A Proposal To Use Ethernet PHYs for RPR

September 2001

Rhett Brikovskis – Lantern Communications
Henry Hsiaw – NEC
Vittorio Mascolo – Alcatel
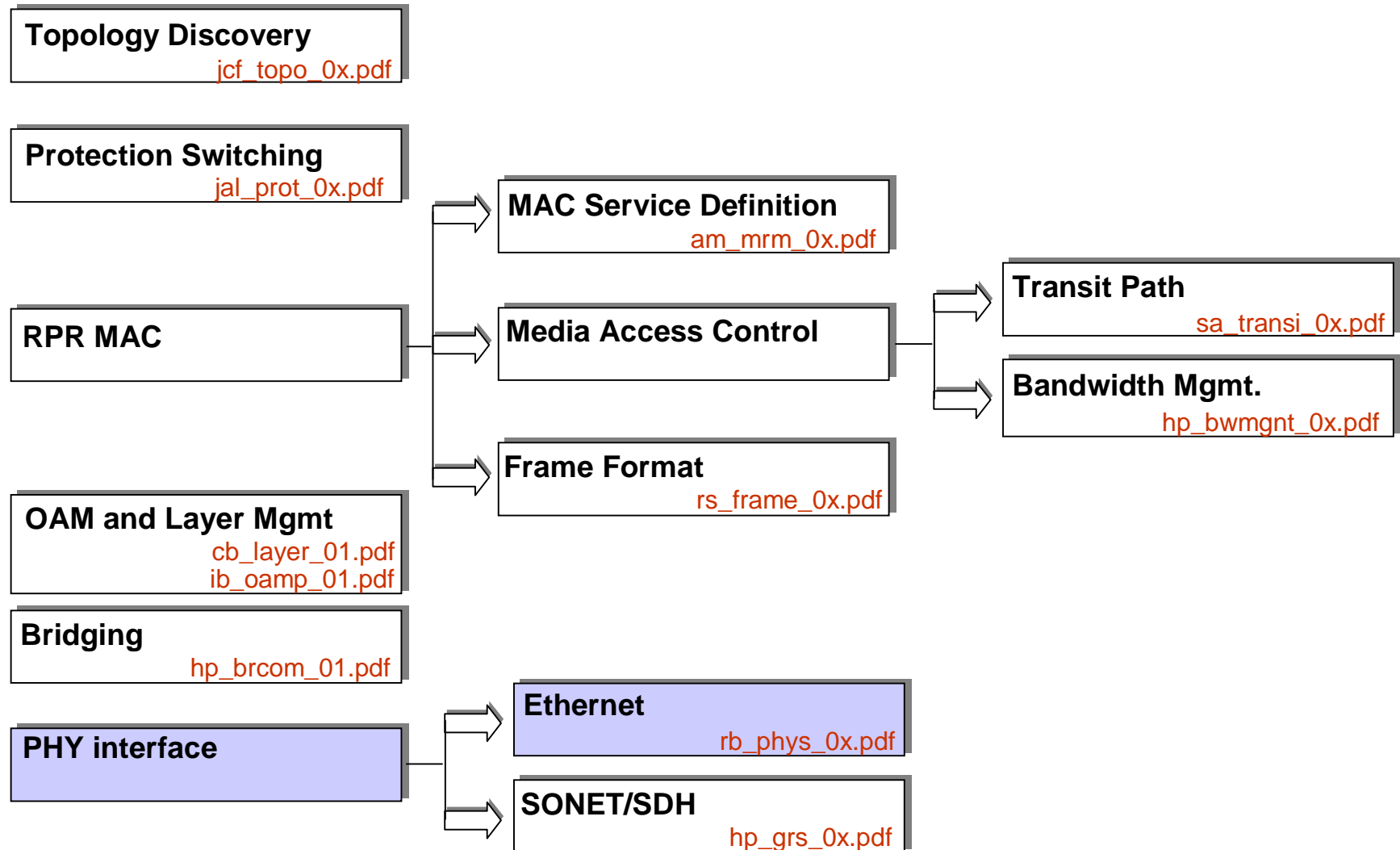Robin Olsson – Vitesse
Harry Peng – Nortel Networks
Raj Sharma – Luminous Networks
Frederic Thepot – Dynarc

