WHITE PAPER

Ethernet Automatic Protection Switching (EAPS)

Table of Contents

Introduction

The need for business continuity has placed a greater demand on today's data networks – redundancy and reliability are imperative and the network must be able to support them. The network infrastructure must be able to achieve a high availability environment and continuous access to resources. For this reason the networking industry has relied on the Spanning Tree Protocol (STP) in large Layer 2 networks to provide a certain level of redundancy. However, STP has proven inadequate to provide the level of resiliency required for real-time and mission critical applications. It is important to note that the entire industry has recognized that a new technology is needed to replace STP and many vendors are in the process of developing pre-standard technologies to meet that requirement.

Technology Overview

Ethernet Automatic Protection Switching (EAPS) is Extreme Networks' solution for fault-tolerant Layer 2 ring topologies. EAPS is responsible for a loop-free operation and a sub-second ring recovery. This revolutionary technology provides end users with a continuous operation usually only available in voice networks. While EAPS provides an advanced function, it does so with radical simplicity. The real strength of EAPS comes from its ability to integrate into existing and new networks to solve real business issues. EAPS can be built using Ethernet, WDM, vDSL, and WAN technologies, or any combination thereof. Furthermore, EAPS is native to all "i" based Extreme switches, making it readily available and not requiring expensive hardware upgrades.

Definition of EAPS Terms

A ring is made up of two or more switches. One of the nodes on the ring is designated as master (S1) as shown in Figure 1. The two ring ports on the Master node are configured as primary port (P) and secondary port (S) respectively. All other nodes on the ring (S2-S6) are designated as transit, which are also configured with their respective primary and secondary ports.

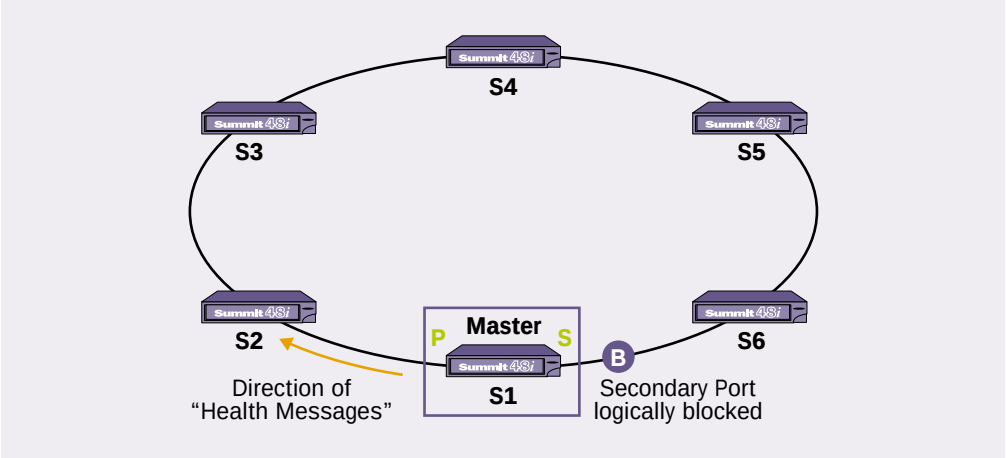

Figure 1: EAPS Ring Elements

An EAPS domain is configured to protect a group of data-carrying virtual local area networks (VLANs), called protected VLANs as shown is Figure 2. There could be multiple EAPS domains running on the same switch, each with its unique control VLAN. Similarly, many domains can co-exist on the same ring protecting different sets of VLANs.

A control VLAN is created per EAPS domain. This control VLAN is for the purpose of sending and receiving EAPS messages.

