Allied Telesis™

EPSRing™

AlliedWare Plus™
OPERATING SYSTEM

Ethernet Protection Switched Ring (EPSRing™)

Introduction

Ethernet Protection Switched Ring is a protection system that prevents loops within Ethernet ring-based topologies. EPSR offers rapid detection and failover recovery rates of less than 50 milliseconds, a rate that is equivalent to that provided by circuit-switched equipment.

EPSR's rapid recovery rate enables disruptions in service to go unnoticed in voice, video, or data (Triple-Play). This speedy recovery makes EPSR a more effective alternative to slower spanning tree options when creating resilient Ethernet networks.

It is important to prevent data loss in converged service networks, especially as Triple-Play data is all delivered through the same access network. EPSR's simple and reliable design ensures high reliability, and therefore very high uptime in service networks.

This white paper provides an overview of the Allied Telesis Ethernet Protection Switched Ring technology, as a technical solution for ensuring reliability in an Ethernet-based access network.

Contents

EPSR Overview

IP over Ethernet is now a well-proven technology in the delivery of converged services. Ethernet-based Triple-Play services have become an established commercial reality world-wide, with service providers offering advanced voice, video, and data packages to their customers.

Network infrastructure must be highly available and perform well under heavy loads for service providers to meet both service level agreements and the expectations of customers for a seamless multimedia experience. Business enterprises also demand high availability in the Local Area Network (LAN) to run multiple applications like surveillance, automated control, video streaming and voice over IP (VoIP) right alongside data and Internet access.

EPSR reduces network downtime

The key to keeping networks available is extremely fast failover in the event of link or node failure.

Many Ethernet networks utilize Spanning Tree Protocol (STP) and Rapid Spanning Tree Protocol (RSTP) for preventing loops and assuring backup paths are available. Both protocols are slow to respond to network failures—30 seconds or more. Today's networks need a technology that performs better than either STP or RSTP.

EPSR is Allied Telesis' premier solution for providing extremely fast failover between nodes in a resilient ring. EPSR enables rings to recover within as little as 50ms, preventing a node or link failure from affecting customer experience, even with demanding applications such as IP telephony and streaming video.

EPSR can protect more complex topologies than just a single ring. With the SuperLoop extensions, EPSR can protect a network consisting of any number of rings with multiple points of contact.

Allied Telesis' EPSR is extremely flexible and interoperates with other standard Ethernet functions.

Simple and reliable

EPSR is a simple protocol, and therefore reliable.

▶ There is no negotiation of the roles that nodes play in the ring, they are statically configured as either Master or Transit, and that is it.

▶ The protocol does not need to discover neighbors, and establish relationships with those neighbors.

▶ The topology of an EPSR domain is just a single ring that uses a few simple packet types to check if a link is up or down. A single health check token can be sent out by the Master on its primary port, and there is a single, definite path that this token will follow to reach the Master's Secondary port.

▶ Almost all the decision making is done by the Master. Transit nodes just let the Master know what they have seen (links go down or links come up) and then await instructions from the Master.

Flexible

▶ SuperLoop provides flexibility in the number of interconnected rings on which EPSR can be used.

▶ EPSR is not tied to any particular type of Ethernet interface—it works on copper and fiber, at all data speeds, and can work on aggregated links.

▶ Nested VLANs (Q-in-Q) can be carried on an EPSR ring.

▶ EPSR can be used in a ring with a mixture of link types and link speeds.

Cost-effective and scalable

▶ EPSR is not confined to the high-end Allied Telesis equipment, but is implemented on lower-end equipment, as it does not need specialised hardware.

▶ There are no limits on the number of nodes in an EPSR ring.

▶ There are no limits on the distance between nodes on an EPSR ring.

▶ EPSR provides a highly reliable core network connection over 1 meter or 100 kilometers.

EPSR Key Concepts

Terminology

Before continuing with a discussion of EPSR, here are the definitions of some of the terms that will be used:

EPSR domain
An EPSR Domain is a collection of data VLANs, a control VLAN, and the associated switch ports, defined on a set of switches connected in a ring.

Ring port
Each node in an EPSR ring has two ports connecting it to the ring. These are referred to as the node's Ring ports.

Primary port
The Primary port of the Master node is the Ring port that is always active and forwarding. It determines the direction of traffic flow.

Secondary port
The Secondary port of the Master node remains active, but blocks all data VLANs from operating until ring failover.

Transit nodes
The Transit nodes operate as conventional Ethernet bridges, but with the additional capability of running the EPSR protocol. This protocol requires the Transit nodes to forward the Healthcheck messages from the Master node and respond appropriately when a ring fault is detected.

Master node
The Master node controls the ring operation. It issues Healthcheck messages (also known as Hello messages) at regular intervals from its Primary port and monitors their arrival back at its Secondary port—after they have circled the ring.

Under normal operating conditions the Master Node's Secondary port is always in the blocking state to all data VLAN traffic.

