Ethernet Tutorial: Standards and Technology; Status and Trends

Jonathan Thatcher
Chief Technologist, World Wide Packets
Chair, Ethernet in the First Mile Alliance
Agenda – Part I (of VI)

Ethernet -- The Big Picture

Ethernet 101
- IEEE 802.3 Context and Standards Process
- A Brief History of Networking

High Level Overviews
- Gigabit Ethernet (GbE)
- 10 Gigabit Ethernet (10GbE)
- Data Terminal Equipment (DTE) Power via Media Dependent Interface (MDI)
- Ethernet in the First Mile (EFM)
Agenda – Part II (of VI)

Digging Deeper

10 Gigabit Ethernet
- Technology Overview
- Applications
- LAN / WAN PHYs; Optics; Layers

Ethernet In The First Mile
- Technology Overview
- Operations, Administration, & Management (OAM)
- Point to point (P2P)
- Ethernet over unclassified copper (EDSL; EFMCu)
- Point to multi-point (P2MP; EPON)
Agenda – Part III (of VI)

Technology Comparison

Resilient Packet Ring (RPR; 802.17)
- Technology Overview
- Structure
- Access
- Fairness
- Protection
- Comparison
Agenda – Part IV (of VI)

Fiber and Optics

Technology
- Product implementation vs. sublayers
- Optics 101
- Challenges in high speed (low cost) optics
- Changes in specification methodology

Putting Down The Fiber
- Fiber recommendations
- Cost of fiber infrastructure
- Alternative Examples:
  - Microtrenching
  - Microconduit
Agenda – Part V (of VI)

Trends and Influences

- Towards Simplification
- Towards higher speed; lower cost vs. Moore’s Law
- Ethernet to the rescue in the Access Space
- QOS and OAM can be and must be solved
- Economic models can support “True Broadband Services”
- Distractions or complements
- Federal regulation and policy will be the single greatest influence on technology development
- Investment as a positive feedback system
Agenda – Part VI (of VI)

Related Organizations

- Ethernet in the First Mile Alliance (EFMA)
- 10 Gigabit Ethernet Alliance (10GEA)
- Optical Internetworking Forum (OIF)
- Fibre Channel (FC)
The Big Picture

Ethernet Basics

Standards Process
Robert Metcalfe’s Drawing

of the first Ethernet design
How CSMA/CD Works – Party Line

- Is anyone on line?
  - If yes, try again later
  - If no, ring the address you want to talk with

- Did anyone else try to get on “at the same time” you did?
  - If yes, try again later
  - If no, you own the media
Ethernet Basics, and Maturation

10BASE2 or 10BASE5 (Coax Cable, Bus Topology, 1985)

Collision (CSMA/CD resolution)

UTP

Half-Duplex 10BASE-T (Star Topology, UTP cable, 1990)

Bridge/Switch

Full Duplex 10/100BASE-T (1992/1993)

Collision-Free

UTP

UTP

UTP

CSMA/CD:

Carrier Sense Multiple Access with Collision Detection

Source: Luke Maki, Boeing Corporation, 2002
Ethernet Basics, and Maturation

10BASE2 or 10BASE5 (Coax Cable, Bus Topology, 1985)

Collision (CSMA/CD resolution)

Half-Duplex 10BASE-T (Star Topology, UTP cable, 1990)

Source: Luke Maki, Boeing Corporation, 2002
Ethernet Basics, and Maturation

10BASE2 or 10BASE5 (Coax Cable, Bus Topology, 1985)

Collision (CSMA/CD resolution)

UTP

Half-Duplex 10BASE-T (Star Topology, UTP cable, 1990)

Dedicated Media

UTP

UTP

UTP

Collision-Free

Full Duplex 10/100BASE-T (1992/1993)

Repeater

Collison (CSMA/CD resolution)

UTP

UTP

Bridge/Switch

UTP

UTP
Ethernet: **Layer 1 & 2 of the OSI Stack**

**OSI Layer Model**
- Application
- Presentation
- Session
- Transport
- Network
- Datalink
- Physical

**802.3 Layer Model**
- Logical Link Control
- MAC Control (Opt)
- Media Access Control
- Reconciliation Sublayer
- Physical Coding Sublayer
- Physical Medium Attachment
- Physical Medium Dependent

(XG) MII

MDI

Media
802 Overview & Architecture

802.2 Logical Link Control
IEEE Std 802.2, ISO 8802-2-1989

802.1 MAC Bridging
IEEE Std 802.1D;1990

Ethernet

802.3 CSMA/CD
IEEE Std 802.3-1990

802.4 TBUS
IEEE Std 802.4-1985

802.5 TRING
IEEE Std 802.5-1985

802.6 DQDB
IEEE Std 802.6-1990

802.9 ISLAN
IEEE Std 802.9-1990

802.11 WLAN
IEEE Std 802.11-1997

802.12 DPAP
IEEE Std 802.12-1998

802.14 CATV
IEEE Std 802.14-1999

802.15 WPAN
IEEE Std 802.15-2003

802.16 BWA
IEEE Std 802.16-2004

802.17 RPR
IEEE Std 802.17-2005

802.7 Broadband TAG (BBTAG)
IEEE 802.7-1989

802.8 Fiber Optic TAG (FOTAG)
IEEE 802.8-1987
### IEEE 802 Working Groups

<table>
<thead>
<tr>
<th>802.1</th>
<th>Higher Layer LAN Protocols Working Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>802.2</td>
<td>Logical Link Control Working Group (Inactive)</td>
</tr>
<tr>
<td><strong>802.3</strong></td>
<td>Ethernet Working Group</td>
</tr>
<tr>
<td>802.4</td>
<td>Token Bus Working Group (Inactive)</td>
</tr>
<tr>
<td>802.5</td>
<td>Token Ring Working Group (Inactive)</td>
</tr>
<tr>
<td>802.6</td>
<td>Metropolitan Area Network Working Group (Inactive)</td>
</tr>
<tr>
<td>802.7</td>
<td>Broadband TAG (Inactive)</td>
</tr>
<tr>
<td><strong>802.8</strong></td>
<td>Fiber Optic TAG (Disbanded)</td>
</tr>
<tr>
<td>802.9</td>
<td>Isochronous LAN Working Group (Inactive)</td>
</tr>
<tr>
<td>802.10</td>
<td>Security Working Group (Inactive)</td>
</tr>
<tr>
<td><strong>802.11</strong></td>
<td>Wireless LAN Working Group</td>
</tr>
<tr>
<td>802.12</td>
<td>Demand Priority Working Group (Inactive)</td>
</tr>
<tr>
<td>802.13</td>
<td>Not Used</td>
</tr>
<tr>
<td>802.14</td>
<td>Cable Modem Working Group (Inactive)</td>
</tr>
<tr>
<td><strong>802.15</strong></td>
<td>Wireless Personal Area Network (WPAN) Working Group</td>
</tr>
<tr>
<td>802.16</td>
<td>Broadband Wireless Access Working Group</td>
</tr>
<tr>
<td>802.17</td>
<td>Resilient Packet Ring Working Group</td>
</tr>
</tbody>
</table>
**System Model – Switched Ethernet**

- **SLA and QoS**
  - 802.1p Priority Queuing
  - 802.1Q VLAN Tagging
- **Buffering**
  - Fair Weighting
  - Bandwidth Shaping
- **Packet Forwarding Engine**
  - Switching Table
- **STA Station**

**Media Access Control and Physical Layer Interface**
- 802.3 CSMA/CD

**Networks**
- WAN MAN
- EAN
- LAN

**MAC and PHY**
- MMD
- MDIO
### The Ethernet Packet

<table>
<thead>
<tr>
<th>OCTETS</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Preamble</td>
</tr>
<tr>
<td>1</td>
<td>Start of Frame Delimiter</td>
</tr>
<tr>
<td>6</td>
<td>Destination Address</td>
</tr>
<tr>
<td>6</td>
<td>Source Address</td>
</tr>
<tr>
<td>2</td>
<td>Length / Type Field</td>
</tr>
<tr>
<td>46–1500</td>
<td>MAC Client Data</td>
</tr>
<tr>
<td>OCTETS</td>
<td>Pad</td>
</tr>
<tr>
<td>4</td>
<td>Frame Check Sequence</td>
</tr>
</tbody>
</table>

OCTETS WITHIN FRAME TRANSMITTED TOP TO BOTTOM; LSB to MSB
**Ethernet Packet + VLAN Tag**

<table>
<thead>
<tr>
<th>7 OCTETS</th>
<th>Preamble</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 OCTET</td>
<td>Start of Frame Delimiter</td>
</tr>
<tr>
<td>6 OCTETS</td>
<td>Destination Address</td>
</tr>
<tr>
<td>6 OCTETS</td>
<td>Source Address</td>
</tr>
<tr>
<td>4 OCTETS</td>
<td><strong>VLAN Tag (802.3ac)</strong></td>
</tr>
<tr>
<td>2 OCTETS</td>
<td>Length / Type Field</td>
</tr>
<tr>
<td>46 –1500 OCTETS</td>
<td>MAC Client Data</td>
</tr>
<tr>
<td>4 OCTETS</td>
<td>Pad</td>
</tr>
<tr>
<td>4 OCTETS</td>
<td>Frame Check Sequence</td>
</tr>
<tr>
<td>Extension</td>
<td></td>
</tr>
</tbody>
</table>

OCTETS WITHIN FRAME TRANSMITTED TOP TO BOTTOM; LSB to MSB
IEEE 802 Overview

- IEEE
- IEEE-SA Standards Association
  - Standards Board
    - RevCom Review Committee
    - NesCom New Stds. Committee
  - IEEE 802 Sponsor Group
    - IEEE 802.3 Working Group
      - IEEE P802.3af Task Force
      - IEEE P802.3ah Task Force
Process in Summary

- Call for interest
- Write and get PAR approved
  - Define the objectives
  - Answer 5 criteria
- Brainstorm, recruit proposals & ideas
- Cut-off new proposals & adopt base-line or “core proposal”
- Write; review; refine & approve drafts
- Publish

Note: ALL TECHNICAL VOTES MUST PASS BY 75%
The 5 Criteria

1. Broad Market Potential
   Broad set(s) of applications // Multiple vendors, multiple users balanced cost, LAN vs. attached stations

2. Compatibility with IEEE Standard 802.3
   Conformance with CSMA/ CD MAC, PLS // Conformance with 802.2

3. Distinct Identity
   Substantially different from other 802.3 specs/ solutions
   Unique solution for problem (not two alternatives/ problem)
   Easy for document reader to select relevant spec

4. Technical Feasibility
   Demonstrated feasibility; reports -- working models
   Proven technology, reasonable testing // Confidence in reliability

5. Economic Feasibility
   Cost factors known, reliable data // Reasonable cost for performance expected // Total Installation costs considered
Other Things Ethernet…

There is a strong cultural history to:

- Leave the MAC alone
- Provide 10X performance at 3-4X the cost
- Minimize number of PHYs per media type
- Develop a standard that guarantees interoperability == “plug and play”
- Spec 10e-12 BER;
  - *Expect better than 10e-15*
- Attempt to achieve 100% consensus
IEEE 802.3 Ballot Process

- RevCom & IEEE Standards Board
  - Approved
  - 802 (LMSC or Sponsor)
    - Approve
    - 802.3 Working Group
      - Approve
      - 802.3** Task Force (e.g., 802.3ah == EFM)
        - Approve
          - Approved Standard

Return with comment
# 10GbE Sponsor Ballot Results

<table>
<thead>
<tr>
<th>Voters: 109</th>
<th>D4.0</th>
<th>D4.1</th>
<th>D4.2</th>
<th>D4.3</th>
<th>D5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>76%</td>
<td>79%</td>
<td>83%</td>
<td>85%</td>
<td>87%</td>
</tr>
<tr>
<td>Abstain</td>
<td>8%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td>Approve</td>
<td>82%</td>
<td>82%</td>
<td>86%</td>
<td>88%</td>
<td>96%</td>
</tr>
</tbody>
</table>
NOT YOUR FATHER’S ETHERNET
Enterprise Networking

- Dumb terminals
  - attached to mainframes
- Star wired
- Relatively short distances
- High reliability
- Easy to maintain
- Lowest cost (?)
- Mission critical

Note: IEEE 802 formed in 1980
Ethernet – CSMA/CD

- Carrier sense multiple access with collision detection
  - Simplex operation
- Shared media (taps)
- Relatively short distance
- Low reliability
- Difficult to maintain
- Difficult to upgrade
- Lowest cost (?)
- Applications?

