PHY OAM Baseline Proposal

Contributors :

Hiroshi Suzuki	Cisco System
Sanjeev Mahalawat	Cisco System
Ben Brown	AMCC
Rich Taborek	Tartan Technology

Supporters:

Yukihiro Fujimoto	NTT
Yannick Le Goff	France Telecom
Onn Haran	Passave
Ariel Maislos	Passave
Martin Nuss	Internet Photonics
Ulf Jonsson	Ericsson
Olli-pekka Hiironen	Nokia
Satoshi Obara	Fujitsu
Don Alderrou	Intel
Richard Cam	PMC-Sierra
Thomas Joergensen	Vitesse
Fred Mohamadi	Transpectrum
Shawn Rogers	TI

Introduction

- **PHY OAM refinements based on feedback at Scotland meeting**

- **Especially addressing GE compatibility, and adding more technical details**

Scope of PHY OAM

- **PHY OAM is applied for Full duplex fibers, P2P and P2MP**

- **PHY OAM is optional enhancements for EFM OAM**

- **PHY OAM is activated via MAC OAM Negotiation**

Functional Partitioning
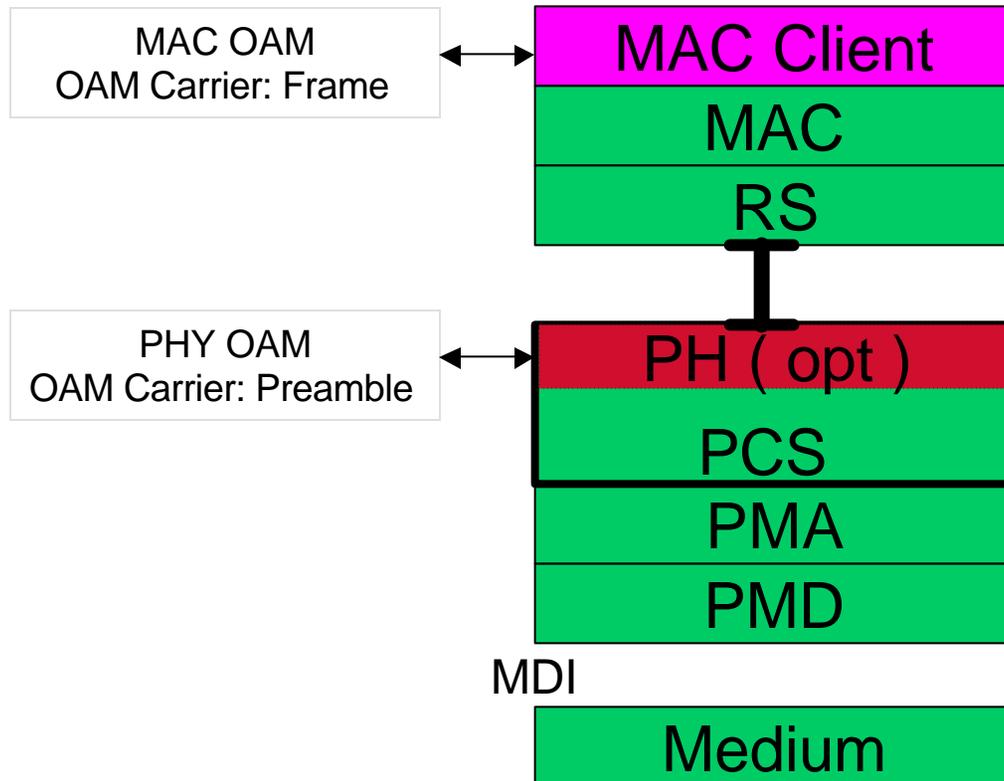## Ethernet OAM Baseline in May2002

Feature	MAC OAM Data transported in frame	PHY OAM State signaled in preamble
Event indication with diagnostics	✓	
Immediate non-intrusive signaling of PHY faults		✓
MAC layer ping/loopback	✓	
Lower layer ping/loopback		✓
Link monitoring	✓	

Different functions may be handled by different transport.

No functional overlap!