Components of a complete RPR Proposal

Topology Discovery
  jcf_topo_0x.pdf

Protection Switching
  jal_prot_0x.pdf

RPR MAC

→ **MAC Service Definition**
  am_mrm_0x.pdf

→ **Media Access Control**
  → **Transit Path**
    sa_transi_0x.pdf
  → **Bandwidth Mgmt.**
    hp_bwmgnt_0x.pdf

→ **Frame Format**
  rs_frame_0x.pdf

OAM and Layer Mgmt
  cb_layer_01.pdf
  ib_oamp_01.pdf

Bridging
  hp_brcom_01.pdf

PHY interface
→ **Ethernet**
  rb_phys_0x.pdf
→ **SONET/SDH**
  hp_grs_0x.pdf

Overview

- Objectives

- Overview of 10 Gigabit Ethernet Physical Layer

- RPR P-SAP Interface

- RPR Reconciliation Sublayer

- 10 Gigabit Ethernet PHYs

- RPR MAC-PHY Mapping

- Summary

Objectives

- Use a universal MAC-PHY logical interface compatible with various PHY types

- Define a Reconciliation Sublayer (RS) for RPR similar to that defined by P802.3ae

- Other than the RS, support the P802.3ae Physical Layer interfaces and sublayers with no changes.

- Support all seven LAN and WAN PHYs that P802.3ae is specifying.

- This proposal specifically addresses 10 GbE, but is intended to be extensible to other Ethernet speeds.

Overview of 10 Gigabit Ethernet

- P802.3ae (10 GbE) standard is scheduled for Mar 2002 completion.

- Standard is currently at Draft 3.2.

- Latest draft can be ordered at
  http://standards.ieee.org/catalog/IEEE802.3.html

Overview of 10 GbE – SUBLAYERS

- **10 GbE Physical Layer consists of the following sublayers:**
  - Reconciliation Sublayer (RS), specific to the 802.3 MAC-PLS interface.
  - 10 Gigabit Media Independent Interface (XGMII) – optional.
  - XGMII Extender Sublayer (XGXS) and 10 Gigabit Attachment Unit Interface (XAUI) - optional XGMII extender.
  - PHY [including Physical Coding Sublayer (PCS), optional WAN Interface Sublayer (WIS), Physical Medium Attachment (PMA) sublayer, and Physical Medium Dependent (PMD) Sublayer] – 7 variants specified.
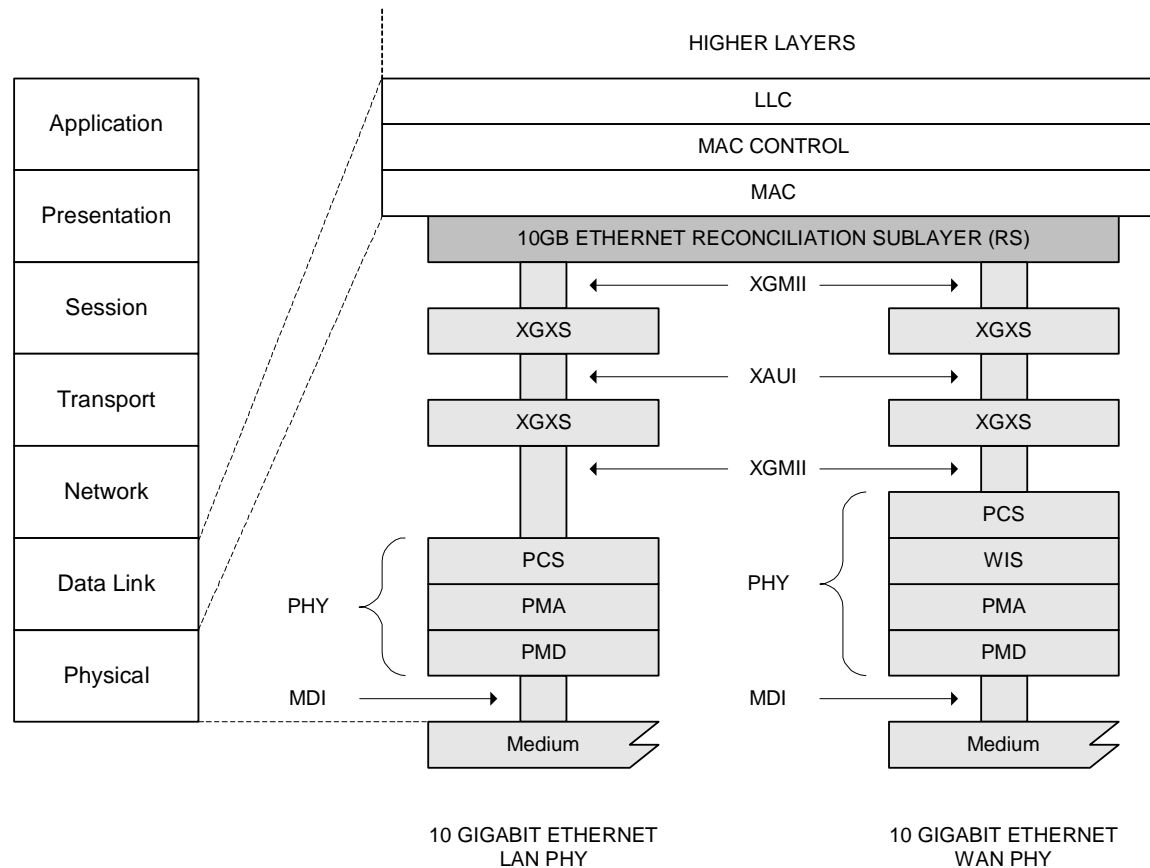 - Medium Dependent Interface (MDI) – not specifically defined.