True of all immature shared media topologies
Enterprise Networking

- Dumb terminal emulation cards in PCs
- Still mission critical
- Enter LOTUS 1-2-3

Sneakernet
Enterprise Networking

- Dumb terminal emulation cards still in PCs (mission critical)
- Ethernet cards also (PC-based SW becoming mission critical)
- > 2x the work
- < ½ the reliability
- > 2x the expense
Ethernet Hubs

- CSMA/CD – Half Duplex
  - Star wired
  - Point-to-point only
  - No shared media
  - But, protocol behaves like shared media

- Increased distance
- Higher reliability
- Easier to maintain
- Easy upgrade path
- Higher cost
Switched Ethernet

- Full Duplex
  - No collisions!
  - Star wired
  - Point-to-point only
  - No shared media
  - Transmitter does not monitor Rcvr

- Increased distance
- Highest reliability
- Easiest to maintain
- Easiest to upgrade
- Higher cost
- Higher performance
Fiber Optic Inter-Repeater Link

Fiber Optic Link Segments

Repeater Set

10BASE-T Link Segments

Repeater Set

MAU

DTE

FOIRL
Fiber Optic Inter-Repeater Link

- 10BASE-F  Clauses 15-18
- Star Wired;
- Distance
  - 10BASE-FP: 1 km; Half Duplex
  - 10BASE-FB: 2 km; Half Duplex
  - 10BASE-FL: 2 km; Half or Full Duplex
  - Other distances apply with multiple segments
- 850 nm LED; 62.5/125 MMF
- BER 10e-9
- 802.3d-1987 (9.9)
- 10 December 1987 (IEEE)
10BASE-T

- Inexpensive media
- Inexpensive ports
- Installation ease

circa 90
Enterprise Networking

- Dumb terminals gone
- Emulators built into PC SW for legacy applications
- Mainframes on FDDI rings
- Wide area connection via T1 lines
- Serious application of shared storage
- Serious DB applications
Fast Ethernet – 100BASE-X

Introduction of multi-speed topologies
**Fast Ethernet – 100BASE-X**

**IEEE 802.3u**
- Pretty much a shift in decimal place from 10BASE-T
- CSMA/CD + Full Duplex
- Cat 3… Cat 5 Copper Technology (100BASE-T)
- **Optical technology from FDDI (100BASE-FX)**
  - 2 km over MMF
  - (10 km over SMF)
- Introduces high speed aggregation between switches

*Sorry Token Ring*
100BASE-FX

26.2 Functional Specifications

- The 100BASE-FX PMD (and MDI) is specified by incorporating the FDDI PMD standard, ISO/IEC 9314-3: 1990, by reference...
- Total of 2 pages (excluding PICS)

Characteristics

- Star Wired (not counter-rotating ring)
- 1310 nm LED over 62.5/125 MMF
  - 50 MMF SMF with laser outside std
- NRZ: Bit Transition = 1; No Transition = 0
- 100 Mbps data rate; 10e-8 BER
- 125 Mbps using 4B/5B encoding line rate
Gigabit Ethernet – 1000BASE-X

Extension of multi-speed topologies
Gigabit Ethernet

IEEE 802.3z

- CSMA/CD + Full Duplex
- Carrier Extension
- Serial technology from Fibre Channel
  - 1000BASE-CX  copper, Twin-ax, generally unused
  - 1000BASE-SX  850 nm, MMF
  - 1000BASE-LX  1310 nm, SMF/MMF
  - Uses 8B/10B code

IEEE 802.3ab

- Support of CAT-5 (CAT-5E) cable: 1000BASE-T

Sorry ATM
Fast Ethernet to GigE Upgrade

See: http://www.10gea.org/Tech-whitepapers.htm
**1000BASE-T**

**IEEE 802.3ab**

- Supports both full & half duplex (CSMA/CD)
  - But, no one uses CSMA/CD mode at 1 Gig

- 1000Mbps Ethernet service over 100 meters of same Category 5 links ANSI/TIA/EIA-568-A. 100BASE-T.

- Same auto-negotiation system as 100BASE-TX
  - Enable PHYs capable of both 100 and 1000 Mbps

- Specifications for field testing of twisted pair cabling system with the additional test parameters for FEXT (ELFEXT)
250 Mbps Bi-Directional on Each Pair
The Challenge: NEXT & FEXT
Gigabit Ethernet Beyond Campus

- IEEE 802.3z specifies 5km over SMF

- Transceivers extended distance & bandwidth:
  - 10 km, 1310 nm, SMF immediately (LX++)
  - 40 km, 1550 nm, within 1 year (proprietary, common pkg)
  - 100 km within 2 years
  - 4 Gbps using 802.3ad and WDM in 3 yrs (> 40 km)

- Ownership significantly less than cost of T1/ATM/SONET…
  - Spokane school district (GigE to every school over fiber)
  - CANARIE project (see www.canarie.ca)

- Spawns new market segments
  - Yipes, Telseon, OnFiber…
  - Grant County, WA; Provo, UT; Jacksonville, FL….
Link Aggregation

IEEE 802.3ad

- Ability to take $N$ links between common nodes – point-to-point – and aggregate a subset as virtual link

- Ideal for intermediate speeds....

- Ideal for TDM & WDM – *non-standard* – solutions

- Utilization of the $N \times$ Serial concept
  - Started in HIPPI for 10Gig
    - 12 x 1 Gig parallel optics
    - circa 1994?
10 Gigabit Ethernet

IEEE 802.3ae

- **MAC: It’s Just Ethernet**
  - Maintains 802.3 frame format & size
  - Full duplex operation only
  - Throttled to 10.0 for LAN PHY or 9.58464 Gbps for WAN PHY

- **PHY: LAN & WAN PHYs**
  - LAN PHY uses simple encoding mechanisms to transmit data on dark fiber & dark wavelengths
  - WAN PHY adds a SONET framing sublayer to utilize SONET/SDH as layer 1 transport

- **PMD: Optical Media Only**
  - 850 nm on variety of MMF types (28m…) to 300m
  - 1310 nm, 4 lambda, WDM to 300 m on MMF; 10 km on SMF
  - 1310 nm on SMF to 10 km
  - 1550 nm on SMF to 40 km
10 Gigabit Ethernet

- Supports dark wavelength and SONET/TDM with unlimited reach
- Several coding schemes – 64b/66b; 8B/10B; scramblers
- Three optional interfaces: XGMII; XAUI; XSBI
- Extension of MDIO interface
- Continues Ethernet’s reputation for cost effectiveness & simplicity – goal 10X performance for 3X cost
- Standard ratified in June 2002
- Business and economic success TBD

Sorry Who?
Overview of DTE Power
P802.3af DTE Power

- AKA “Power over Ethernet”
- Provides up to 13W to a connected device
  - IP phone
  - Web cam
  - Wireless access point
  - Security, lighting, HVAC controls
  - Enables many new types of devices
- Supports 10, 100, 1000BASE-T
  - Power over signal pairs or
  - Power over “idle” pairs
- Eliminates the need for AC power to devices
  - No “wall warts”
  - No expensive AC power wiring for wireless access points
P802.3af DTE Power

- Power supply equipment
  - Powered hub or switch OR
  - Mid-span insertion unit
- Allows for flexible UPS strategies
- Provides “discovery” of DTE-capable device
  - Power only applied when proper “signature” is detected
  - Will not harm legacy equipment
  - Works with existing 2 or 4 pair cable plant

- Project Status
  - Task force formed January 2000
  - Draft in working group ballot now
  - Published standard early 2003
  - Broad industry support
P802.3af DTE Power

First “world-wide” standard for power distribution

- IP Phone
- The Ethernet shaver!
Overview of Ethernet in the First Mile
Ethernet in the First Mile
GbE LX vs. Single Fiber P2P

GbE LX

MAC/Switch

SERDES

Laser Driver

Post-Amp

TIA

Laser FP

Photo-diode

SMF, 5km

P2P Focus!

EFM

MAC/Switch

SERDES

Laser Driver

Post-Amp

TIA

Laser TBD

Photo-diode

SMF, 10km
P2P Focus

Optical Element

SMF, 10km

Optical Filter

Laser
DFB/FP/VCSEL

Photodiode
PIN/APD

TX
RX

TX/RX
P2MP (EPON) Downstream

- Downstream channel uses true broadcast.
- 802.3 Frames extracted by MAC addresses.

OLT = Optical Line Terminal
ONU = Optical Network Unit
P2MP (EPON) Upstream

- Upstream time slicing
- No collisions
- No packet fragmentation
EFM Copper (Unclassified)
Hybrid Fiber/Copper

- Next-generation, high-speed architectures
  - EFM copper for the last 700 to 800 meters
  - Minimum 10 Mbps – higher if possible
  - High bandwidth for entertainment – client/server
  - For stepwise buildout to work, EFMCu must support next-gen applications

Source: EFMA 2002
OAM Operations

- General Communications Mechanism
- Link Monitoring
- Remote Failure Indication
- Remote Loop-back
- Data Link Layer Ping
- Capability Discovery
New Concepts in Current Projects

- Powering devices over UTP-5
- Variable data rate MAC
- Embedded Framer within PCS
- Use of SONET as Layer 1 transport
- Embedded BERT within PCS
- High speed differential, multi-lane, bus (XAUI)
- Use of WDM
- Extend link length to 40 km
- Single fiber, full duplex PHY
- Support of unclassified twisted pair
- OAM
- Extended temperature operation
- Extension into Metro, Backbone, and Access Spaces
10 Gigabit Ethernet in Detail
10 GbE Layer Diagram

Media Access Control (MAC)
Full Duplex

10 Gigabit Media Independent Interface (XGMII) or
10 Gigabit Attachment Unit Interface (XAUI)

- WWDM LAN PHY (8B/10B)
- Serial LAN PHY (64B/66B)
- Serial WAN PHY (64B/66B + WIS)

- WWDM PMD 1310 nm
- Serial PMD 850 nm
- Serial PMD 1310 nm
- Serial PMD 1550 nm
- Serial PMD 850 nm
- Serial PMD 1310 nm
- Serial PMD 1550 nm

- LX4
- SR
- LR
- ER
- SW
- LW
- EW
IEEE P802.3ae Objectives

- Preserve 802.3 Ethernet frame format
- Preserve 802.3 min/max frame size
- Full duplex operation only
- Fiber cabling only
- 10.0 Gbps at MAC-PHY interface
- LAN PHY data rate of 10 Gbps
- WAN PHY data rate of ~9.29 Gbps
802.3ae Detailed Objectives

- Preserve the 802.3/Ethernet frame format at the MAC client service interface
- Meet 802 functional requirements, with the possible exception of hamming distance
- Preserve minimum and maximum FrameSize of current 802.3 standard
- Support full-duplex operation only
- Support star-wired local area networks using point-to-point links and structured cabling topologies
- Specify an optional media independent interface
- Support proposed standard P802.3ad (link aggregation)
- Support a speed of 10.000 Gbps at the MAC/PLS service interface
802.3ae Detailed Objectives

- Define two families of PHYs
  - A LAN PHY, operating at a data rate of 10.000 Gbps
  - A WAN PHY, operating at a data rate compatible with the payload rate of OC-192c/SDH VC-4-64c

- Define a mechanism to adapt the MAC/PLS data rate to the data rate of the WAN PHY

- Provide physical layer specifications which support link distances of:
  - At least 65 m over MMF
  - At least 300 m over installed MMF
  - At least 2, 10, and 40 km over SMF

- Support fiber media selected from the second edition of ISO/IEC 11801 (802.3 to work with SC25/WG3 to develop appropriate specifications for any new fiber media)
# 802.3ae to 802.3z Comparison

## 1 Gigabit Ethernet
- CSMA/CD + Full Duplex
- Carrier Extension
- Optical/Copper Media
- Leverage Fibre Channel PMDs
- Reuse 8B/10B Coding
- Support LAN to 5 km

## 10 Gigabit Ethernet
- Full Duplex Only
- Throttle MAC Speed
- Optical Media Only
- Create New Optical PMDs from Scratch
- New Coding Schemes
- Support LAN to 40 km; Use SONET/SDH as Layer 1 Transport
“Running Ethernet over WANs may sound like a nice idea in principle, but it’s tough to pull off in practice. One of the fundamental rules about Ethernet is that the faster the network runs, the smaller the network gets.

At 10 Gbps, you end up with a very small network indeed – extending a couple of hundred yards over multimode fiber, max.”

WRONG
## 802.3ae

<table>
<thead>
<tr>
<th></th>
<th>LAN</th>
<th>MAN</th>
<th>RAN</th>
<th>WAN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LAN PHY</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>WAN PHY</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Bandwidth/Distance Evolution

- **Ethernet**
- **Fast Ethernet**
- **Gigabit Ethernet**
- **10 Gigabit Ethernet**

Distance (kilometers):
- 0.1
- 1
- 10
- 100
- 1000

Bandwidth (Mbps):
- 1
- 10
- 100
- 1000
- 10,000
## PMD Distances Supported

<table>
<thead>
<tr>
<th>Fiber</th>
<th>62.5 MMF</th>
<th>50 MMF</th>
<th>2000</th>
<th>SMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHz*km</td>
<td>160</td>
<td>400</td>
<td>200</td>
<td>500</td>
</tr>
<tr>
<td>850 nm SR/SW</td>
<td>28m</td>
<td>69m</td>
<td>300m</td>
<td>-</td>
</tr>
<tr>
<td>1310 nm LR/LW</td>
<td>35m</td>
<td>86m</td>
<td>300m</td>
<td>-</td>
</tr>
<tr>
<td>1310 nm ER/EW</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10 km</td>
</tr>
<tr>
<td>1310 nm LX4</td>
<td>300m</td>
<td>240</td>
<td>-</td>
<td>10 km</td>
</tr>
<tr>
<td>1310 nm @500MHz*km</td>
<td>-</td>
<td>300m</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
10 GbE Applications

10GBASE-LX4 or -LR Campus Link

10GBASE-LW or -EW Metro Link

10GBASE-LR, -ER Metro Link

10GBASE-LR, -ER, -LW, or -EW Metro Link

Enterprise A

Enterprise B

Enterprise C

Server Farm

Campus X

Campus Y
10 GbE in the LAN

10 GbE in:
SP data centers & enterprise LANs
- Switch-to-switch
- Switch-to-server
- Data centers
- Between buildings

Diagram:
- Campus A
- Campus B
- Internet
- Extranet
- Data Center
- Server Farm
- Switches
- 10GbE connections
10 GbE in the MAN over DWDM

Enterprises:
- 10 GbE enables server-less buildings remote backup disaster recovery

Service Providers:
- 10 GbE enables dark wavelength Gigabit services at costs less than T3 or OC-3
10 GbE in the MAN
over Dark Fiber
10 GbE in the WAN

- Attachment to the optical cloud
- Compatibility with the installed base of SONET STS-192c/SDH VC-4-64c
Layer Model

OSI Reference Model Layers

Application
Presentation
Session
Transport
Network
Data Link
Physical

P802.3ae LAYERS

Higher Layers
LLC
MAC Control
MAC
Reconciliation Sublayer (RS)

64B/66B PCS
PMA
PMD

MDI = Medium Dependent Interface
XGMII = 10 Gigabit Media Independent Interface
PCS = Physical Coding Sublayer

PMA = Physical Medium Attachment
PMD = Physical Medium Dependent
WIS = WAN Interface Sublayer

10GBASE-R
10GBASE-W
10GBASE-X
# Device Nomenclature

<table>
<thead>
<tr>
<th>Device</th>
<th>Logic</th>
<th>Optics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8B/10B PCS</td>
<td>64B/66B PCS</td>
</tr>
<tr>
<td>10GBASE-SR</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-SW</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-LR</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-LW</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-ER</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-EW</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10GBASE-LX4</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Logic: 8B/10B PCS, 64B/66B PCS, WIS
Optics: 850nm Serial, 1310nm WWDM, 1310nm Serial, 1550nm Serial
10GBASE-X