This is to prevent data loops forming within the ring. This port, however, operates in the forwarding state for the traffic on the control VLAN. Loops do not occur on the control VLAN because the control messages stop at the Secondary port, having completed their path around the ring.

Control VLAN
The function of the control VLAN is to monitor the ring domain and carry the packets that perform its signalling functions. The control VLAN carries no user data.

Data VLAN
The data VLANs carry the user data around the ring. Several data VLANs can share a common control VLAN.

Healthcheck messages
The Master node issues Healthcheck messages at regular periods from its Primary port as a means of checking the condition of the EPSR network ring. A failover timer (also known as Hello timer) is set each time a Healthcheck message leaves the Master node's Primary port. The Master node continues sending Healthcheck messages and increments the Hello Sequence number with each message.

If the failover timer expires before the Master node's Secondary port receives the transmitted Healthcheck message, the Master node assumes that there is a fault in the ring, and implements its fault recovery procedures.

Pre-forwarding

The term for the state of an Ethernet interface that was previously down, but that is now blocked and awaiting a control message to unblock.

Ring-Down-Flush

When a Master node detects an outage somewhere in the ring, it sends a Ring-Down-Flush message to its nodes telling them to delete the entries in their forwarding databases.

Ring-Up-Flush

Once a fault in the ring or node has been rectified, the Master sends a Ring-Up-Flush message to the Transit nodes telling them to change their port states from blocking to forwarding (if necessary) and to delete entries in their forwarding databases. This restores normal conditions and allows data to flow again.

EPSR ring recovery in a nutshell

The most important feature of EPSR is fault recovery. EPSR has two ring recovery mechanisms:

1. Slow recovery—a ring break is detected by loss of periodic Healthcheck packets.
2. Fast recovery—nodes either side of a broken link inform the Master, and it performs a recovery procedure immediately.

The structure of the protocol is focused on enabling these recovery processes, especially the fast recovery, to operate reliably, without false positives.

Slow Recovery - Master node polling fault detection

This method uses the Healthcheck process (as described on page 5).

This can be a relatively slow detection method, because it depends on how often the node sends Healthcheck messages.

Fast Recovery - Transit node unsolicited fault detection

To speed up fault detection, EPSR Transit nodes directly communicate when one of their interfaces goes down. When a Transit node detects a fault at one of its interfaces, it immediately sends a Link-Down message over the link that remains up. This notifies the Master node that the ring is broken and causes it to respond immediately. This is the mechanism that can achieve ring recovery as quickly as 50ms.

When a node or link fails, EPSR detects the failure rapidly and responds by unblocking the blocked port so that data can flow around the ring.

This is the prime method of fault detection. The Healthcheck method is a backup for this mechanism.

How EPSR Works

EPSR operates on physical rings of switches, not on meshed networks. In EPSR, each ring of switches forms an EPSR domain. Each EPSR domain has a single Master node and the other switches are called Transit nodes. Each node connects to the ring via two Ethernet ports.

On the Master node one port is configured to be the Primary port and the other is the Secondary port. When all the nodes in the ring are up, EPSR prevents loops by blocking the data VLANs on the Master's Secondary port.

The Master node does not need to block any port on the control VLAN because loops never form on the control VLAN. This is because the Master node never forwards any EPSR messages that it receives in the control VLAN.

The following diagram shows a simple single ring with all the switches in the ring up. The ring comprises one Master node and a number of Transit nodes.

Although a physical ring can have more than one domain, each domain must operate as a separate logical group of VLANs and must have its own Master node. This means that several domains may share the same physical network, but must operate as logically separate VLAN groups.