PHY layer defect indication and health check is important for service providers

EFM OAM Layers in 802.3

MAC OAM OAM Carrier: Frame	↔

MAC Client
MAC
RS

PHY OAM OAM Carrier: Preamble	↔

PH (opt)
PCS
PMA
PMD

MDI

Medium

PH = Preamble Handler function for PHY OAM
(New Clause)

# PHY OAM Function in PCS

1. **Preamble Handler (PH, PHY OAM function) inside PCS**

   PCS State Machine, GMII & MII, RS Unchanged

   PH has direct access to all PHY failure / status registers

2. **1000BASE-X and G/MII Compatible**

   PH now can synch with PCS tx_even alignment timing

3. **MDIO/MDC used as Management Interface**

   Standard interface

4. **Only one new clause" PH"**

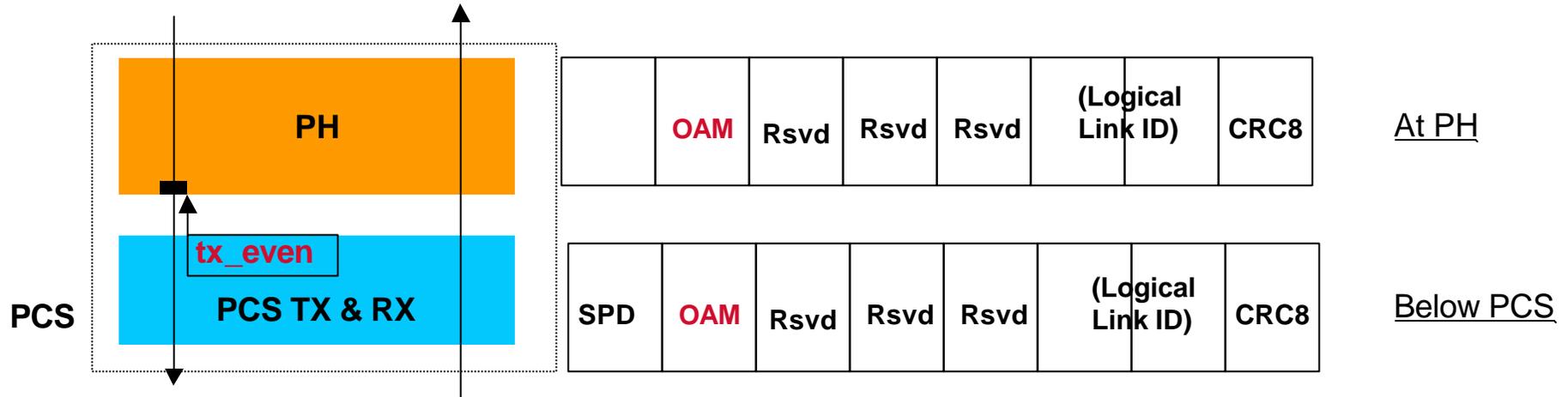   Include both interface specs for 1000BASE-X and 100BASE-X

