OAM in Frames

Denton Gentry, Dominet Systems
Shimon Muller, Sun Microsystems
Dan Romascanu, Avaya Inc
David Law, 3Com
K. Seto, Hitachi Cable
Craig Easley, Extreme Networks
Kevin Daines, World Wide Packets
David Horne, Intel
Steve Jackson, Hatteras
Walt Kutzavitch, Philips
Vincent Bemmel, Alloptic
Ryan Hirth, Terawave Communication

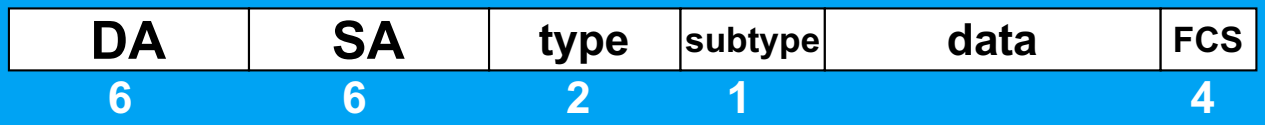
Send comments to: denny@dominetsystems.com

Summary of proposal

- Functionality in MAC Control layer
- OAM in Frames
  - Send statistics from clause 30
  - Link monitor sends one frame per second
  - Failure events also send stats
- Independent of PHY
  - Works with existing PHYs
  - No additional burden for future PHYs
- Base on Slow Protocol (Annex 43B)
  - Limit number of frames/sec (5 now, can increase if needed)
  - 802.1D compliant bridges do not propagate

Slow Protocols

- Defined in IEEE Std 802.3 Annex 43B
  - Defined by Link Aggregation task force
  - Ethertype: Slow_Protocols_Type (88-09)
  - Up to 10 subtypes possible (2 consumed so far)
- One of 16 reserved multicast addresses
  - Does not propagate through IEEE Std 802.1D bridges
- Restrictions
  - no more than 5 pkts/sec (currently)
  - Recommended small (<128 byte) frames
- Suitable for firmware impl (ex: LACP)



Encapsulation in Slow Protocol

- Simple encapsulation
  - 1 byte code

00	PING Request
01	PING Response
02	Link Monitor
03	etc…

DA	SA	type	subtype	code	data	FCS
6	6	2	1	1		4

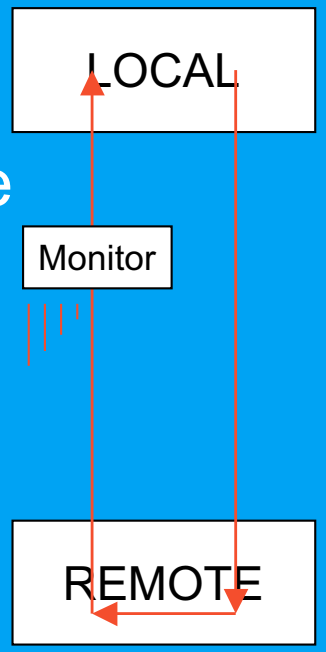
Link Monitoring

- Send stats from Clause 30
  - Encoded as type,length,value
  - Type from Annex 30 arcs
    - Start with tuple after csmacdmgmt.
  - Define vendor extension mechanism
    - If we don't, they'll each choose a different mechanism
    - Distinguish via OUI?
  - Doesn't extend to arbitrary MIB variables
    - SNMP MIBs depend on SNMP semantics
- Periodic announcement is the key mechanism
  - Could also allow queries for additional information

<statType, statLen, statValue>

Supplement to SNMP

- OAM intended as supplement to SNMP
  - Store stats from remote end
  - SNMP can query them later, even after failure
- Received stats stored in oRemoteEntity
  - New object class in Clause 30
  - Prepend source MAC address
    - Needed for shared networks & PONs



MACaddr1: <stat1><stat2><stat3>
MACaddr2: <stat1><stat2>
MACaddr3: <stat1><stat2><stat3>