Simple single ring with all the switches in the ring up
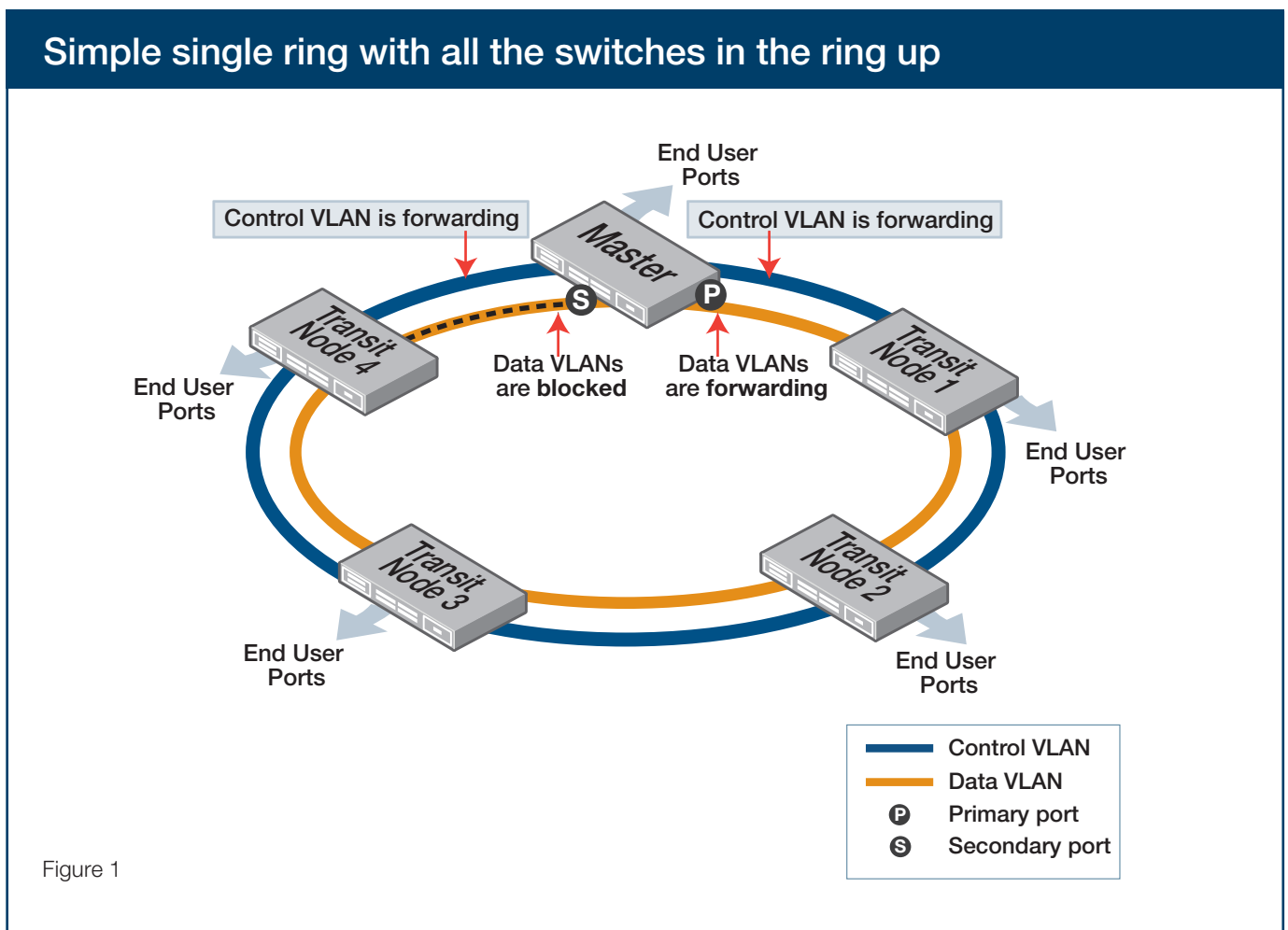
Figure 1

Normal ring operation

Once EPSR is configured on the switches, the following steps complete the EPSR ring:

1. The Master node creates an EPSR Healthcheck message and sends it out the Primary port. This increments the Master node's Transmit Healthcheck counter.
2. The first Transit node receives the Healthcheck message on one of its two Ring ports and sends the message out its other Ring port.

> **Note** Transit nodes never generate Healthcheck messages, only receive and forward them with their switching hardware. This does not increment the Transit node's Transmit Health counter.

3. The Healthcheck message continues around the rest of the Transit nodes.
4. The Master node eventually receives the Healthcheck message on its Secondary port. Because the Master received the Healthcheck message on its Secondary port, it knows that all links and nodes in the ring are up.
5. When the Master node receives the Healthcheck message back on its Secondary port, it resets the Failover timer. If the Failover timer expires before the Master node receives the Healthcheck message back, it concludes that the ring must be broken.

The Master node does not send that particular Healthcheck message out again. If it did, the packet would be continuously flooded around the ring. Instead, the Master node generates a new Healthcheck message when the next Healthcheck cycle begins.

Fault Detection and Recovery

Recovering from a fault

Fault in a link or a Transit node

When the Master node detects an outage somewhere in the ring, using either the slow or fast detection methods described on page 6, it restores traffic flow by:

1. Declaring the ring to be in a Failed state.
2. Unblocking its Secondary port, which enables data VLAN traffic to pass between its Primary and Secondary ports.
3. Flushing its own Forwarding Database (FDB) for the two Ring ports.
4. Sending an EPSR Ring-Down-Flush-FDB control message to all the Transit nodes, via both its Primary and Secondary ports.