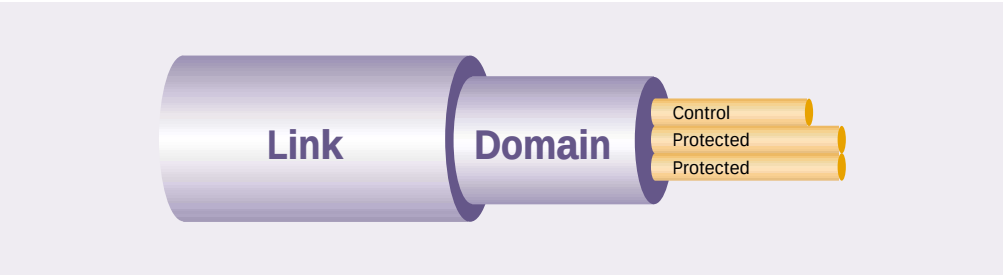

Figure 2: Domain and VLAN Relationship

Normal Operation

EAPS operates by declaring an EAPS domain on a single ring. Any VLAN that warrants fault protection is configured on all ring ports in the ring, and is then assigned to an EAPS domain. On that ring domain, a node is designated as the master, and one of its two ring ports is designated as the primary port and the other as the secondary port. The master node blocks the secondary port for all non-control traffic belonging to this EAPS domain, thereby avoiding a loop on the ring. Layer 2 switching & learning mechanisms operate as normal. The control VLAN is not blocked at the master secondary port and control traffic is allowed to flow through. The master sends out periodic poll packets from its primary port on the control VLAN to be received on the secondary port, thus ensuring that the ring is up.

Fault Detection

When a fault occurs on the ring as shown in Figure 3, the master detects it either by missing poll packets or by special fault-detection packets (traps) generated by the nodes that detect the fault. Upon learning of a fault, the master unblocks its secondary port allowing protected VLAN traffic through.

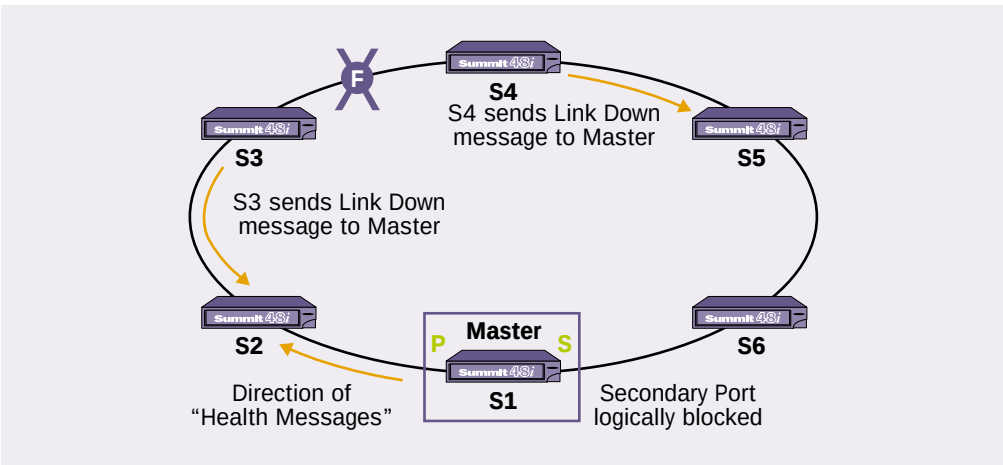

Figure 3: Fault Detection

Fault detection is accomplished in one of two ways:

Trap message sent by a transit switch

When a transit switch detects any of its ring ports losing link, it immediately sends a "link-down" message to the master on its good link via the control VLAN. When the master receives this link-down message, it immediately declares failed state, and opens the logically blocked protected VLANs on the secondary port. It also flushes its forwarding database, or FDB, and sends a "flush FDB" message to all other transit switches on the ring via the control VLAN. The other nodes on the ring need not be aware of the fault; they simply flush their FDB on all VLANs belonging to this domain. The destination switching decisions are then re-learned following the normal Layer 2 learning mechanisms.

Polling

Polling is the failsafe method for ring recovery. If for any reason the traps from the transit nodes do not reach the master node for immediate recovery, the polling mechanism will force a recovery in a few seconds.

During normal operation, the master node sends out a "health" packet every hellotime milliseconds on the control VLAN. If the ring is complete, the master will receive the packet on its secondary port (control VLAN is not blocked on this port). When the master receives the health packet, it resets its failtimer and remains in complete state.

If the master doesn't get the health packet before failtimer times out, it declares failed state, and performs same operations as describes above, which are: unblock (open) the logically blocked protected VLANs on the secondary port; flush FDB; send "flush FDB" message to all transit switches.

Fault Restoration

The Master continues to send "health" messages out on its primary port even if the state is failed (i.e. the ring is broken). As long as there is a break in the ring, the master's failtimer will keep timing out, and it will remain in the failed state.

