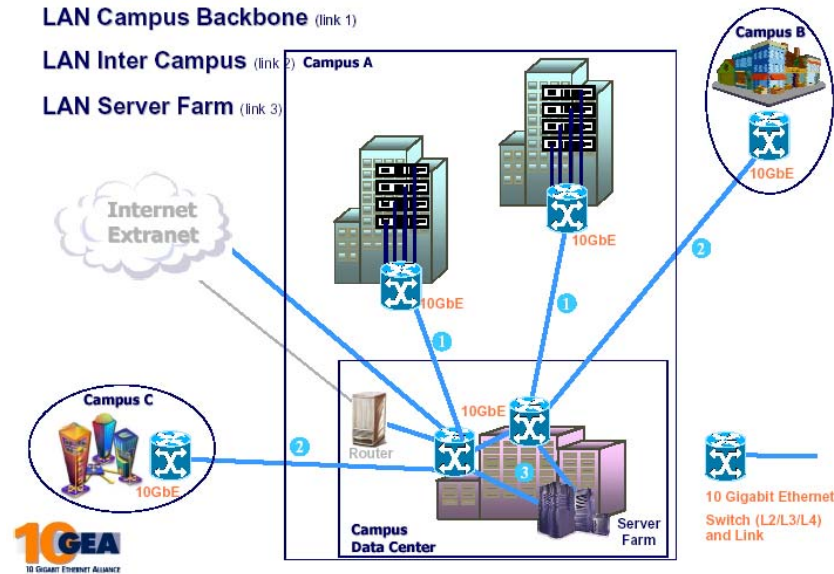


10GbE 规范是 IEEE802.3ae, 现时为 Draft 5.0 版本, 2002  
10GEA = 10 Gigabit Ethernet Alliance (<http://www.10gea.org>)

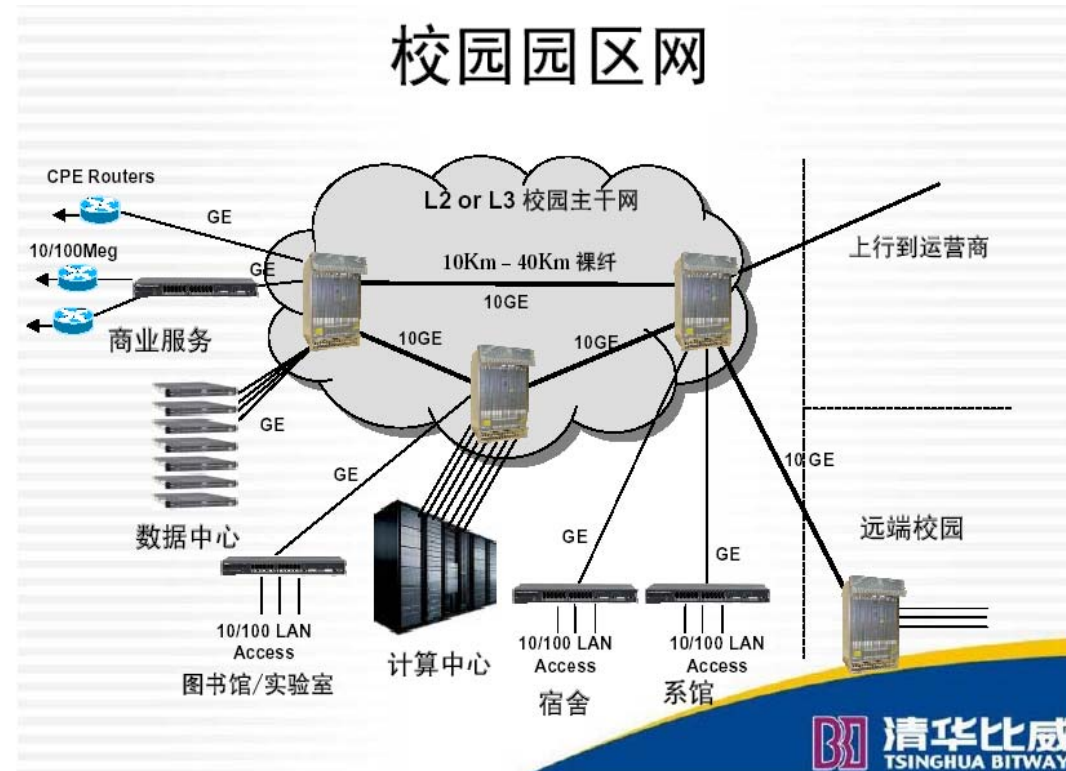
## 10GbE 的典型应用:

1 – 传统的 LAN 应用: 提升接入的带宽, 1GbE 的链路会被 10GbE 取替 : 例: Campus LAN 校园的局域网前面提升, Inter-Campus Link 校园之间的互连, Server Farm 提供内容寄存的服务商的连接等.

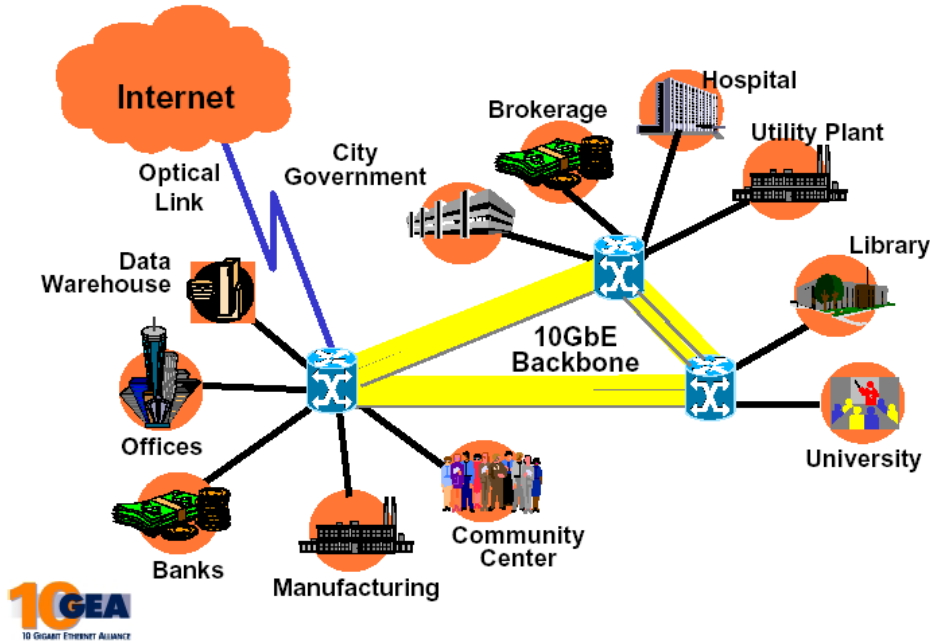
### Traditional LAN Applications:



## 校园园区网



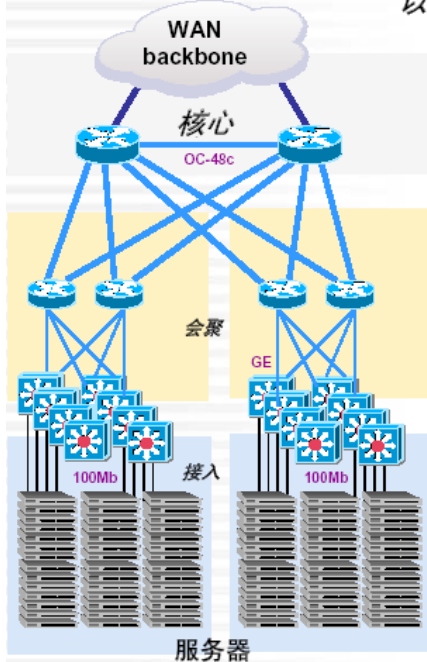
2 – MAN 城域网的应用: 本地接入网 Local Loop 将会由多种不同的方法提升现有的带宽, 如 B-ISDN, xDSL, FTTH (Fiber To The Home), 1-Mile Gigabit Ethernet 等, 使各用户能够实现高速连网, 多种应用变成可能, 如: 远程学校, 多媒体的娱乐(如: HDTV 40Mbps), 电子商务, 电子家居等. 10GbE 提升城域网骨干的数据交换能力, 以往使用 SDH 程控交换机, 在成本, 安装, 运维管理, 可扩展能力与寿命都比传统的 SDH 要好.



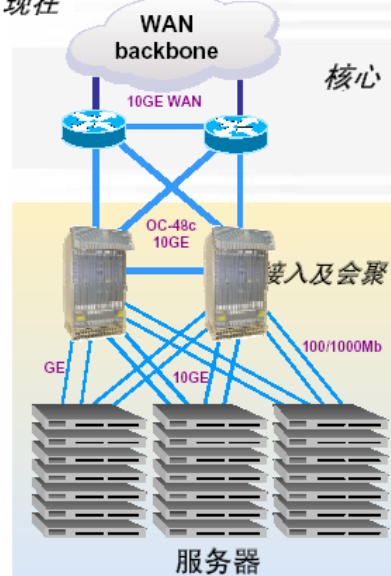
## 简化的下一代数据中心

以前

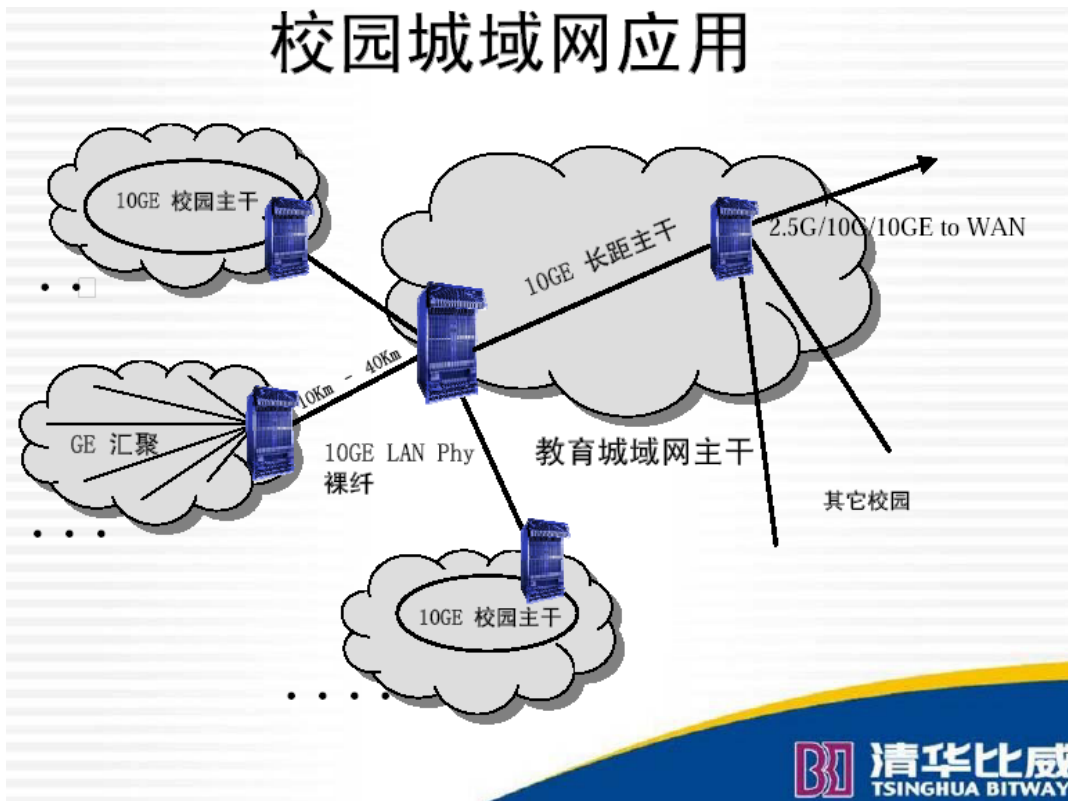
现在



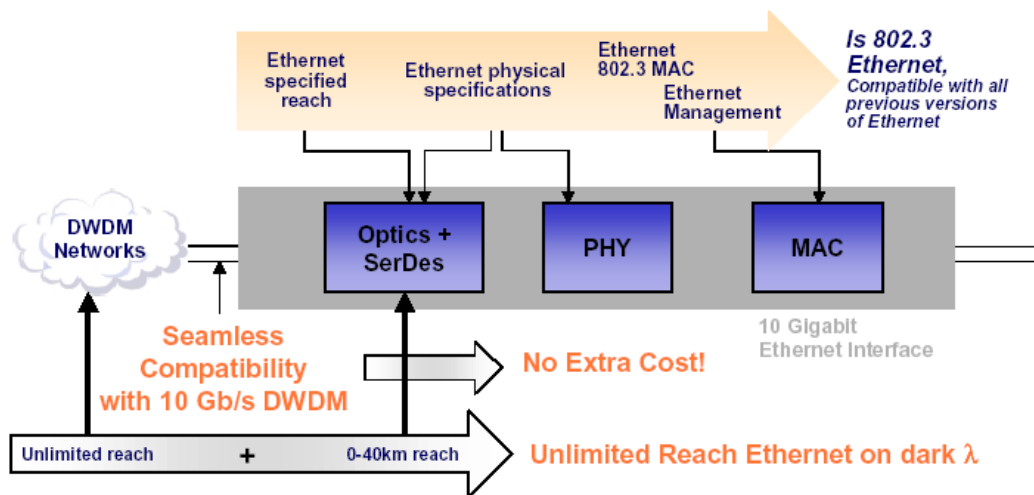
- 大容量及端口密度
- 简化了的运作
- 减少了网络层次
- 减少了设备数量
- 减少了机房空间
- 减少了功耗
- 简化了系统平台
- 减少了设备费用及管理费用 (TCO: CapEx & OpEx)
- 高度的未来投资保护



# 校园城域网应用



3 – WAN 广域网的应用: IEEE802.3ae 规范 10GbE 有两种物理层接口, 分别应用在 LAN 的 10GBase-R (LAN PHY – 数据速率为 10.3125Gb/s)与应用在 WAN 的 10GBase-W (WAN PHY – 数据速率为 9.953Gb/s). 10GBase-W 的数据流能插入 SONET STS-192c 或 SDH VC-4-64c 的载体 Payload, 再使用 SONET/SDH DWDM 光传输网传输, 实现无缝接入。



**10 GbE empowers Ethernet to DOMINATE beyond LAN, with unlimited reach at no extra cost!**



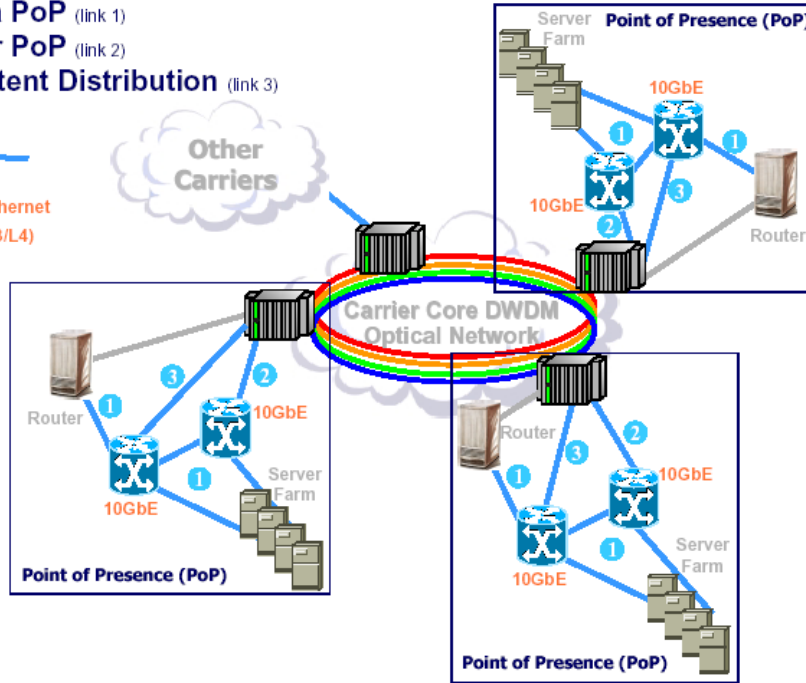
应用范畴：营运商机房 PoP(Point of Presence)内的互连，营运商之间机房的互连，营运商与 Content Provider 内容提供商的互连等。

## Carriers & Service Providers Applications:

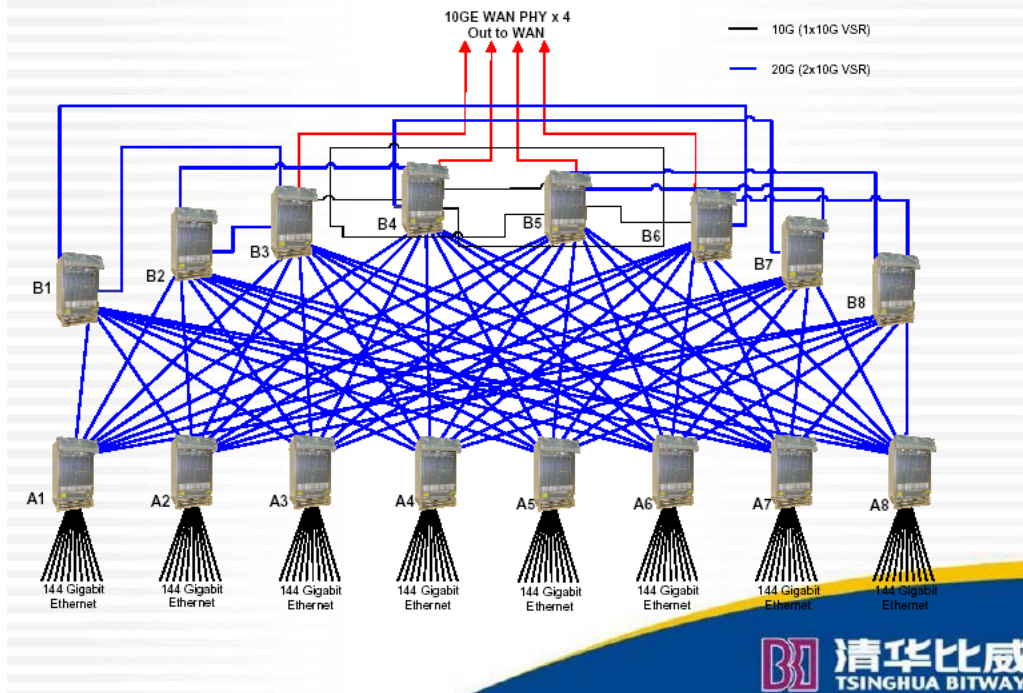
- WAN Intra PoP (link 1)
- WAN Inter PoP (link 2)
- WAN Content Distribution (link 3)