PH location in PCS

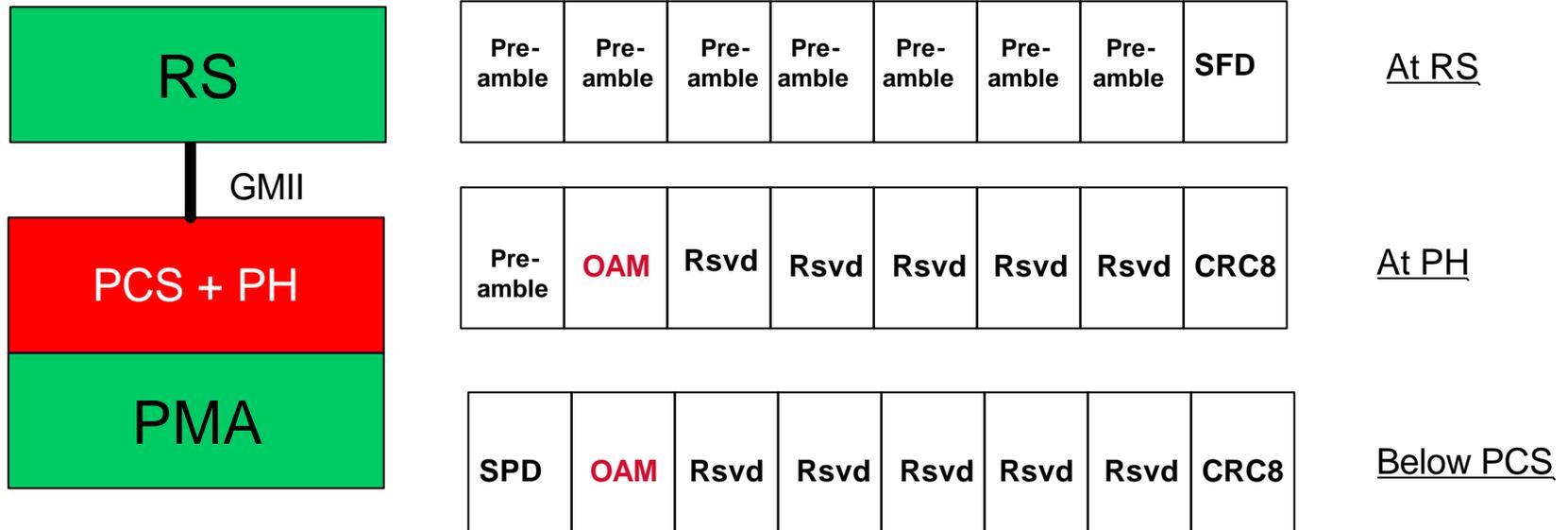**GMII/MII**

PCS
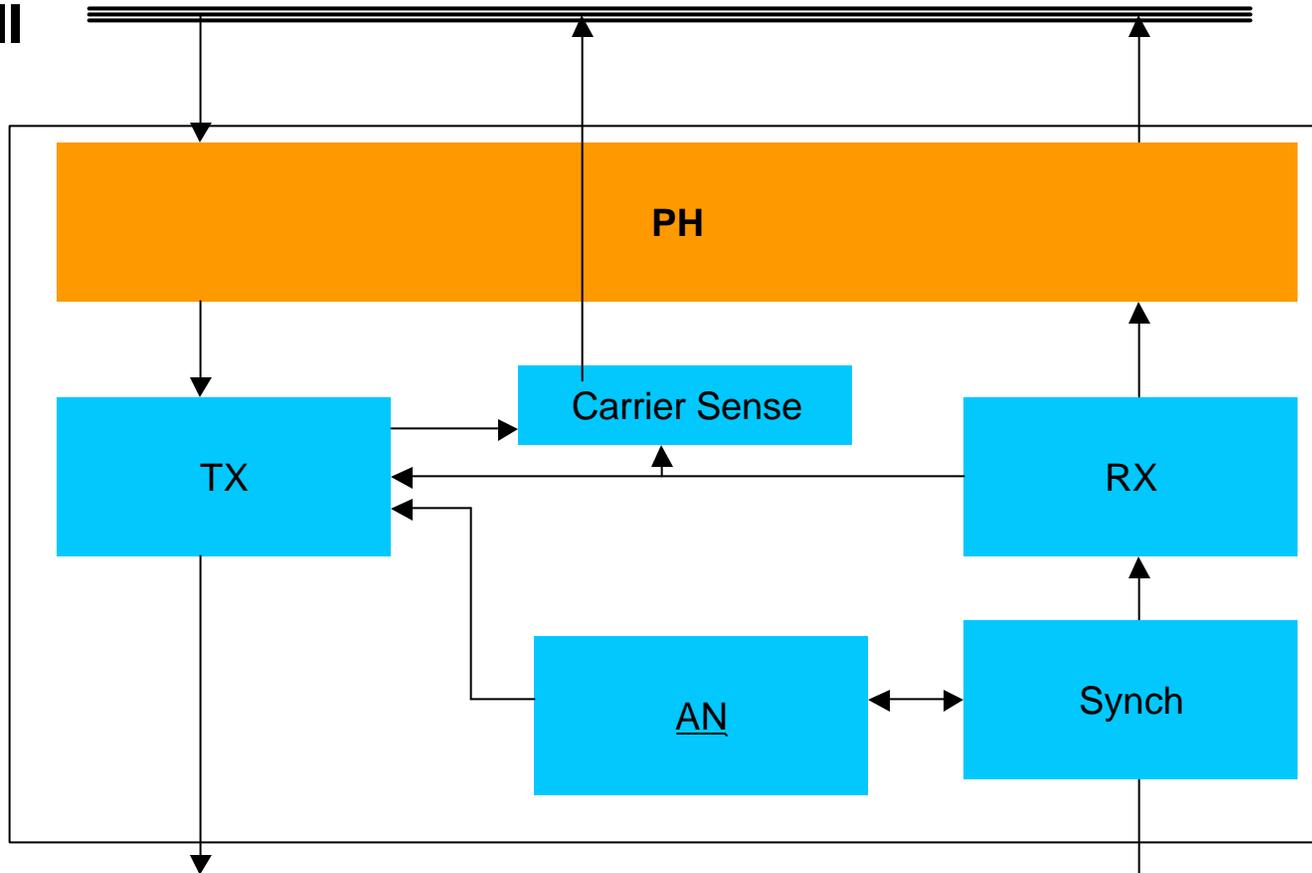
1) No change in PCS state machine.
2) PH has direct access to all PCS registers

1000BASE-X Compatibility

			OAM	Rsvd	Rsvd	Rsvd	(Logical Link ID)	CRC8	At PH

SPD	OAM	Rsvd	Rsvd	Rsvd	(Logical Link ID)	CRC8	Below PCS

PCS

PH

tx_even

PCS TX & RX

- PH has direct access to **tx_even** register in PCS.