Link Monitor

- Link Monitoring stats to send defined by attribute
  - A SEQUENCE of statistics
- Hook for authentication:
  - Include an authenticationState attribute
    - Allows external authentication mechanisms to affect OAM
  - Include two attributes defining stats to send
- Allows minimal information before authentication
  - Maximal information after authentication

aFramesReceivedOK
aFramesTransmittedOK
aSymbolErrors (new)

Nonauthenticated

aFCSErrors

Authenticated

aFramesReceivedOK
aFramesTransmittedOK
aSymbolErrors (new)

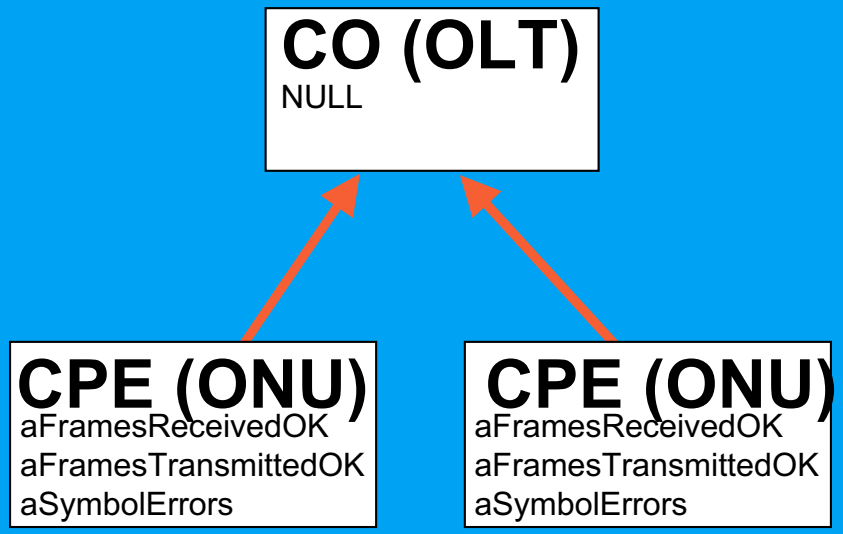
Link Monitor & Authentication

- Some environments are security conscious
  - Only enable a port if authenticated (802.1x, etc)
- Premium SLAs might include monitoring proviso
  - After authentication more information can be sent
- 802.3 should not invent another auth protocol
  - The world has enough of them already
  - We'd just reinvent 802.1x / EAP anyway
- OAM helps ensure link functionality
  - If link no worky, authentication no worky
  - Need limited OAM before authentication
- No SETs
  - Security threat only of leaking information

No Master/Slave Relationship

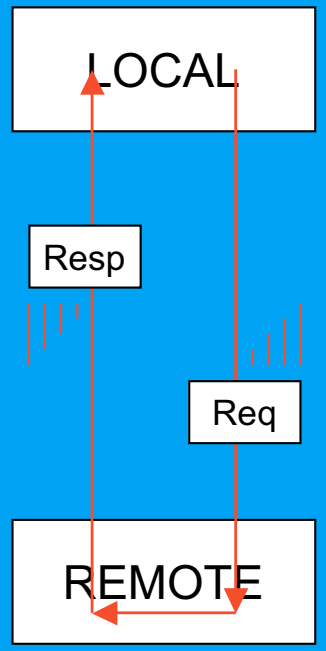
- No inherent Master/Slave relationship
  - Link Monitor stats defined by a variable
    - Configure OLT not to send stats to CPE
  - Do not embed master/slave relationship into 802.3 spec
    - 802.3 covers more than one market space

aFramesReceivedOK
aFramesTransmittedOK
aSymbolErrors (new)