10 Gigabit Ethernet  
Switch (L2/L3/L4)  
and Link



## 超级计算中心试例 (1152 无阻塞服务器组)



## 为什么选择 10GbE?

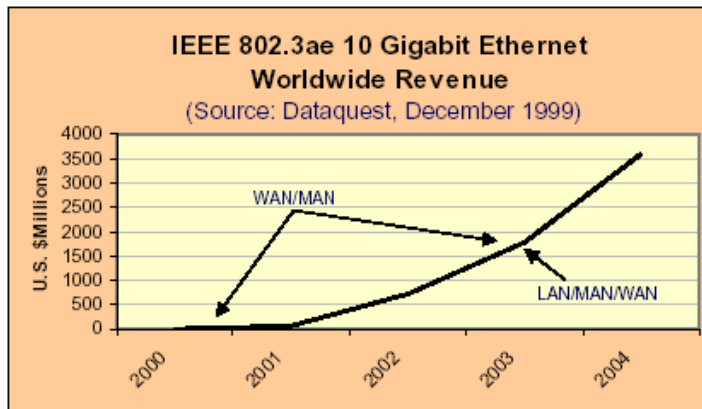
### OC-192 POS与10GE LAN/WAN PHY的对比:

#### OC-192 POS

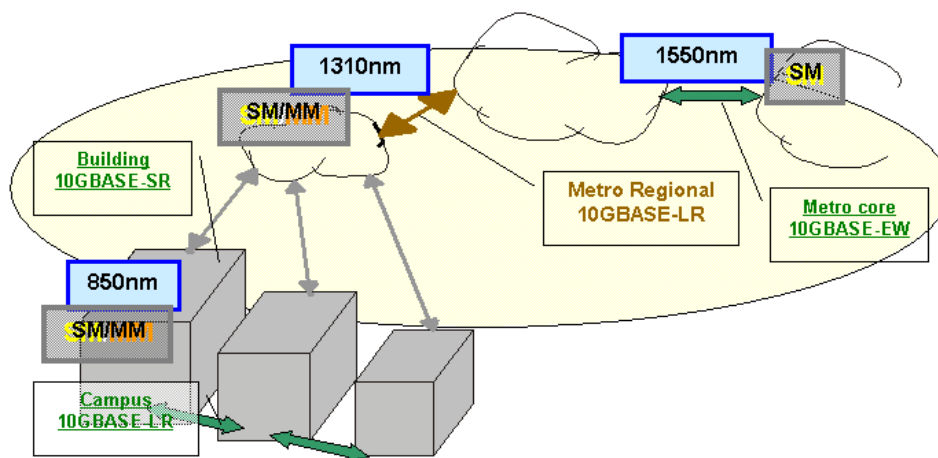
- 只支持OSI第三层
- 只支持IP (但是90% 的IP包是在以太网上产生的)
- 不能直接传输以太网 (ITU X.86, T1X1)
- PPP 封装(每个IP包多加8个字节)
- 丰富的检测和维护能力
- 费用高

#### 10GE LAN/WAN PHY

- 支持OSI第二层 (交换Switching)和第三层(路由Routing)
- 支持所的路由协议 (RIP, OSPF, BGP)
- 实现端到端的以太网直接传输
- 丰富的检测和维护能力
- 费用低



按估计, 到 2004 年, 将有超过 800,000 个 10GbE 接口, 10GbE MAN/WAN 市场将达 US\$1.8B, 10GbE LAN 市场将达 US\$3.6B. 从技术角度估计, 10GBase-LR 的城域网互连将会是热点, 10GBase-LR 实现 10KM 的 10Gb/s 连接, 直接解决现有接入网瓶颈的问题 (注: 10GBase-ER 实现 40KM 的连接, 但是技术上实现比较困难, 首批多个 10GbE SerDes, 设备制造厂都偏向支持 10GBase-LR)



**不同种类的 10GbE:**

10Gbs/s 以太网有两类，分别是串行的 10GBase-S/L/E-R/W 与 4 路并行 WDM (波分复用)的 10GBase-LX4

**10GBase-S/L/E-R/W:**

- S = Short Wavelength, Short Range, 850nm
- L = Long Wavelength, Long Range, 1310nm
- E = Extended-Long Wavelength 1550nm, Extra-Long Range

**W = WAN PHY 广域网物理层**, 9.95328Gb/s 码率, 与 SONET STS-192c 及 SDH VC-4-64C 兼容, 9.95328Gb/s 的以太网物理层能够数据直接插入在 SONET STS-192c 或 SDH VC-4-64C Payload 载体里, 再使用 DWDM 或 SDH/SONET 光 / 传输网作传送, 使 10G 以太网无缝接入 SDH, 再不需要像以前在接入前调整速率.

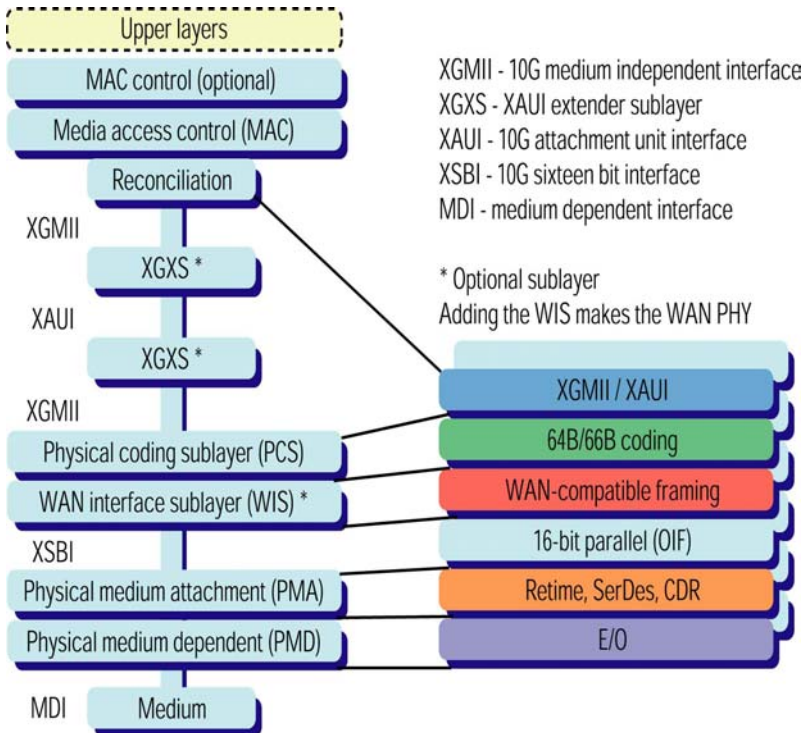
**R = LAN PHY 局域网物理层**, 10.3125Gb/s 码率

所以串行的 10GbE 共有 6 种, 分别是 10GBase-SR, 10GBase-LR, 10GBase-ER 与 10GBase-SW, 10GBase-LW, 10GBase-EW 6 种.

在串行 10GBase-R, 10GBase-W 中, 使用 64B/66B 编码方法, 而 4 路并行的光 WDM 确 10GBase-LX4 则沿用与千兆以太网相同的 8B/10B 编码方法. 通过 64B/66B 编码, 原来的 10Gb/s, 会变成 10.3125Gb/s (=10Gb/s X 66/64). 所以 10GBase-R 的速率是 10.3125Gb/s. 但是因为 10GBase-W 要与 SONET/SDH 无缝兼容, 所以在 10GBase-W 的定制里加入了 WIS (WAN Interface Sublayer)协议层, 主要功能将 10GbE LAN 应用中的 IDLE 码除掉, 使速率降到 9.95328Gb/s, 与 SDH VC-4-64c 与 STS-192c 兼容.

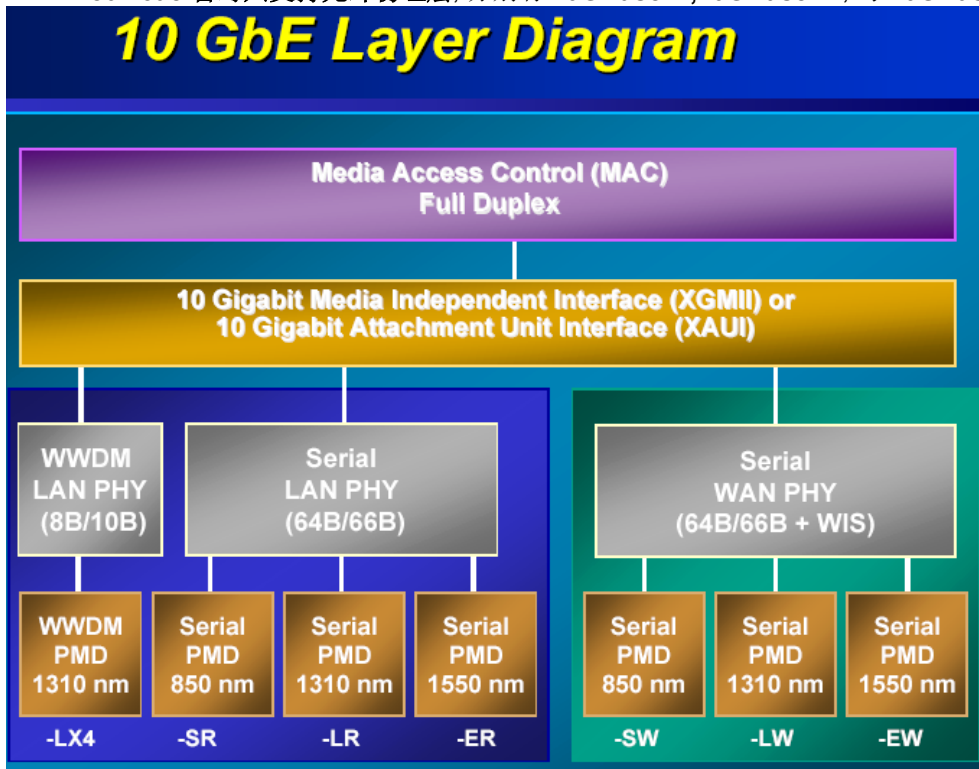
	Media λ nom. [nm] / mode / max. distance	Rate, Phy Type	Rate, Phy Type
Line Bit Rate		10.3125Gb/s	9.953Gb/s
10GBASE-S... (short λ)	850 / MM / 200m	...-R standard: 64b/66b coding	... -W standard: 64b/66b coding; Data and format compatible with WAN ( SONET STS-192c and SDH VC-4-64c .)
10GBASE-L... (long λ)	1310 / SM / 10km		
10GBASE-E... (Extra long λ)	1550 / SM / 40km		
10GBASE-X, i.e. 10GBASE-LX4	1310, 4x / MM: <300m	-LX4: 8b/10b coding; 3.125Gb/s per link. No WAN compatibility. Probably a smaller market.	
	1310, 4x / SM: <10km		

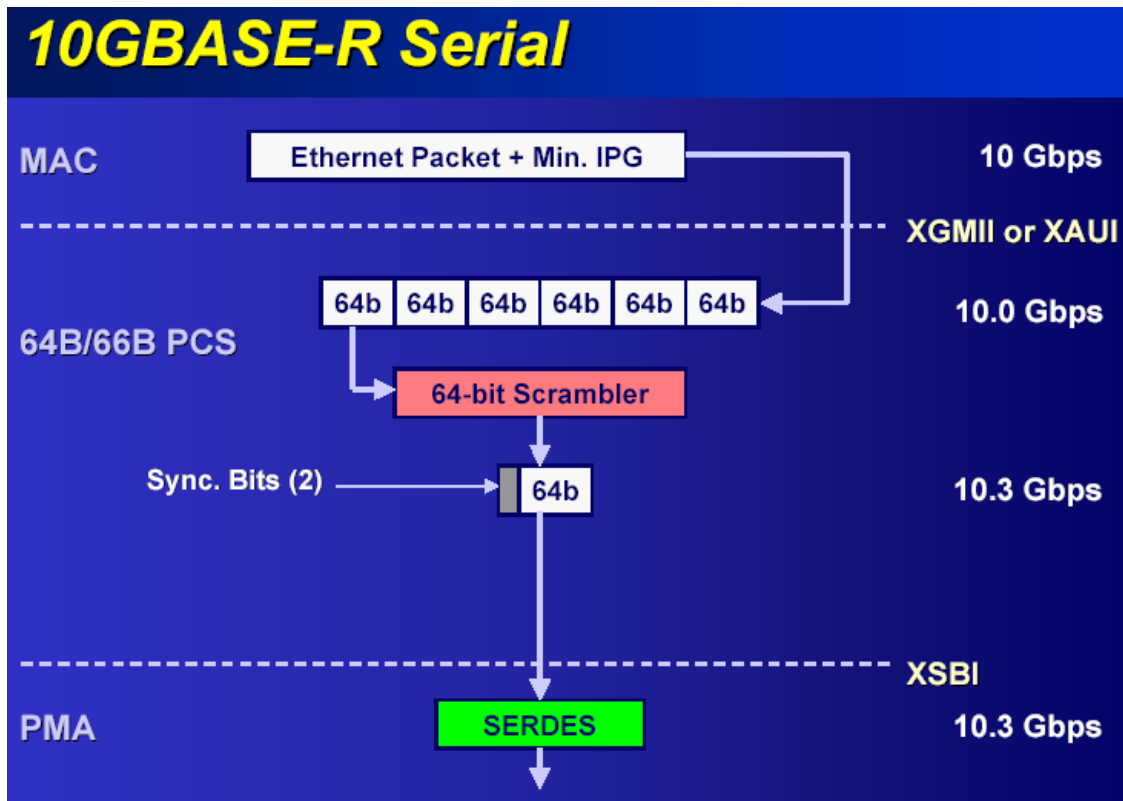
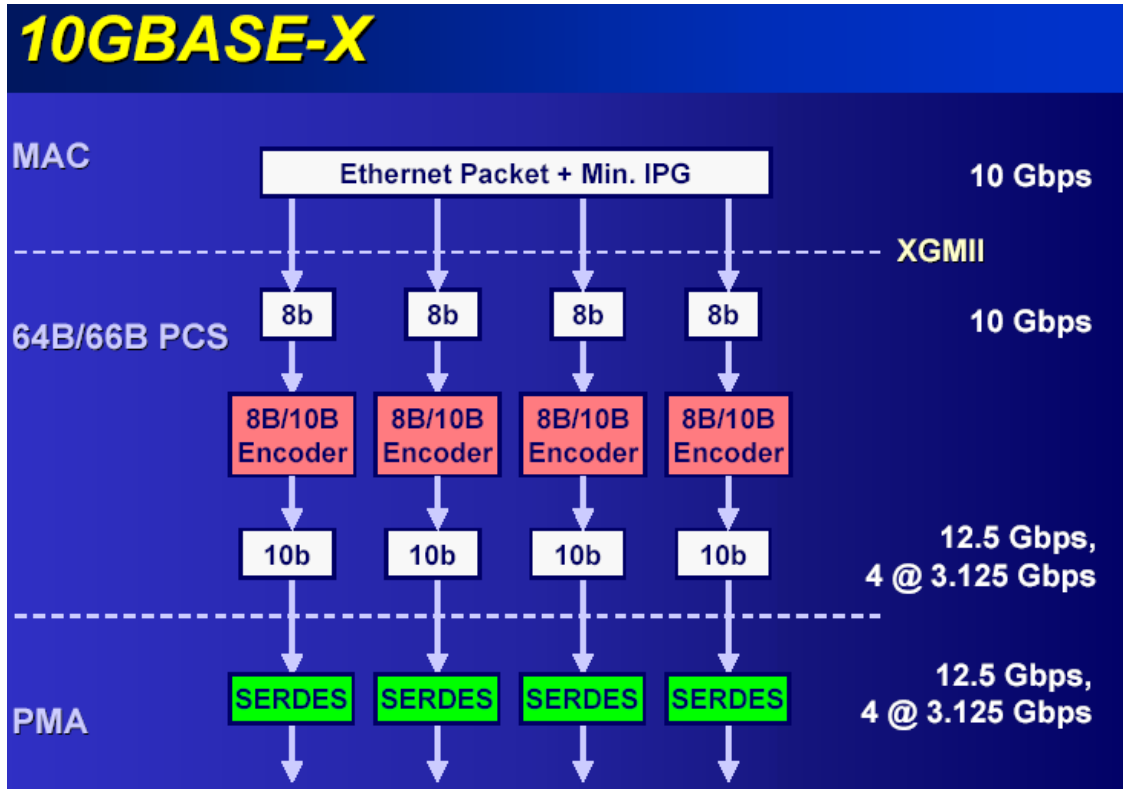
**10GbE IEEE802.3ae 对照 OSI 参考模型的介绍:**