MAC

Ethernet Packet + Min. IPG

8b

8b

8b

8b

8B/10B Encoder

8B/10B Encoder

8B/10B Encoder

8B/10B Encoder

10b

10b

10b

10b

12.5 Gbps, 4 @ 3.125 Gbps

PMA

SERDES

SERDES

SERDES

SERDES

12.5 Gbps, 4 @ 3.125 Gbps

10 Gbps

10 Gbps

10 Gbps
10GBASE-R Serial

MAC

Ethernet Packet + Min. IPG

64B/66B PCS

64b 64b 64b 64b 64b 64b

64-bit Scrambler

Sync. Bits (2)

64b

PMA

SERDES

XGMII or XAUI

10 Gbps

10.0 Gbps

10.3 Gbps

XSBI

10.3 Gbps
The 10 Gigabit Ethernet LAN

- Faster: 10X
- Further: 40 km (expect proprietary extensions or WAN)
- Format: No change; same size packet
- Management: Consistent

Simple, Predictable, Elegant
10GBASE-W Serial

- **MAC**
- **Ethernet Packet + Min. IPG**
- **Extra IPG**
- **64B/66B PCS**
- **64-bit Scrambler**
- **Sync. Bits (2)**
- **Simplified SONET Framer**
- **WIS**
- **PMA**
- **SERDES**

- **XGMII or XAUI**
- **XSBI**
- **10 Gbps**
- **9.29 Gbps**
- **9.95 Gbps**

Extra IPG Dumped
Interfaces

- **XGMII (10G Media Independent I/F)**
  - 4 byte-wide lanes with 1 control bit per lane

- **XAUI (10G Attachment Unit I/F)**
  - Extends XGMII reach (3” vs. 20”)
  - 4 differential lanes at 3.125 Gbps

- **XSBI (10G Sixteen-Bit Interface)**
  - Based on the OIF SFI-4 interface
  - 16 differential signals at 622-645 Mbps
**XGMII Extender**

- **XGXS - XAUI - XGXS** blocks can be used to extend the XGMII with any PHY
- With LAN WWDM, the PHY-side XGXS & the 8B/10B PCS+PMA simplified to a re-timer
The 10 Gigabit Ethernet LAN

- Faster: 10X
- Further: 40 km
  - expect proprietary extensions on WAN
- Format: No change; same size packet
- Management: Consistent

Simple, Predictable, Elegant
‘Path,’ ‘Line,’ ‘Section’

- Path(s)
- Line
- Section

Note: A Line can be longer than two sections

**PCS Frame:**

Viewed as 9 x 17280 Octets

PCS Frame = STS-192c Frame

STS-192c = Synchronous Transport Signal – level 192, c = concatenated

Transmission order: Top to bottom, row-by-row, left to right
**Payload Capacity – 9.58464 Gbps**

PCS Frame = STS-192c Frame

9 rows

- **576 octets**
- **17280 octets**

(STS-192c) Envelope Capacity

(STS-192c) SPE

Payload Capacity

16640 octets

9 rows

Path Overhead column

Section

Transport Overhead

Line

Fixed Stuff

PCS Frame = STS-192c Frame

IDLE packet IDLE packet IDLE

IDLE packet IDLE

STC-192c = Synchronous Transport Signal – level 192, c = concatenated

SPE = Synchronous Payload Envelope
**Path Overhead and “Fixed Stuff”**

Defined overhead octets (F2, H4, Z3-5), unused by 10GE WAN PHY (set to zero)

“Fixed Stuff” columns provide compatibility with SONET/SDH byte-interleaving and concatenation rules (set to zero)
10GBASE-W Is SONET Friendly

SONET friendly does NOT mean SONET compliant...

- SONET frame (bits) are SONET compliant
  - No Layer 2 bridging required
  - Overhead will be interoperable with existing equipment

- Does NOT
  - Meet SONET jitter requirements
  - Match the ITU grid

- Does NEED a PHYSICAL layer conversion
Test Patterns

- Required – Built in
  - Pattern A seed: 0x3C8B44DCAB6804F
  - Pattern B seed: 0x3129CCCCF3B9C73
  - High Frequency Test Pattern (101010…)
  - Low Frequency Test Pattern (111110000011111…)
  - Mixed (+/- K28.5… = 11111010110000010100…)
  - PRBS31 G(x) = 1 + x 28 + x 31

- Required – Build in not required
  - CJPAT

- Other
  - CRPAT
Summary of 10 Gigabit Ethernet

- **MAC**
  - It’s just Ethernet
    - Maintains 802.3 frame format and size
    - Full duplex operation only

- **PHY**
  - LAN PHY uses simple encoding mechanisms to transmit data on dark fiber & dark wavelengths
  - WAN PHY adds a SONET framing sublayer to enable transmission of Ethernet on SONET transport infrastructure

- **PMD**
  - Support distances from 65m on installed MMF to 40km on SMF
  - No copper solution proposed
    - But, behind the scenes work starts on XAUI based....
Ethernet First Mile in Detail
802.3ah Task Force Objectives

- Support subscriber access network topologies:
  - Point-to-multipoint on optical fiber
  - Point-to-point on optical fiber
  - Point-to-point on copper

- Provide a family of physical layer specifications:
  - 1000BASE-LX extended temperature range optics
  - 1000BASE-X >= 10km over single SM fiber
  - 100BASE-X >= 10km over SM fiber
  - PHY for PON, >= 10km, 1000Mbps, single SM fiber, >= 1:16
  - PHY for PON, >= 20km, 1000Mbps, single SM fiber, >= 1:16
  - PHY for single pair non-loaded voice grade copper distance >=750m and speed >=10Mbps full-duplex
  - PHY for single pair non-loaded voice grade copper distance >=2700m and speed >=2Mbps full-duplex
802.3ah Task Force Objectives

- Support far-end OAM for subscriber access networks:
  - Remote Failure Indication
  - Remote Loopback
  - Link Monitoring

- Optical EFM PHYs to have a BER better than or equal to 10^-12 at the PHY service interface

- The point-to-point copper PHY shall recognize spectrum management restrictions imposed by operation in public access networks, including:
  - Recommendations from NRIC-V (USA)
  - ANSI T1.417-2001 (for frequencies up to 1.1MHz)
  - Frequency plans approved by ITU-T SG15/Q4, T1E1.4 and ETSI/TM6

- Include an optional specification for combined operation on multiple copper pairs
**OAM Overview**

- **Operations, Administration, and Maintenance**
  - Mechanisms for monitoring link operation; link and network health; and fault isolation
  - Data conveyed in 802.3 “Slow Protocol Frames” between two ends of a single link
- **No capability for station management, bandwidth allocation, or provisioning**
  - Vendor specific extensions supported
- **Applicable to all Ethernet PHYS**
  - Slow protocol allows implementation in software

*Fills major requirement to reduce EFM OpEx*
MDI = Medium Dependent Interface
XGMII = 10 Gigabit Media Independent Interface
PCS = Physical Coding Sublayer
PMA = Physical Medium Attachment
PMD = Physical Medium Dependent
WIS = WAN Interface Sublayer
**OAM Ping**

**Operation**
- Local end sends an ping request protocol data unit (PDU) to remote end
- PDU may contain data
- Remote end returns a ping response PDU
**OAM Frame Loopback**

**Operation**
- Local end sends loopback control PDU requesting remote end to go into loopback for a prescribed period of time
- Local end sends arbitrary data frames
- Remote end returns data frames
- Frame BER equals bit BER to high probability when bit BER is better than $10^{-6}$
Frame Errors vs. Bit Errors

- Assume errors are Poisson distributed in time
  - e.g., system dominated by white, Gaussian noise
  - ignores burst noise
- FER = BER if probability of >1 bit errors over the length of the frame is small
  - depends on BER & frame length
  - depends on acceptable probability for FER $\neq$ BER
- Sample calculation:
  - 30kb frame
  - acceptable probability $\leq 1\%$
  - $\Rightarrow$ BER $\leq 5 \times 10^{-6}$

Source: John Ewen, JDSU 2002
OAM: Other Functions

- Sends limited link status flags with each PDU
  - Local / Remote Fault
  - Dying Gasp
  - Alarm Indication

- Status PDU

- Event notification PDU

- Variable request and response PDUs
  - Transfer via variable containers for Ethernet attributes; objects and packages
Point-To-Point Overview

- 4 New Links (6 PMDs)
  - Standardizes 100 Mbps 10km dual fiber
    - Based on FDDI
  - Standardizes 1 Gbps, 10km dual fiber
    - Based on existing 10km parts available
  - Adds 100 Mbps single fiber
    - Based on TTC’s TS-1000 specification
  - Adds 1 Gbps single fiber
    - New

- No changes to PMA; PCS; or MAC
  - Excepting simplex operation for OAM
## Optical PMD Summary Sheet

<table>
<thead>
<tr>
<th>Port Type</th>
<th># Fibers</th>
<th>SMF (km)</th>
<th>MMF (m)</th>
<th>λ Tx (nm)</th>
<th>λ Rx (nm)</th>
<th>Rx Sen (dBm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000BASE-EX</td>
<td>2</td>
<td>&gt;10</td>
<td>&gt;500</td>
<td>1260-1360</td>
<td>1260-1360</td>
<td>-20</td>
</tr>
<tr>
<td>1000BASE-BX-OLT</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1480-1500</td>
<td>1260-1360</td>
<td>-20</td>
</tr>
<tr>
<td>-BX-ONU</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1260-1360</td>
<td>1480-1500</td>
<td>-20</td>
</tr>
<tr>
<td>100BASE-LX</td>
<td>2</td>
<td>&gt;10</td>
<td>-</td>
<td>1260-1360</td>
<td>1260-1360</td>
<td>-25</td>
</tr>
<tr>
<td>100BASE-BX-OLT</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1480-1580</td>
<td>1260-1360</td>
<td>-30</td>
</tr>
<tr>
<td>-ONU</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1260-1360</td>
<td>1480-1600</td>
<td>-30</td>
</tr>
<tr>
<td>1000BASE-PX-OLT-A</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1480-1500</td>
<td>1270-1360</td>
<td>-26</td>
</tr>
<tr>
<td>-ONU-A</td>
<td>1</td>
<td>&gt;10</td>
<td>-</td>
<td>1270-1360</td>
<td>1480-1500</td>
<td>-25</td>
</tr>
<tr>
<td>-OLT-B</td>
<td>1</td>
<td>&gt;20</td>
<td>-</td>
<td>1480-1500</td>
<td>1270-1360</td>
<td>-29</td>
</tr>
<tr>
<td>-ONU-B</td>
<td>1</td>
<td>&gt;20</td>
<td>-</td>
<td>1270-1360</td>
<td>1480-1500</td>
<td>-25</td>
</tr>
</tbody>
</table>
EFM Copper Introduction

- Ethernet in the First Mile Copper (EFMC)
  - Brings native Ethernet to the “First Mile” (ex. Last Mile) twisted-pair access network

- Why do we need it?
  - Existing Ethernet PHYs designed for engineered wiring
  - Public access network originally designed for voice-only, not data
  - FCC requirements for spectrum compatibility & EMI not met by existing Ethernet PHYs
  - Existing DSLs optimized for non-Ethernet protocols
Loop Length Distribution Graphs

Distribution of 26 AWG Loop Lengths from 1990 DLC Loop Survey

Distribution of Non-Loaded 26 AWG Loop Lengths From 1983 Loop Survey
PSTN Loop Plant

- Multiple pairs wrapped tightly together in each binder
- Binders fan out as they extend toward subscribers
- "Bridge Taps" occur where stubs are left unconnected
- In-building wiring also a factor
Transmission Characteristics

- **Attenuation**
  - Loss increases with frequency

- **Crosstalk**
  - Predominant impairment in loop plant
  - Interference from same type of service on other pairs in binder (self-crosstalk), or other types of service (alien-crosstalk)

- **POTS/ISDN overlay**
  - POTS (0-25 KHz) or ISDN (0-138 KHz) may be operating on same pair
Band Plans for Different Services

Band plan definitions administered by regulators to help endure operation of different services in same binder.
Near-End Crosstalk (NEXT)

- Attenuation to Crosstalk Ratio (ACR) gives measure of SNR
- ACR approaches 0 for many EFM cable types at 3kft, 2MHz

Strong Tx signal Kills weak Rx
Crosstalk: FEXT and NEXT

- **FEXT: Far-End X-Talk**
  - Caused by transmitter operating on another pair in binder, at opposite end from receiver
  - Crosstalk level attenuated by loop attenuation

- **NEXT: Near-End X-Talk**
  - Caused by transmitter operating on another pair in binder, at same end as receiver
  - No loop attenuation; higher level than FEXT

- NEXT more problematic; commonly handled by using FDM to split upstream and downstream
Channel Capacity

Theoretical maximum bitrate depends on available bandwidth, noise level

\[ C = \int \log_2 \left( \frac{1 + s(f) \times |H(f)|^2}{N(f)} \right) df \]

- \( C \) – theoretical bitrate capacity
- \( s(f) \) – signal PSD, watts/Hz vs. freq.
- \( N(f) \) – noise PSD at receiver
- \( H(f) \) – loop loss vs. freq.

Channel capacity increases with bandwidth and signal PSD, decreases with loop loss, noise

- Noise includes –174 dBm/Hz thermal noise & crosstalk
Self-Disturber Rate vs. Reach

Simulation update from March Presentation [1].

- Loop model refined
- Reduced total power on 4 midrange speeds to comply with composite VDSL Mask

![Graph showing 100BASE-Cu Raw Bitrate, Symmetric Traffic, 50 Self Disturbers Only, -140dBm/Hz. 26AWG](image)
Regulatory Issues

- Loop Unbundling
  - Loops in a binder may be operated by different Telcos
  - Crosstalk from pairs operated by one company will affect performance on pairs operated by another

- Spectral compatibility
  - Spectral limits and deployment guidelines to ensure fair use of binder resources
  - Mandated by national regulators (FCC, etc.)