The EPSR fault recovery process

3 Secondary port transitions to forwarding

2 Master node state changes to FAILED

5 Send Ring-Down-Flush FDB message

5 Send Ring-Down-Flush FDB message

4 Flush FDB on ring ports

6 Transit nodes respond to Ring-Down-Flush FDB and flush FDB ring port entries

6 Transit nodes respond to Ring-Down-Flush FDB and flush FDB ring port entries

1 Link goes down

— Control VLAN
— Data VLAN
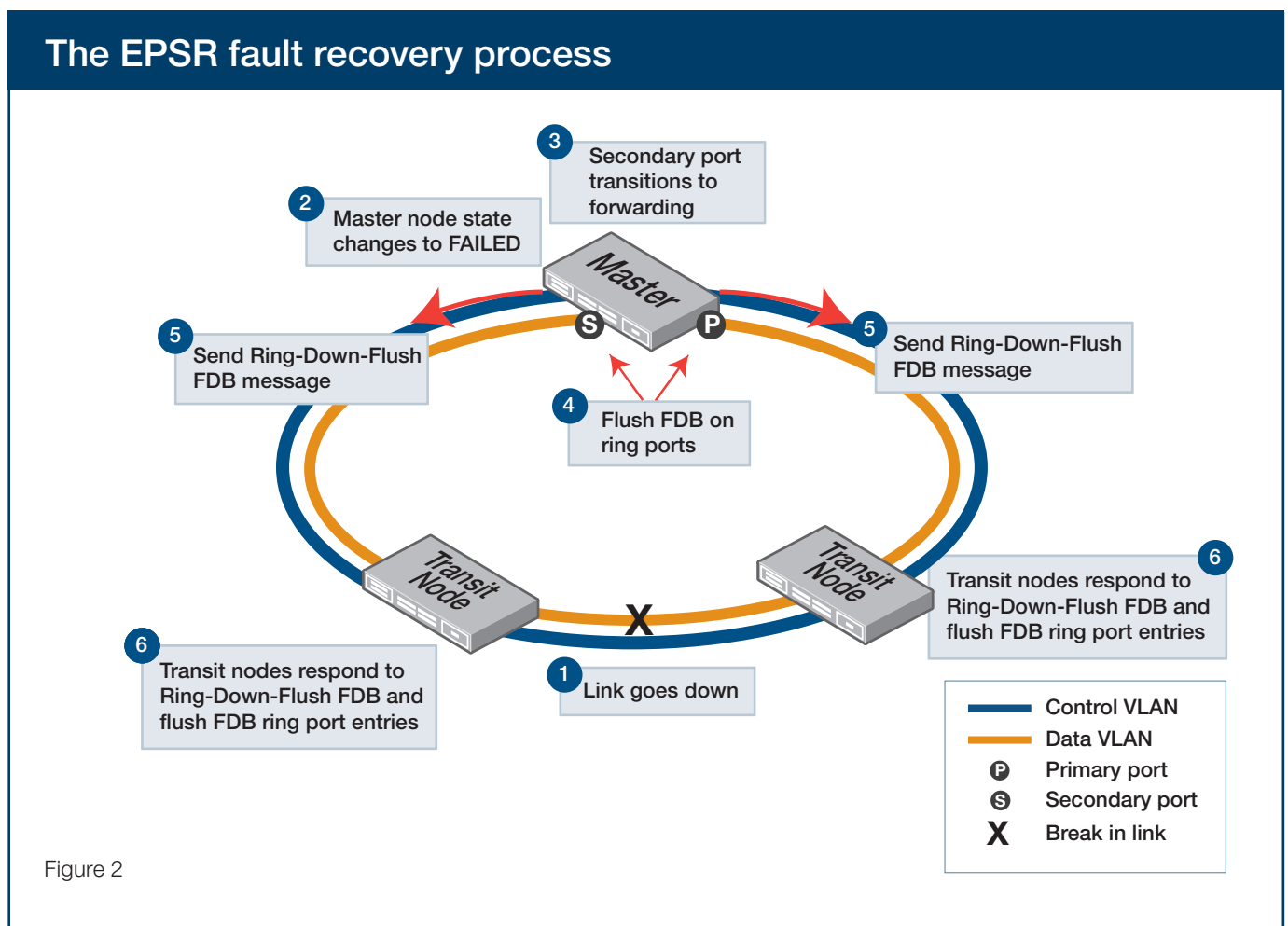Ⓟ Primary port
Ⓢ Secondary port
X Break in link

Figure 2

The Transit nodes respond to the Ring-Down-Flush-FDB message by flushing their forwarding databases for each of their Ring ports. As the data starts to flow in the ring's new configuration, the Master and Transit nodes re-learn their Layer 2 addresses. During this period, the Master node continues to send Healthcheck messages over the control VLAN. This situation continues until the faulty link or node is repaired.

Fault in the Master node

If the Master node goes down, the Transit nodes simply continue forwarding traffic around the ring. Their operation does not change.

The only observable effects on the Transit nodes are that:

▶ they stop receiving Healthcheck messages and other messages from the Master node.

▶ the Transit nodes connected to the Master node experience a broken link, so they send Link-Down messages. When the Master node is down these messages are simply dropped.

Neither of these symptoms affects how the Transit nodes forward traffic. Once the Master node recovers, it continues its function as the Master node.