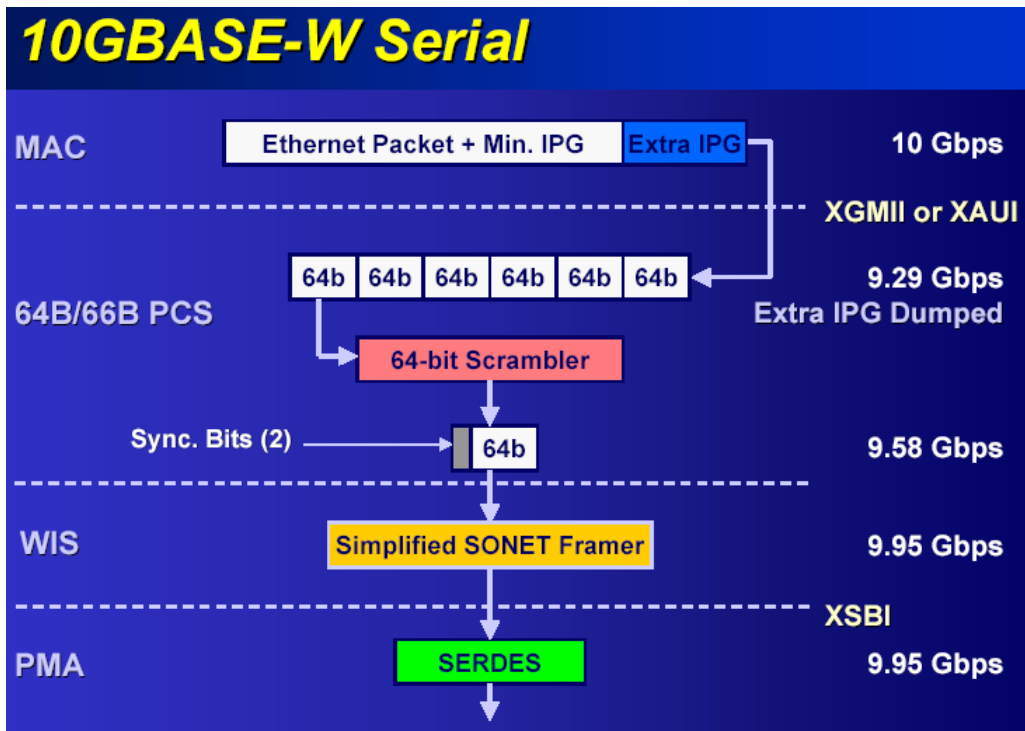
**第 1 层 - PMD Physical Media Dependent 物理媒介依赖层:**

IEEE P802.3ae 暂时只支持光纤物理层, 分别有 10GBase-R, 10GBase-W, 与 10GBase-LX 3 种.









**第 2 层- PMA Physical Media Attachment (SerDes – Serializer/De-Serializer 串并处理) :**

**PMA 的主要功能:**

接收端(从 PMD 物理层接收数据, 通然后发送到 PCS 编码层/WIS 层)的功能:

- 1- 从输入 PMD 的串行数据中**恢复时钟**, 并将时钟传送到 PCS 编码层或 WIS 层
- 2- 将 10Gb/s **串行数据并行化**, **变成 16bit 并行**, **速率为 644.53Mbps** (16bit X 644.53Mb/s =10.351Gb/s)LVDS 数据(时钟与数据分别传送), 通过 XSBI(10G Sixteen Bit Interface)接口, 发送到 PCS 编码层或 WIS 广域网接口子层
- 3- 提供链路状况信息

发射端(从 PCS 编码层或 WIS 层接收数据, 然后发送到 PMD 层)的功能:

- 1- 提供时钟给 PCS 编码层或 WIS 层
- 2- 从 PCS 编码层或 WIS 层所接收的 16bit 并行数据进行串行化, 变成 10Gb/s 串行数据发送到 PMD 层

**PMA 层的串行化与并行化过程: (业内一般使用 SerDes 串并处理器实现串行化与并行化过程)**

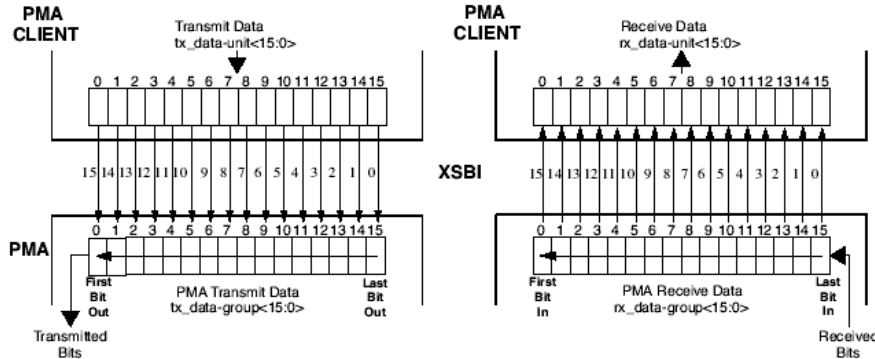


Figure 51-2—XSBI physical interface mapping into PMA service interface

为什么数据要串行化?

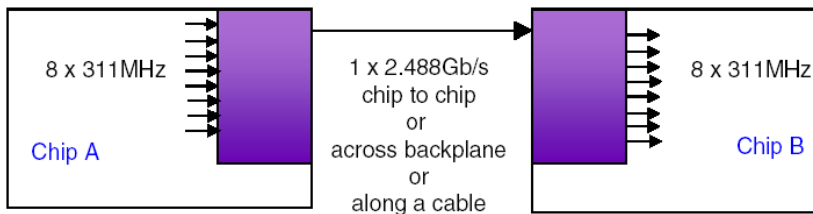


What is a Serdes

Serializer/Deserializer -- reduces lines on boards (chip to chip) and across backplanes; increases speed of signal through cables

- Serializer (mux) -- gathers many low speed parallel signals, converts them to one high speed serial signal
- Deserializer (demux) -- receives high speed serial signal, recovers the clock and data, reconfigures it into many low speed parallel signals
- Signal speeds depend on the version of Serdes
  - Serial high speeds -- 622Mb/s, 1.0 - 3.125 Gb/s, 1.22 - 3.18 Gb/s, or 10Gb/s
  - Parallel low speeds -- 311MHz or 156MHz

Example shows a Serdes macrocell in Chip A transmitting a 2.5Gb/s signal to the Serdes macrocell in Chip B. The rest of the device operates at 311MHz.

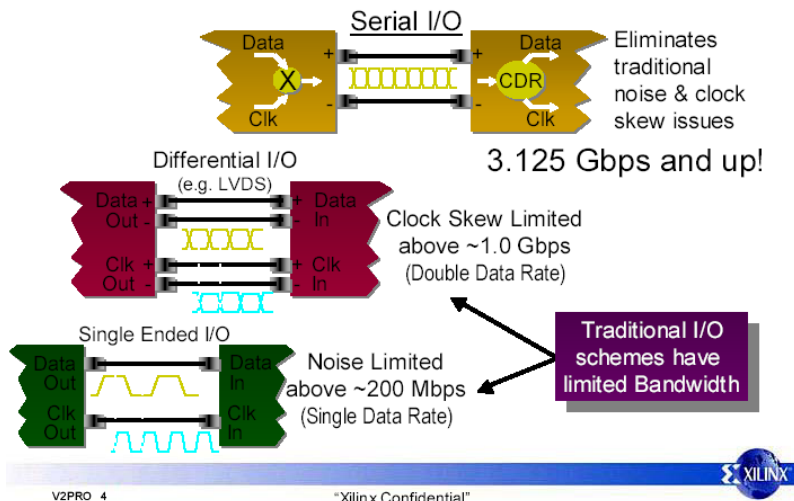


现今在高速的数据传送上, 非常流行串行数据传输的概念. 不单应用在 GbE 与 10GbE 上, 新定制的 Multi-Gigabit I/O 总线, 如 Infiniband, PCI Express 等无不使用 串行+ 嵌入时钟 方法来突破在传统并行 I/O 的两个主要缺陷:

1 - 由于 PCB 阻抗不匹配所导致的微小信道间的传输时差, 使并行的数据不能同时到达接收端. 而当数据的速度越来越快时, 这个在多个并行信道间的时间偏差就相对地显得过大而不能接收, 并且因为信道数目多的时候, 而为了控制信道间的传输时差, 使得设计时非常复杂 - 而串行数据就能免掉信道间偏差的问题.

2 - 传统的并行数据传输, 时钟是与数据是独立地传送的, 往往数据与时钟之间因传输而产生时差 Skew, 导致时钟是与数据不能同时到达接收端. 同样, 嵌入时钟就能免掉数据与时钟间偏差的问题.

Advantages of Serial Connectivity



V2PRO 4

"Xilinx Confidential"



**WAN Interface Sublayer WIS (只在 10GBase-W 中使用):**

在 10GBase-W WAN PHY 里, 在 PCS 64B/66B 编码层与 PMA 串并处理层之间, 加插了 WIS 层. 而 WIS 层的主要功能:

- 1- 提供 9.95328Gb/s 数据速率, 可插入在 SONET STS-192c SPE Synchronous Payload Envelope 里, 或 SDH VC-4-64c 的 Virtual Container 里.
- 2- 支持 SONET/ SDH 的帧格式 Framing(例如:加入/去除 Section, Line Path Overhead, 加入/检查 B1, B2, B3), 扰码 Scrambling, 与错误检测 Error Detection 功能

注: B1 为 BIP-8 (Bit Interleaved Parity, 8Bit), 用途为检测 Regenerator Section Overhead 里在传输时有否出错, B2 为 BIP-N (Bit Interleaved Parity, N X 24Bit), 用途为检测 Multiplex Section Overhead 里在传输时有否出错, B3 为 BIP-8 (Bit Interleaved Parity, 8Bit) 用途为检测 Path Overhead 里在传输时有否出错, 在计算这写字节 Octects 的 Parity("0", "1"的对称性)得知 SDH 帧在传输时有否出错.

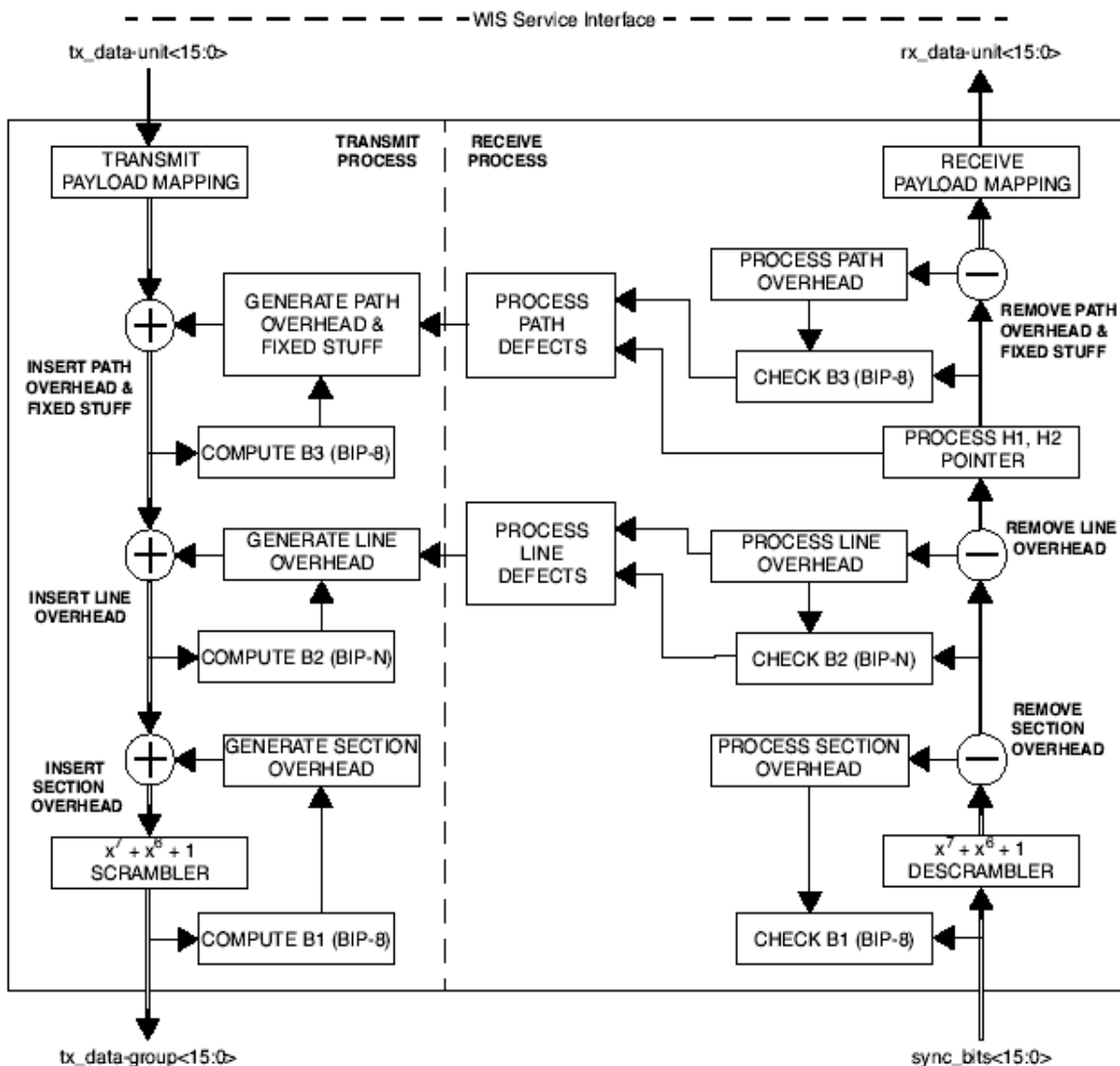


Figure 50-3—WIS Transmit and Receive processes

以下是标准的 STS-192 或 VC-4-64c 帧格式 (Path Overhead, Fixed Stuff 与 Payload):

COLUMN	1	2	64	65	66	67	68	69	16703	16704	
ROW 1	PATH OVERHEAD	FIXED STUFF	FIXED STUFF	0	1	2	3	4	-----	16-638	16-639
ROW 2				16-640	16-641	16-642	16-643	16-644	-----	33-278	33-279
ROW 3				-----	-----	-----	-----	-----	-----	-----	-----
ROW 4				-----	-----	-----	-----	-----	-----	-----	-----
ROW 5				-----	-----	-----	-----	-----	-----	-----	-----
ROW 6				-----	-----	-----	-----	-----	-----	-----	-----
ROW 7				-----	-----	-----	-----	-----	-----	-----	-----
ROW 8				116-480	116-481	116-482	116-483	116-484	-----	133-118	133-119
ROW 9				133-120	133-121	133-122	133-123	133-124	-----	149-758	149-759

PATH OVERHEAD
FIXED STUFF
PAYLOAD CAPACITY

以下是标准的 STS-192 或 VC-4-64c 帧格式 (Section Overhead 与 Section Overhead):

COLUMN	1	2	3	192	193	194	195	384	385	386	387	576				
Section Overhead	ROW 1	A1	A1	A1	-----	A1	A2	A2	A2	-----	A2	J0	Z0	Z0	-----	Z0
	ROW 2	B1														
	ROW 3															
Line Overhead	ROW 4	H1	H1	H1	-----	H1	H2	H2	H2	-----	H2	H3	H3	H3	-----	H3
	ROW 5	B2	B2	B2	-----	B2	K1					K2				
	ROW 6															
	ROW 7															
	ROW 8															
	ROW 9	S1							M1							

Synchronous Payload Envelope

NOTE—The Section and Line Overhead comprise 5184 octets in total (576 x 9). Only 1349 octets are defined for the WIS. Octets that are undefined and unused by the WIS are indicated as blank boxes in this figure.

注: SDH/SONET 每秒传送 8000 帧, 所以传输速率为  $[(16704 + 576) \times 9 \times 8] \times 8000 = 9.95328\text{Gb/s}$

WIS 层生成 SDH/SONET 格式的流程如下:

- 1 - 产生与加入 Path Overhead 并计算 B3
- 2 - 产生与加入 Line Overhead 并计算 B2
- 3 - 产生与加入 Section Overhead, 扰码并计算 B1

因为 SDH/SONET 传输次序是 MSB 到 LSB (即 Bit1 到 Bit8), 而 Ethernet 传输次序则是 LSB 到 MSB (即 Bit0 到 Bit7), 为了 MAC 所使用的 Frame Checksum 仍然有效, 故 WIS 层在收发 PCS 层或 PMA 层的数据时, 需要进行 Bit Re-Labeling.