- ANSI T1.417
  - U.S. standard for spectral compatibility
  - Requires demonstration of compatibility with widely-deployed “basis systems”
Overview / Intro of DSL Technologies

- **DSL – Digital Subscriber Line**
  - Use of twisted-pair access loops for the transmission of wideband digital signals
  - Operates up to 12 MHz bandwidth (e.g., VDSL)

- Various DSLs
  - **HDSL** – symmetric, T1 carriage, no POTS overlay
  - **ADSL** – asymmetric, POTS overlay, medium-long loops
  - **VDSL** – symmetric & asymmetric, short loops, high speed
EFM Copper: Based on DSL Technologies

- EFM copper PHYs use DSL modulation techniques
- Leverages years of work on DSL modulation development
- Ensures spectral compatibility
  - And thus legality of deployment
DSL Modulation Techniques

Two broad categories:

- **DMT** – Discrete Multitone Modulation
  - Large number of narrowband, orthogonal, modulated carriers

- **QAM** – Quadrature Amplitude Modulation
  - Single wideband, modulated carrier

Both types commonly used in various DSL standards
EFMC: An Evolutionary Improvement over Existing DSL

EFM simplifies, specifies, mandates interoperability

- Simplified protocol layers
- Reduces configuration, provisioning options
- IEEE 802.3 Ethernet tradition ensures interoperability
- Two Ethernet port types vs. a myriad of non-interoperable DSL types
EFM Protocol Streamlining

- Current typical DSL protocol stack a byzantine collection
  - Built to accommodate services that were never deployed
  - Result is additional costs for needless provisioning, configuration, and maintenance
Protocol Streamlining (cont’d)

Typical IP connection begins and ends on Ethernet

- Flexibility of ATM unutilized; complexity unnecessary
- New DSL systems will strip out intermediate sublayers, move to native Ethernet on DSL
Work In Progress (cir 9/02)

Ethernet First Mile Task Force Copper:

working to select line code for long reach from between DMT and QAM

- “…omahony_copper_1_0702.pdf” as the basis for the line code evaluation criteria.
- “…limit proposals for consideration regarding the long reach objective to those based on “artman_copper_1_0702.pdf” and “jackson_copper_1_0702.pdf”
EPON Overview

- Point-to-multipoint fiber network
- High bandwidth: 1 Gbps shared
- Low cost Ethernet + low cost fiber plant
- Minimizes use of fiber, CO feeders, and transceivers
- Passive optical infrastructure
- Fiber-to-the-home/building/business applications
- Suitable for voice, data, and video services
Optical First Mile

**Point-to-Point Ethernet**
- N or 2N fibers
- 2N optical transceivers

**Curb Switched Ethernet**
- 1 trunk fiber
- Minimum fiber/space in CO
- 2N+2 optical transceivers
- Electrical power in the field

**Ethernet PON (EPON)**
- 1 trunk fiber
- Minimum fibers/space in CO
- N+1 optical transceivers
- No electrical power in field
- Drop throughput up to trunk rate
- Downstream broadcast (video)
EPON System Architecture

EPON is typically deployed as a tree or tree-and-branch topology, using passive 1:N optical splitters.
Example: EPON Network

The **Optical Line Terminal** (OLT) resides in the central office (PoP, local exchange). This is typically an Ethernet switch or media converter platform.

The **Optical Network Terminal** (ONT) resides at or near the customer premise. The ONT can be located on the curb/outside, in a building or at a subscriber residence. This unit typically has an 802.3ah WAN interface and an 802.3 subscriber interface.
EPON Configuration

- Single fiber point-to-multipoint
- Full-duplex mode (no CSMA/CD)
- Subscribers see traffic only from head end, not from each other. Headend permits only one subscriber at a time to transmit using TDMA protocol
- Flexible optical splitter architectures
- 1490 nm downstream, 1310 nm upstream
Ethernet PON can be deployed in an Ethernet access platform, with both point-to-point and point-to-multipoint access cards.
Multipoint Control Protocol (MPCP)

- EPON uses Multipoint Control Protocol (MPCP) to control Point-to-Multipoint (P2MP) fiber network
- MPCP performs \textit{bandwidth assignment, bandwidth polling, auto-discovery process and ranging}, and is implemented in the MAC control layer
- New 64 byte MAC control messages are introduced. GATE and REPORT are used to assign and request bandwidth. REGISTER messages are used to control the auto-discovery process
- MPCP provides hooks for network resource optimization:
  - ranging is performed to reduce slack
  - reporting of bandwidth requirements by ONTs for DBA
  - optical parameters are negotiated to optimize performance
ONT and OLT Operation

ONT

- Performs auto-discovery process which includes ranging, assignment of logical link IDs, assignment of bandwidth
- Synchronizes to OLT timing through timestamps on the downstream GATE MAC control message
- Receives GATE message and transmits in permitted time period

OLT

- Generates time stamped messages to be used as global time reference
- Generates discovery windows for new ONTs, and controls registration process
- Assigns bandwidth and performs ranging
EPON Downstream

- Physical broadcast of 802.3 Frames
- 802.3 Frames extracted by logical link ID in preamble
- 64 byte GATE messages sent downstream to assign bandwidth
EPON Downstream: GATE Message

OLT

MAC Control Client

Generate GATE message

Start
Stop

MAC Control

Clock register

Timestamp GATE message

Start
Stop

MAC

PHY

ONU (1 of N)

MAC Control Client

MA_CONTROL.request(GATE)

MA_CONTROL.indication(GATE)

MA_DATA.request(…)

MAC Control

Write registers

Start
Stop

TS

Clock register

Slot Start register

Slot Stop register

Laser ON/OFF

Upstream Data Path

MAC

PHY

INSCOPE
EPON Upstream

- Upstream control managed by MPCP protocol
- Time slots contain multiple 802.3 Ethernet frames
- 64 byte REPORT Message sends ONU state to OLT
- No collisions
- No packet fragmentation
EPON Upstream: REPORT Message

OLT

MAC Control Client

MAC Control

Clock register

RTT register

Measure Round-Trip Time

ONU (1 of N)

MAC Control Client

MAC Control

Generate REPORT message

MA_CONTROL.indication(REPORT)

TS

TBD

TBD

MA_CONTROL.request(REPORT)

TS

TBD

TBD

Timestamp

REPORT message

Clock register

TBD

TBD

MAC

PHY

MAC

PHY
Round Trip Time (RTT) Measurement

1. OLT sends GATE at T1
2. ONU receives GATE and sets its clock to T1
3. ONU sends REPORT at T2
4. OLT receives REPORT at T3
5. OLT calculates RTT = T3 – T2

\[ RTT = (T3 - T1) - (T2 - T1) = T3 - T2 \]

** based on OLT clock; *** based on ONU clock
Work in Progress (cir 9/02)

Ethernet First Mile Task Force P2MP:

- Creating sublayers for P2MP that support inherent downstream broadcast and P2P emulation
- Working to resolve architectural issues with the 802.3 layer stack
- Investigating possible support of L2 security
- Investigating possible use of forward error correction (FEC) to simplify P2MP optics
Hybrid Fiber/Copper

• Next-generation, high-speed architectures
  – EFM copper for the last 700 to 800 meters
  – Minimum 10 Mbps – higher if possible
  – High bandwidth for entertainment – client/server
  – For stepwise buildout to work, EFMCu must support next-gen applications

Source: EFMA 2002
Bandwidth vs. Time

- Ethernet (2000X in 12 yr)
- Modem (47X in 17 Yr)
- DSL (13X in 12.5 yr)
From Copper to Fiber

This chaos cannot be resolved by some central authority
IEEE 802.17
aka Resilient Packet Ring
aka RPR
aka ?Ethernet Loop?
RPR Overview

- Dual counter-rotating ring topology
- Frame-based transmission *(jumbo support)*
- Defines a Layer 2 protocol
  - Support for Unicast/Multicast/Broadcast
  - Familiar 48-bit MAC addresses
- Native support for QoS
  - 4 classes: Reserved, high, medium, low
  - Fair access to available (unreserved) capacity
- Fast fail-over *(sub 50ms)*
- Dynamic topology discovery
- Use 802.3 and SONET PHY technology
It Came from the MAN…

- Targeted at SONET Metro rings
  - “SONET Reliability at Ethernet Cost”

- How are costs lowered?
  - Spatial reuse (unicast)
  - Both fibers carry traffic (SONET is active/standby)
  - Multiple traffic classes allow TDM
  - Ethernet “goodness”

- How is reliability maintained?
  - Maintains the two-ring topology
  - Protocol supports 50ms fail-over for failing links/stations
  - Same protocol supports plug-and-play
RPR: A System View

Station

Outer Ringlet

Inner Ringlet

Span
An RPR Data Frame

<table>
<thead>
<tr>
<th>Field</th>
<th>Length</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPR HEADER</td>
<td>2 Octets</td>
<td></td>
</tr>
<tr>
<td>DESTINATION MAC ADDRESS</td>
<td>6 Octets</td>
<td></td>
</tr>
<tr>
<td>SOURCE MAC ADDRESS</td>
<td>6 Octets</td>
<td></td>
</tr>
<tr>
<td>PROTOCOL TYPE FIELD (TBD)</td>
<td>2 Octets</td>
<td></td>
</tr>
<tr>
<td>PAYLOAD</td>
<td>m Octets</td>
<td></td>
</tr>
<tr>
<td>FCS</td>
<td>4 Octets</td>
<td></td>
</tr>
</tbody>
</table>

TTL

<table>
<thead>
<tr>
<th>Bit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>TTL</td>
</tr>
<tr>
<td>6</td>
<td>MODE</td>
</tr>
<tr>
<td>5</td>
<td>RI</td>
</tr>
<tr>
<td>4</td>
<td>PRI</td>
</tr>
<tr>
<td>3</td>
<td>IOP</td>
</tr>
</tbody>
</table>
RPR Header Fields

- **TTL (8 bits)**
  - Time To Live
  - Set to number of hops to destination
  - Decremented when forwarded by node
  - Allows for 255 nodes on ring

- **MODE (3 bits)**
  - Frame type

---

<table>
<thead>
<tr>
<th>Mode Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
</tr>
<tr>
<td>1</td>
<td>Reserved</td>
</tr>
<tr>
<td>2</td>
<td>Reserved</td>
</tr>
<tr>
<td>3</td>
<td>Steering only data</td>
</tr>
<tr>
<td>4</td>
<td>Protection Control</td>
</tr>
<tr>
<td>5</td>
<td>Control</td>
</tr>
<tr>
<td>6</td>
<td>Fairness</td>
</tr>
<tr>
<td>7</td>
<td>Data</td>
</tr>
</tbody>
</table>
RPR Header Fields (cont.)

- RI(1 bit) - Ringlet Identifier
  - Origination ringlet

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Clockwise ringlet</td>
</tr>
<tr>
<td>1</td>
<td>Counterclockwise ringlet</td>
</tr>
</tbody>
</table>

- IOP(1 bits) – In/Out Profile
  - Used for medium priority traffic
  - Out of profile traffic treated as low priority

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Out of profile</td>
</tr>
<tr>
<td>1</td>
<td>In profile</td>
</tr>
</tbody>
</table>
RPR Header Fields (cont.)

- PRI(3 bits) – Priority

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>High priority</td>
</tr>
<tr>
<td>0-6</td>
<td>Low priority</td>
</tr>
</tbody>
</table>

- Entire 3-bit priority used by MAC client
  - For transmit scheduling
  - For receive processing
Overview of an RPR MAC

MAC Control Sublayer

MAC Fairness Control Unit

Drop Logic

MFU of mate

Ringlet Input

Ringlet Output

Policer/Shaper

M: rate monitor

STOP_LP/MP/HP

MAC Client
More About the MAC

- Not compatible with Ethernet!
- RPR MACs come in pairs
- RPR MAC can hide or expose the dual-ring nature
  - If exposed, the MAC client can choose which ring to send a frame on
  - Otherwise, the MAC makes the decision
RPR Traffic Classes

- **Reserved (A₀)**
  - Guaranteed rate and tightly bounded delay/jitter

- **High (A₁)**
  - Committed rate with controlled delay/jitter
  - Subject to capacity restoration

- **Medium (B)**
  - Committed rate + burst capability
  - In profile/out of profile (excess MP)
  - eMP subject to RPR-FA (Fairness Algorithm)

- **Low (C)**
  - Best effort
  - Subject to RPR-FA
**RPR Ring Access**

- **Forwarding**
  - 1 or 2 transit buffers (HP & LP/MP)

- **Policing**
  - Each node has maximum total add rate
  - And an add rate for each traffic class (A,B,C)
  - Implemented with token buckets
  - Communicate status back to MAC client

- **Dynamic shaping**
  - Nodes can make use of the excess or recovered bandwidth
  - Utilizes the RPR-FA algorithm
Access Rules

- **HP Transit Has Frame?**
  - Yes: HP Transit
  - No: LP Transit

- **LP Transit Nearly Full?**
  - Yes: LP Transit
  - No: Add HP/MP?

- **Add HP/MP?**
  - Yes: HP/MP Add
  - No: (Add_rate + forward_rate) > (max_rate – reserved_rate)?

- **(Add_rate + forward_rate) > (max_rate – reserved_rate)?**
  - Yes: Trigger Flow Control
  - No: Add MP/LP?

- **Add MP/LP?**
  - Yes: Update FA
  - No: P: Policing Engine

- **LP Transit Has Frame?**
  - Yes: LP Transit
  - No: Add MP/LP?