Enhanced Recovery

A Transit node port enters the Pre-forwarding state when the Ring port becomes physically available. Enhanced Recovery can speed a node's recovery from the Pre-forwarding state to full forwarding.
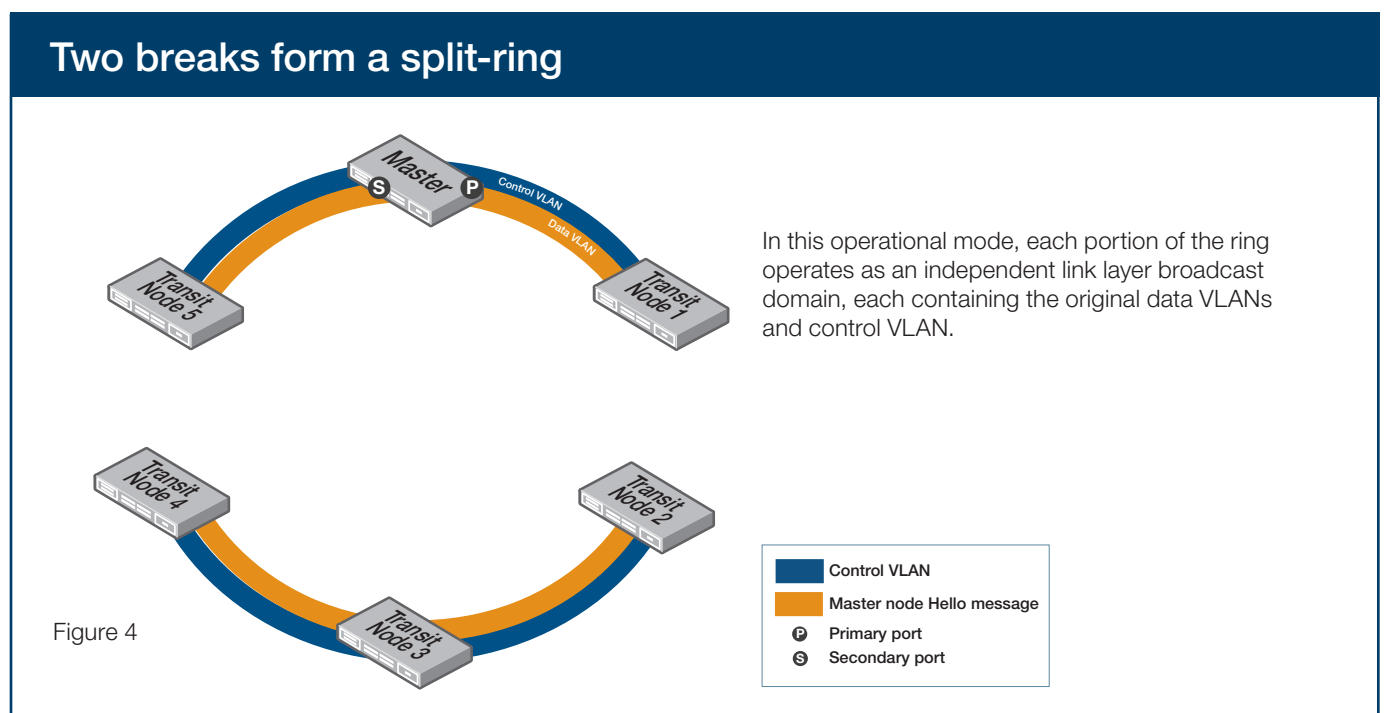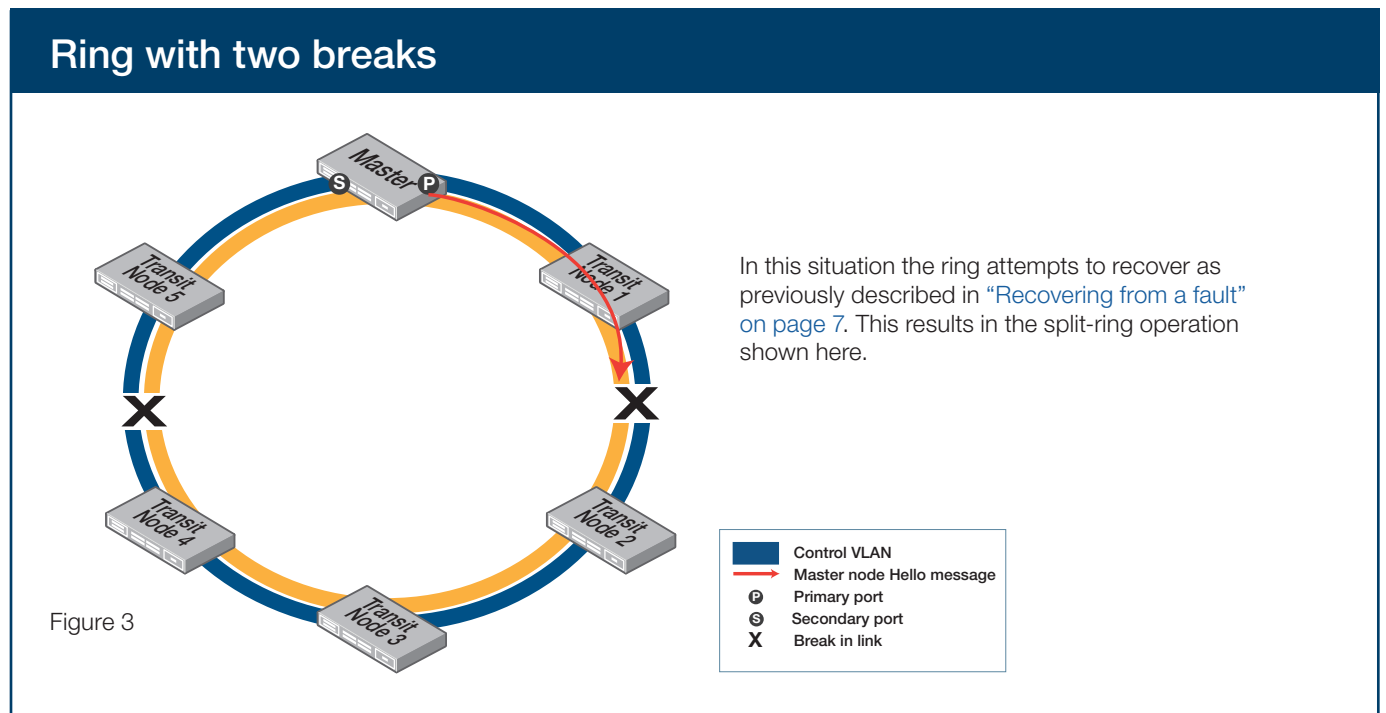
With Enhanced Recovery, the Transit node port can exit the Pre-forwarding state without the entire ring becoming complete. It does this in one of two ways:

