan 1000 members port 1/1/1-2 tagged
vlan 1001 members port 1/1/1-2 tagged
vlan 2000 members port 1/1/1-3 tagged
vlan 2001 members port 1/1/1-3 tagged

/ERP ring node #3
! ERP:
erp-ring 1 port1 1/1/1 port2 1/1/2 service-vlan 1000 level 1
erp-ring 2 sub-ring-port 1/1/3 service-vlan 2000 level 1
erp-ring 1 rpl-node port 1/1/1
erp-ring 1 wait-to-restore-timer 1
erp-ring 1 enable
erp-ring 2 enable
! VLAN:
vlan 1000 members port 1/1/1-2 tagged
vlan 1001 members port 1/1/1-2 tagged
vlan 1001 members port 1/1/18 tagged
vlan 2000 members port 1/1/1-3 tagged
vlan 2001 members port 1/1/1-3 tagged

/ERP ring node #7
! ERP:
erp-ring 2 port1 1/1/1 port2 1/1/2 service-vlan 2000 level 1
erp-ring 2 enable
! VLAN:
vlan 2000 members port 1/1/1-2 tagged
vlan 2001 members port 1/1/1-2 tagged
vlan 2001 members port 1/1/7 tagged

/ERP ring node #8
! ERP:
erp-ring 2 port1 1/1/1 port2 1/1/2 service-vlan 2000 level 1
erp-ring 2 rpl-node port 1/1/1
erp-ring 2 wait-to-restore-timer 1
erp-ring 2 enable
! VLAN:
vlan 2000 members port 1/1/1-2 tagged
vlan 2001 members port 1/1/1-3 tagged

Alcatel·Lucent Enterprise