802.1aj Two port MAC Relay status

John Messenger

December 2007
Two port MAC Relay

- Industry recognises that full 802.1 bridges are sometimes unnecessarily complex
- TPMR (802.1aj) attempts to provide a simpler relay function than a VLAN bridge
- PAR granted December 2004
  - Initial draft 0.0 May 2005
  - Draft 1.0 July 2005
  - Draft 1.1 August 2005
  - Draft 1.2 November 2005
  - Draft 1.3 May 2006
  - Draft 1.4 June 2006
  - Draft 2.0 January 2007
  - Draft 2.1 May 2007
  - Draft 2.2 October 2007 (Working group ballot)
- Interest from other standards bodies including MEF and DSL Forum
- This presentation represents a personal view of the status following the November 2007 meeting of 802.1 and of how a TPMR might be used.
TPMR topics

- Zero configuration option – should work out-of-the-box
- Topologies
- Link maintenance
- Discovery
- Management
- A TPMR is a bridge
- Forwarding
- MAC types
- Loopback
- Link status propagation
TPMRs can be deployed singly, or in a chain

A typical application might be as a demarcation device (NID)

A TPMR has exactly two ports
  - Each port can be by Ethernet or any MAC or emulated MAC which supports the 802.1 Internal Sublayer Service

Protection is not supported in the draft standard

Management using SNMP over Ethernet
  - Envisaged to be from an intelligent device which proxies the TPMR’s managed objects into its own MIB
For Ethernet links, 802.3ah EFM OAM may be employed
  - This provides an indication of link up/down
    - Ethernet MAC link down indication is notoriously unreliable

802.3ah also provides
  - Link status change information
  - Link statistics including errored seconds etc.
  - Managed object access, which is NOT used in TPMR

E-LMI (MEF UNI Phase 2) was considered, but is not suited as a link maintenance protocol
  - Intended for CE to retrieve status and service attributes from the network
  - Includes UNI and per-EVC configuration and status information

Other MACs and emulated MACs can use their own protocol
A mechanism is required to allow discovery of TPMRs, so that the managing device knows what to manage.

Mandatory CFM (802.1ag) is the primary discovery method:
- At least a level 0 MIP is required in TPMR
- Attached bridge or station can use Linktrace to discover connectivity of attached TPMR chain
- All TPMRs in a chain can be found, but a method is needed to know when the end of the chain has been reached
- CFM tells you what kind of device it is (uncertain if TPMR defined)

LLDP (802.1ab) may be used for further probing:
- LLDP support is optional

Ethernet EFM OAM (802.3ah) could have been chosen for Ethernet links, but is harder to use for chain discovery.
Management

- SNMP over Ethernet, without IP, is mandatory
  - SNMP over IP was rejected because of the desire to avoid IP address management and NMS interaction with individual TPMRs
  - 802.3ah EFM/OAM was rejected because of concerns over scalability to a chain and lack of “Set” capability
  - CORBA was considered too much of a stretch given that nothing else in 802.1 uses it

- Management is required to be supported on at least one of the data ports on the TPMR

- SNMP over Ethernet is specified in RFC4789
  - Untagged frames are used
  - Management VLAN option not yet discussed much

- SNMP over other transports is not precluded
  - For example, traditional SNMP over IP is allowed
Discovery is used first to find what to manage

How remote management is done is not specified, but perhaps
- Retrieved objects are incorporated into the managing device’s MIB
- Incorporation into Interface MIB objects is a possibility

Which ports can be used to manage the device?
- A management block is provided to prevent access from the customer port
- This block can be turned on and off by management

Unspecified issues:
- Is access provided by authentication, to allow a device which is installed the wrong way round to be “recovered” remotely?
- In a device with different port types, which port then?
A TPMR is a kind of 802.1 Bridge

- Only two ports
- No MAC address learning
- No VLAN tagging, but can be priority aware
- No Spanning Tree
  - BPDUs require special treatment (see later slide)
General idea is to be transparent to protocols the TPMR does not implement
- But some protocols are filtered out by the MAC, e.g. Pause
- Transparent to BPDUs
- Transparent to LACP (despite the layering violation)
- One reserved address will be terminated by the TPMR and used for the LAN Status Propagation Protocol

No modification of user data frames (e.g., tagging)

Multiple queues are optional
- Extract priority from Q-tag and 802.1ae LinkSec tag
  - Only for integrity-protected frames – unencrypted
  - Recognise L2 control protocols and place in fastest queue
    - Typically BPDUs

Otherwise like 802.1d/Q
- Note that MRP (802.1ak) needs special handling in a Q-bridge
Loopback (not part of 802.1aj)

- Per-link loopback on Ethernet with 802.3ah EFM/OAM
  - Invoked by SNMP to previous hop
  - Beware that EFM OAM loopback discards returned frames

- Multi-hop CFM-based (802.1ag) loopback
  - Uses a special loopback frame
  - Can contain arbitrary data inside a TLV
  - Non-intrusive, in that user data continues
  - Stateless
  - Issues include VLAN-non-awareness

- Stateful per-VLAN loopback is not supported
  - CFM rejected this idea as not sufficiently useful at resolving data-driven errors
  - Could be provided using an EFM/OAM extension invoked from the previous hop using SNMP
LAN Status Propagation

- 802.1aj incorporates a link loss forwarding function called LAN Status Propagation
- Where TPMRs are placed in between existing bridges, they could interfere with protection and restoration processes
- LAN Status Propagation ensures that changes in connectivity are signalled to the bridges at the ends of a chain of TPMRs
- The signalling is compatible with existing 802.1 bridges
The protocol is aimed at links over which RSTP may be running. It’s important that when the connectivity is announced, it is available.