- PH transmits Frame with PCS "**tx_even**" alignment timing.

- Preserves 8 byte preamble, while **achieving no change on PCS state machine / GMII.**

- At PCS TX IPG could vary between 11 to 13 bytes with average of 12 (Not an issue with Full Duplex Operation)
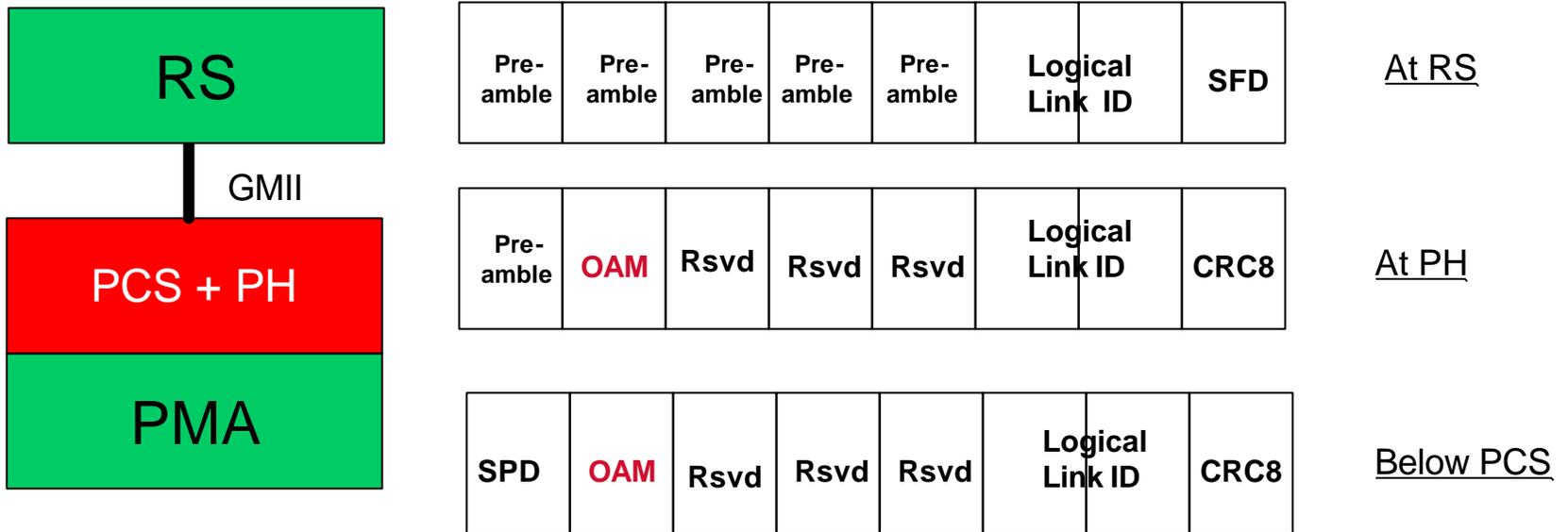
PH Function (P2P)

RS							

GMII

PCS + PH

PMA

Pre-amble	Pre-amble	Pre-amble	Pre-amble	Pre-amble	Pre-amble	Pre-amble	SFD

At RS

Pre-amble	OAM	Rsvd	Rsvd	Rsvd	Rsvd	Rsvd	CRC8

At PH

SPD	OAM	Rsvd	Rsvd	Rsvd	Rsvd	Rsvd	CRC8

Below PCS

- TX: PH replaces 2nd Preamble byte with PHY OAM, calculates CRC8 over 6-bytes of Preamble, replaces SFD with CRC8 value and passes it to PCS; In the absence of Ethernet Frames, 8-byte OAM octets are sent.

- RX: PH checks CRC8 on preamble, if correct, then process OAM byte, PH converts OAM Preamble to standard Preamble (7 bytes of 55 and 1 Byte SFD) and sends it to RS; 8-byte OAM octets are terminated in PH.

PH Function (P2MP)

Pre-amble	Pre-amble	Pre-amble	Pre-amble	Pre-amble	Logical Link ID	SFD

At RS

Pre-amble	OAM	Rsvd	Rsvd	Rsvd	Logical Link ID	CRC8

At PH

SPD	OAM	Rsvd	Rsvd	Rsvd	Logical Link ID	CRC8

Below PCS

**RS** — **GMII** — **PCS + PH** — **PMA**

•TX: PH receives Preamble with Logical Link ID from RS, PH replaces 2nd Preamble byte with PHY OAM, calculates CRC8 over 6-bytes of Preamble, replaces SFD with CRC8 value and passes it to PCS; In the absence of Ethernet Frames, 8-byte OAM octets are sent.