- **P: Policing Engine**
- **FA: Fairness Algorithm**
RPR Fairness (RPR-FA)

- Defined at the MAC layer
  - Supplemented by MAC client
- Uses source-based weighted fairness
  - Divide the available bandwidth among nodes
  - Nodes may be weighted to get more or less than their “fair share”
- Applies only to LP/eMP traffic
- Goals
  - Reclaim unused committed BW
  - Fast response
  - High BW utilization
  - Stability
  - Scalability
RPR Fairness (RPR-FA)

Components at each station
- Determine congestion
  - Monitor the outgoing link rate
  - Watchdog timer for LP/MP packets
  - LP transit buffer reaches threshold
- Calculate an advertisement rate
  - \( \frac{\text{Add\_rate}}{\text{node\_weight}} \)
  - If congested, advertises rate (Type A message) to upstream node
    - The upstream node may advertise its own rate if it is also congested, forward this rate, or forward a null rate
- Determine the station’s allowed rate
  - Based on advertised rate of most congested node
  - Multiplied by stations weight
**Extended RPR Fairness**

- Handled by MAC client
- Uses Type B fairness messages
  - Broadcast to all nodes
- Allows all choke points to be simultaneously tracked
  - Leads to better spatial reuse
  - Supports virtual destination queues
  - Allows unlimited traffic for frames that are in front of a choke point
  - Requires only that each FA rule between source and destination is obeyed
Ring Protection

- Wrapping vs. Steering
  - This was a major sticking point
  - The compromise was “Do Both”

- Steering
  - Mandatory part of standard
  - “Steers” frames away from failed links
  - Uses protection messages to advertise failures
  - More frames may be dropped

- Wrapping
  - Optional in standard
  - All traffic is wrapped around when a station detects a failure in its neighbor
  - Fewer dropped frames
Steering Example – A to C
Wrapping Example – A to C
Physical Layer

- There is no RPR PHY!
- The standard defines reconciliation layers for:
  - 1 Gig Ethernet – GMII
  - 10 Gig Ethernet – XGMII, XAUI
  - SONET/SDH at 155Mbps to 10Gbps
Scenario 3: RPR Vision
## RPR to SONET Comparison

<table>
<thead>
<tr>
<th></th>
<th>RPR</th>
<th>SONET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair access to ring bandwidth</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>High bandwidth efficiency on dual-ring topology</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Full FCAPS* with LAN-like economics</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Controlled latency and jitter</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>50-millisecond ring protection</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Optimized for data</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Cost-effective for data</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

*fault-management, configuration, accounting, performance, and security
## Fair Comparison?

<table>
<thead>
<tr>
<th>Feature</th>
<th>RPR</th>
<th>SONET</th>
<th>Ethernet</th>
<th>FCAL**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair access to ring bandwidth</td>
<td>✓</td>
<td>?</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>High bandwidth efficiency on dual-ring topology</td>
<td>✓</td>
<td>✓?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Full FCAPS* with LAN-like economics</td>
<td>✓</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Controlled latency and jitter</td>
<td>✓</td>
<td>✓</td>
<td>✓?</td>
<td>✓</td>
</tr>
<tr>
<td>50-millisecond ring protection</td>
<td>✓</td>
<td>✓</td>
<td>✓?</td>
<td>✓</td>
</tr>
<tr>
<td>Optimized for data</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cost-effective for data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

*But, are rings the way of the future, or a simply a means to replace SONET in the metro?*

*fault-management, configuration, accounting, performance, and security  
**Fiber-Channel Arbitrated Loop*
RPR Conclusion

- Frame based like Ethernet
- Supports a familiar topology to offer data services (SONET ring) ...and Ethernet can't?
- Spatial Reuse which Ethernet doesn't need!
  - Like SSA and dual Ring FC
  - Unlike SONET
- Provides a layer-2 standard to address QoS and reliability which Ethernet can do with much greater flexibility
- Not Ethernet

But, does RPR offer sufficient benefit over Ethernet?
Transceivers, Fibers, and Issues with Optics
OSI Layer Stack Mapping

**Data Format**
- GMII: N x 8 bit
- TBI: 10 bit
- 1 bit
- MDI: 1 bit

**Typical 1 Gig Link**
- Protocol
- Coding
- SERDES
- XCVR
- Media

**OSI Reference**
- Application
- Presentation
- Session
- Transport
- Network
- Data Link
- Physical

**Proposed 10 Gig Link**
- Protocol
- Coding
- SERDES
- XCVR
- Media

**Typical 1 Gig Link**
- GMII: N x 8 bit
- TBI: 10 bit
- 1 bit
- MDI: 1 bit

**Proposed 10 Gig Link**
- XGMII: 4 x 8 bit
- XAUI: 4 x 1 bit
- MDI: 1 bit or 4 x 1 bit or
Example: 1 Gig Partitions

Protocol with “Integrated SERDES”

“MAC” & “PCS”

“GLM”

“10B”

“1x9” GBIC

SFF

SFP
1GbE: Typical Implementation

802.3 Layer Model

- Logical Link Control
- MAC Control (Opt)
- Media Access Control
- Reconciliation Sublayer

(XG)MII

- Physical Coding Sublayer
- Physical Medium Attachment
- Physical Medium Dependent

MDI

Media

circa ‘01
Typical 1 Gigabit Optical XCVRs

Pin in Hole

Pluggable

1996

1997

2001

2002
Early 10 Gigabit Optical XCVRS

Seen at Optical Fiber Conference:
• XENPACK; FTRX (300 pin MSA)
• XXP; XPAK; XFP; SFP (@10 Gig!)
Multimode vs. Single Mode Cost

The vast majority of the cost difference is in the size of the target!

Challenge: Control mechanical tolerances over temperature
Distance: Attenuation & Modal Bandwidth Issues of Fiber

Fiber distances are primarily impacted by:

- **Attenuation (850 >> 1300 >> 1550)**
  - *The amount of loss per meter of optical power*

- **Bandwidth * Distance Product**
  - *Modal Dispersion*
    - *62.5 MMF > 50 MMF >>> SMF*

- **Chromatic Dispersion**
  - *850 >> 1300 < 1550 for “standard SMF”*
  - *1310 nm is the “zero dispersion wavelength”*
Fiber Attenuation

Wavelength in nanometers

Early 1970s
Mid 1970s
Early 1980s

More Loss
Less Loss
Effects of Dispersion

Optical power at fiber input
850 nm Oxide VCSEL @ 1.25 GBd

...and end of 600 m of 62.5 micron multimode fiber
Modal Dispersion

- The net speed of light is a function of the path (mode)
  - The smaller the core of the fiber, the fewer the number of modes that will propagate
  - Single mode fiber (SMF) has only one mode and therefore no modal dispersion (e.g., railroad)
  - Multi-mode fiber (MMF) “profiles” are doped so that all paths take about the same time. Index at center of fiber “slows down” low order modes
Chromatic Dispersion

- Speed is a function of color ($\lambda$)
- Spectral width ($\Delta\lambda$) is a measure of the source’s color range.
- Chromatic dispersion is reduced by controlling the source’s $\Delta\lambda$
  - Example: Use of a narrow linewidth source (e.g., DFB laser)
  - Example: Low $\alpha$ (chirp) laser: Small change in $\lambda$ as laser turns on and off (modulates)
  - Example: External modulation (reduces chirp)
How Is 10 km Achieved When the 802.3z LW SMF Spec. Is 5 km?

- **Limit 1:** Link budget = Minimum optical power output - Minimum receive sensitivity
  - A portion of the link budget is allocated to fiber loss (attenuation)
  - Use simple photodiode

- **Limit 2:** Receiver Dynamic Range = Maximum - Minimum optical power into receiver

<table>
<thead>
<tr>
<th></th>
<th>802.3z</th>
<th>New FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx min (dBm)</td>
<td>-19</td>
<td>-20</td>
</tr>
<tr>
<td>Tx Min (dBm)</td>
<td>-11</td>
<td>-9.5</td>
</tr>
<tr>
<td>Budget (dB)</td>
<td>8</td>
<td>10.5</td>
</tr>
<tr>
<td>Fiber Alloc.</td>
<td>2.5</td>
<td>5</td>
</tr>
<tr>
<td>Attn (dB/km)</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Distance (km)</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Rx/Tx max</td>
<td>-3</td>
<td>-3</td>
</tr>
<tr>
<td>Dynamic Rng</td>
<td>16</td>
<td>17</td>
</tr>
</tbody>
</table>

- 802.3z set objectives to achieve 3 km; some members objected to greater Rx dynamic range
How Is 20 to 50 km Achieved with 1300 nm LW?

- Increase the launch power
  - closer to the eye safety limit
- Increase the sensitivity of the receiver (APD)
- Increase the dynamic range of the receiver

<table>
<thead>
<tr>
<th></th>
<th>802.3z</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx min (dBm)</td>
<td>-19</td>
<td>-22</td>
</tr>
<tr>
<td>Tx Min (dBm)</td>
<td>-11</td>
<td>0</td>
</tr>
<tr>
<td>Budget (dB)</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Fiber Alloc.</td>
<td>2.5</td>
<td>20?</td>
</tr>
<tr>
<td>Attn (dB/km)</td>
<td>0.5</td>
<td>0.5?</td>
</tr>
<tr>
<td>Distance (km)</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>Rx/Tx max</td>
<td>-3</td>
<td>&gt;2</td>
</tr>
<tr>
<td>Dynamic Rng</td>
<td>16</td>
<td>&gt;24</td>
</tr>
</tbody>
</table>
How is 100 km Achieved with 1500 nm?

- Increase the launch power
  - Eye safety virtually no problem at 1550 nm
- More Rx sensitivity
- More Rx dynamic range or engineer link to bound attenuation
- Control the $\Delta \lambda$:

<table>
<thead>
<tr>
<th></th>
<th>802.3z</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx min (dBm)</td>
<td>-32</td>
<td>-32</td>
</tr>
<tr>
<td>Tx Min (dBm)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Budget (dB)</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Fiber Alloc.</td>
<td>25?</td>
<td>25?</td>
</tr>
<tr>
<td>Attn (dB/km)</td>
<td>0.25?</td>
<td>0.25?</td>
</tr>
<tr>
<td>Distance (km)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Rx/Tx max</td>
<td>&gt;1</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Dynamic Rng</td>
<td>&gt;33</td>
<td>&gt;33</td>
</tr>
</tbody>
</table>
Gigabit Ethernet Fiber Issues

- Differential Modal Dispersion (DMD)
  - FDDI Grade Multimode Fiber
  - Defect in center of fiber
  - Causes pulse splitting
  - Not specified in fiber
- Distance reduced for 850 nm from objective
- Fixed with an “offset patch cord” for 1310 nm
  - Single mode launch
  - Offset from center by 17 to 23 microns for 62.5 MMF
  - Offset from center by 10 to 16 microns for 50 MMF
Example DMD from NIST

DMD from 50 m sections of
1.17 km 62.5/125 μm MMF #1
@ 850 nm
10G Ethernet Fiber Issues

- Polarization Modal Dispersion (PMD)
  - Single mode fiber
  - Two polarities of light propagation travel over single mode fiber at different velocities
  - Variation varies over time
  - Specified as a probability with a maximum delay

- Extremely important at high speeds and long distances (e.g., 100km at OC-768)
  - 40 km at 10 Gig not an issue
    - 95% probability will not exceed 16 ps
1 Gig Stressed Rx Eye Definition

Figure 52–14—Required characteristics of the conformance test signal at TP3

Robust, difficult to create and validate
10 Gig Stressed Rx Eye Definition

Jitter histogram (at waveform average, may not be at waist)

Vertical eye closure histograms (at time-center of eye)

Approximate OMA (difference of means of histograms)

Less robust; substantially easier to create and validate

Figure 52-11—Required characteristics of the conformance test signal at TP3
Stressed Eye – Lone Bit Pattern

Bob Zona, Intel
850 nm OMA

Figure 52–3—Triple tradeoff curve for 10GBASE-S
(informative figure)
OMA vs. Optical Power (Sample)

-3
-3.5
-4
-4.5
-5
-5.5
-6
-6.5
-7
-7.5
-8

Pavg (dBm)

0.35
0.45
0.56
0.71

OMA (mW)

Operation Conditions with Current Specification
Area Opened With OMA
1G and 10G Test Points (TP)

TP1: SERDES Out
TP2: TX Out
TP3: RX In
TP4: SERDES In

1550 nm Tx spec’ed at TP3 (chirp...)
10G Ethernet Optics Issues

- Pushing the low cost technology envelope
- Problems with test and measurement
  - Created “best of breed,” modern test method
    - BER jitter masks
  - Test equipment was simply not good enough
    - Yesterday’s “fat” is today’s specification
    - Testing indicated high percentage of “false negatives”
  - Changed methodology for 10G Serial
    - Time and Dispersion Penalty (TDP)
10G Jitter Masks – Almost

This scheme is still used for 10GBASE-LX4
10GbE Modified Tx Eye Mask

Figure 52–8—Transmitter eye mask definition
Block Diagram for LX4 PMD
# 10GBASE-LX4 Tx Specifications

<table>
<thead>
<tr>
<th>Description</th>
<th>62.5 μm MMF, 50 μm MMF, 10 μm SMF</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmitter Type</td>
<td>Longwave Laser</td>
<td></td>
</tr>
<tr>
<td>Signaling speed per lane (nominal)</td>
<td>3.125 +/- 100 ppm</td>
<td>GBd</td>
</tr>
<tr>
<td>Lane wavelengths (range)</td>
<td>1269.0 - 1282.4</td>
<td>nm</td>
</tr>
<tr>
<td></td>
<td>1293.5 - 1306.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1318.0 - 1331.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1342.5 - 1355.9</td>
<td></td>
</tr>
<tr>
<td>Trise/Tfall (max. 20-80 % response time)</td>
<td>120</td>
<td>ps</td>
</tr>
<tr>
<td>Side-mode suppression ratio (SMSR), (min)</td>
<td>0.0</td>
<td>dB</td>
</tr>
<tr>
<td>RMS spectral width (max)</td>
<td>0.62</td>
<td>nm</td>
</tr>
<tr>
<td>Average launch power, four lanes (max)</td>
<td>5.5</td>
<td>dBm</td>
</tr>
<tr>
<td>Average launch power, per lane (max)</td>
<td>-0.5</td>
<td>dBm</td>
</tr>
<tr>
<td>Optical Modulation Amplitude (OMA), per lane (max)</td>
<td>750 (-1.25)</td>
<td>μW (dBm)</td>
</tr>
<tr>
<td>Optical Modulation Amplitude (OMA), per lane (min)</td>
<td>237 (-6.25)</td>
<td>μW (dBm)</td>
</tr>
<tr>
<td>Extinction Ratio (min)</td>
<td>3.5</td>
<td>dB</td>
</tr>
<tr>
<td>Average launch power of OFF transmitter, per lane (max)</td>
<td>-30</td>
<td>dBm</td>
</tr>
<tr>
<td>RIN_{12} (OMA)</td>
<td>-120</td>
<td>dB/Hz</td>
</tr>
</tbody>
</table>
## 10GBASE-LX4 Rx Specifications