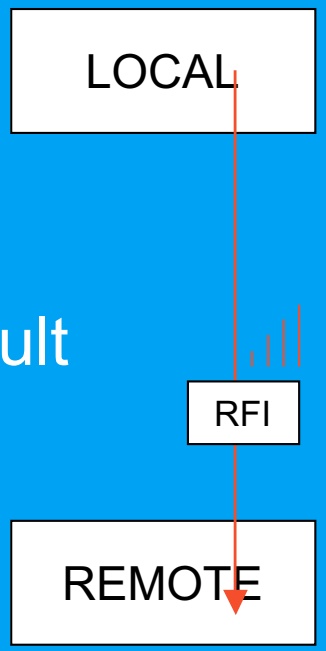
L2 Ping

L2 Ping satisfies Remote Loopback objective
  - Send request, get response
  - Non modal (mix ping with regular traffic)
- Intended for connectivity test
  - Limited number of packets/sec.
- Not intended as throughput test
  - Best done at L3, where the services run
- Not intended as BERT test
  - Symbol & FEC error count is measure of link quality
  - High bit rate ping = more expensive implementation



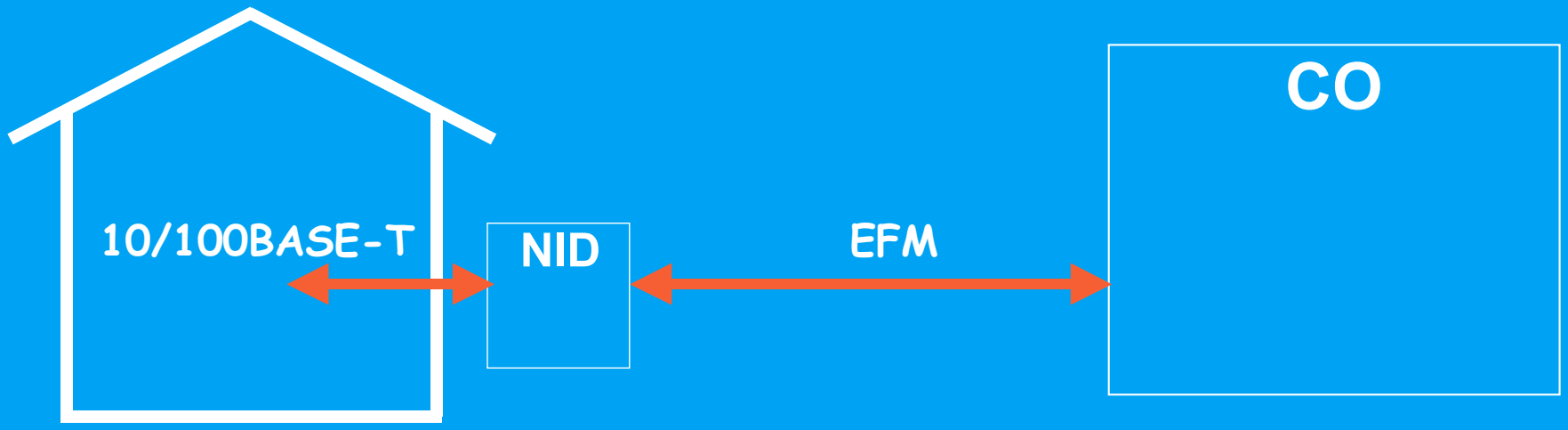
Remote Fault & new PHYs

- Most PHYs provide binary RFI indication
  - This is good.
- Access market requires more
  - Troubleshooting performed at CO
  - Subscriber has little expertise
  - Truck roll to subscriber is expensive
- Use OAM facility for this
  - Send OAM packets with information about fault
  - PHY continues to send frames one direction
  - Alternative is complex error handling in PHY



Deployment model

- Demarc should be a bridge or L3 device
  - Has to transfer between dissimilar speed links
  - EFM <-> 10/100/1000 or 802.11, for example
  - Ethernet is a packet network
    - The demarc should understand packets.
- Demarc must not propagate reserved multicasts



What it is not

Note a few things not supported:

- No SETs
  - OAM does not modify configuration of remote
  - Ethernet links configure themselves locally
- Not a full-fledged management facility
  - OAM strives only to maintain link integrity
    - Even with an “unmanaged” device at one end
  - Managed devices must include a management protocol
- Not routable
  - Messages transit only a single link
    - Possible to define a forwarding proxy; out of 802.3 scope
  - Not intended to manage/provision entire infrastructure

Why Frames?