• RX: PH checks CRC8 on preamble, if correct, process OAM byte, replaces CRC8 byte with SFD and passes the Preamble to RS. If CRC8 is bad, disregard OAM byte and set RX_ER to indicate ERROR in Preamble; 8-byte OAM octets are terminated in PH.

How is PHY OAM Implemented?

- PH implements 20B FIFO Queue
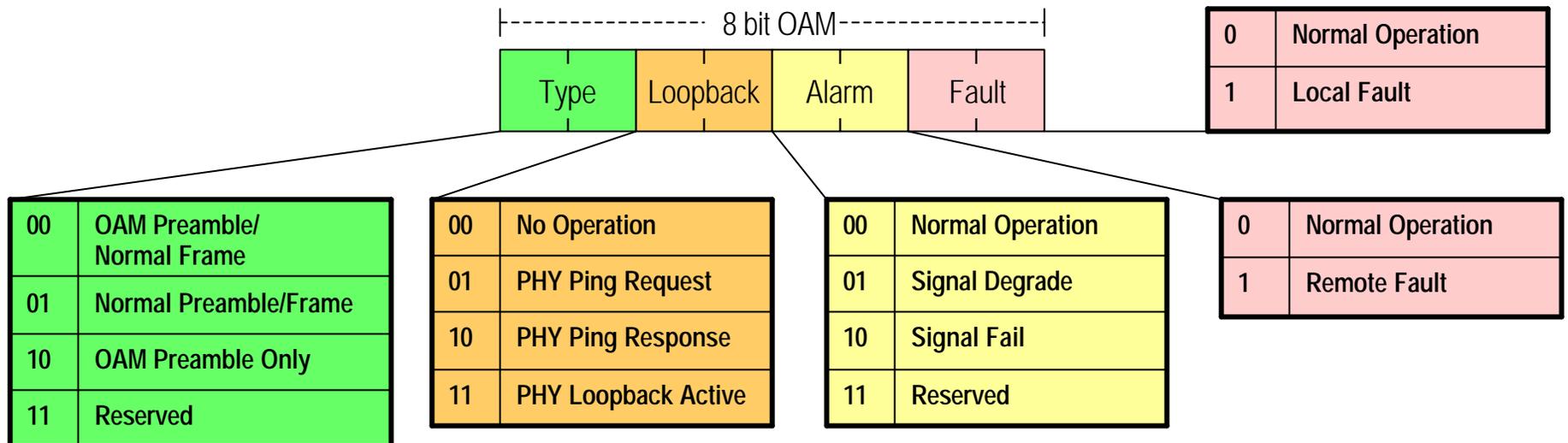- Normal Preamble/SFD substituted with OAM Preamble
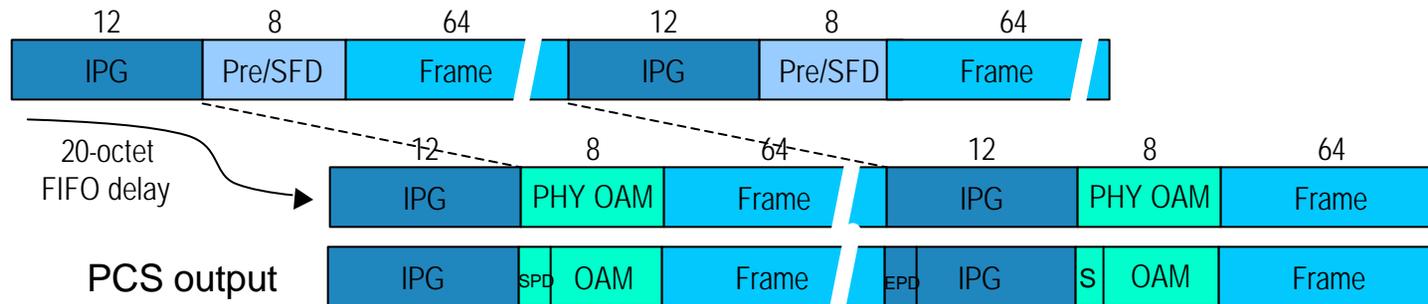
- PH *may* replace IPG octets with PHY OAM octets
- IPG (idles) substituted with PHY OAM octets
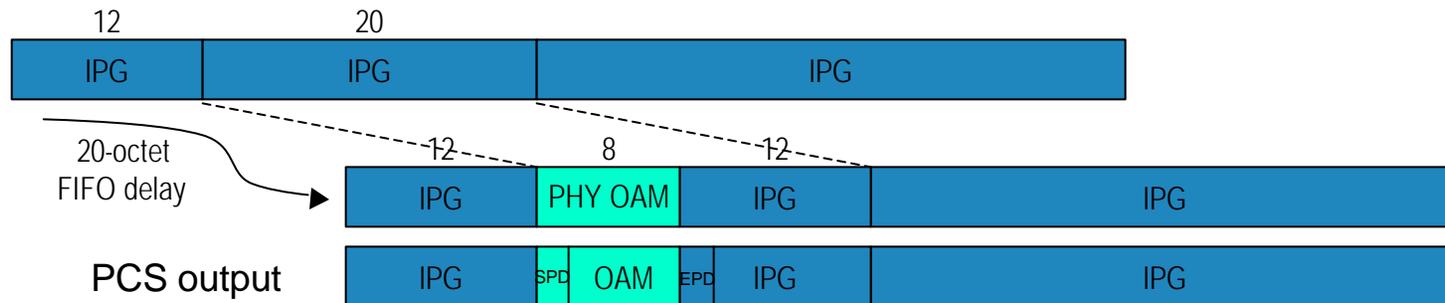
- PHY OAM octets ARE NOT FRAMES, do NOT PASS THE PH, and have NO IMPACT ON MAC FUNCTION
- SPD and EPD are PCS framing delimiters