When the broken link is restored, the master gets its "health" message back on its secondary port, and declares the ring to be complete. It will then perform the standard ring complete operations: logically blocking (closing) the protected VLANs on the secondary port and flushing the FDB on all transit switches.

During recovery, from the time the link goes up on the transit switch until the master detects ring complete state, the transit node must not begin forwarding traffic until the master secondary port in blocked. Otherwise, a temporary loop may occur due to having all ports forwarding traffic on the ring. To rectify this condition, EAPS takes the following steps on the transit node:

- Put all the protected VLANs on the repaired port in a blocked state
- Remember which port has been temporarily blocked
- Set its state to preforwarding

When the master node detects the ring is up via its polled "health" message, it sends a "flush FDB" message to all the transit switches. When the transit switches receive this "flush FDB" message, they perform the following steps:

- Flush FDBs on protected VLANs
- If the state is set to preforwarding, begin forwarding on all the protected VLANs on that port

Continuous Operation

While sub-second fault detection and recovery is good enough for some applications, it is not reliable enough for others when operating alone. Some applications rely on a higher-level protocol to retransmit and recover from a fault, therefore selecting a new path and getting redirected. However, applications that do not rely on acknowledgement from the remote end, like multicast, need an intelligent network protocol to help with fast recovery.

EAPS does just that. EAPS adds intelligence to the network to help multicast streams get redirected around a broken link with blazing speed, resulting in an uninterrupted multicast service. This is the type of traffic that usually runs over a university distance-learning program, corporate voice-over-IP network, or service provider video broadcast. With real-time and mission critical applications such as these, EAPS is the only choice for non-stop operation. All other protocols cause multicast clients to timeout or hang. Not only is the interruption noticeable but it also requires user intervention to get the service restarted. EAPS reduces overall business interruption and improves availability.

Multiple EAPS Domains Per Ring

Remember, each EAPS domain has its own master node, its own unique control VLAN, and its own group of protected VLANs. Different EAPS domains could have their masters on the same switch or on different switches. Furthermore, multiple EAPS domains may coexist on the same ring. This feature allows EAPS to take advantage of available resources and bandwidth on the ring, called spatial reuse. It provides the flexibility to control each group of VLANs independently, therefore utilizing ring

bandwidth more efficiently. For instance, blocking the secondary port on a master node in one domain renders that link useless (standby), but forwarding on that link in a different domain takes advantage of that bandwidth. In addition, a domain may contain VLANs with clients in close proximity allowing more direct paths between nodes and controlling the direction of traffic flow.

VLANs in Multiple EAPS Domains (Multiple Rings)

A data VLAN could span 2 rings interconnected by a common switch as shown in Figure 4. Each ring is configured with a master node, an EAPS domain, and a control VLAN. The data VLAN that is spanning both rings would be added as a protected VLAN to both EAPS domains.