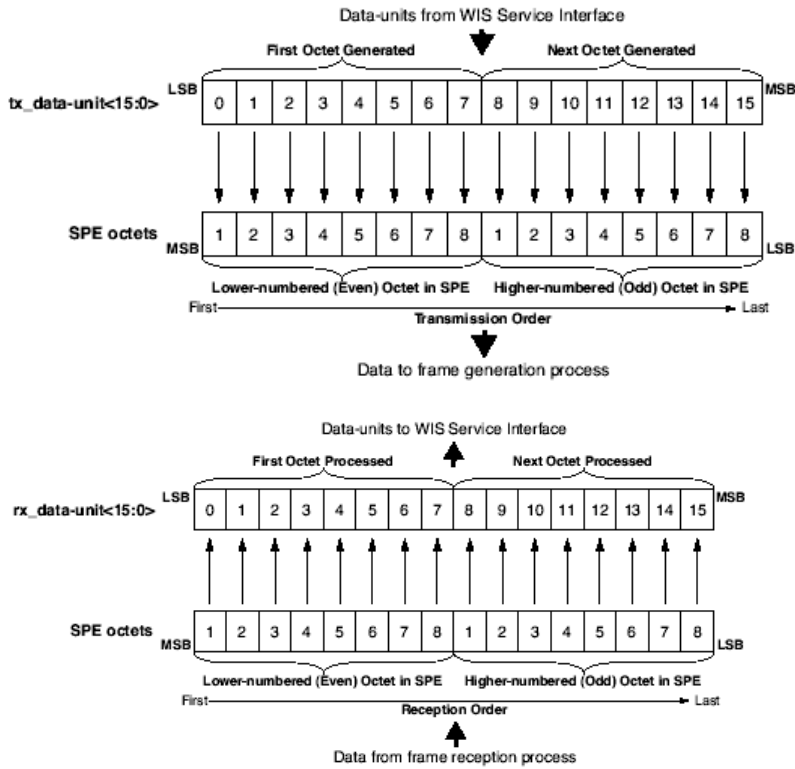


Figure 50-6—Receive bit relabeling function

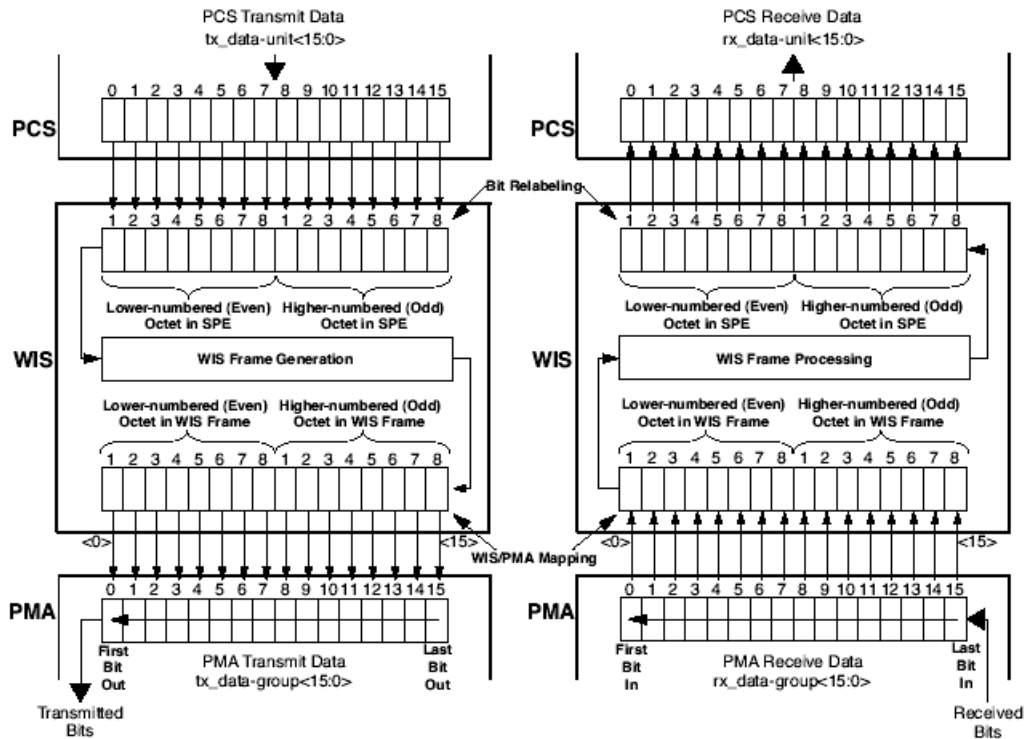
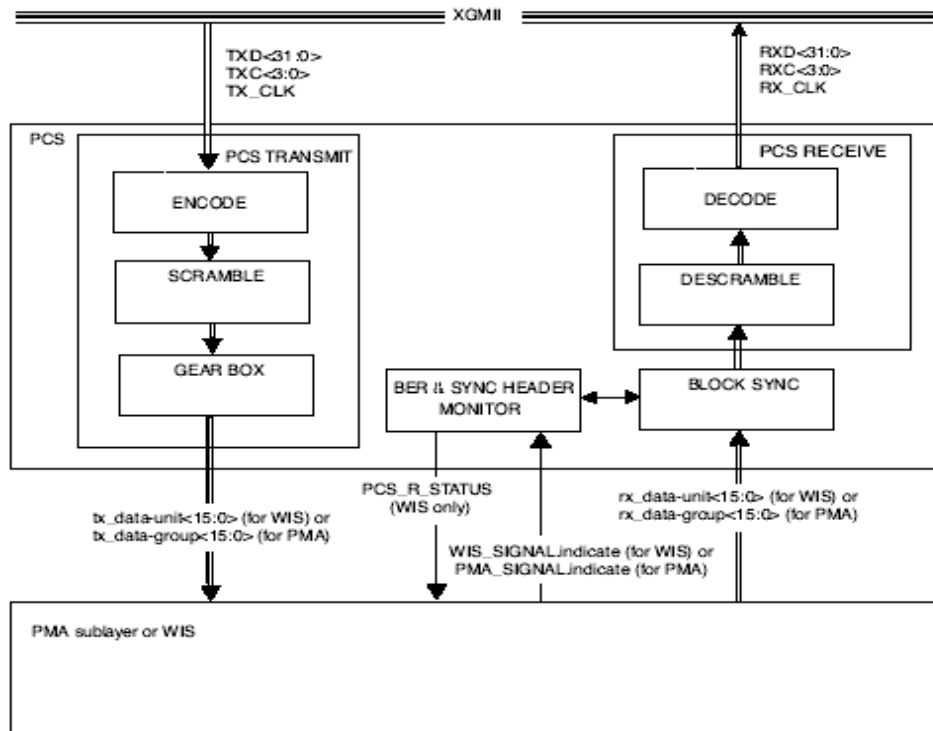


Figure 49-3—10GBASE-W PHY Transmission order

**第3层- Physical Coding Sublayer PCS (64B/66B 编码) :**

串行 10GBase-R(LAN PHY)与 10GBase-W (WAN PHY) PCS 编码层主要负责:

- 1- PCS Transmit (PCS 帧传送) - PCS 帧传送有两种模式:
  - a. 正常模式 Normal Mode : **64B/66B 编码**时加入 2-bit 同步帧头(Synchronization Header), 分别是“01”=“Data”字符, “10”=“Control”控制字符, 目的: 保证数据前 2-bit 一定有信号的跳变, 从而让接收器恢复时钟作同步. 然后 Scrambling (注: 同步帧头不需扰码), 再通过 Gearbox 输出 16bit 数据流到 PMA (LAN PHY)或 WIS(WAN PHY)层
  - b. 抖动模式 Jitter Mode : 预定的抖动码型嵌入在数据流中传送, 监测误码率
- 2- PCS Receive (PCS 帧接收) (Block Sync 块同步, De-Scramble, 66B/64B 解码)
- 3- Block (Frame) Synchronization 帧同步对准 Alignment
- 4- Bit Error Monitoring 误码监测, 支持少于  $10^{-12}$  BER 误码率
- 5- 为了使 10GbE 以太网能与 SONET/SDH 兼容, PCS 层将 10Gb/s 码速的 XGMII 码流中除掉 IDLE 帧, 将速率降为 9.95328Gb/s, 与 SONET/SDH 兼容. 反之, 从 WIS 层接收 9.95328Gb/s 的码流, PCS 层则加上 IDLE 帧, 使码速变成 10Gb/s, 发送到 MAC 层.
- 6- 支持 Full Duplex MAC



在串行 10GBase-R, 10GBase-W 中, 使用 64B/66B 编码方法, 以下是使用 64B/66B 编码的目的:

- 1- 64B/66B 编码使原由发射器的数据, 经过编码后, 产生足够的跳变边沿, 让接收器能从跳变中恢复时钟作同步
- 2- 检测单个或多个传输数据误码变得更容易, 尤其当单个误码成为数据块时, 只要计算块内的 CRC Checksum 便能得知有没有误码
- 3- 加入 2-bit 同步帧头 Synchronization Header, 让接收器能将实现数据块定时对准 Block Alignment

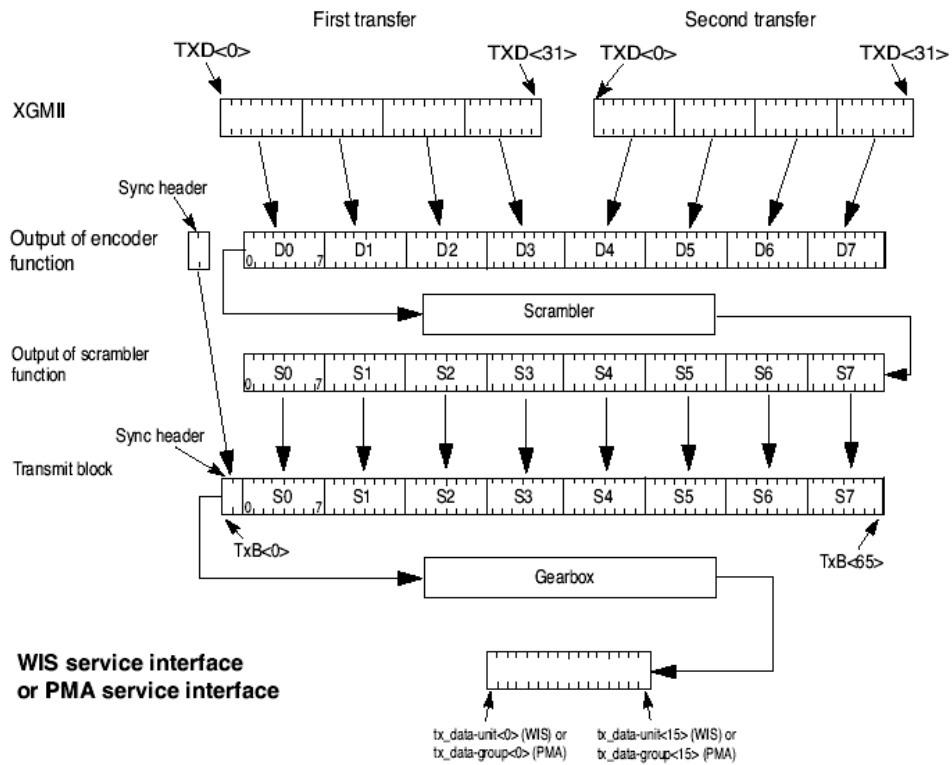


Figure 49-5—PCS Transmit Bit Ordering

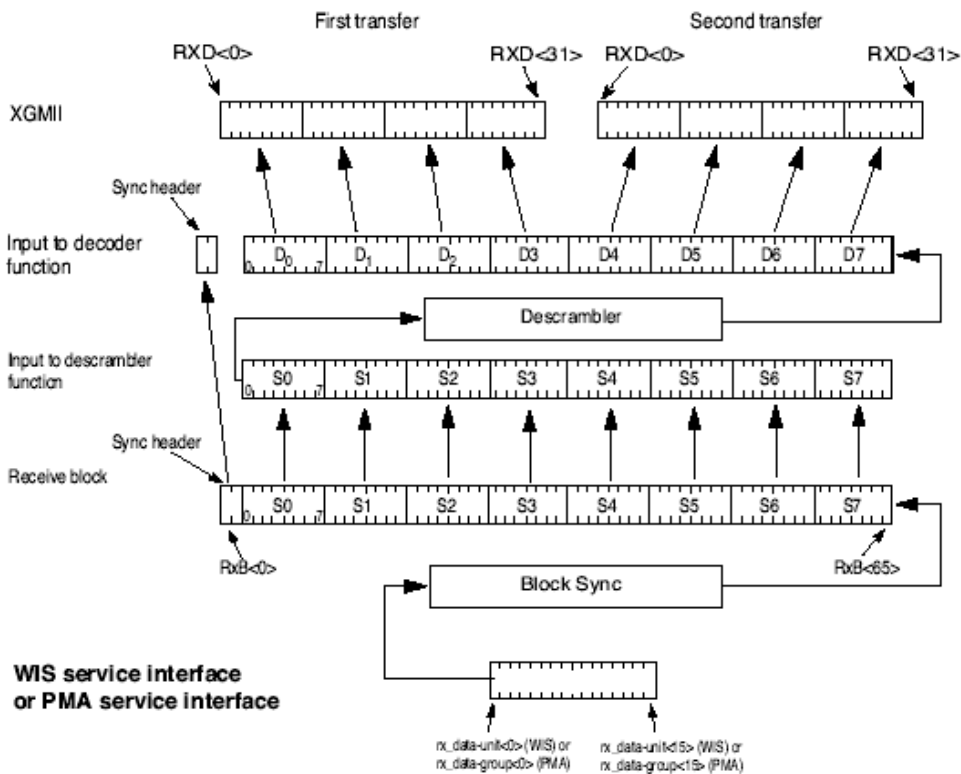
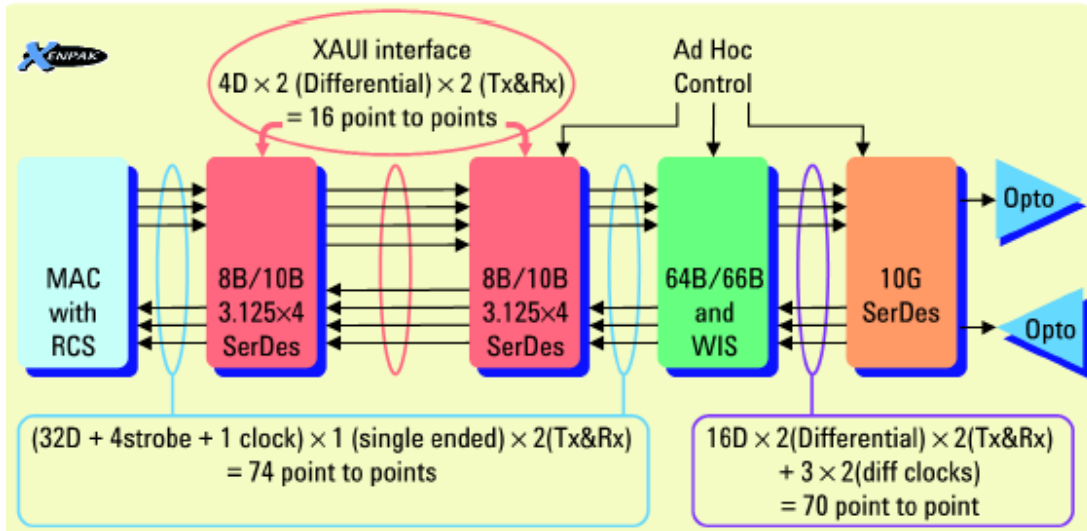


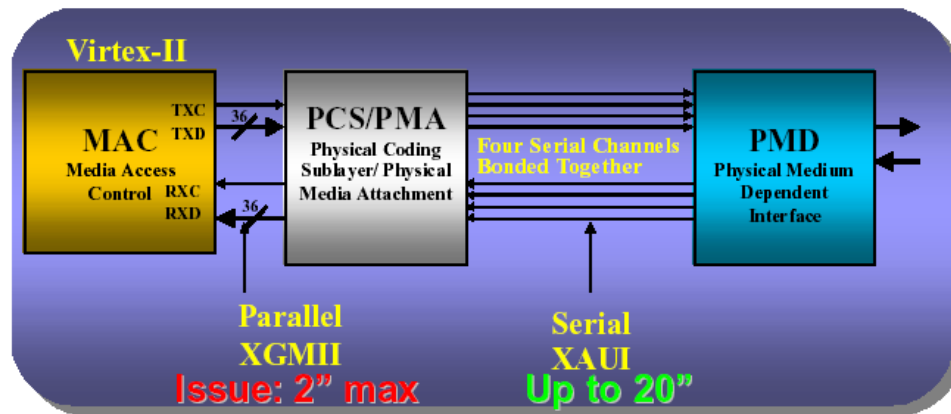
Figure 49-6—PCS Receive Bit Ordering

**XGMII 10 Gigabit Media Independent Interface (X = 10 in Roman) :**

连接 10Gb/s MAC 与 10Gb/s PHY, 是 Chip-to-Chip 芯片间的传输接口方式, 在 IEEE802.3ae 规范中定义为可选的接口. 实际上, 因为 **XGMII 的极限传输距离只有 7cm (3 inch)**, 不能传输很远, 并且实现 **XGMII 时需要 74 腿(32 路发数据+32 路收数据+14 路控制线)**的芯片, 使芯片的 PIN-COUNT 增加, 使制造的难度与成本也增加. 所以 IEEE802.3ae 规范同时也定义 XAUI 接口, XAUI 芯片间数据传送, 绝大部分生产串并 IC SerDes (Serializer / Deserializer)的厂家, 都会将 PCS, PMA 与 MAC 层的功能合并到同一芯片实现, 继而采用 XAUI 来取替 XGMII 作为芯片间数据传送的接口方式.