Overview of 10 GbE – LAYER DIAGRAM

- ## 10 GbE Layer Diagram

OSI REFERENCE
MODEL LAYERS

802.3ae LAYERS

HIGHER LAYERS

Application	LLC
Presentation	MAC CONTROL
	MAC
Session	10GB ETHERNET RECONCILIATION SUBLAYER (RS)

XGMII
XGXS ◄ XGMII ► XGXS
XAUI
XGXS ◄ XAUI ► XGXS
XGMII
◄ XGMII ►

PCS
WIS
PCS
PHY { PMA } PHY { PMA }
PMD
PMD

MDI ► MDI ►

Medium
Medium

10 GIGABIT ETHERNET
LAN PHY

10 GIGABIT ETHERNET
WAN PHY

LLC = LOGICAL LINK CONTROL
MAC = MEDIA ACCESS
CONTROL
MDI = MEDIUM DEPENDENT
INTERFACE
PCS = PHYSICAL CODING
SUBLAYER
PHY = PHYSICAL LAYER
ENTITY
PMA = PHYSICAL MEDIUM
ATTACHMENT
PMD = PHYSICAL MEDIUM
DEPENDENT
WIS = WAN INTERFACE
SUBLAYER
XGMII = 10 GIGABIT MEDIA
INDEPENDENT INTERFACE

OSI layers: Application, Presentation, Session, Transport, Network, Data Link, Physical

RPR P-SAP (MAC-PHY) Interface

- Define a logical P-SAP interface between the RS (PHY) and the MAC similar to the 802.3 MAC-PLS interface.

- The interface is defined as a set of service primitives:
  - PHY_DATA.request
  - PHY_DATA.indicate
  - PHY_DATA_VALID.indicate
  - PHY_LINK_OK.indicate
  - PHY_READY.indicate