1. When entering the Pre-forwarding state, the Transit node sends a Link-Forward-Request message and waits for a response from the Master node. When the Master receives this message, it sends a special Healthcheck message. If the Master does not receive the Healthcheck back within a given period, the Master sends a Permission-Link-Forward message to the Transit node. The Transit node can then take the port from Pre-forwarding to forwarding.

2. If the Transit node doesn't receive a Permission-Link-Forward message within a given period, it makes the decision that the Master is not reachable and starts forwarding anyway.

Without Enhanced Recovery, the Transit node port waits in the Pre-forwarding state until it receives the Ring-Up-Flush message from the Master. This occurs when the Master receives back its Healthcheck messages, and the ring is declared complete.

Managing rings with two breaks

The EPSR Enhanced Recovery feature automatically restores a link when a ring has suffered two breaks. Consider the network shown below:

Ring with two breaks

In this situation the ring attempts to recover as previously described in "Recovering from a fault" on page 7. This results in the split-ring operation shown here.

Legend:
- Control VLAN
- Master node Hello message
- Ⓟ Primary port
- Ⓢ Secondary port
- X Break in link

Figure 3

Two breaks form a split-ring

In this operational mode, each portion of the ring operates as an independent link layer broadcast domain, each containing the original data VLANs and control VLAN.

Legend:
- Control VLAN
- Master node Hello message
- Ⓟ Primary port
- Ⓢ Secondary port
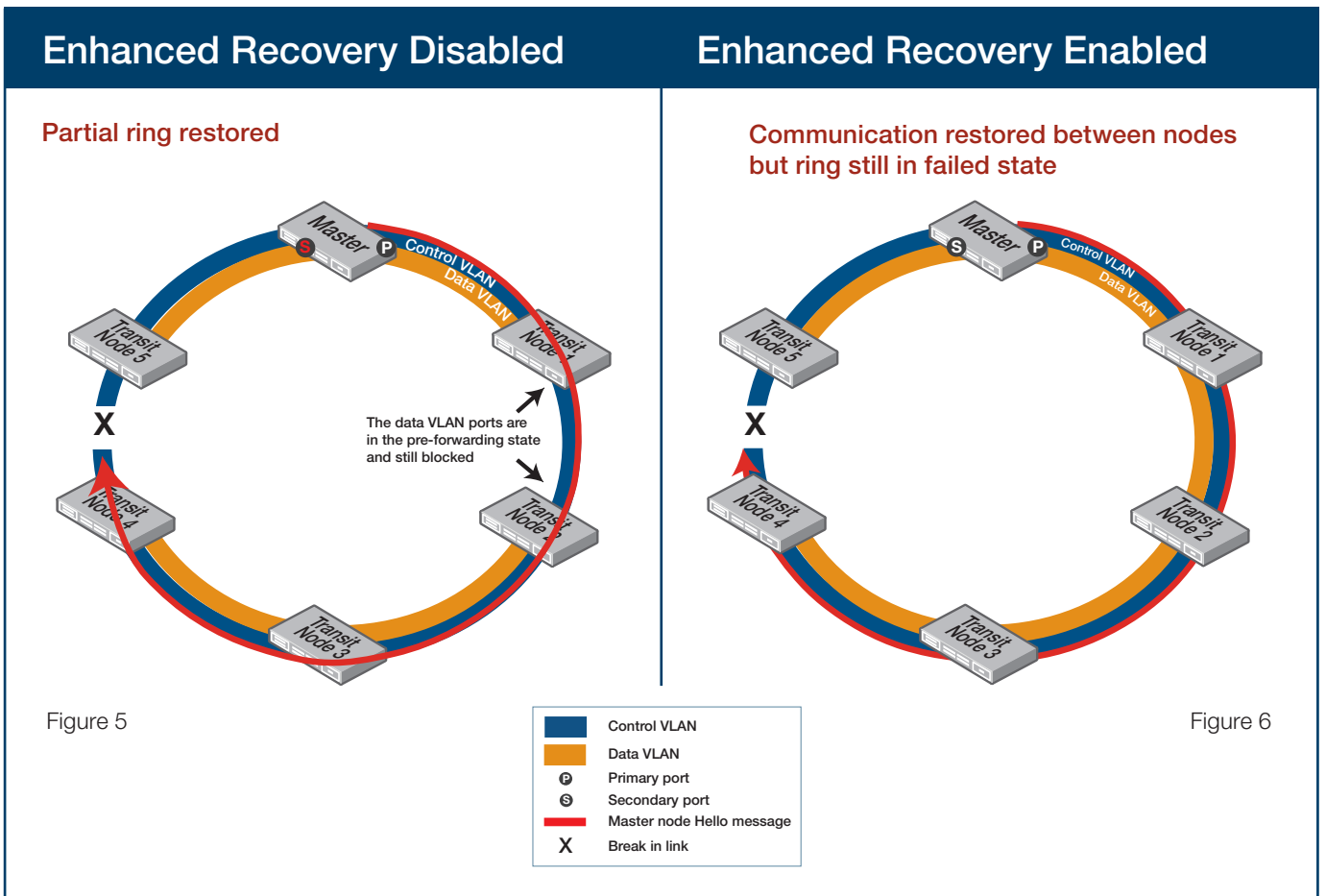
Figure 4