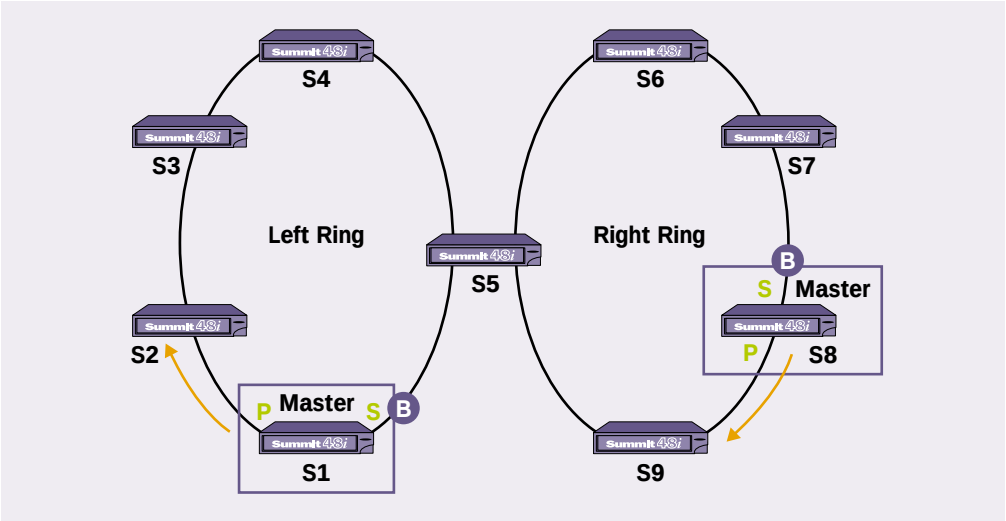

Figure 4: VLANs in multiple EAPS Domains

Figure 5 shows a more complex setup, where 2 switches, sharing a common link, interconnect the 2 rings. This setup adds resiliency between the two rings in case one of the common switches fails. A problem arises if the common link breaks. The master of each ring will open their respective secondary port. A protected VLAN spanning both rings will have a super-loop caused by this break (S1-S2-S3-S4-S5-S6-S7-S8-S9-S10-S1). Future code releases will address super-loop protection. In the meantime, either use STP in conjunction with EAPS or make sure the common link never fails unless a common switch fails (S5 or S10). One way to improve link uptime is to place S5 and S10 at the same site and aggregate multiple links between them, using different modules if possible.

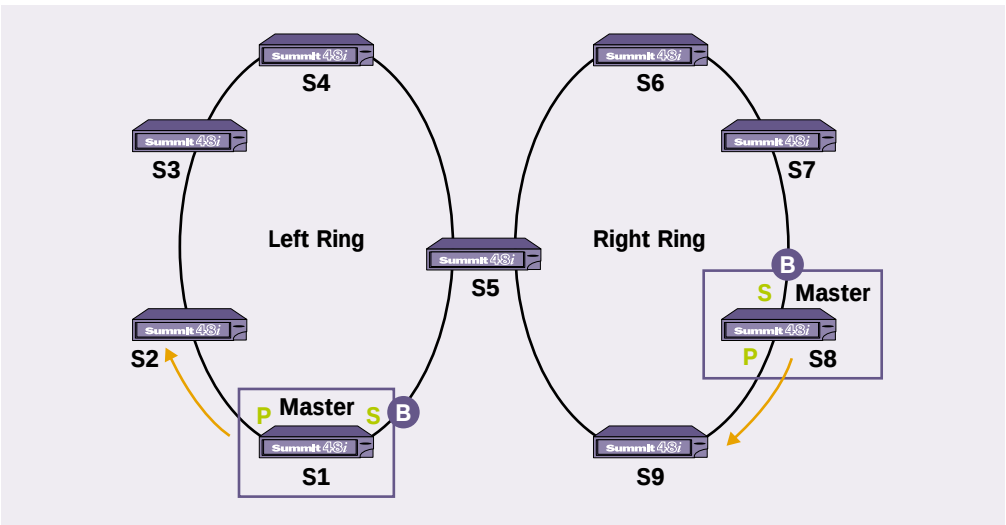

Figure 4: VLANs in multiple EAPS Domains

Design Considerations

EAPS was created to solve slow recovery times inherent to STP, in essence replacing STP in ring topologies. Although STP and EAPS use a similar mechanism to avoid network loops, EAPS provides much more control, resiliency and flexibility. When designing an EAPS network follow these “best practices” guidelines to achieve the desired results:

- EAPS is a Layer 2 resiliency protocol
- Designed for ring and interconnected ring topologies
- Can coexist with a Layer 3 protocol like VRRP, ESRP, OSPF
- Can coexist with STP – Layer 2
- Can be used in the core or at the edge
- A ring can be built with as few as 2 switches using EAPS
- There is no theoretical maximum on the number of switches on the ring
- Multiple EAPS domains can coexist on a single ring
- Multiple EAPS domains can be defined on a single node
- Only one master can be defined per domain
- An EAPS domain can be defined on only one ring (can not cross rings)
- A maximum of 64 EAPS domains can be defined on a single switch
- A maximum of 64 EAPS domains can be defined on a single ring
- A maximum of 4,096 EAPS VLANs can be defined on a switch
- Both protected and control VLANs are counted towards the maximum VLAN limit
- Works with many technologies, like Ethernet (10, 100, 1000), WDM, vDSL, WAN
- Master node selection should be based on least busiest link (standby secondary port)
- EAPS requires all “i” based switches
- EAPS requires a full Layer 3 license on every switch
- User must configure the control VLAN to use Quality of Service profile QP8
- The control VLAN should not carry data traffic or be assigned an IP address

3585 Monroe Street Santa Clara, CA 95051-1450 Phone 408.579.2800 Fax 408.579.3000
Email info@extremenetworks.com Web www.extremenetworks.com

 L-WP-EAPS-203