- The interface is octet-based.
  - Similar to the bit-based interface of 802.3

- The interface <u>does not</u> operate at a constant data rate.
  - The MAC frame does not include IPG—inserted by the RS.
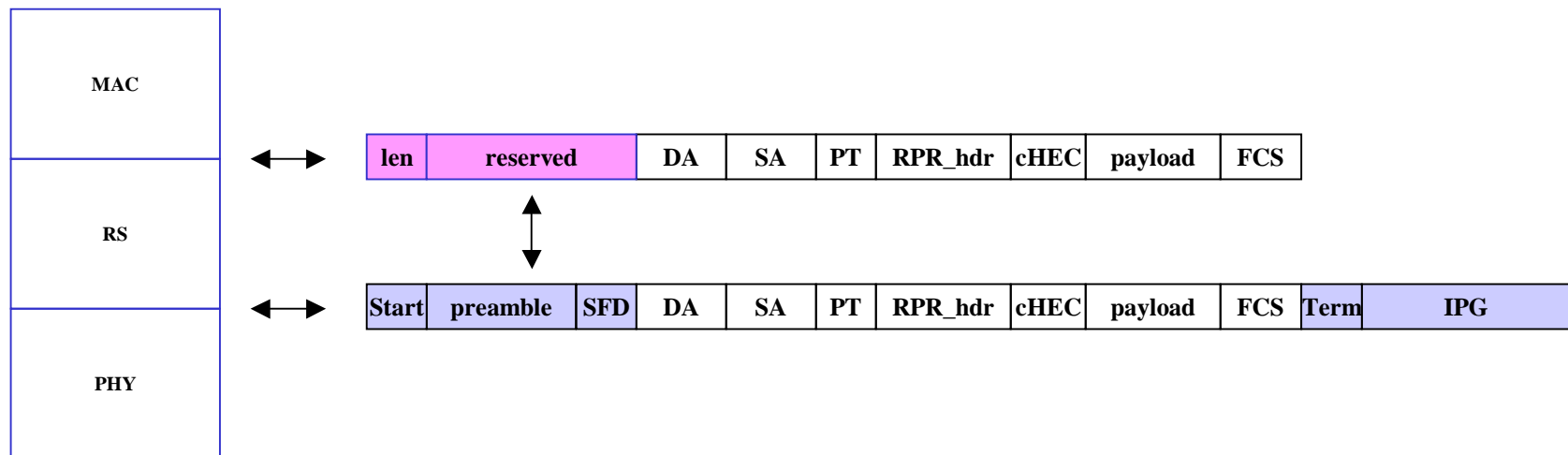
RPR P-SAP Service Primitives

- ## The RPR P-SAP service primitives:
  - PHY_DATA.request(OUTPUT_UNIT)

    Defines the transfer of an octet of data from the MAC to the RS.

    OUTPUT_UNIT={octet_of_data, DATA_COMPLETE}
  - PHY_DATA.indicate(INPUT_UNIT)

    Defines the transfer of an octet of data from the RS to the MAC.

    INPUT_UNIT={octet_of_data}
  - PHY_DATA_VALID.indicate(DATA_VALID_STATUS)

    Indicates whether the parameter of PHY_DATA.indicate contains valid data.

    DATA_VALID_STATUS={VALID, NOT_VALID}
  - PHY_LINK_OK.indicate(LINK_STATUS)

    Indicates whether the PHY indicates that the link is OK.

    LINK_STATUS={OK, FAIL, DEGRADE}
  - PHY_READY.indicate(READY_STATUS)

    Indicates whether the PHY is ready to accept a new MAC frame.

    READY_STATUS={READY, NOT_READY}

RPR RS For 10 GbE

- The Reconciliation Sublayer maps a logical RPR MAC frame to an Ethernet-compatible physical frame, and adds IPG:

MAC
RS
PHY

len	reserved	DA	SA	PT	RPR_hdr	cHEC	payload	FCS

Start	preamble	SFD	DA	SA	PT	RPR_hdr	cHEC	payload	FCS	Term	IPG

RS Inputs and Outputs

- **Reconciliation Sublayer Inputs and Outputs:**

P-SAP Service Primitives   **XGMII Signals**

Reconciliation Sublayer

PHY_READY.indicate   TXD<31:0>
                                                                         TXC<3:0>
                                                                         TX_CLK

PHY_DATA.request

PHY_DATA.indicate

                                                                         RXD<31:0>
PHY_DATA_VALID.indicate   RXC<3:0>
                                                                         RX_CLK