Recovery when one break is restored

Figure 5 below shows a ring with the link between nodes 1 and 2 restored. At this point the ring's behaviour depends on whether EPSR Enhanced Recovery has been enabled.

Enhanced Recovery disabled

With the enhanced recovery feature disabled, the ports either side of the restored link will remain in a pre-forwarding state. From a user's perspective, the ring will remain as shown in the split state shown in Figure 5 .

Enhanced Recovery enabled

With the Enhanced Recovery feature enabled, switch nodes 1 and 2 can detect the restored link, and will place all their Ring ports in the forwarding state. Although the ring remains in the "failed" state because of the remaining break; communication between the nodes is restored. The network then operates as shown in Figure 6.

Enhanced Recovery Disabled

Partial ring restored

The data VLAN ports are in the pre-forwarding state and still blocked

Figure 5

Enhanced Recovery Enabled

Communication restored between nodes but ring still in failed state

Figure 6

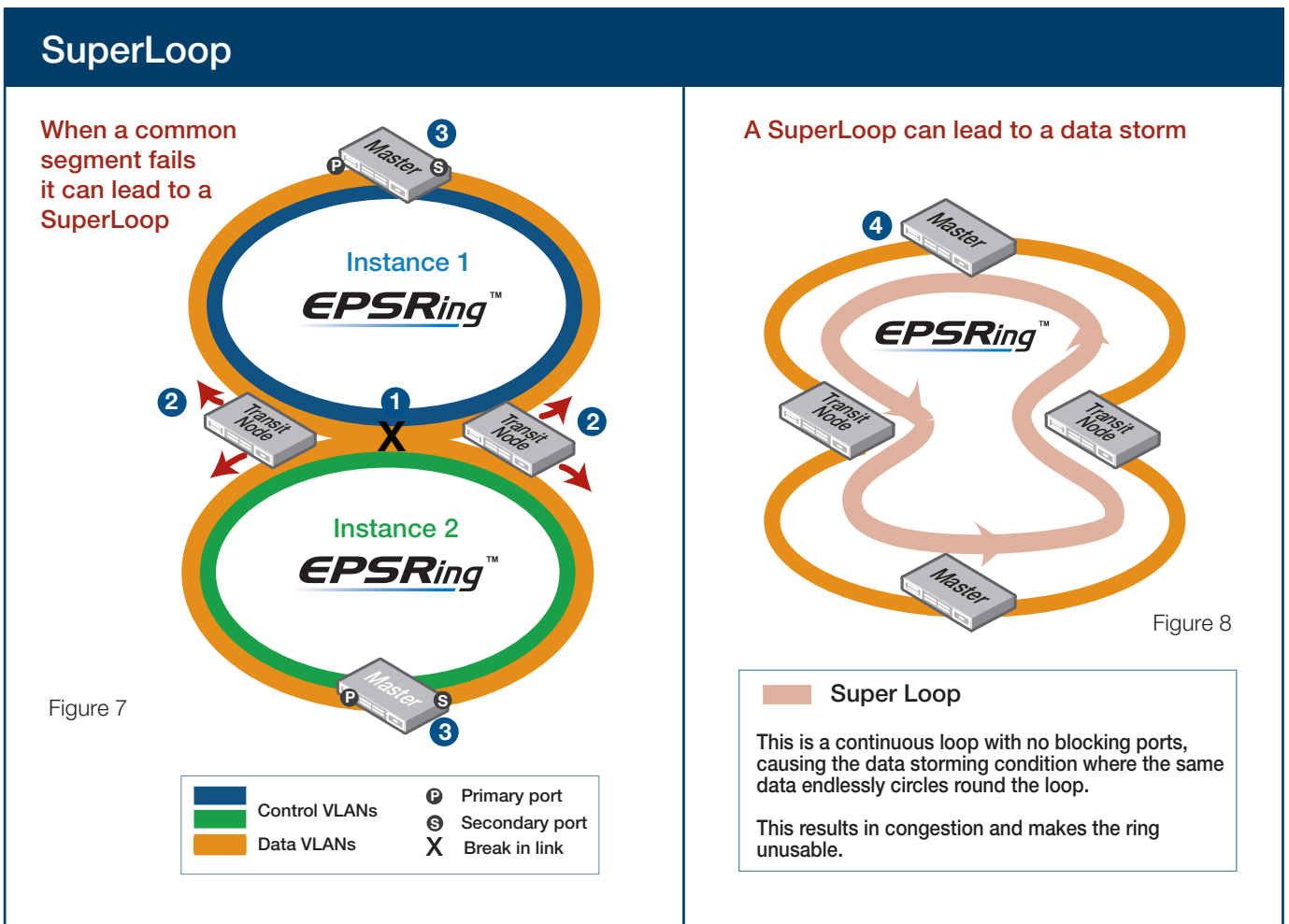|	
▬	Control VLAN
▬	Data VLAN
Ⓟ	Primary port
Ⓢ	Secondary port
▬	Master node Hello message
X	Break in link