<table>
<thead>
<tr>
<th>Description</th>
<th>62.5 μm MMF 50μm MMF</th>
<th>10 μm SMF</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signaling speed per lane (nominal)</td>
<td>3.125 ± 100 ppm</td>
<td>3.125 ± 100 ppm</td>
<td>GBd</td>
</tr>
<tr>
<td>Lane wavelengths (range)</td>
<td>1269.0 - 1282.4</td>
<td>1269.0 - 1282.4</td>
<td>nm</td>
</tr>
<tr>
<td></td>
<td>1293.5 - 1306.9</td>
<td>1293.5 - 1306.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1318.0 - 1331.4</td>
<td>1318.0 - 1331.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1342.5 - 1355.9</td>
<td>1342.5 - 1355.9</td>
<td></td>
</tr>
<tr>
<td>Average receive power, four lanes (max)</td>
<td>5.5</td>
<td>5.5</td>
<td>dBm</td>
</tr>
<tr>
<td>Average receive power, per lane (max)</td>
<td>-0.5</td>
<td>-0.5</td>
<td>dBm</td>
</tr>
<tr>
<td>Return loss (min)</td>
<td>12</td>
<td>12</td>
<td>dB</td>
</tr>
<tr>
<td>Receive sensitivity (OMA), per lane</td>
<td>37.4 (-14.25)</td>
<td>32.7 (-14.85)</td>
<td>μW (dBm)</td>
</tr>
<tr>
<td>Stressed receive sensitivity (OMA)(^{\dagger,}) per lane</td>
<td>93 (-10.3)</td>
<td>45 (-13.5)</td>
<td>μW (dBm)</td>
</tr>
<tr>
<td>Vertical eye closure penalty(^{\ddagger,}) per lane</td>
<td>3.6</td>
<td>1.0</td>
<td>dB</td>
</tr>
<tr>
<td>Receive electrical 3 dB upper cutoff frequency, per lane (max)</td>
<td>3750</td>
<td>3750</td>
<td>MHz</td>
</tr>
</tbody>
</table>
# 10GBASE-L Tx Specifications

**Table 52–12—10GBASE-L transmit characteristics**

<table>
<thead>
<tr>
<th>Description</th>
<th>10GBASE-LW</th>
<th>10GBASE-LR</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signaling speed (nominal)</td>
<td>9.95328</td>
<td>10.3125</td>
<td>GBd</td>
</tr>
<tr>
<td>Signaling speed variation from nominal (max)</td>
<td>± 20</td>
<td>± 100</td>
<td>ppm</td>
</tr>
<tr>
<td>Center wavelength (range)</td>
<td>1260 to 1355</td>
<td></td>
<td>nm</td>
</tr>
<tr>
<td>Side Mode Suppression Ratio (min)</td>
<td></td>
<td>30</td>
<td>dB</td>
</tr>
<tr>
<td>Average launch power (max)</td>
<td>0.5</td>
<td></td>
<td>dBm</td>
</tr>
<tr>
<td>Average launch power(^a) (min)</td>
<td>-8.2</td>
<td></td>
<td>dBm</td>
</tr>
<tr>
<td>Launch power (min) in OMA minus TDP(^b)</td>
<td>-6.2</td>
<td></td>
<td>dBm</td>
</tr>
<tr>
<td>Optical Modulation Amplitude(^c) (min)</td>
<td>-5.2</td>
<td></td>
<td>dBm</td>
</tr>
<tr>
<td>Transmitter and dispersion penalty (max)</td>
<td>3.2</td>
<td></td>
<td>dB</td>
</tr>
<tr>
<td>Average launch power of OFF transmitter(^d) (max)</td>
<td>-30</td>
<td></td>
<td>dBm</td>
</tr>
<tr>
<td>Extinction ratio (min)</td>
<td>3.5</td>
<td></td>
<td>dB</td>
</tr>
<tr>
<td>RIN(_{12})OMA (max)</td>
<td>-128</td>
<td></td>
<td>dB/Hz</td>
</tr>
<tr>
<td>Optical Return Loss Tolerance (max)</td>
<td>12</td>
<td></td>
<td>dB</td>
</tr>
<tr>
<td>Transmitter Reflectance(^e) (max)</td>
<td>-12</td>
<td></td>
<td>dB</td>
</tr>
<tr>
<td>Transmitter eye mask definition {X1, X2, X3, Y1, Y2, Y3}</td>
<td>{0.25, 0.40, 0.45, 0.25, 0.28, 0.40}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# 10GBASE-L Rx Specifications

## Table 52–13—10GBASE-L receive characteristics

<table>
<thead>
<tr>
<th>Description</th>
<th>10GBASE-L</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signaling speed (nominal)</td>
<td>10.3125</td>
<td>GBD</td>
</tr>
<tr>
<td>10GBASE-LR</td>
<td>9.95328</td>
<td>GBD</td>
</tr>
<tr>
<td>Signaling speed variation from nominal (max)</td>
<td>± 100</td>
<td>ppm</td>
</tr>
<tr>
<td>Center wavelength (range)</td>
<td>1260 to 1355</td>
<td>nm</td>
</tr>
<tr>
<td>Average receive power(^a) (max)</td>
<td>0.5</td>
<td>dBm</td>
</tr>
<tr>
<td>Average receive power(^b) (min)</td>
<td>-14.4</td>
<td>dBm</td>
</tr>
<tr>
<td>Receiver sensitivity (max) in OMA(^c)</td>
<td>0.055 (-12.6)</td>
<td>mW (dBm)</td>
</tr>
<tr>
<td>Receiver Reflectance (max)</td>
<td>-12</td>
<td>dB</td>
</tr>
<tr>
<td>Stressed receiver sensitivity (max) in OMA(^d, e)</td>
<td>0.093 (-10.3)</td>
<td>mW (dBm)</td>
</tr>
<tr>
<td>Vertical eye closure penalty(^f) (min)</td>
<td>2.2</td>
<td>dB</td>
</tr>
<tr>
<td>Stressed eye jitter(^g) (min)</td>
<td>0.3</td>
<td>UI pk-pk</td>
</tr>
<tr>
<td>Receive electrical 3 dB upper cutoff frequency (max)</td>
<td>12.3</td>
<td>GHz</td>
</tr>
</tbody>
</table>
The Challenge: Putting Down the Fiber
Fiber Recommendations

Outside the building? Install SMF

- Consider higher grade fiber if:
  - Longer distances
  - Potential for upgrade to DWDM

Inside building

- Jumpers? Don’t care; buy with equipment
- Vertical and horizontal
  - Easy to re-pull? 2000 MHz*km MMF good to 10 Gig
  - Expensive to re-pull? SMF or Hybrid SMF/MMF
  - Still not sure? Safe bet is SMF
Infrastructure Issues

- Cost to build out fiber infrastructure high (CapEx)
  - Labor costs are not declining (greatest % in USA)
    - Installation technologies will evolve and optimize for specific solutions
      - *Micro Trenching*
      - *Blown Fiber*
  - Equipment makes up 25 to 33%
    - Equipment will rapidly drop in cost; increase in performance; will be replaced a much greater rate than traditional telephony
    - Infrastructure must not impede this advance
  - Fiber, enclosures, batteries, etc. unlikely to decline
  - Next infrastructure must be future-proof!
    - 100 Mbps → 1 Tbps → ???

- No reason to delay – no large decreases in sight
  - Sin to not be filling open ditches with conduit (if not fiber)
Distribution Costs

Distribution of Hub Capital

- Fiber to hub materials: 50%
- Fiber to hub labor: 0%
- Hub cabinet material: 7%
- Hub cabinet labor: 0%
- Hub splicing material: 8%
- Hub splicing labor: 2%
- Hub battery backup material: 8%
- Hub battery backup labor: 1%
- Hub electronics material: 19%

Labor: 8 + 2 + 19 = 29%
Distribution of Home Capital

- home splicing material: 0%
- home splicing labor: 3%
- home drop material: 4%
- home drop labor: 7%
- hub to home fiber material: 26%
- hub to home fiber labor: 7%
- home cabinet material: 5%
- home cabinet labor: 5%
- home battery backup material: 21%
- home battery backup labor: 0%
- home electronics material: 0%
- home electronics labor: 0%

Labor: 5 + 7 + 7 + 26 + 3 = 48%
Network Cost Modeling
Deployment Cost Distribution

Corning, June ‘02
Network Cost Modeling

Deployment Cost Distribution

Budgetary Pricing

Corning, June ‘02
Traditional Fiber Builds

Street Cutting

Excavation
Traditional Fiber Builds, con’t

Vault Placement

“Temporary” Restoration
Micro-Trench

- Up to 4 Cables per Cut
- Low Intensity Construction
- Non-Destructive Installation
- Rapid Deployment
- Improved Agency Acceptance
MTC Technique

Cutting the Micro-Trench
- Shallow Depth Trench
- Narrow Width Cut--10mm
- Fully protected in Hardened Space

Cleaning the Cut
- Power Washer Clean
- Air Pressure Dry
Hold Strip and Thermal Seal

- ½” Polyfoam Hold down Rod
- 7/16” EPDM Sponge Rubber Thermal Seal
MTC Technique (con’t)

Sealing the Cut

- Hot Bitumen Sealant
- Silica Grout Seal

- Low Impact to Traffic
- Installed quickly
- Flexible, Durable
Micro Trench Construction (MTC)

What is MTC?
- Shallow Depth Trench
- Fiber Payload Encased in Fully Protected, Hardened Space
- Can Deploy more than 1,000 feet per day per crew

Why MTC?
- Traditional “Carrier Class” Depth Cost Prohibitive to Address Last Mile Development
- Other Alternatives (Sewer/Gas lines) Too Complex for Wide Adoption
- Match Solution to Application
Illustrative Example of Build Costs

- Trenching
- Robotics
- Directional Boring
- MTC

MTC Less Than All Other Options
MTC Advantages

• Fastest Fiber Deployment/Delivery Method Available Permitting Through Construction
• Delivers Access and Point-to-Point Fiber Solutions
• Delivers Fiber At Wire Line Prices
• Minimal Disruption To Pedestrian and Traffic Flow
• Survivable and Diverse Entry Topologies
• Very Rapid Repair and/or Restoration
Blown Fiber Microconduit
A - The Concept

- The Fibreflow system itself consists of dedicated channels of micro-tubes enclosed in a protective jacket designed to suit a range of environments both indoors and outdoors.

- Fibre unit bundles are then blown down the tubes on demand.

- When your customers ask for a connection, small optical fibre units are blown into the micro-tubes, without the need to splice.

- Branching can be done anywhere along the route by cutting into the protective jacket and connecting the existing micro-tube to a branch micro-tube using a permanent or push/pull connector.

- The Fibreflow solution can provide fibre optic links all along the network on a “Just in time” basis

- Fibreflow can be laid: within existing telecommunications ducting, within other utilities connections, as direct bury or over head.

Emtelle, June 2002
Sales Generation

- Why gamble on Dark Fibre?
- Saleable capacity with no more street digs
- Innovative solution capable of winning new contracts
- Numerous order winning features and benefits
- Back up support to deliver cutting edge solutions
- Assists utilisation of unemployed fibre in legacy networks
- Access customer with greater ease
- Ease of response to changing customer demands
- Point to Point Fibre product offering
- Dedicated fibre path offering
- Fibre can be upgraded with minimum customer interruption

Emtelle, June 2002
Profit Generation

- Lower Network Costs
- Efficient use of Capital
- Reduced installation costs
- Reduced space required all along the network
- Reduced Access charges
- Reduced number of splices between POP and customer
- Elimination of Outdated Fibres in Existing Networks
- Elimination of Unused Fibres in New Networks
- Maintenance Costs Reduced
- Cheaper closures and Connectivity Products
- Reduced fibre costs in the short and long term

Emtelle, June 2002
A

Savings

Emtelle, June 2002
Trends and Influences
**Trends and Influences**

- **Towards Simplification**
  - Towards Higher Speed; Lower Cost vs. Moore’s Law
  - Ethernet to the Rescue in the Access Space
  - QoS and OAM Can Be and Must Be Solved
  - Economic Models Can Support “True Broadband Services”
  - Distractions or Complements
  - Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
  - Investment as a Positive Feedback System
The Pythagorean Paradigm

- The planets, sun, moon, and stars move in perfectly circular orbits;
- The speed of the planets, sun, moon, and stars in their circular orbits is perfectly uniform;
- The Earth is at the exact center of the motion of the celestial bodies.
Plato’s Homework Problem

Plato gave his students a major problem to work on. Their task was to find a geometric explanation for the apparent motion of the planets, especially the strange retrograde motion.

One key observation:
As a planet undergoes retrograde motion (drifts westward with respect to the stars), it becomes brighter.
Ptolemaic System
Problem Solved Mathematically
Ptolemy’s Epicycles
And then....