PHY_LINK_OK.indicate
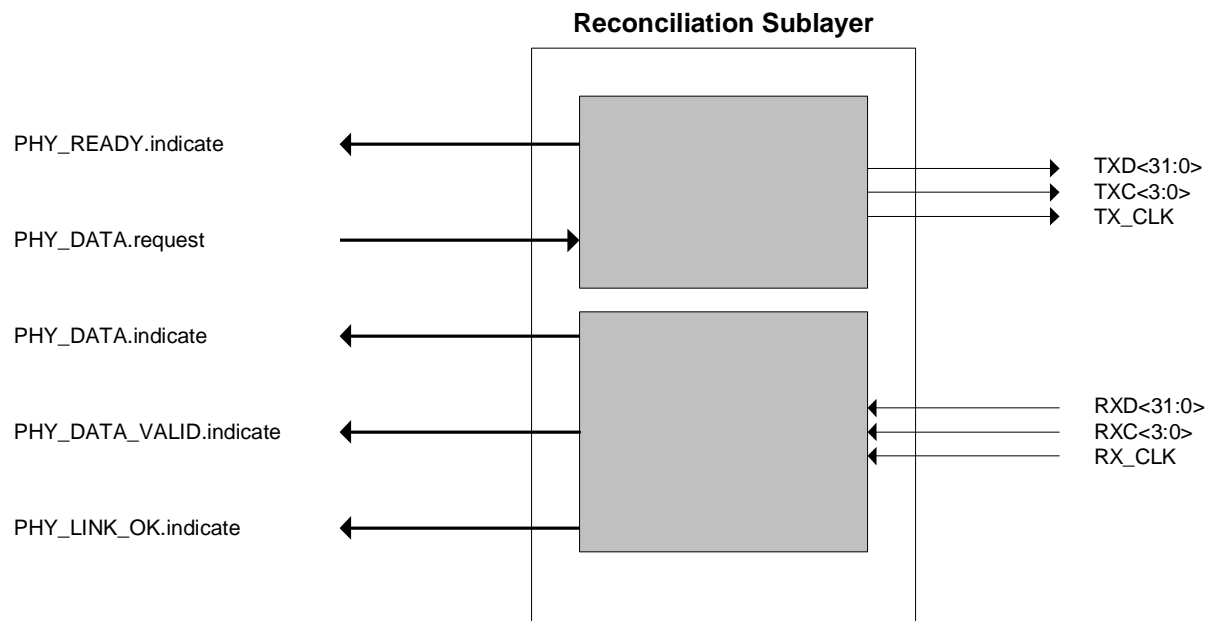
RS Functions

- **The RS retains the following functions specified by P802.3ae, Clause 46:**

  - Converts the logical P-SAP service primitives to/from electrical signals at the XGMII.

  - Map the first octet of Preamble to a Start control character, and align it to "lane 0" on the XGMII.

  - Map the first octet of IPG following a packet to a Terminate control character.

- **For RPR, add/modify the following RS functions:**

  - Generate IPG according to the rules specified in 802.3 (for ethernet, interframe gap period is generated by the MAC).

  - Modify the Link Fault Signaling behavior of the RS to allow dual-simplex operation

RS Link Fault Signaling

- **Modify the Link Fault Signaling behavior—currently intended for a duplex point-to-point link:**
  - P802.3ae treats the link as a full-duplex media—a fault in one direction needs to be signaled in the opposite direction, and MAC data transmission stops.
  - Layers in the PHY may be able to recognize link faults, and generate a Local Fault signal when faults exist. As defined by 10 GbE, when a Local Fault condition is received by the RS, it stops sending MAC data and continuously generates Remote Fault status in the other direction.
  - Similarly, when a Remote Fault condition is received, the RS stops sending MAC data and generates continuous Idle in the other direction.

- **For RPR, redefine the RS fault behavior**
  - RS receives Local Fault conditions and relays the information to the RPR MAC.
  - Opposite direction is unaffected.
  - Remote Fault signal is never generated—handled by MAC messaging.
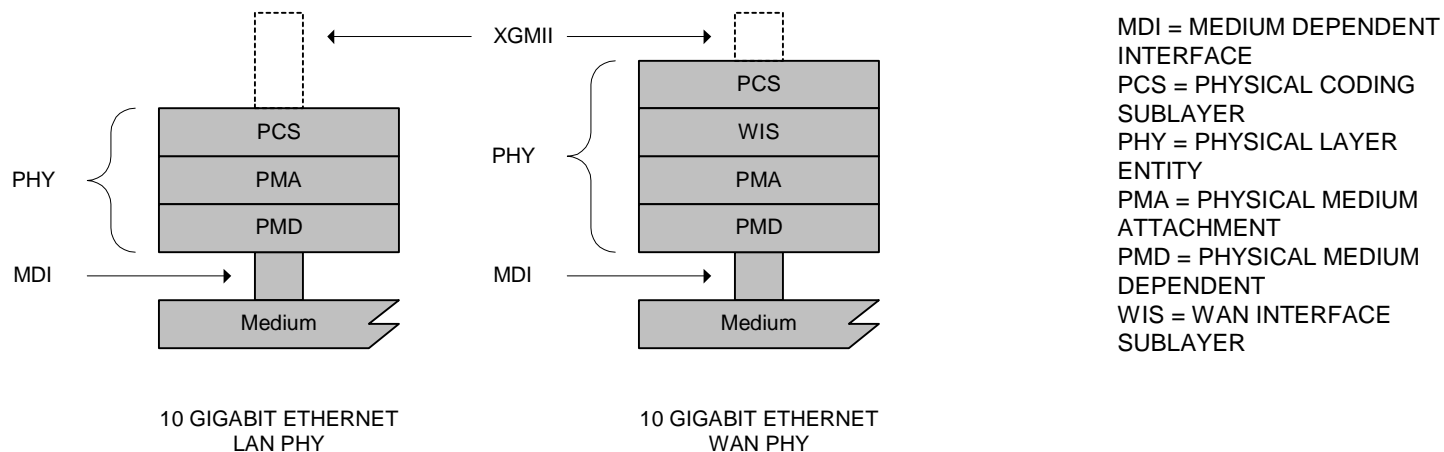
10 GbE PHYs –XGXS/XAUI