Another aim is to keep links up when possible, to allow management traffic to a maximal subset of TPMRs in a chain.

Basic idea:
- For loss: send a loss message; if unacknowledged, blip the link;
- For add: send an add message; if unacknowledged, blip the link;
- Blipping the link alerts the next layer that connectivity has changed.

The protocol communicates only changes, not states.
- Link down is not periodically sent.

Optimisations are possible when a TPMR already knows that its neighbour doesn’t speak the LAN Status Propagation Protocol (LSPP)
New Connectivity (general case)

**Key**
- OperUp
- OperDown but locally
- MAC_Enabled
- MAC_Enabled FALSE
- MAC_Operational, but reporting OperDown to clients

**Timers**
- $T_d =$ blip time (longish)
- $T_r =$ retransmit time
- $T_w =$ ACK wait time

Figure 23-5—New connectivity with MAC status notification
In the more general case, TPMRs don’t know whether their link partners speak LSPP and must try it with timeouts.

To illustrate this, consider TPMR2 and its rightward neighbours.

MAC recovers in TPMR1 and TPMR2: MAC_Operational detected.

Both TPMRs enter ⌹ to avoid creating loops.

As TPMRs 2-4 don’t know if their neighbours speak LSPP:
- TPMR2 sends “add” to its right neighbour and starts $T_r$ and $T_w$.
- TPMRs 3 and 4 forward the “add”, ack leftward, and start $T_w$.
- Received acks cancel $T_w$ in TPMRs 2 and 3.
- $T_r$ may repetitively expire in TPMR2, triggering retransmission of “add”.
- $T_w$ will expire in TPMR4, triggering rightwards “blip” and leftwards “confirm”.
- Confirm cancels $T_r$ in TPMR2.
Loss of Connectivity

Key

- OperUp
- OperDown but locally MAC_Enabled
- MAC_Enabled FALSE
- MAC_Operational, but reporting OperDown to clients

Timers

- $T_d$ = blip time (longish)
- $T_r$ = retransmit time
- $T_w$ = ACK wait time

Figure 23-6—Connectivity failure
In the more general case, TPMRs don’t know whether their link partners speak LSPP and must try it with timeouts.

To illustrate this, consider TPMR2 and its rightward neighbours.

Link fails between TPMR1 and TPMR2: MAC_Operational FALSE

Both TPMRs enter.

As TPMRs 2-4 don’t know if their neighbours speak LSPP:
- TPMR2 sends “loss” to its right neighbour and starts $T_r$ and $T_w$
- TPMRs 3 and 4 forward the “loss”, ack leftward, and start $T_w$
- Received acks cancel $T_w$ in TPMRs 2 and 3
- $T_r$ may repetitively expire in TPMR2, triggering retransmission of “loss”
- $T_w$ will expire in TPMR4, triggering rightwards “blip” and leftwards “confirm”
- Confirm cancels $T_r$ in TPMR2
Questions and Answers

- How does LAN state propagation protocol interact with CFM?
  - This question is particularly relevant if RSTP is not being used
  - LSPP tells you “something changed”; CFM on top confirms connectivity
  - CFM could be used underneath to control MAC_Operational

- At initialisation of the system, will the TPMR learn whether its neighbours speak LSPP?
  - When TPMR comes up, its MACs will come up
  - That should trigger “add” messages going out
    - If there are responses, then there’s an LSPP neighbour
    - If not, there isn’t
  - So as long as initialisation is carefully handled, the TPMR will know its neighbours’ LSPP capability
  - This means response to link loss can be immediate
Questions and Answers

- Can one assume that same link partner is present after blipping the link?
  - Even if the link partner changes, the protocol will still work
  - Once you learn your link partner doesn’t speak the protocol, you won’t learn they do if they are changed.
Use with telco-style protection

- LSPP is designed for RSTP-based protection systems
- It uses single events to signal a change
  - “loss” or “add” messages to partners speaking LSPP
  - “blipping” the link to partners who don’t speak LSPP
- These events can trigger protection switching
  - “blipping” will automatically cause switching
  - Receipt of LSPP “loss” messages could do so
- However nothing in LSPP marks a link as down
  - Other events might trigger a switch back to this bad link
    - Examples include a break on the other link or revertive switching back
    - Needs an additional mechanism to verify recovery (e.g. slow CFM)
    - Mark the link as bad until you know it’s good
- LSPP doesn’t have the high traffic load of fast CFM
Thank You

JMessenger@advaoptical.com