MAC 层功能, 与 PCS/PMA 层功能分别由不同的 IC 实现, 数据需要通过 XGMII 接口从 MAC 层发送到 PCS 层:

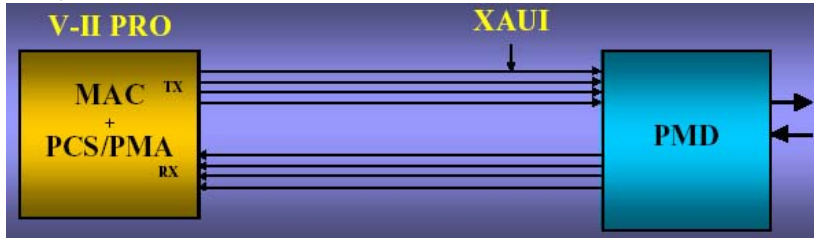


**XAUI (10 Gigabit Attachment Unit Interface) :**

连接 10Gb/s MAC 与 10Gb/s PHY, 是 Chip-to-Chip 芯片间的传输接口方式. 主要设计目的是为了扩展 10Gb/s MAC 与 10Gb/s PHY 的在 FR4 板材上实现达 **50cm(20 inch)**的传输距离, 更适合于 Chip-to-Chip 或背板高速信号的传诵. **XAUI 是 4 路平行的差分信号, 每路以 3.125Gb/s 的速度传送, 并且时钟是嵌入在数据里的, 使实现高速长 距离的数据传输比 XGMII 更容易.**

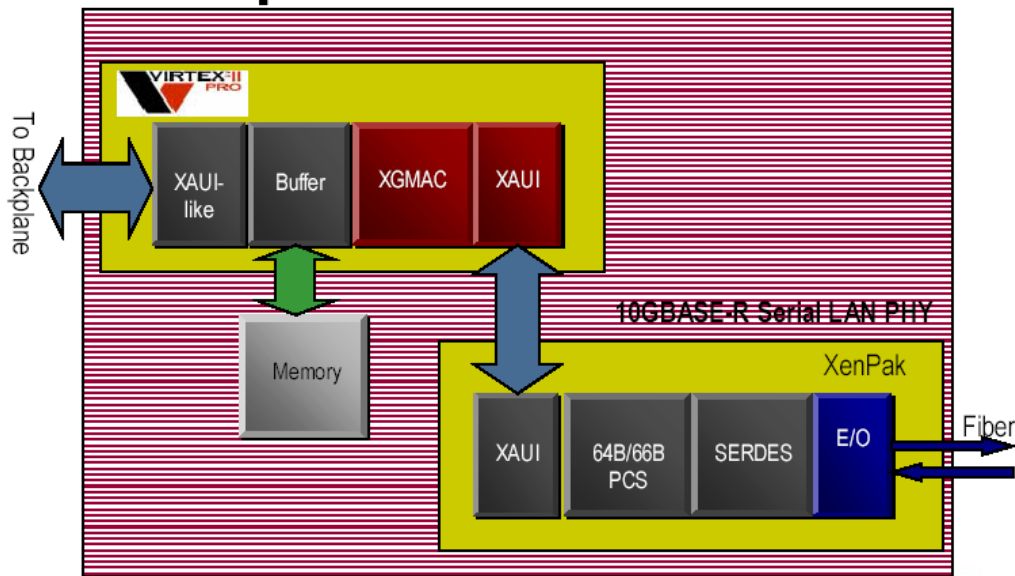


MAC 层功能与 PCS/PMA 层功能由同一的 IC 实现, 数据只需要从 MAC/PCS/PMA 芯片通过 XAUI 接口传送到 PMD 芯片即可以, 因为 XAUI 的关系, 可以实现到 50cm 的传输距离 – 适用于芯片间的传输, 或是背板的长距离传输:

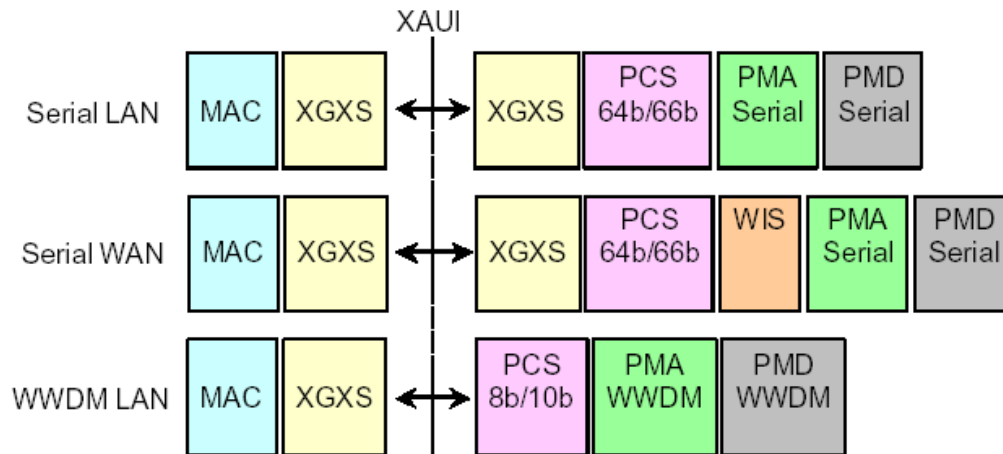


例:Xilinx 的 Virtex II Pro FPGA 支持 XAUI 接口, 将数据从 MAC 层传送到 PCS 层:

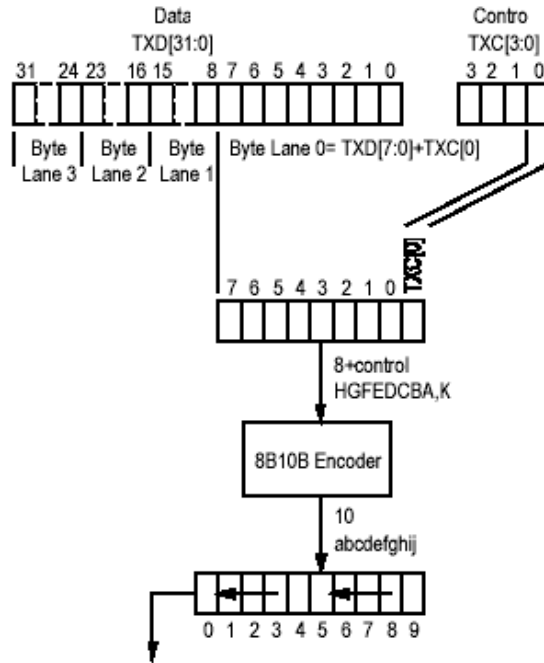
## Simple 10GbE Line Card



注: XGXS (XAUI eXtender Sublayer)负责将 XGMII 32bit 的码流, 分成 4 路 (4-Lane), 加入 Control Code, 然后 8B/10B 编译成 XAUI 码流,



# Mapping from XGMII to XAUI



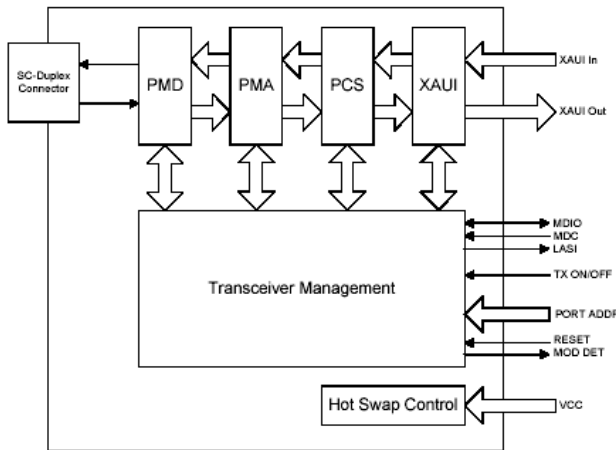
3 种 10GbE 中采用的接口方式对比:

Name	Number of Lines	Data Rate Per Line	Number of Clock and Control Lines	Useful Transmission Distance	Comments
XGMII 10Gb/s Media Independent Interface	32 TX 32 RX	312.5 Mbps	4 Tx Ctrl 1 Tx Clk 4 Rx Ctrl 1 Rx Clk	< 7cm (3")	74 Signal Lines
XSBI 10Gbps Serial Bus Interface	16 TX 16 RX LVDS	622 Mbps	1 Tx Clk 1 Rx Clk	<20 cm (8") + 1 connector	64 Signal lines
XAUI 10Gbps Attachment Unit Interface	4 TX 4 RX Differential	3.125Gbps	None	>50 cm (20") + 2 connectors	16 signal lines

**什么是 Xenpak? XPAK?**

业内为了加速 10GbE 的发展，一些公司提倡了 version 一些 MSA (Multiple Source Agreement), 协议中定制一套标准的 Plug&Play Module 模块形式，较流行的有 Xenpak 与 XPAK 两种，目的是使多间公司能够按照这些标准，制造 Pluggable 10GbE Transceiver Module，而模块形式采用 XAUI 为标准的 MAC 层到 PCS 层的输入/输出方法，另一端已经是 MDI 光纤的输入/输出，使真正制造 10GbE 单板与路由器 / 交换机系统的厂商能够：1)集中精力研发板内，板间与系统级的设计与调试；2) 市场上有更多的 Pluggable 10GbE Transceiver Module 选择，也同时缩短了总体开发时间与产品上市时间；3)使维修有关光发射器部分更快更方便—缩短了 MTTR(Mean Time To Repair)及相关成本。

Fig 11. Functional Diagram of Typical XENPAK Style Transceiver



标准 Xenpak 可插式 10GbE 光收发器：

This diagram illustrates the standardized architecture of a Xenpak 10GbE transceiver module. It is divided into several functional sections:
 

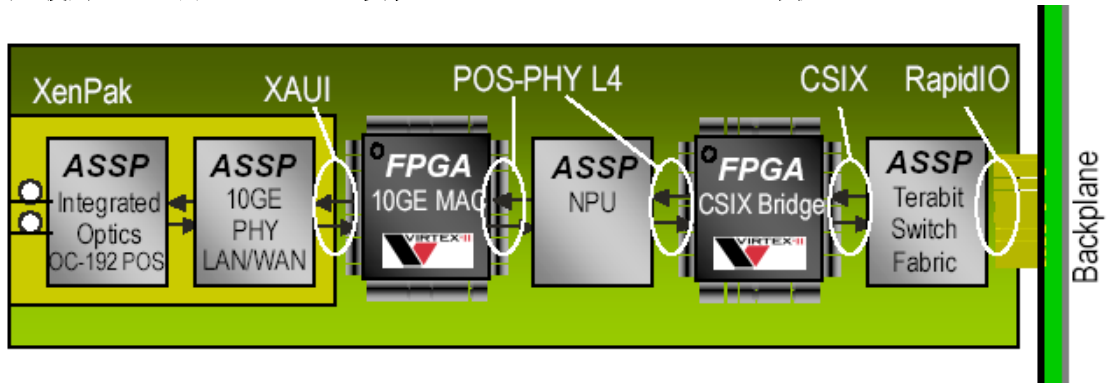
- Control, PSUs, Monitors**: Management and power sections on the left.
- MAC with RCS**: Media Access Control with Retention Control Sublayer.
- 8B/10B SerDes**: The interface between the MAC and the optical layer.
- SFP based electrical connected technology**: The central interface connecting the module to the system board.
- 8B/10B SerDes**: The interface between the system board and the optical layer.
- MDIO**: Management Data Input/Output interface.
- 64B/66B and WIS**: Wireline Interface Sublayer.
- 10G SerDes**: The high-speed data interface.
- Opto**: Optical transmitters and receivers.
- SCD connector technology**: The physical connector on the right.

 A list of features is provided below the diagram:
 

- Allows multiple sources of all IEEE defined PMD types
- All port types interchangeable via front panel pluggability
- Proven connectorized optical interface
- Proven electrical interface
- Standardized, control, PSUs, monitors
- Supports very high physical port densities (8 per 19 inch card)
- Extensive design focus on EMI and thermal management



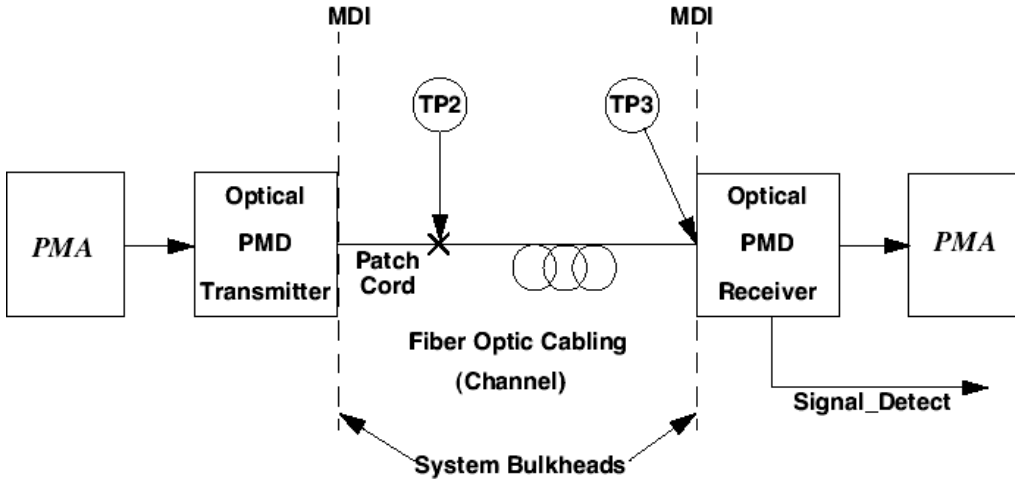
例：使用 Xilinx 的 Virtex-II Pro 实现 10GbE LAN/WAN Line Card 单板：



(注：10GbE 单板设计与调试中，也采用 POS-PHY L4(即 SPI-4 Phase II), CSIX, Rapid IO 等总线. Tek TLA700 提供对所有以上总线的支持包，提供反汇编与多总线系统级的调试功能)

**IEEE802.3ae 10GBase-R/W 以太网物理层测试:**

IEEE802.3ae Draft 5.0 暂时只支持光载体. 对应 IEEE802.3ae 测试分两个层次:



**1—测试光发射器(Test Point 2)与接收器 (Test Point 3)的光电参数指表:**

使用示波器, 对应 Test Point 2 测试发射器的:

- 发射器眼图模板测试 Transmitted Eye Mask
- 光调制幅度测试 OMA Optical Modulation Amplitude
- 消光比测试 ER Extinction Ratio (Min 最小值)
- 平均光功率测试 Average Launch Power (Min 最小值与 Max 最大值)

使用光频谱仪: 对应 Test Point 2 测试发射器的:

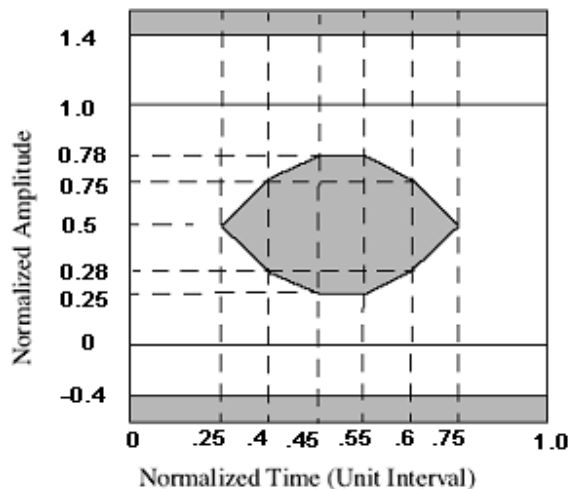
- 中心波长测试 Center Wavelength
- RMS 频谱宽度测试 RMS Spectral Width

**10GbE 光发射器特性测试:**

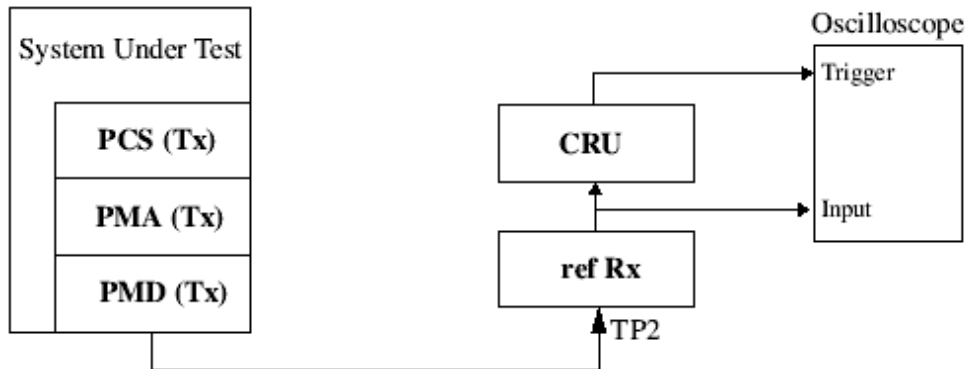
**发射器眼图模板测试 Transmitted Eye Mask :**

对应以下任何一个码型测试发射器的眼图是否能够通过模板测试 :

- 1-  $2^{23} - 1$  PRBS 伪随机码型
- 2- 有效的 10GBase-W 或 10GBase-R 信号
- 3- OC-192c 或 VC-4-64c 信号
- 4- IEEE802.3ae 规范定义的 Test Patterns



IEEE802.3ae 规范进行眼图模板测试时的仪器设置:



测试眼图模板时的 4 项注意:

- 1- 测试眼图模板时候, IEEE802.3ae 规范要求必须使用 OCR(或 CRU, Clock Recovery Unit)光时钟恢复单元从数据里恢复时钟作为触发源, 并且所使用的 **OCR 带宽必须 < 4MHz, 而且滚降特性必须是 -20dB/decade.** (注: Agilent 所提供的时钟恢复模块 83494A, OCR 带宽是 4.5MHz 并不满足规范, 选件 option 103 亦只能够支持 10.3125Gb/s LAN PHY 速率的时钟恢复)

IEEE802.3ae 规范要求 OCR 的频响与滚降特性, 主要原因是模仿 10GBase-W/R 接收器的 CRU 时钟恢复单元受抖动影响的特性. 在 IEEE802.3ae 规范里, 对接收器的抖动容限测试范围如下:

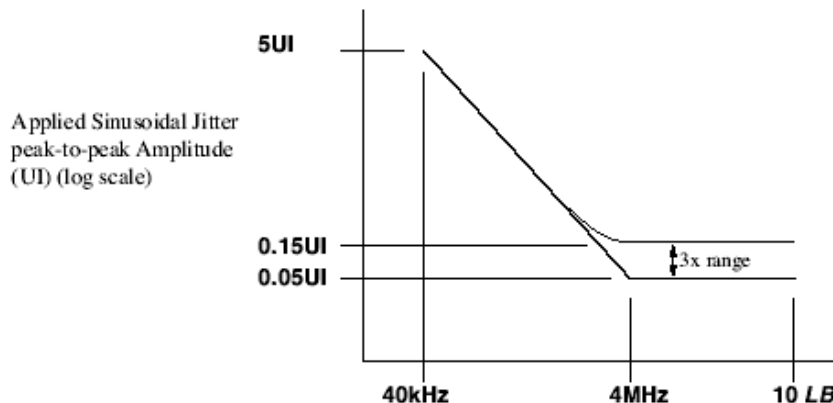


Figure 52-4—Mask of the sinusoidal component of jitter tolerance (informative)

按研究, 10GBase-W/R 接收器的 CRU 时钟恢复单元对高频的抖动不太敏感, 导致系统出现故障往往是因为低频的抖动或信号漂移, 使接收器的 CRU 里的 PLL 失锁. 所以规范在测试 10GBase-W/R 接收器的抖动容限特性时, 只会测试接收器在 40KHz 至 4MHz 的抖动容限.