Portrait of Copernicus
Before 1584 AD - Tobias Stimmer
Network General Guide to Communication Protocols

OSI Layers

- Application 7
- Presentation 6
- Session 5
- Transport 4
- Network 3
- Logical Link 2
- Physical 1

Total Network Visibility™
Courtesy Network General Corporation
Complexity Resolved

Layer 2
Logical Link

Ethernet
Token Bus
Token Ring
FDDI
DQDB
ISDN
SONET
PPP
ATM
Frame Relay
SMDS
Modems
Complexity Resolved (again)

Layer 3

Network

IP

IPX

DLSw

COFP

CLNP

DRP

IDP

DDP

VIP
Convergence == Simplicity

Resolving Network Complexity
from the Bottom Up
Teenagers Set Up Networks for FUN

100 Mb/s Ethernet network set up for evening of gaming
In the future, network market segments will not be defined strictly by geography.
Ethernet ‘Trucks’
SONET ‘Ferry’

The Legacy Network
‘Bridges’ to the Future
Just a Bridge Too Far…
From Copper to Fiber

Copper

Free Space Optics

Wireless

Ethernet

FSAN

xDSL

APON

Fiber

Business Models
Customer Usage
Regulation
Competition
Economics
Technology
Applications
Culture

This chaos cannot be resolved by some central authority
The Interconnect Dilemma:

**Too Many Alternatives!**

Source: Intel, 2001
Trends and Influences

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must Be Solved
- Economic Models Can Support “True Broadband Services”
- Distractions or Complements
- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
- Investment as a Positive Feedback System
Towards Moore’s Law

At 10Gig – We are definitely pushing the limit of “low cost” technology doing full speed serial implementations

- Optical: Relatively easy for 100 & 1000
  - Borrowed 100 from FDDI
  - Borrowed 1000 from Fibre Channel
  - Created “our own” for 10 Gig
- Copper: Pushing the limit at 1000
- Test and measurement not keeping up

But – WDM will likely provide ability to meet or exceed requirements for several number of years
## 10 to 1 Gig Price-Performance

<table>
<thead>
<tr>
<th>Component</th>
<th>Ratio</th>
<th>Basis (Supercom '02)</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>40 : 1</td>
<td>10GBASE-LR : 1000BASE-LX</td>
</tr>
<tr>
<td>Optics</td>
<td>20 to 30 : 1</td>
<td>10GBASE-LR : 1000BASE-LX</td>
</tr>
<tr>
<td>SERDES</td>
<td>40 : 1</td>
<td>Single SERDES; (1 Gig Quad/Octal/Integrated SERDES much greater ratio)</td>
</tr>
<tr>
<td>NIC</td>
<td>10 : 1</td>
<td>10GBASE-LR : 1000BASE-LX</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000BASE-LX (seemed unreasonably high)</td>
</tr>
</tbody>
</table>

*Usual inflection point and objective for economic feasibility is 3 – 4 : 1 for a 10X speed upgrade*
10 GbE Price/Performance

- Ethernet Pricing Model

Dollars per Gigabit of Bandwidth

$6,000

$5,000

$4,000

$3,000

$2,000

$1,000

L3 modular 100 Mbps fiber

L3 modular 1000 Mbps fiber

10 GbE

2000 2001 2002 2003 2004

10GigE actual

Greg Collins, Dell’Oro 5/01
10 GbE Price/Performance

- SONET/SDH Pricing Model

Dollars Per Gigabit of Bandwidth

- OC-3 (155 Mbps)
- OC-12 (622 Mbps)
- OC-48 (2.5 Gbps)
- OC-192 (10 Gbps)

Greg Collins, Dell’Oro 5/01
1 Gig E Technology Directions

High Speed Serial
- Early: BiCMOS; BiPolar; GaAs
- Mature: CMOS

Optical
- Early: 850 nm CD Lasers; 1310 nm FP Lasers
- Mature: 850 nm VCSEL Lasers; 1310 nm FP Lasers (1310 nm VCSELs soon?)

Packages
- Early: OLM
- Mature: SFP; Integrated MAC/PHY/SERDES
10 Gig E Technology Directions

High Speed Serial
- Now: SiGe
- Future: CMOS (2003 - 2004?)

Optical
- Now: 850 nm VCSEL Lasers; 1310 & 1550 nm DFB Lasers
- Future: 850 & 1310 nm VCSEL Lasers; 1550 ?

Packages
- Now: XENPAK (XAUI); FTRX (300 pin)
- Future: {XXP; XPAK; XFP; SFP}?
1 & 10 Gig Availability vs. Standard

- HSSG FORMED
- Task Force Formed
- PAR Approved
- WG Ballot
- LMSC Ballot
- STD!

1999
2000
2001
2002

1999
2000
2001
2002

- Preliminary 1 Gig Optics Available
- Preliminary 10 Gig Optics Available
- 1 Gig Interoperability
- 10 Gig Interoperability

Draft 1
Draft 1.1
Draft 2
Draft 2.1
Draft 3
Draft 3.1
Draft 3.2
Draft 3.3
Draft 4
Draft 4.1
Draft 4.2
Draft 4.3
Draft 5

Preliminary 1 Gig Optics Available
Preliminary 10 Gig Optics Available
1 Gig Interoperability
10 Gig Interoperability
40 Gig Next?

Moore’s Law

Megabits per second

10BASE-T
Fast Ethernet
Gigabit Ethernet
10 Gigabit Ethernet


10 100 1000 10000 100000
100 Gig Next?


Megabits per second

100000

10000

1000

100

10


10 BASE-T

Fast Ethernet

Gigabit Ethernet

10 Gigabit Ethernet

Moore’s Law
Reasons for 40

- It would be a whole lot easier than 100
  - Not as technically aggressive as Moore’s law
  - 10 Gig was nearly torture; it would be nice to have a break

- We have multiple ways to do it
  - 4 lambdas at 10 Gig each with 802.3ad link aggregation
  - SONET Style OC-768

- Many SONET people believe Ethernet and SONET should walk together into the sunset….
Reasons for 100

- Economics limiting R&D investment
  - Current economy delaying uptake of 10
  - More time required for essential research

- Longer cycle (inter-speed) provides opportunity for cost reduction cycles
  - Reduces overlap in concurrent design projects
  - Improves ROI on principal technology investment

- Longer cycles encourage competition
Gigabit bandwidth is needed to balance Intel® Pentium® 4 Processor Performance.
10G Short Distance Interconnects

Inter-Chip
Chip-to-Chip/Card

Inter-Board
Blade-to-Blade
Intra-Cabinet

Inter-Cabinet
Rack-to-Rack
Box-to-Box

Inter-Facility, Enterprise,
Site-to-Site, VSR,
Data Center-to-Data Center

Copper Serial Bus

Copper Backplane

Optical Backplane

Standard Optical

Source: Intel
I/O Architecture Evolution

Signaling Rate GHz

8.33 MHz

UP TO 66 MHz

UP TO 800 MHz

1GHz Parallel Bus Limit

>12 GHz Copper Signaling Limits

80s 90s 00s

- ISA
- PCI
- VESA
- EISA
- MCA
- AGPx
- PCIx
- HL
- HT
- R I/O
- HT
- HL

Third Generation I/O Architecture

- Full Serial
- Point-to-point
- Max Bandwidth/Pin
- Scalable >10 GHz
- Flexibility
- Multiple market segment

Source: Intel
OK, But What Speed’s Next?

- Too early to tell
- Highly likely that IEEE 802.3 will wait until:
  - Recovery of the market
  - 10 Gig is available at better price-performance
    - Lessons from 10 GbE not yet known
  - Ethernet in the First Mile (802.3ah) is complete (or nearly complete)
    - EFM will drive demand for 10G and higher in the backbone and core
Trends and Influences

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must Be Solved
- Economic Models Can Support “True Broadband Services”
- Distractions or Complements
- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
- Investment as a Positive Feedback System
Ethernet in the Access Space

Did I mention Ethernet in the First Mile?

What were people thinking when they built out the WAN without EFM?
Backbone Glut or Access Dearth?
Emerging End-to-End Ethernet

Integrated Services Video – Voice – Data

New World Order

**Backbone**  
Continent-to-Continet  
Coast-to-Coast  
all over Fiber at 10 Gbps & up

**Metro**  
City-to-City – Town-to-Town  
all over Fiber at 1 Gbps → 10 Gbps

**Access**  
to the Optically Fibered World  
“First Mile / Last Mile”  
56 kbps → 1 Gbps

**LAN**  
Desktop-to-Desktop – Floor-to-Floor  
10 Mbps → 1 Gbps
Optical Ethernet Capabilities: Long Reach

- **10GbE (802.3ae)**
  - **1000BaseCX**
    - Copper: 25m, 65m, 220m, 275m, 300m, 500m, 550m, 5km, 10km, 40km
  - **1000BaseT (802.3ab)**
    - Copper: 25m, 65m, 275m, 300m, 500m, 550m, 5km, 10km, 40km

- **10GbE (802.3ae)**
  - **1000BaseSX**
    - 850nm: 62.5 MMF 160 MHz-km Modal BW
    - 50 or 62.5 MMF, 400 or 500 MHz-km Modal BW
  - **1000BaseLX**
    - 1300nm: 50 or 62.5 MMF, 400 or 500 MHz-km Modal BW

- **10GbE (802.3ae)**
  - **1000BaseLX**
    - 1300nm: 9 SMF
  - **1000BaseSX**
    - 850nm: 50 MMF 500 MHz-km Modal BW
    - 50 MMF 400 MHz-km Modal BW
  - **10GbE (802.3ae)**
    - Balanced Copper: 4 Pr Cat5 UTP

Source: Luke Maki, Boeing Corporation, 2002
Trends and Influences

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- **QoS and OAM Can Be and Must Be Solved**
- Economic Models Can Support “True Broadband Services”
- Distractions or Complements
- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
- Investment as a Positive Feedback System
HAVE YOU CONSIDERED THE SERVICES YOUR NOC WILL PROVIDE?

- Monitoring – Core & Distribution Networks
- Call Center
- Change Management
- Technical Team Support
- Knowledge Base
- Capacity Planning
- Security
- Customer Care
- Contingency Planning
- Asset Management / Control / Configuration
- Trouble-Ticketing
- Expedient Problem Escalation and Resolution
Other Key Management Issues

Service and Design
Release
Resolution
Supplier Management
Control Cost Savings and Containment
Security Management
Availability and Contingency Mgmt.
Service Level Management
Service Reporting
Capacity Management
Testing of New Technology
Design of Change and Release Timelines

Risk Assessment
Rollback and Contingency Mgmt.
Plans for Actual Release Roll-outs
Incident Management
Escalation Management
Problem Management
SLA (Service Level Agreement) Management
OLA (Operation Level Agreement) Management
Reporting on Actual Performance vs. Contract Terms
Asset and Configuration Mgmt.
Change Management
Monitor and Maintain Configuration Baselines

Source: Scott Alldridge – CEO, IP Services, June 2002 -- Reformatted
What Is an OSS/BSS
(Operational and Business Support System)

Tools that allow the System Operator to:

- Take an Order from a Customer
- Fulfill that Order for Services
- Bill the Customer for the Services
- Take Care of Complaints Through Customer Care
- Manage the Network to Provide Quality of Service
- Ensure the Network Can Meet the Future Needs as You Add Customers
Aren’t ATM/SONET/SDH Better Than Ethernet for QoS?

This is the **Wrong Question**

Get over it!

- Ethernet owns the ends
  - You can’t improve QoS with some other technology in the middle

- Right question
  - What do we need to do to have reliable and verifiable service level agreements?
    - 1. Inexpensive, high bandwidth pipes
    - 2. Service class management
    - 3. OAM&P
Optical Ethernet
Deficiencies and Mitigation

- Fault Protection/Restoration Times
- Providing QoS
- Performance Monitoring and Fault Management
- Scalable OA&M Capabilities

These five slides derived from: Luke Maki, Boeing Corporation, 2002
Optical Ethernet Deficiencies

- **Fault Protection/Restoration Times**
  - > 1 second (industry likes 50 ms)
  - Contributors to restoration time:
    - Original 802.1D Spanning Tree can take up to 50 seconds
    - Aggregate link failover ‘one second or less’ per 802.3ad

- **Mitigation**
  - Spanning Tree improvements via 802.1s and 802.1w, bringing convergence to 1 second
  - Actual aggregate link failover is being achieved in 100 ms or less
Optical Ethernet Deficiencies

- **Providing QoS**
  - Over-provisioning bandwidth (higher network cost)
  - CoS on aggregate traffic flows does not necessarily get applied where needed in the network
  - Spanning Tree does not distribute traffic on available capacity

- **Mitigation**
  - Low cost of Ethernet allows for over-provisioning
  - 802.1s will enable better utilization of links otherwise unused under 802.1D
**Optical Ethernet Deficiencies**

- **Performance Monitoring & Fault Management**
  - Gigabit Ethernet (and less) provide NO overhead for performance monitoring, alarms, etc.
  - SNMP monitoring can be ‘after the fact’

- **Mitigation**
  - The 10GbE WAN interface provides a limited set similar to SONET
  - The Ethernet First Mile Task Force is working proposals to mitigate the issues
Optical Ethernet Deficiencies

- Other OAM&P Capabilities
  - Single-ended maintenance
  - Loopback testing
  - Flow-through provisioning
  - Integrated operations support systems
  - Capacity planning and management
  - Service level agreements

- Mitigation
  - EFM working on Layer 2 “OAM” features
  - Provisioning / OSS / BSS not Ethernet
  - Expect solutions from 802.1 and IETF
Ethernet QoS & OAM Summary

- Ethernet does not prohibit QoS
  - Ethernet compliant equipment can (and does) support CoS, QoS, and provisioning
  - QoS is solved above the Ethernet MAC

- Ethernet EFM project’s OAM resolves issues with link diagnostics and management
  - But, only on a single link basis
  - IETF solution required for end-to-end diagnostics management (not 802.3’s job)
Trends and Influences

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must Be Solved
- **Economic Models Can Support “True Broadband Services”**
- Distractions or Complements
- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
- Investment as a Positive Feedback System
## Household Budget

### Bandwidth Budget

<table>
<thead>
<tr>
<th>Service</th>
<th>Budget Per Unit</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDTV Video per channel</td>
<td>20 Mbps</td>
<td></td>
</tr>
<tr>
<td>2 Channels of SDTV</td>
<td>8 Mbps</td>
<td></td>
</tr>
<tr>
<td>Web Surfing</td>
<td>10 Mbps</td>
<td></td>
</tr>
<tr>
<td>Games</td>
<td>2 Mbps</td>
<td></td>
</tr>
<tr>
<td>Phone</td>
<td>.064 Mbps</td>
<td></td>
</tr>
</tbody>
</table>