EPSR SuperLoop Prevention

In a network that consists of multiple EPSR rings, it is desirable for adjacent rings to have at least two points of contact, to avoid their junction becoming a single point of failure. As a result of the rings touching at more than one point, they share a common segment. This is the section of network where both rings share the same links. If two adjacent rings share the same set of data VLANs and their common segment fails, a loop forms. This loop is known as a SuperLoop.

The sequence of events without EPSR-SLP as shown in Figure 7 and Figure 8:

1. The common link goes down.
2. The Transit nodes at each end of the common link send Link Down messages to both Master nodes.
3. The Master nodes both unlock their Secondary ports, which results in a loop.
4. Data circulates continuously around this loop, congesting the network.

SuperLoop

When a common segment fails it can lead to a SuperLoop

Instance 1
EPSRing™

Instance 2
EPSRing™

Figure 7

Control VLANs	Ⓟ Primary port
Data VLANs	Ⓢ Secondary port
	X Break in link

A SuperLoop can lead to a data storm

Figure 8

■ **Super Loop**

This is a continuous loop with no blocking ports, causing the data storming condition where the same data endlessly circles round the loop.

This results in congestion and makes the ring unusable.

How does EPSR SuperLoop Prevention work?

EPSR SuperLoop Prevention (EPSR-SLP) is an enhancement to the original EPSR.

EPSR-SLP prevents SuperLoops forming in the following way:

1. A priority is assigned to each EPSR ring between 0 and 127, with 1 representing the lowest priority and 127 the highest. A priority of 0 (the default setting) applies the functionality of no SuperLoop prevention.
2. It ensures that common segment Transit nodes send Link Down messages only to the Master of the highest priority ring.
3. When a link in a common segment goes down, only the Master of the highest priority ring opens its Secondary port, because this is the Master node that will receive the Transit node's Link Down message.

Conclusion

Ethernet is becoming the universal data transport medium in the world. With the diversification of the applications for which Ethernet data transport is used, the requirements on the performance of Ethernet grow. In particular, the application of Ethernet to the transport of real time video and voice communications has required that Ethernet provides extremely fast fault recovery mechanisms. The failover in the event of link or node failure needs to be so fast as to be barely perceptible to the human eye or ear.

Moreover, a highly fault-tolerant Ethernet network must also continue to provide the key advantages of Ethernet—cost effectiveness, simplicity, flexibility.

Allied Telesis' EPSR provides an effective solution to address this challenge. It monitors the ring's domain and maintains operational functions, to detect faults and respond immediately to achieve ring recovery as quickly as 50ms. It enables network engineers to manage multiple data VLANs using a simple, reliable protocol. It is flexible, cost-effective, scalable, and greatly improves network resiliency. Allied Telesis' EPSR ensures Ethernet ring networks can handle the stress of providing voice, video, and data services.

About Allied Telesis

For nearly 30 years, Allied Telesis has been delivering reliable, intelligent connectivity for everything from enterprise organizations to complex, critical infrastructure projects around the globe.

In a world moving toward Smart Cities and the Internet of Things, networks must evolve rapidly to meet new challenges. Allied Telesis smart technologies, such as Allied Telesis Management Framework™ (AMF) and Enterprise SDN, ensure that network evolution can keep pace, and deliver efficient and secure solutions for people, organizations, and "things"—both now and into the future.

Allied Telesis is recognized for innovating the way in which services and applications are delivered and managed, resulting in increased value and lower operating costs.

Visit us online at **alliedtelesis**.com

Allied Telesis™   **NETWORK SMARTER**