如此, 示波器的 OCR 只会容许 4MHz 以下的抖动进入, 因为示波器的 OCR 的抖动容限与接收器的 CRU 是完全一样的, 所以, 若果抖动过大而导致示波器的 OCR 不能恢复时钟的话, 即意味接收器的 CRU 同样也不能恢复相同的信号. 所以为什么规范要求示波器必须配备 OCR, 并且带宽必须 < 4MHz, 而且滚降特性必须是 -20dB/decade.

[泰克的 80C08B 模块提供选件 CR1, 不单支持 10GbE 所有速率\(10.3125Gb/s 与 9.95328Gb/s\)的时钟恢复, OCR 更完全按照 IEEE802.3ae 规范的频响与滚降特性, 实现时钟恢复.](#)

- 2- 测试超长波长的信号时，如：10GBase-EW(可达 40KM 传输距离)，规范里要求接收器的灵敏度能够接受最低的信号功率为-11.3dBm。所以测试模块的噪声必须很低，达到一定的信噪比，测试的结果才能准确，而显示出来的信号噪声是真属于信号的，而不是模块的噪声。

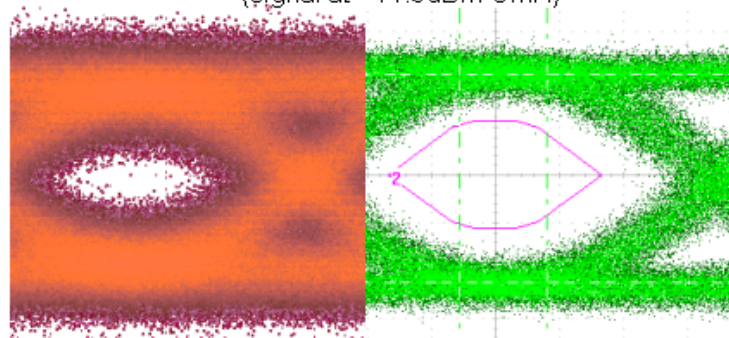
泰克的 80C08B 模块噪声低达 -24dBm, 提高测试的准确度, 重复性, 及减少因为测试模块噪声而造成信号通过不了模板测试.

Agilent 所提议的 86105B+opt 101 15GHz O/E 模块, 不是专为 10GbE, 只有-12dBm 灵敏度, 基本上 86105B 模块自身的噪声会将信号掩盖, 导致眼图变的非常糟糕. 同样, 86115B+opt 410, 则只有-7dBm 灵敏度, 根本连 10GBase-EW 最低的信号功率-11.3dBm 都测不到.

**An older module originally designed for OC-192**

**Tektronix 80C08B**

(signal at -11.3dBm OMA)



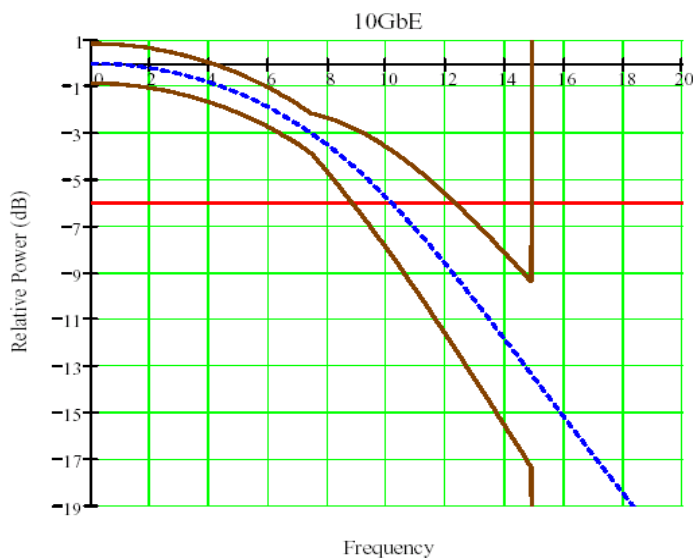
- 3- 测试眼图模板时，需要使用 **Optical Reference Receiver** 光参考接收器 (O/E + Filter)，信号首先经光电转换，然后经 4th Order Bessel-Thosom 滤波器过滤，按 IEEE802.3ae 以太网测试所设计规范要求，此滤波器的带宽要 7.5GHz，而其传递函数如下：

$$H(y) = \frac{105}{105 + 105y + 45y^2 + 10y^3 + y^4}$$

而

$$y = 2.114p; \quad p = \frac{j\omega}{\omega_p}; \quad \omega_p = 2\pi f_p; \quad f_p = 7.5GHz$$

测试 10GbE 所使用的 4th Order Bessel-Thosom 滤波器的频响如下：



注意: **CSA8000+80C08B** 能够自动测试与统计消光比 **ER Extinction Ratio** 的最小值与平均光功率 **Average Launch Power** 的最小与最大值, 都能从眼图上测试

4- 因为 10GbE 有不同的物理层, 分别使用不同的光波长与光纤, 所以测试模块必须能够支持多个光波长 (分别: 850nm, 1310nm 与 1550nm), 进行眼图模板测试, 光调制幅度测试, 光功率测试与消光比测试.

**泰克的 80C08B 模块同时支持 850nm(10GBase-S), 1310nm(10GBase-L) 与 1550nm(10GBase-E) 3 种不同物理层. Agilent 所提议的 86105B/86115B 只支持单模光纤(1000nm-1600nm), 而 86102U+opt 203 只能支持多模光纤(750- 860nm), 但 86102U+opt 203 又不支持 9.95328Gb/s WAN PHY.**

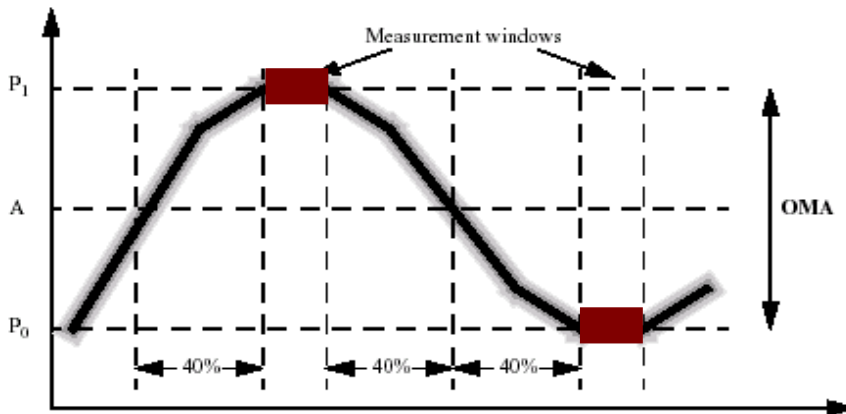
**光调制幅度测试 OMA Optical Modulation Amplitude :**

OMA 光调制幅度的定义: 在最通常的“1”与“0”位置上测量两者的光功率差, 用公式表示:

$$OMA_{dBm} = 10 \log \left( \frac{High - Low}{1mW} \right)$$

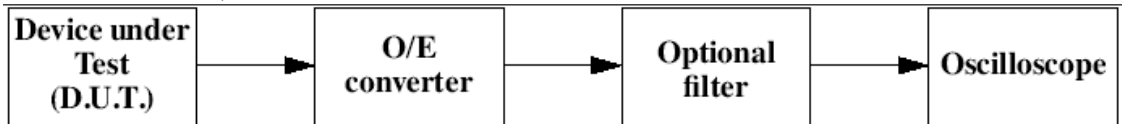
或是

$$OMA = High - Low$$



测量信号在“高”状态时中间 20%位置的平均光功率(上图用颜色表示出来的位置), 然后测量信号在“低”状态时中间 20%位置的平均光功率, 两者相减就能得出 OMA 值.

注意: 测试 OMA 时, 按 IEEE802.3ae 规范要求必须使用



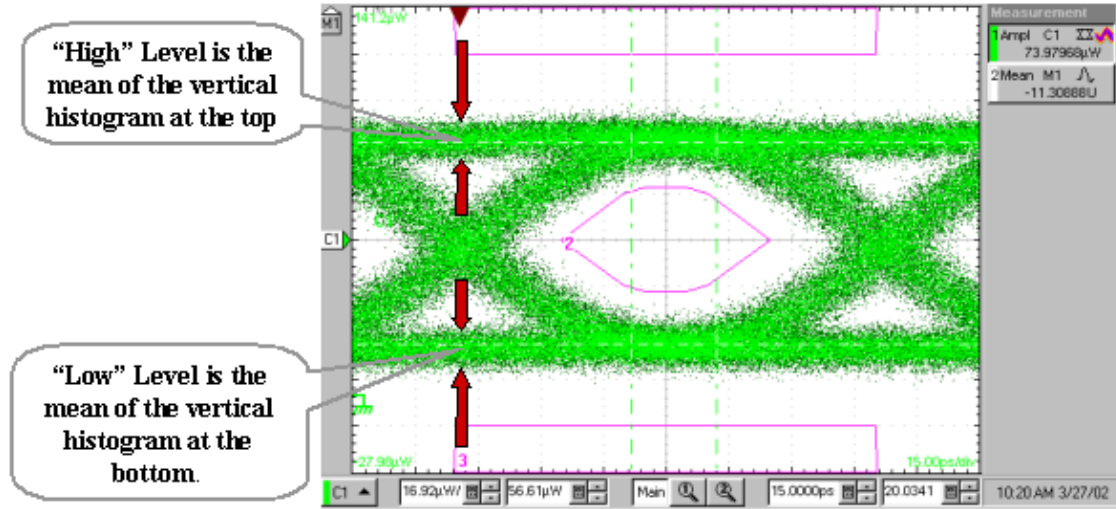
**Figure 52-5—Recommended test equipment for measurement of optical modulation amplitude**

注: 规范要求测量 OMA 时, 需要使用 Square Wave 方波, 连续的 4 个至 11 个“1”, 接着相同数目的“0”, 而测试 WAN PHY 时使用 452MHz 方波, 测试 LAN PHY 时使用 1289MHz 方波.

因为测量 OMA 时, 不能使用 Live Traffic 进行, 必须使用以上提到的方波, 所以测试器件没有问题, 但是在测试系统 OMA 时便显得非常不方便. 所以, 业内有一些 OMA 的近似测量方法:

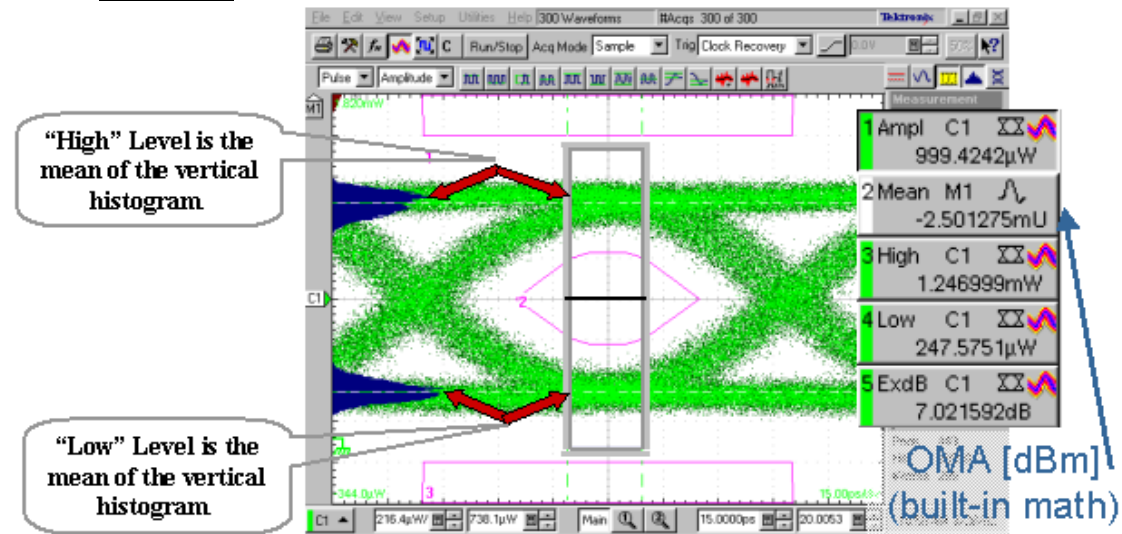


近似方法 1:



OMA 近似测量方法 1: 量度 "小" 眼睛的 "高" 与 "低" 功率, 相减所得出的值与正式的 OMA 值非常近似. 通常在测试 "Stressed Receiver Sensitivity" 受压的接收器灵敏度时, 采用这种近似方法. (缺点: 当噪声很大时, 其与正式的 OMA 差异较大而不准)

近似方法 2:



OMA 近似测量方法 2: 量度眼睛中间±20%位置的功率, 相减所得出的值与正式的 OMA 值非常近似, 但是对比近似测量方法 1 所得出的 OMA, 比较没有这么容易受到噪声的影响.

无论如何, 对应发射器, 还是需要测试 OMA + ER + AOP:

- 发射器眼图模板测试 Transmitted Eye Mask
- 光调制幅度测试 OMA Optical Modulation Amplitude
- 消光比测试 ER Extinction Ratio (Min 最小值)
- 平均光功率测试 Average Launch Power (Min 最小值与 Max 最大值)

**10GbE 光接收器特性测试:**

测试接收器特性时,使用到的仪器如: 泰克都没有提供 – 所以销售策略是:

客户没有提出要测试接收器特性时,千万不要提到有关测试接收器的事宜. 若客户主动提出了, 请跟客户解释相关的测试方法, 并且提出泰克的 [CSA8000 + 80C08B + opt CR1](#) 比 [Agilent DCA 86100B + \(83494A+opt 103\) + \(86105B+opt101\)/\(86115B+opt410\) + \(86102U+opt203\)](#)的方案更全面, 更简单, 成本要低得多.

10GbE 光接收器所要测试的参数如下:

对应 Test Point 3 测试接收器的, 使用多种的仪器测试:

**Stressed Receiver Sensitivity in OMA (Max)** 受压的接收器灵敏度测试(用 OMA 来表达 – 测量最大值)

注: 泰克只提供光示波器(眼图仪), 有关的正弦波调制源, 时钟源, 码型发生器, 激光源, 光衰减器, 泰克都没有提供, 但是 **Agilent** 就有提供.