- Support the P802.3ae XGMII Extender Sublayer (XGXS) and 10 Gigabit Attachment Unit Interface(XAUI) with no changes.

PHY = PHYSICAL LAYER ENTITY
XAUI = 10 GIGABIT ATTACHMENT
UNIT INTERFACE
XGMII = 10 GIGABIT MEDIA
INDEPENDENT INTERFACE
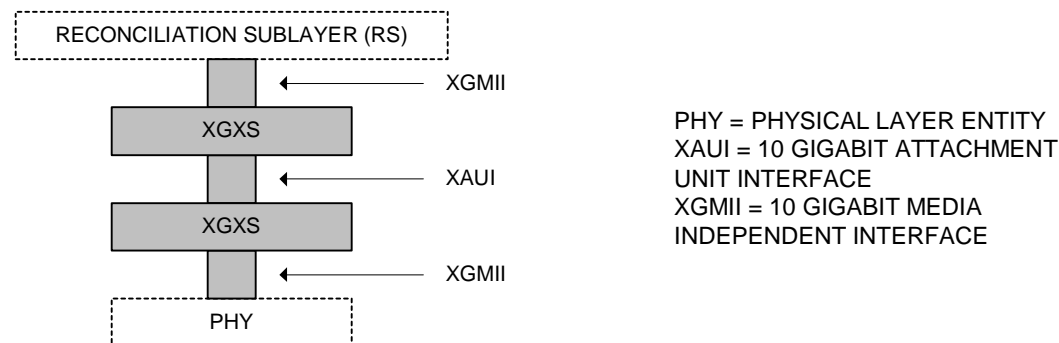
10 GbE PHYs – LAN/WAN PHYS

- **P802.3ae Clauses 48-53 define the sublayers that implement two families of PHYs:**
  - "LAN PHYs" operating at a data rate of 10.0 Gbps.
  - "WAN PHYs" operating at a data rate and format compatible with SONET STS-192c and SDH VC-4-64c.

10 GIGABIT ETHERNET LAN PHY

10 GIGABIT ETHERNET WAN PHY

MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER ENTITY
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
WIS = WAN INTERFACE SUBLAYER

10 GbE PHYs – PHY VARIANTS

- **Seven PHY variants are defined:**

Description	Reach/Fiber	10 GbE Designation	
		LAN PHY	**WAN PHY**
850 nm serial	~85 m/MMF	10GBASE-SR	10GBASE-SW
1310 nm serial	10 km/SMF	10GBASE-LR	10GBASE-LW
1550 nm serial	40 km/SMF	10GBASE-ER	10GBASE-EW
1310 nm WDM	10 km/SMF ~300 m/MMF	10GBASE-LX4	-

MAC-PHY Mapping – Direct Mapping

- RPR typically has more than one PHY connected to a MAC. Need to define a MAC to PHY (to media) mapping.