PHY OAM Byte : Bit definitions

8 bit OAM

Type	Loopback	Alarm	Fault

0	Normal Operation
1	Local Fault

00	OAM Preamble/ Normal Frame
01	Normal Preamble/Frame
10	OAM Preamble Only
11	Reserved

00	No Operation
01	PHY Ping Request
10	PHY Ping Response
11	PHY Loopback Active

00	Normal Operation
01	Signal Degrade
10	Signal Fail
11	Reserved

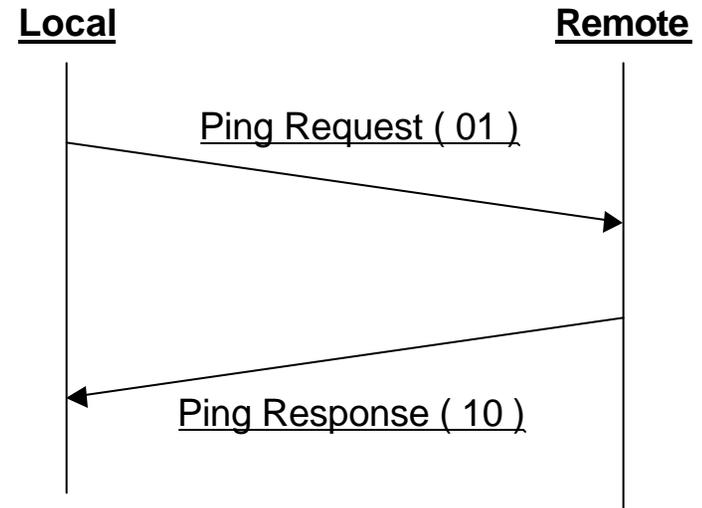| 0	Normal Operation
1	Remote Fault

PHY-Loopback only for P2P

Signal fail: Loss of Signal from RX Optical module / Loss of Synch

Signal degrade: RX-CRC8 error counter threshold exceeded

PHY Ping

Local   **Remote**

00	No Operation
01	PHY Ping Request
10	PHY Ping Response

Ping Request (01)

Ping Response (10)

- **PING provides PHY Layer Availability Check**
    - **Local sends Ping-Request and wait for Response.**
    - **Remote sends Ping-Response upon Request**
    - **Local can re-send Ping-Request after time out, or after receiving a response**