1 - **Stressed Receiver Sensitivity in OMA (Max)** 受压的接收器灵敏度测试(用 OMA 来表达 – 测量最大值):

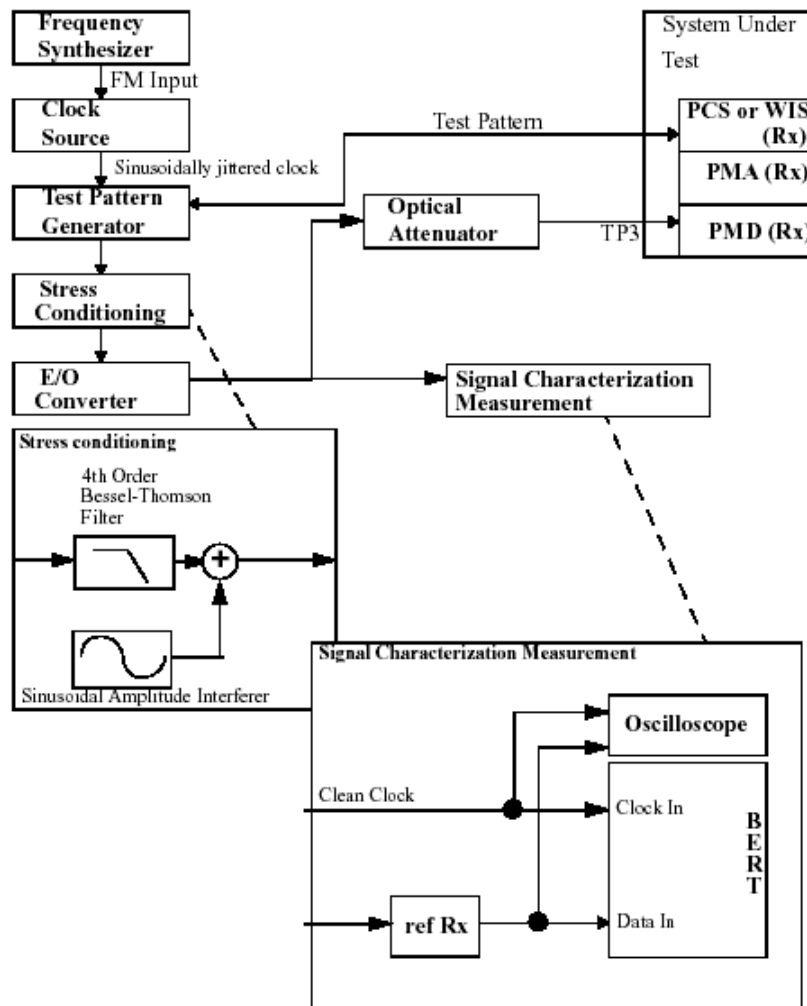
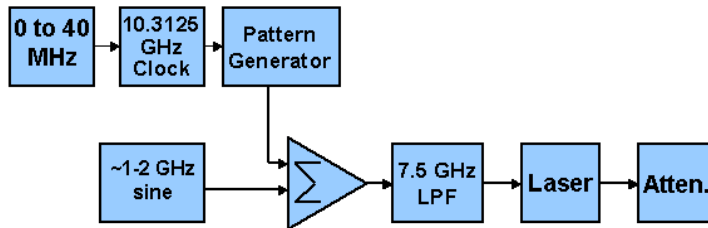
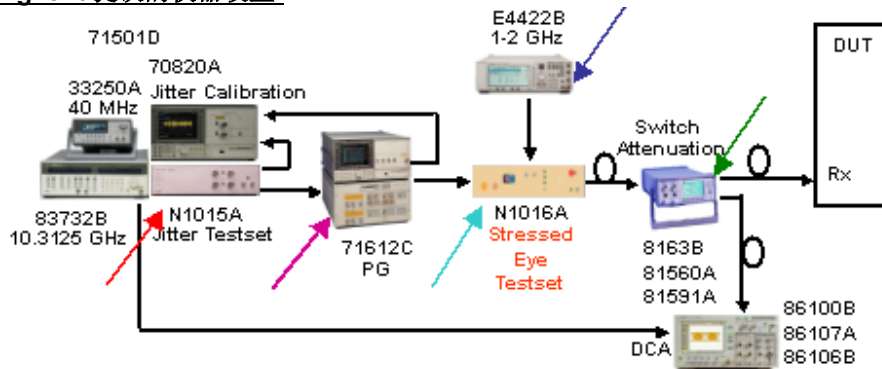


Figure 52-10—Stressed receiver conformance test block diagram



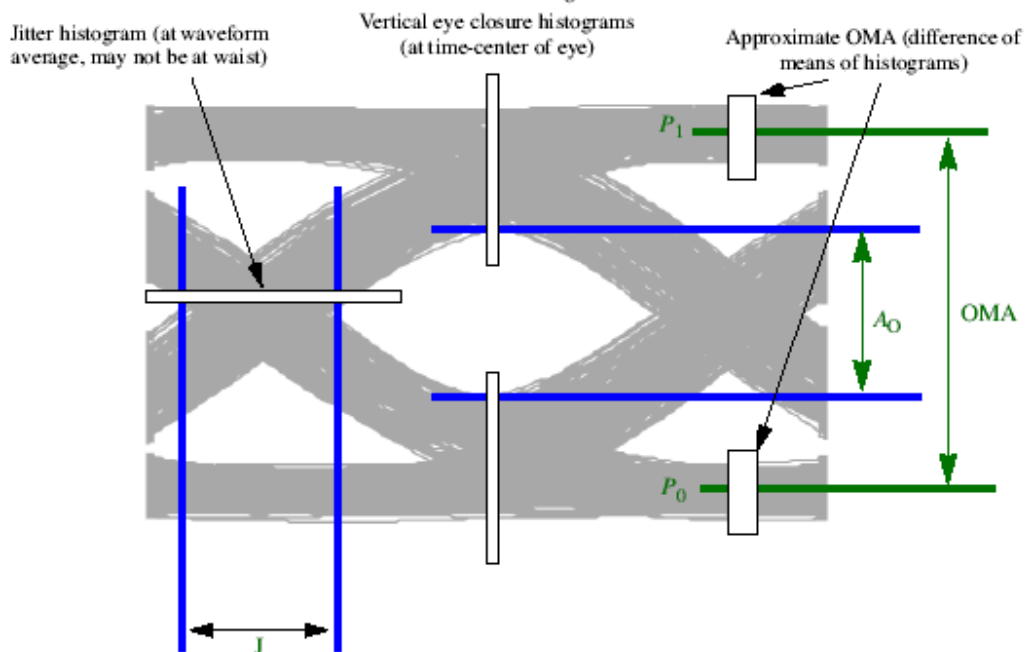
Agilent 提议的仪器设置:



测试方法:

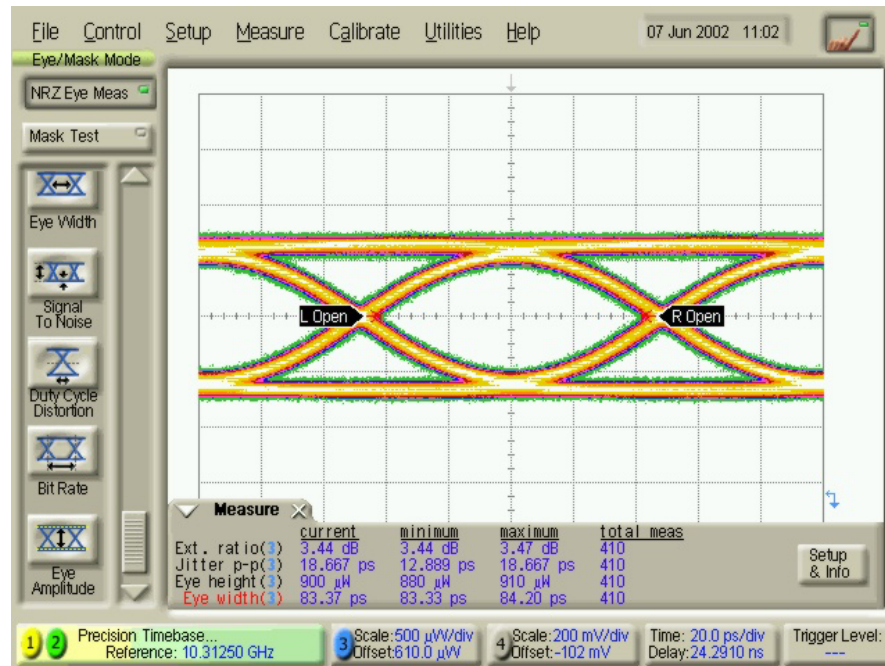
- 1- 码型发生器发送 >PRBS10 (即  $2^{10}-1$ ) 码流到接收器, 正弦波调制干扰源与正弦波抖动发生器先不开, 在受压的眼图测试仪上设定消光比 ER (若 10GBase-S 或 10GBase-E 设定消光比为 3dB; 若 10GBase-L 设定消光比为 3.5dB)
- 2- 测量 OMA
- 3- 加入 4<sup>th</sup> Order Bessel Thomson 滤波器, 测量 VECP

$$\text{Vertical eye closure penalty [dB, optical]} = 10 \times \log \frac{OMA}{A_0}$$

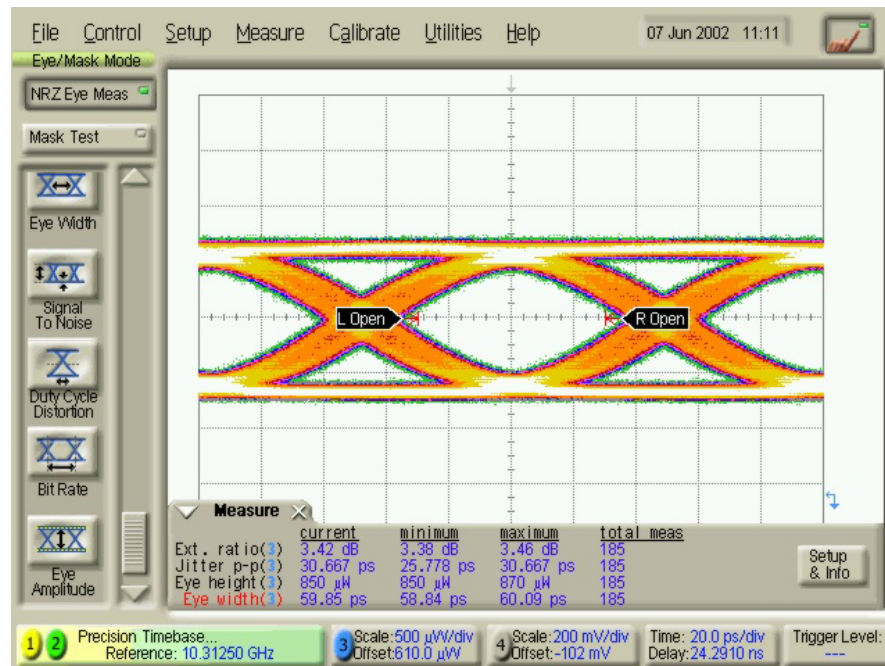


加入 4<sup>th</sup> Order Bessel Thomson 滤波器, 是测试因为 ISI 而导致的 VECP, 而所产生的 VECP 不应少于总的 66.7% , 其余的 33.3%VECP 应由正弦波调制干扰源与正弦波抖动发生器所导致.

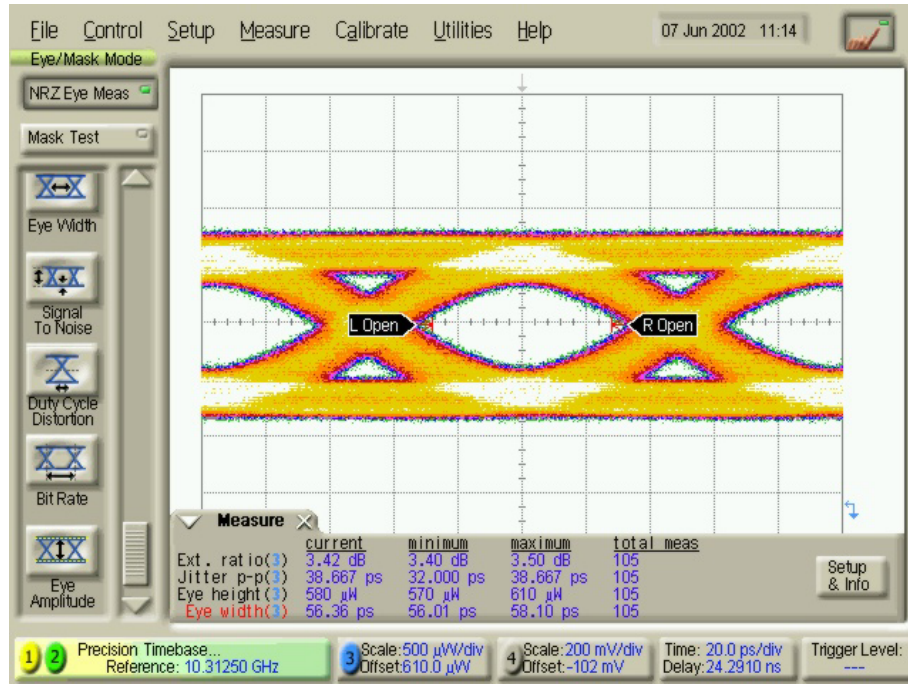
注: 按照规范, 10GBase-S 最小的 VECP 是 3.5dB; 10GBase-L 最小的 VECP 是 2.2dB, 10GBase-E 最小的 VECP 是 2.7dB, 而 10GBase-S/L/W 最小的 Stressed Eye Jitter 都是 0.3UI Pk-Pk



- 4- 加入正弦波抖动发生器, 测量 VECP. (正弦波抖动发生器的加入, 会增加水平方向的信号抖动)



- 5- 加入正弦波调制干扰源，测量 VECP。调节正弦波抖动发生器与正弦波调制干扰源的输出，使 VECP 达到 10GBase-S 规范里的 3.5dB，10GBase-L 规范里的 2.7dB 10GBase-E 规范里的 2.7dB。若接收器有高线性特性，所加入正弦波调制干扰源并不会影响 OMA，但会影响  $A_0$  (所以也影响 VECP) — 加入正弦波调制干扰源减小垂直方向的眼眸开度。



- 6- 使用光衰减器，将光信号减到规范里容许的最小功率，[10GBase-S 为-7.5dBm](#)，[10GBase-L 为-10.3dBm](#)，[10GBase-E 为-11.3dBm](#)。
- 7- 按照以下对接收器抖动容限的规范模板，调节正弦波抖动发生器的输出幅度与频率，由 40kHz 一直调到 40MHz (Loop Bandwidth = 4MHz)，监测 BER。

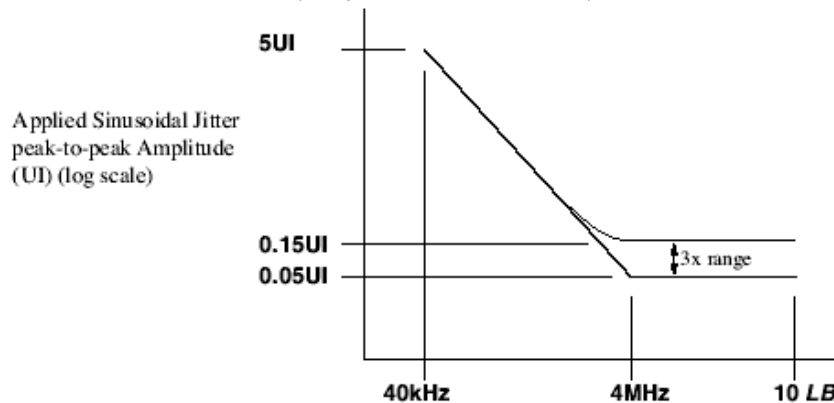


Figure 52-4—Mask of the sinusoidal component of jitter tolerance (informative)

- 8- 判断在按照抖动容限的规范模板调节正弦波抖动发生器的输出幅度与频率时，BER 有否超过  $10^{-12}$ 。BER 的测量可以观察 PCS 层内置的 BER 监测功能，对误码的计算，或是采用误码仪。

注：在整个受压的接收器灵敏度测试过程中，示波器起的作用是验证输入到接收器的信号是否如规范所要求的 OMA 与抖动的状况下。按规范要求，在最小的 OMA 与最大的抖动

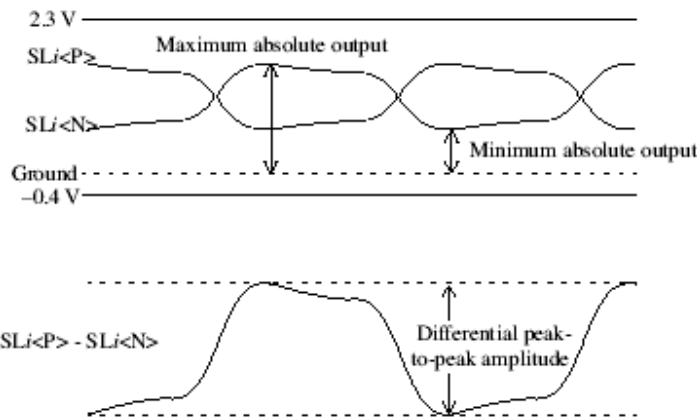
情况下，接收器的抖动都应该少于  $10^{-12}$ 。其他仪器如：正弦波抖动发生器与正弦波调制干扰源都是用作产生最坏的状况，由此而判断接收器在最坏状况下有否产生过多的误码。

注：因为 Agilent 所提供的 O/E 模块，就算最好的 86105B 其灵敏度都只有-12dbm, 而 10GbE 规范里容许的最小光功率为 10GBase-S 为-7.5dBm, 10GBase-L 为-10.3dBm, 10GBase-E 为-11.3dBm, 在测试 10GBase-E 时，86105B 从根本上是不能分辨信号与噪声的，所以并不适合测试 10GBase-E. (以上图片所提到的 86106B 甚至只有-7dbm, 连 10GBase-S 都测不到.)