- MAC could be mapped directly to the corresponding PHY on a single RPR ringlet, but this isn't always optimum for RPR applications…
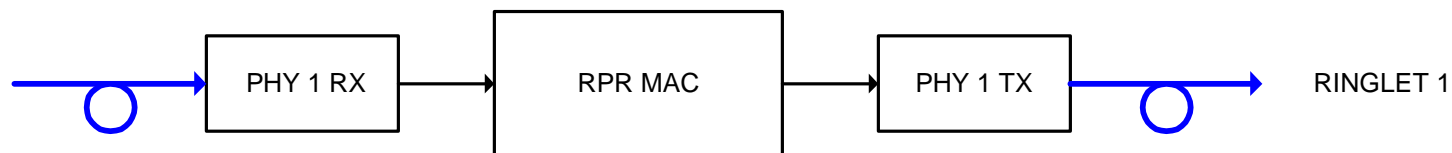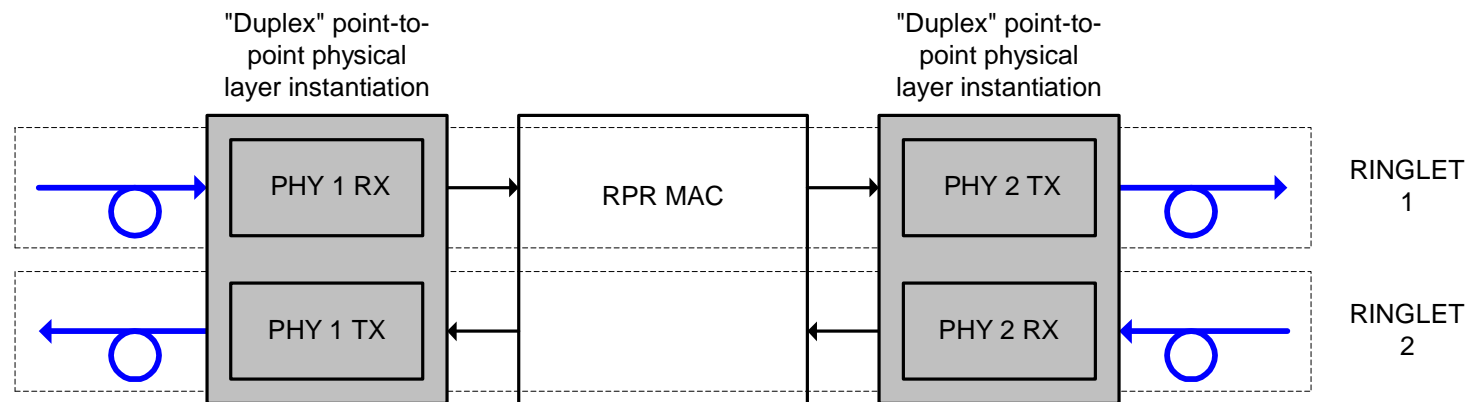
Direct Mapping Problems

- **Direct mapping creates several significant issues for RPR applications:**

  - Generally, this implies that the link to the preceding RPR station and the subsequent RPR station will use a single PHY type (same wavelength, same reach, etc). Extending this to the entire ringlet, all segments within a ringlet must use the same PHY type.

  - Most PHYs are intended for duplex operation—some include fault-processing and error-reporting capability that will be lost using direct mapping (ie, near-end and far-end status for the 10 GbE WAN PHY).

RPR MAC-PHY Mapping

- Propose the following optional MAC-PHY mapping to resolve the direct-mapping problems:

  - Retains duplex behavior and duplex fault-detection capabilities.

  - Allows mixing PHY types (different reaches) on different links.

  - <u>Not</u> proposing support for mixed PHY data rates.

- Implies pairs of opposing ringlets.

"Duplex" point-to-point physical layer instantiation

"Duplex" point-to-point physical layer instantiation

| PHY 1 RX | RPR MAC | PHY 2 TX | RINGLET 1 |

| PHY 1 TX | | PHY 2 RX | RINGLET 2 |

PHY NUMBERING, RING NUMBERING, AND RING DIRECTIONS ARE SHOWN FOR REFERENCE ONLY

Summary

- Define a P-SAP interface common to all RPR PHYs.

- Define a 10 GbE Reconciliation Sublayer for RPR to map the RPR P-SAP primitives to the P802.3ae PHYs and interfaces.

- Include the 10 GbE Physical Layer in the RPR standard by reference to P802.3ae (excluding the RS).
  - Support the optional XGMII with no changes.
  - Support the optional XGXS/XAUI with no change.
  - Support all seven PHYs with their associated sublayers with no changes.

- Assign RPR MAC-PHY mapping options to support:
  - Direct mapping of a PHY to a ringlet;
  - Mapping such that a PHY is mapped to two ringlets of a single ring segment.