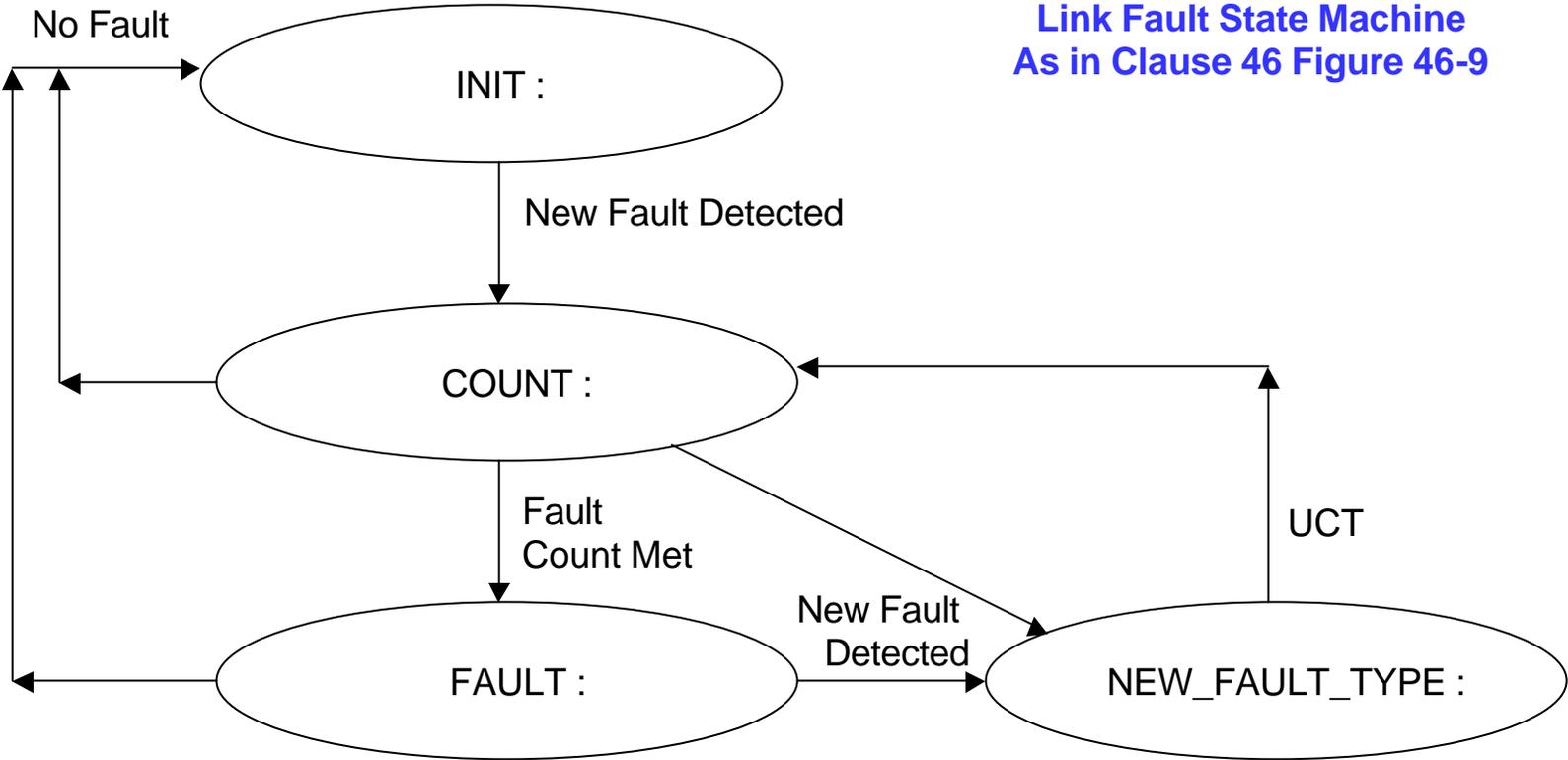- **Provides Link test when upper layers may have failed.**

Link Fault Operation

- **Link Faults include Local Fault (LF) & Remote Fault (RF)**

- **When PHY- Rx path failure detected and when Local Fault is received, PH sends Remote Fault Signal to remote DTE indicate the transmit link of remote DTE is faulty**