- Independent of PHY
  - Works with current and future PHYs
  - Common implementation, maximum leverage
- Suitable for firmware implementation
  - Requires no special hardware
  - Minimizes cost, maximizes flexibility
- The Ethernet Way
  - Ethernet is a packet network
  - OAM in Frames fits the evolutionary path of Ethernet
- Fits easily into existing Ethernet practices
  - Vendors: design practice, switch fabrics, cores, etc
  - Operators: management stations, network monitors, etc

Part 2: Rebuttal

Denton Gentry, Dominet Systems

(an earlier version was circulated for comment, changes have been made)

Rebuttal: Why not the preamble?

You want to send a packet of OAM data. Do you:

a) Send the packet

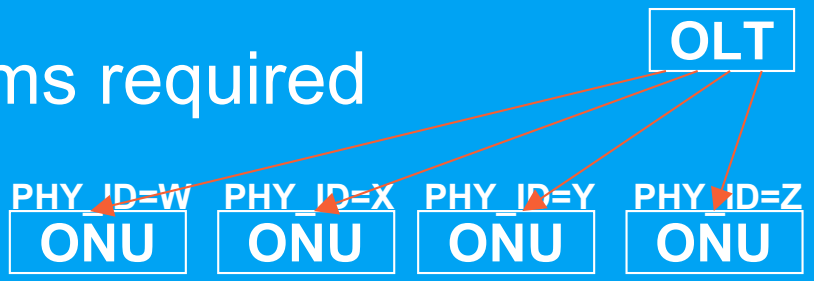
b) HDLC encode the packet: byte stuff the escape character, calculate an HDLC checksum. Segment one byte into each outgoing ethernet preamble (unless there are no outgoing frames in which case you generate imaginary frames). For PON, maintain separate HDLC state per PHY_ID, interleaving multiple HDLC frames on the wire. One byte of HDLC data travels across the net to the receiver. For a PON, examine each PHY_ID to determine whether to reassemble. HDLC decoder puts bytes back together into a frame, checks the HDLC checksum and removes the byte stuffing.

Rebuttal: Copper PHY

- Why do we have a preamble?
  - Provides RX clock synch in half duplex networks
    - 10/100/1000 FD retains preamble, same PHY for HD.
    - Also matches frame spacing between full and half duplex
  - 802.3ah Copper will not be CSMA/CD
    - frame rate will not evenly match 10/100/1000, regardless
- Copper track could discard preambles
  - 7 bytes reclaimed for every packet
  - Preamble consumes B/W with every frame
    - Between 1/2% to 8%, depending on frame size
    - OAM Frames only consume B/W when needed
  - Putting data in the preamble would make this impossible

Rebuttal: ePON

- For PON, multiple HDLC streams required
  - One per PHY_ID



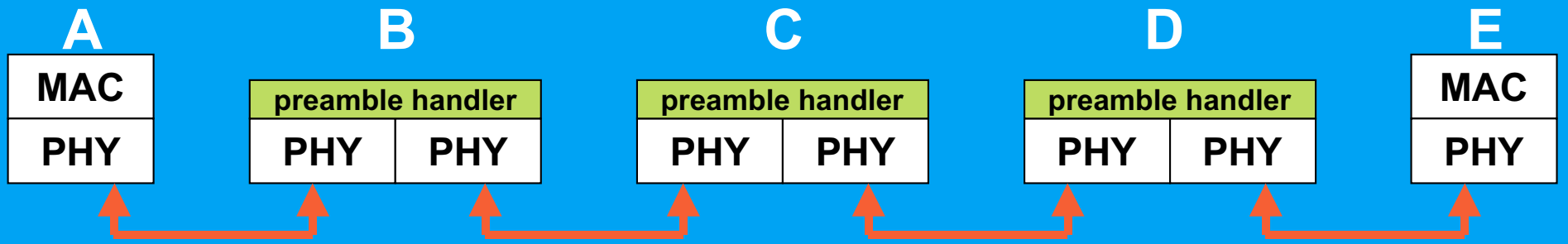
- Data traffic pattern unlikely to align with OAM traffic patter
  - OAM in preamble only if a data frame is going to the proper PHY_ID
  - Dummy frames must be generated for remaining PHY_IDs
  - OAM in Preamble becomes OAM in frames, but badly done