**2 – 对 XAUI 接口的测试：**

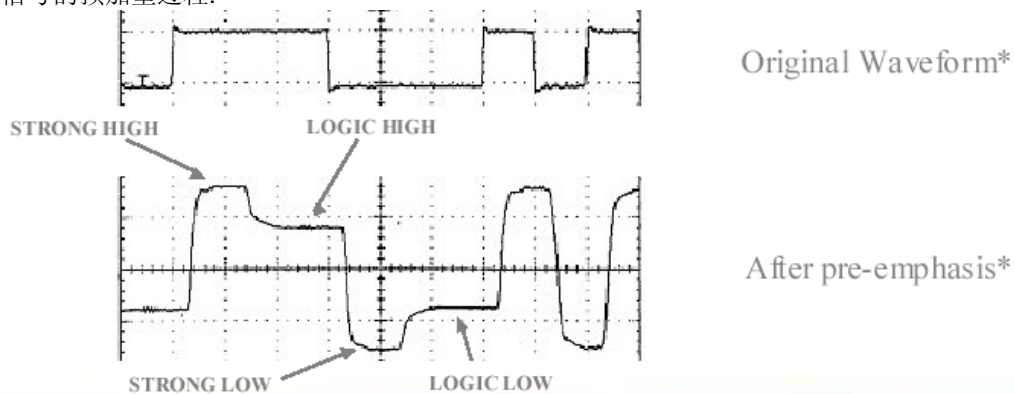
**眼图模板测试 Eye Template Test :**

XUAI 是 4 路 3.125Gbaud/s +/- 100ppm 的数据, UI = 320ps , 驱动器最大差分信号峰峰值输出幅度不能超过 1600mV, 最大绝对输出电压 < 2.3V, 最小绝对输出电压 > -0.4V, 请参看下图:

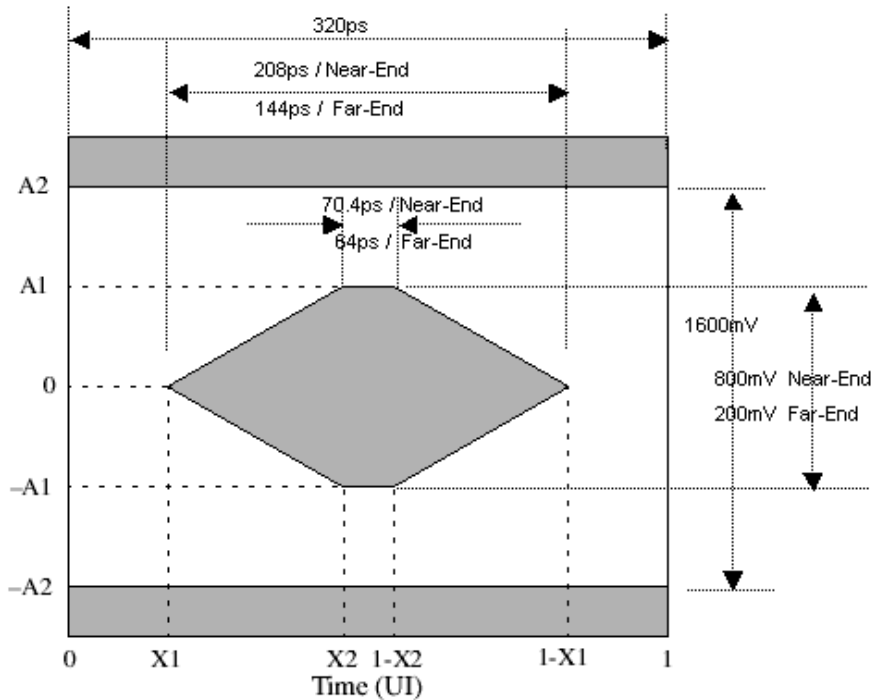
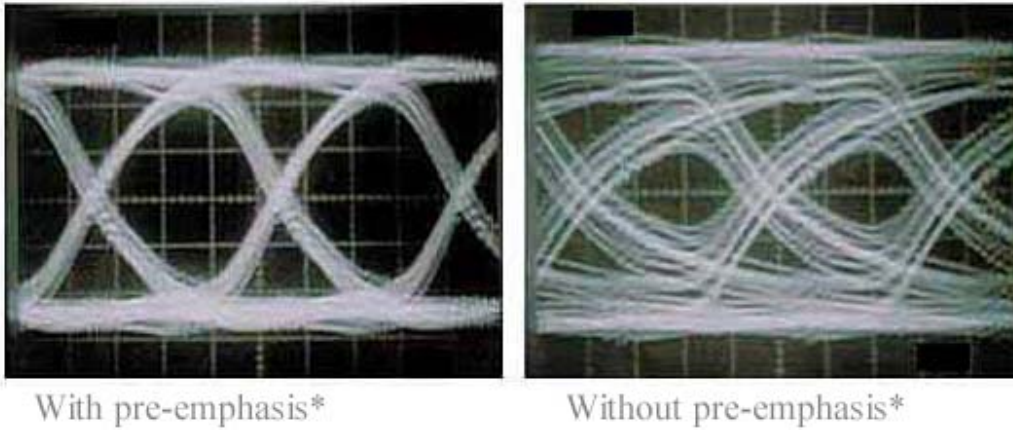


XAUI 模板分近端(Near End)与远端(Far End), 因为业内绝大部分的 SerDes 都有预加重 (Pre-emphasis)功能, 这功能是 SerDes 在发送信号时预先将信号高频部分加强, 目的是抵消信号在传输介质中传送时, 因为介质的有限带宽而使信号高频部分比低频部分相对被衰减的更厉害, 所以预先将信号高频部分加强, 使到远端接收时的信号变得高频低频都均衡. 所以规范只要求 XAUI 满足近端或远端模板便可以, 因为打开预加重, 信号会通过不了近端模板, 但是通过传输后, 若发射器符合规范的话, 应该能通过远端模板. 反之, 若预加重不打开, 若发射器符合规范的话, 应该能通过近端模板, 但经过传输后, 高频部分被衰减, 远端模板就不会通过.

信号的预加重过程:



信号经过与没有经过预加重, 在远端接收器的状况:



测试模板需要注意的事项:

- 1 使用 CJPAT 码型
- 2 采集足够的的数据, 确保误码率少于  $10^{-12}$

**抖动测试 Jitter Test :**

发生器抖动 Transmitter Jitter :

近端 Near End :  $Tj \text{ max} = \pm 56\text{ps} (\pm 0.175\text{UI})$ ,  $Dj \text{ max} = \pm 27.2\text{ps} (\pm 0.085\text{UI})$

远端 Far End :  $Tj \text{ max} = \pm 88\text{ps} (\pm 0.275\text{UI})$ ,  $Dj \text{ max} = \pm 59.2\text{ps} (\pm 0.185\text{UI})$

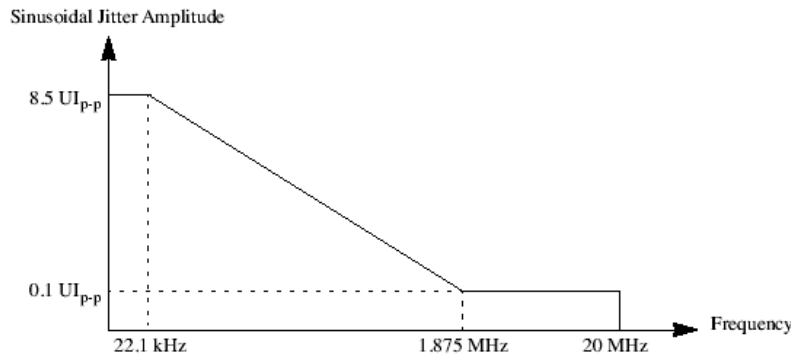
注: 测试近端发生器抖动时, 发生器需要接上  $100\text{Ohm} \pm 5\%$ ; 测试远端发生器抖动时, 发生器需要接上符合标准的互连.

接收器抖动容限 Receiver Jitter Tolerance :

峰峰值接收器抖动容限 pk-pk Receiver Jitter Tolerance =  $208\text{ps} (0.65\text{UI})$

Jitter 可以分成 3 种含量, 分别是 Random Jitter 随机抖动, Deterministic Jitter 固有抖动, Sinusoidal Jitter 正弦抖动

规范要求接收器能容许不少于 176ps(0.55UI)的随机抖动(0.08UI)+固有抖动(0.47UI), 并且能容许不同频率的正弦抖动, 而其变化如下:



因为信号从发生器传送到接收器还会受到互连(如: PCB 板的差分传输线的特性阻抗不匹配, 连接器的衰减等), 或其他因素的影响, 以下是规范对从发射器到接收器的损耗, 时延, 与总抖动的要求:

Table 47-4—Informative XAUI loss, skew and jitter budget

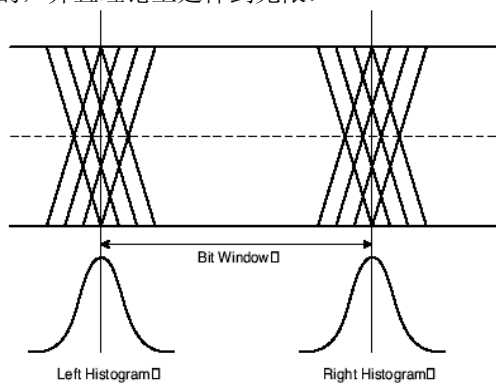
	Loss (dB) <sup>a</sup>	Differential skew (ps <sub>p-p</sub> )	Total jitter (UI <sub>p-p</sub> ) <sup>c</sup>	Deterministic jitter (UI <sub>p-p</sub> ) <sup>c</sup>
Driver	0	15	0.35	0.17
Interconnect	7.5	60	0.20	0.20
Other <sup>b</sup>	4.5		0.10	0.10
Total	12.0	75	0.65	0.47

同样, 测试抖动时也要注意:

- 1 一使用 CJPAT 码型
- 2 一采集足够的的数据, 确保误码率少于 10<sup>-12</sup>

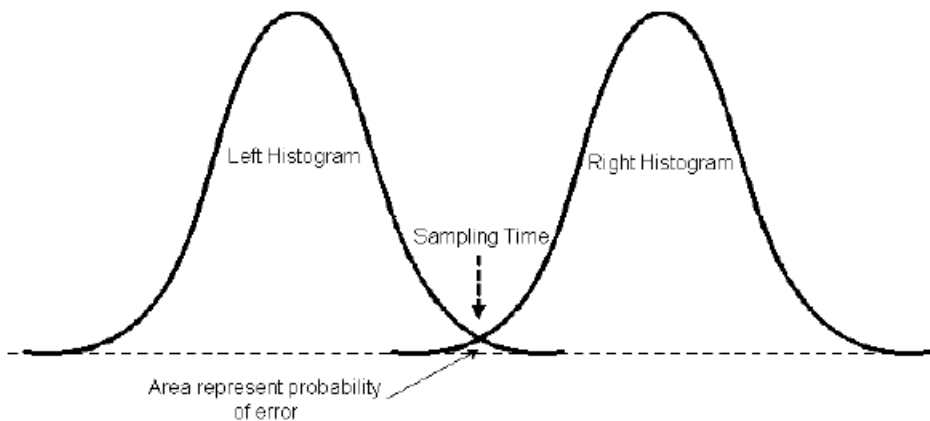
**注: 抖动与误码率的关系—Bathtub Curve 浴盆曲线:**

下图显示的是信号边沿跳变时的时间偏差分布, 一般定义为抖动的概率分布函数, 即信号边沿在某一时刻跳变的概率. 在某一时刻概率分布函数的垂直幅度越大时, 代表信号边沿在那一刻跳变的概率越大, 所以一般情况下, 抖动概率分布函数的最大值通常是出现在信号边沿应该预定跳变的时刻, 即 0 抖动时刻. 若导致抖动的原因完全是随机的话, 抖动的概率分布函数应该是高思正态分布的, 并且理论上延伸到无限.





在接收端使用 BERT 误码仪对眼图进行采样时，假设误码仪采样的位置是眼图的中心位置，即  $t=0.5UI$ ，因此抖动所造成的误码是由于：左边的分布函数代表是信号的“0”的跳变，误码仪(在这里，所扮演的角色等同接收器)把左边分布函数右方大于  $0.5UI$  的偏差的“0”读成“1”，同样，右边的分布函数代表是信号的“1”的跳变，误码仪把右边的分布函数左方大于  $0.5UI$  的偏差的“1”读成“0”，所以左边跟右边分布函数交叉范围内的面积，正代表了由错误解读跳变边沿而出现误码的概率。若将此概率乘以在那时刻出现跳变的概率，其乘积就是误码了。



所以在某一时间位置，由于左边抖动分布函数所造成的误码率应是：

$$BER_{left}(\tau_{sample}, W, \sigma) = \Gamma_{transition} \cdot \int_{\infty}^{\tau_{sample}} JT(\tau, W, \sigma) \delta\tau$$

$\tau_{sample}$  是误码仪的采样时间位置， $JT(\tau, W, \sigma)$  是抖动的概率分布函数， $\tau$  是时间， $W$  = 固有抖动的峰峰值， $\sigma$  是随机抖动的 RMS 值  $\Gamma_{transition}$  是在那一时刻的边沿跳变概率。

$$BER_{right}(\tau_{sample}, W, \sigma) = \Gamma_{transition} \cdot \int_{-\infty}^{\tau_{sample}} JT(\tau - UI, W, \sigma) \delta\tau$$

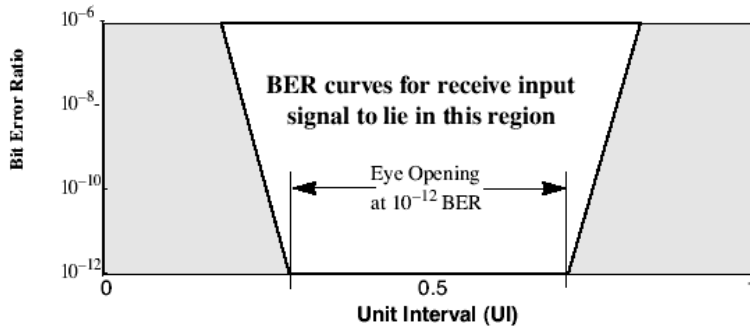
同样，

注：右边的抖动概率分布函数的中心在  $1UI$  的位置，所以是时间是  $\tau - UI$

所以总的误码率 BER 是：

$$BER_{total}(\tau_{sample}, W, \sigma) = BER_{left}(\tau_{sample}, W, \sigma) + BER_{right}(\tau_{sample}, W, \sigma)$$

所以误码仪在不同  $\tau_{sample}$  的位置所得的  $BER_{total}(\tau_{sample}, W, \sigma)$  值都不一样。而一般称  $BER_{total}(\tau_{sample}, W, \sigma)$  vs  $\tau_{sample}$  为 BER Bathtub Curve 浴盆曲线，而在不同  $\tau_{sample}$  采样点扫描眼睛的做法，叫 BERT Scan:



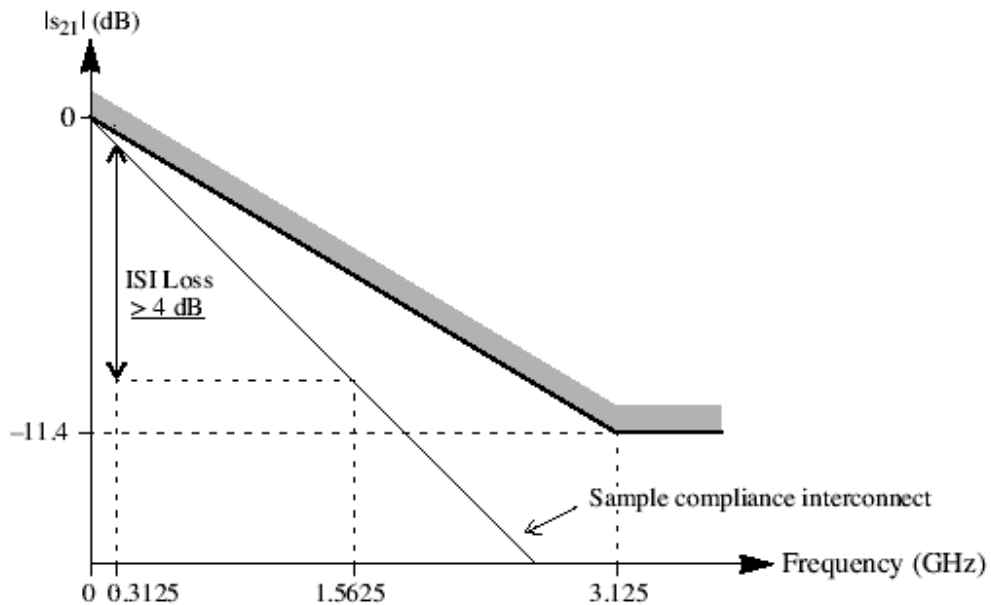
**互连验证 Interconnects Compliance :**

XAUI 的设计目的是使到从 PCS 与 MAC 层之间的传输距离增加到 50cm, 以下是规范对互连的要求:

	Loss (dB) <sup>a</sup>	Differential skew (ps <sub>p-p</sub> )	Total jitter (UI <sub>p-p</sub> ) <sup>c</sup>	Deterministic jitter (UI <sub>p-p</sub> ) <sup>c</sup>
Driver	0	15	0.35	0.17
Interconnect	7.5	60	0.20	0.20
Other <sup>b</sup>	4.5		0.10	0.10
Total	12.0	75	0.65	0.47

注: Other 是指信号的串扰, 噪声等其他因素所造成的损耗与时延.

同时, 需要使用 VNA 对互连在不同频段, 测试其  $S_{21}$  Transmission Magnitude Response, 并需要满足以下规范:



注: ISI Loss 定义为两个频点的  $S_{21}$  之差, 在 312.5MHz 与 1.5625GHz 这两个频点中, 规范要求其  $S_{21}$  之差  $>4\text{dB}$ , 这个必需要使用 VNA 才能测试. 并且当频率  $>3.125\text{GHz}$ ,  $S_{21}$  不能少于  $-11.4\text{dB}$ .

注:

- 1 规范同时要求 PCB 板差分传输线在 100MHz 到 2.5GHz 频率范围内, 其差分阻抗为  $100\Omega \pm 10\%$ .
- 2 规范也要求 Connector 接收器的特性阻抗为  $100\Omega \pm 30\%$