- **Alarms (Signal Fail / Signal Degrade) provides reasons of Rx-Failure**

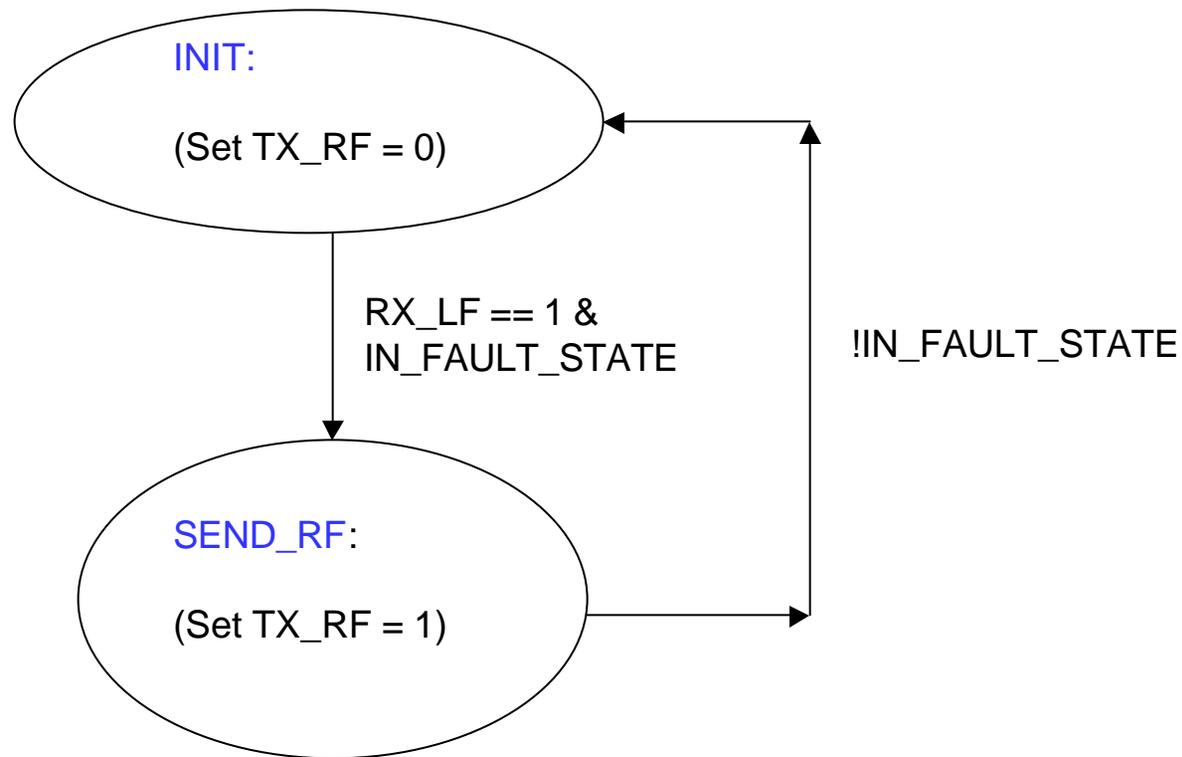- **PH qualifies received Fault Messages via a state machine**

Link Fault Signaling: Receive

Same functionality as 10GBASE-X Link Fault State Machine As in Clause 46 Figure 46-9

No Fault

INIT :

New Fault Detected

COUNT :

Fault Count Met

UCT

FAULT :

New Fault Detected

NEW_FAULT_TYPE :

Link Fault Signaling: Transmit

INIT:

(Set TX_RF = 0)

RX_LF == 1 &
IN_FAULT_STATE

!IN_FAULT_STATE

SEND_RF:

(Set TX_RF = 1)

PHY OAM improvements over MAC OAM

- **Need Physical layer health check / failure isolation**

  - Isolation of Link failure from failures above PHY

  - Current Auto-negotiation in GE does not provide RF for receive link problems

- **Provide faster detection & indication than MAC OAM**

  - Non PHY based OAM needs management entities polling PHY device via MDIO for PHY Failure Detection (No guarantee to meet 10msec detection)

  - Link speed Physical layer failure detection, diagnostics and response

- **No BW impact on data traffic**

  - Important for EPON where 32 ONUs are under OLT

  - Important for bit level defect indications

- **Data traffic transparent**

  - Does not depend upon flow control or queue states

  - Does not hinder user data traffic

Backup

More PHY OAM improvements

- **Extensive PHY layer Health check**
  - Can do health-check Multiple-port PHY chip separated from MAC chip by G/MII
  - Even every-msec heart-beat possible with zero impact on data traffic

- **RX Failure Detection latency examples**
  - Achieve SONET comparable Failure Detection latency < 10msec
  - PHY-OAM is "guaranteed" to detect RX failure immediately < 1msec
  - MAC-OAM needs management entity to "poll" PHY devices via MDIO (in general > 10ms~100ms) and that would require more CPU cycles and increase cost

- **Cost effective OAM implementation**
  - PHY-OAM provides all functionality with small state machine
  - MAC-OAM flexibility and its extensible functions suited for SW implementation, while its HW implementation would be VERY costly and complex

EPON specific PHY OAM improvements

- **Defect Indication (DI) Generation Latency**
  - PHY-OAM is guaranteed to send DI in the next Cycle (in 1~few msec)

    PHY-OAM can send DI even within the same slot after failure detection

    Using Idle Frame generation, it does NOT need to wait the next cycle
  - MAC-OAM likely to lose opportunity to send DI by multiple Cycles (> 10msec)

    DI frames needs to be transmitted multiple times. But MAC-OAM

    can not send consecutive 128 byte DI frames unless big enough BW allocated.

- **BW allocation impact of OAM traffic**
  - PHY-OAM : Zero
  - MAC-OAM

    ex. 128 OAM byte / Cycle time = 128x8/1msec = 1Mbps @LLID

    At least 1Mbps OAM-BW has to be allocated to send consecutive OAM frames, although it will be wasted due to the limit of max number of frame / sec.