- Why not use the station's MAC Address for PHY_ID?
  - Need 6 byte addr plus 1 byte CRC
  - not enough room left if OAM is in the preamble

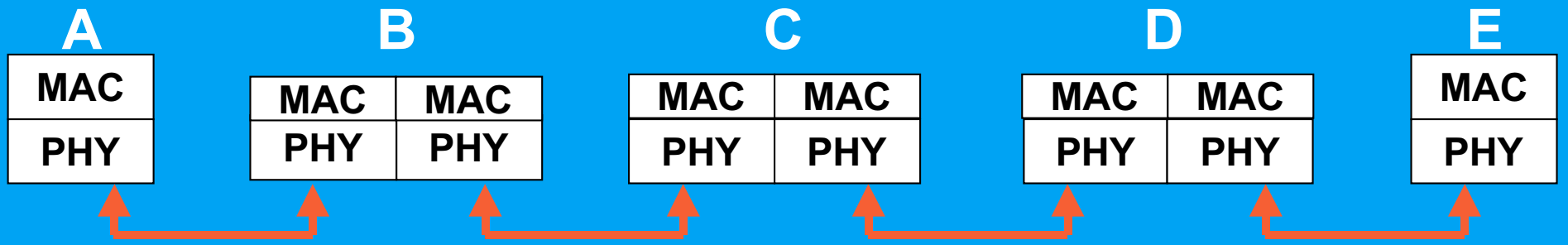
SOP	OAM	HDLC	reserved	PHY_ID	CRC
1	1	1	2	2	1

Rebuttal: Metro

The argument has been made that preamble should be used to manage metro ethernet networks, including optical regenerators.



You need unique PHY_IDs, forwarding rules, loop detection and prevention, etc. PHY_ID assignment no longer per-link. Similar complexity to a new MAC.



With Frames, B, C, & D pass data frames through, but can recognize and replace OAM frames inserted regularly by A & E. To generate an immediate alarm B, C, or D must buffer incoming data, but will “catch up” at the next OAM frame.

Rebuttal: Carrier Networks

Assertion: “Preamble looks like Sonet POH/LOH/TOH”

Rebuttal: “Only if you squint at it just right.”

Sonet overhead is constant bit rate, derived from stratum clock. It can be tunneled as needed, as the bit rates match. Not so with preamble, which varies by the packet size distribution of the subscriber traffic.

Assertion: “Carriers use Out Of Band management”

Rebuttal: Carriers use OOB where it makes sense. Other techniques are used where they make sense.

Circuit based carrier networks use OAM channels (Sonet).

Cell based carrier networks use OAM Cells (ATM)

Packet based carrier networks use OAM packets (Frame Relay)

Summation

- OAM in Frames is best for EFM Copper:
  - Allows maximum B/W utilization
- OAM in Frames is best for EFM ePON:
  - In band, for same reasons MPCP is in band
  - No dynamic assignment of PHY_ID required
- OAM in Frames is best for EFM P2P:
  - No changes required to 1000BASE-X
  - No changes for 100BASE-X (if adopted)
- OAM in Frames is best for Ethernet:
  - Ethernet is a packet network
  - Does not break the functional layering model
  - Backward and forward compatible with all PHYs