**HELP!** Total = 40.064 Mbps
## Coast-to-Coast DVD Movie Transfer*

<table>
<thead>
<tr>
<th>Method</th>
<th>Bit Rate</th>
<th>Minutes</th>
<th>Hours</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modem</td>
<td>56 kbps</td>
<td>13</td>
<td></td>
<td>13 days</td>
</tr>
<tr>
<td>Pony Express</td>
<td></td>
<td></td>
<td></td>
<td>11 days**</td>
</tr>
<tr>
<td>ISDN</td>
<td>128 kbps</td>
<td></td>
<td></td>
<td>5 ½ days</td>
</tr>
<tr>
<td>Cable Modem</td>
<td>1.5 Mbps</td>
<td>11 hrs 36 min</td>
<td>11 hrs 12 min</td>
<td></td>
</tr>
<tr>
<td>T-1</td>
<td>1.54 Mbps</td>
<td>10 hrs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSL</td>
<td>8.5 Mbps</td>
<td>2 hrs 12 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PON OC-12/32</td>
<td>19.4 Mbps</td>
<td>53.6 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35 mph</td>
<td></td>
<td>30 min***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Ethernet</td>
<td>100 Mbps</td>
<td>10.4 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gigabit Ethernet</td>
<td>1000 Mbps</td>
<td>1 min</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ‘The Matrix’ DVD 7.18 GB from New York, NY 10005 – delivered to Beverly Hills, CA 90210

** extrapolated from record: 7 days 17 hrs - approx 2,000 miles from St. Joseph, Missouri to Sacramento, California  Lincoln’s Inaugural Address, March 4, 1861

*** if you live close – no traffic – it’s in stock & there’s no line
Residential Revenue Opportunity

Voice @ $30 / month
Data @ $35 / month
Video @ $45 / month
Other up to $70 / month

$44 Monthly Margin

$66 per month to deliver services

Residential Subscriber

Cost of fiber plant $1,000 amortized* @ 20 years = $9 per month
Cost of electronics $1,500 amortized* @ 7 years = $25 per month
Cost of delivering content per subscriber = $32 per month

*Levelized cost at 8.5%

Total = $66 per month
Business Revenue Opportunity

10 Office Multi-Tenant Business

**Capital Costs**

- Cost of Lateral & Building Entrance – Fiber = $50,000
- Cost of Third-Party voice-switching equipment = $23,785
- Cost of Ethernet Access – 2 Gbps = $x,xxx

Total = $xx,xxx

---

Payoff

$6,450 Monthly Revenue

~1 year
Service Provider Summary

June 7, 2002

Gig-E FTTH, business & farm
Layer-2 transport

Open Access Philosophy

- 12 Internet Service Providers (ISPs)
- 2 Video Service Providers (VSPs)
- 1 Telephone Service Provider
- 1 Security Service Provider
Construction Summary

June 7, 2002

- 7,110 Meters Passed (to-date)
- 6,436 Homes Passed
- 2,289 Customers Lit
- ~30-50 new customers per week

35% overall
Economic Development

June 7, 2002

- 24 New Business Employees
  - 5 new high-tech businesses
- 17 NOC Employees
- 28 other PUD Support
- 25 contract labor (3-5 yr)
- 2 NCESD, K20

96 new jobs!

>$16M Economic Benefit

Fiber optic cable to every home in Grant County!
Lessons Learned

- Build it “once” to every home/business
- Supervision of contract labor
- Multi-vendor interoperability

- Economic catalyst to avoid chicken & egg
  - Video IP Head-end, Telephone IP Gateway
Summary

- Grant County PUD FTTH Project
  - will influence community change by:
    - Removing the access bottleneck
    - Eliminating the impact of distance
    - Removing the barrier to entry
    - Open Access, non-discriminating pricing

- *Digital imagination without limits*
**Trends and Influences**

- Towards Simplification
- **Towards Higher Speed; Lower Cost vs. Moore’s Law**
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must Be Solved
- Economic Models Can Support “True Broadband Services”

**Distractions or Complements**

- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development
- Investment as a Positive Feedback System
Distractions or Complements?

EtherEveryThings
- Chip-to-Chip Communication?
- 60-90 GHz Pt-to-Pt Radio?
- Ethernet Disk Drives?
- Subspace?

EtherKin
- 802.11 -- “Wireless Ethernet?”
- 802.17 – “Ethernet Loops” (RPR)

Other
- Infiniband (NGIO -> Infiniband -> 3GIO -> ?)
- Fibre Channel vs. iSCSI
- Digital Wrappers
- MPLS; VPLS;
Trends and Influences

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must be Solved
- Economic Models Can Support “True Broadband Services”
- Distractions or Complements
  
  **Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development**

- Investment as a Positive Feedback System
Regulatory Impact…

The single, most profound influence on the future of networking will be the acceptance and adoption of the “OPEN ACCESS MODEL” (or NOT)

Jonathan Thatcher; 2/2/2000 :-}
“Behind the existing rules, however, were two unwritten principles.

- First, by separating industries through regulation, government provided a balance of power in which each industry could be set against one another in order for elected figures to raise money from the different camps that sought advantageous regulation.

- Second, by protecting monopolies, the Commission could essentially guarantee that no communications businesses would fail. Repealing these implicit rules was a far less facile affair than promoting competition.”

 Former Chairman of the FCC
Customer as Hostage

Single Use Infrastructure
Who’s Monopoly Is It Anyway?

- Water
- Roads
- Sewers
New Paradigm
Community Access Network

Users

Service Providers

Management

Concentration

Access Portal

World Wide Packets

ESPN

CNN
Customer Choice

Shared Infrastructure

Community Access Network Operations Center

Satellite 1 → Satellite 2 → Satellite 3

Cable 1 → Cable 2 → Cable 3 → Cable 4

Wireless 1 → Wireless 2

ISP 1 → ISP 2 → ISP 3 → ISP 4

CLEC 1 → CLEC 2 → CLEC 3 → CLEC 4 → ILEC
Clash of Paradigms

The Public Network at Bay

<table>
<thead>
<tr>
<th>20th Century</th>
<th>21st Century</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit switched</td>
<td>Packet switched</td>
</tr>
<tr>
<td>Centralized</td>
<td>Decentralized</td>
</tr>
<tr>
<td>Voice driven</td>
<td>Data driven</td>
</tr>
<tr>
<td>Value in metering use</td>
<td>Value apps and services</td>
</tr>
<tr>
<td>Deterministic</td>
<td>Evolutionary</td>
</tr>
<tr>
<td>Monopoly</td>
<td>Competitive</td>
</tr>
</tbody>
</table>

Source: Center for Internet Studies, 8/8/2002, Rex Hughes
**Trends and Influences**

- Towards Simplification
- Towards Higher Speed; Lower Cost vs. Moore’s Law
- Ethernet to the Rescue in the Access Space
- QoS and OAM Can Be and Must Be Solved
- Economic Models Can Support “True Broadband Services”
- Distractions or Complements
- Federal Regulation and Policy Will Be the Single Greatest Influence on Technology Development

- **Investment as a Positive Feedback System**
Bandwidth Driving Revenue

Services Driving Bandwidth

The Future?

HDTV (2) 40 Mbps

HDTV (1) 20 Mbps

SDTV (2) 10 Mbps

SDTV (1) 5 Mbps

Gaming 2 Mbps

Web surfing 56 Kbps
Voice - 6.5 Kbps

Dial-Up 56 Kbps
ISDN 128 Kbps
Cable Modem 1.5 Mbps
ADSL 8.5 Mbps
PON 19.5 Mbps
Gigabit Ethernet 1,000 Mbps

not drawn to scale...
## Valuation – 5 Years Out

<table>
<thead>
<tr>
<th></th>
<th>VDSL</th>
<th>Ethernet</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Customers</td>
<td>100K</td>
<td>100K</td>
</tr>
<tr>
<td>Services</td>
<td>V-V-D</td>
<td>V-V-D</td>
</tr>
<tr>
<td>Install Cost</td>
<td>x</td>
<td>1.2x</td>
</tr>
<tr>
<td>Operational Costs</td>
<td>x</td>
<td>.5x</td>
</tr>
<tr>
<td>Equipment Life</td>
<td>2-3 years</td>
<td>&gt; 5-10 years</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>&lt; 10 Mbps?</td>
<td>1.0 Gbps</td>
</tr>
<tr>
<td>Infrastructure Life</td>
<td>&lt; 3 years?</td>
<td>&gt;30 years</td>
</tr>
</tbody>
</table>
10 Gigabit Ethernet Forecast

Port Shipments (000s)

Source: Dell'Oro Group (5/02)
Worldwide Ethernet Switch Market
10 Gigabit Ethernet Forecast

Manufacturer Revenue ($M)

Source: Dell'Oro Group (5/02)
Worldwide Ethernet Switch Market
10 Gigabit Ethernet Forecast

Manufacturer ASP ($)

Source: Dell'Oro Group (5/02)
Worldwide Ethernet Switch Market
## 10 Gigabit Ethernet Forecast

<table>
<thead>
<tr>
<th></th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manufacturer Revenue ($M)</strong></td>
<td>250</td>
<td>525</td>
<td>1058.8</td>
<td>2290.8</td>
<td>3537.2</td>
</tr>
<tr>
<td><strong>Port Shipments (000s)</strong></td>
<td>10</td>
<td>30</td>
<td>110</td>
<td>340</td>
<td>750</td>
</tr>
<tr>
<td><strong>Manufacturer ASP ($)</strong></td>
<td>25,000</td>
<td>17,500</td>
<td>9,625</td>
<td>6,738</td>
<td>4,716</td>
</tr>
</tbody>
</table>

Source: Dell’Oro Group (5/02)  
Worldwide Ethernet Switch Market
10 Gig Ethernet Externalities

“It’s the Economy Stupid”

- Drivers
  - Ethernet in the First Mile (2003-2004?)
  - Upgrades to Gigabit Enterprise Gear (?)
  - Ethernet over All Optical Networks (?)

- Volume / Price Tail Chasing

- Mainstream Technologies

- Graphical & Video Applications
Related Organizations & Technologies
802.3ah & EFMA Roles

**802.3ah**
- An IEEE task force
- Create the EFM standard
- Address four areas
  - OAM
  - Fiber Point-to-Point
  - Fiber PON
  - Copper

**EFMA**
- An industry alliance
- Support the standards process with resources
- Market the technology
- Host interoperability events
- Proven concept
EFMA Goals

Marketing Goals
- Create industry awareness, acceptance, and advancement of the Ethernet in the First Mile standard and products
- Provide resources to establish and show multi-vendor interoperability through coordinated events

Technical Goals
- Support the Ethernet in the First Mile standards effort conducted in the IEEE 802.3ah Task Force
- Contribute technical resources to facilitate convergence and consensus on technical specifications
Marketing & Technical

Marketing
- Promotion Material
  - First Whitepaper is out
- Speakers Bureau
  - Delivering the message
- Participate in Events
  - Panels & info booths

Technical
- Technical Meetings
  - First two conducted
- EFM Tutorials
  - Broaden understanding
- Inter-op Events
  - Prove products interwork
10 GEA Mission

- Promote industry awareness, acceptance, and advancement of technology and products based on the emerging 10 Gigabit Ethernet standard
- Accelerate industry adoption by driving technical consensus and providing technical contributions to the IEEE 802.3ae Task Force
- Provide resources to establish and demonstrate multi-vendor interoperability of 10 Gigabit Ethernet products
What Is OIF?

- Launched in April of 1998
- Open forum: 320+ members including many of the world’s leading carriers & vendors
- The only industry group bringing together professionals from the packet & circuit worlds

Mission: To foster the development and deployment of interoperable products and services for data switching and routing using optical networking technologies
OIF and Standards Bodies

OIF submissions perform two functions:

- Request standardization of specific OIF recommendations
- Provide informational documents to the target standards group

Established Liaisons With:

- ANSI T1
- IETF
- ATM Forum
- IEEE 802.3ae 10 Gbit Ethernet
- NPF
OIF Technical Committee

- Architecture
  - Services, network requirements, & architectures
- Carrier
  - Requirements and applications
- Signaling
  - Protocols for automatic setup of lightpaths
- OAM&P - Operations, Administration, Maintenance & Provisioning
  - Network management
- Interoperability
  - Interoperability testing
- Physical & Link Layer
  - Equipment & subsystem module interfaces
OIF Implementation Agreements

- SPI-3: System Packet Interface Level 3:
- SPI-4 phase 1: System Physical Interface Level 4
- SPI-4 phase 2: System Packet Interface Level 4
- SFI-4: SERDES/Framer Electrical Interface: Common electrical interface between framers and serializer/deserializer parts for STS-192/STM-64 interfaces
- Very Short Reach (VSR) OC-192 Interface based on 12 fiber Parallel Optics
- Serial OC192 1310 nm Very Short Reach (VSR) Interfaces
- Very Short Reach (VSR) OC-192 Interface based on 4 fiber Parallel Optics
- Serial OC192 850 nm Very Short Reach (VSR) Interfaces
- Etc.
Parallel Optics-Based VSR Interface

10G Framer 10G Optics 10G Optics 10G Framer

16:12 CONVERTER 16:12 CONVERTER

Low speed parallel links

1.24Gbps Optics

Parallel Ribbon Cable (<400m) 62.5um MM

1.24Gbps Optics

VCSEL - Vertical Cavity Surface Emitting Laser (850nm wavelength)
OIF Summary

- Brings together professionals from the data and circuit worlds
- Addressing key issues important to carriers and vendors - carrier group established
- Eight technical documents ratified as implementation agreements
- Optical module interface standards will allow industry to gain needed economies of scale
- Future work expected (NNI) Network-to-Network Interface and richer functionality UNI 2.0
Fibre Channel and SAN 10GFC

- ANSI T11 & IEEE continue to share
  - Ethernet borrowed 1 Gb from FC
  - Fibre Channel 10G borrowed from Ethernet
- One common wire and XCVR technology to leverage economy of scale and one cable plant technology – user runs one type of cable for SAN & LAN
  - Exception is that FC identifies a potentially more “Core SAN” cost effective option of 4-lane short wave optics (4 X 2.5) for 10G SAN solutions before 2004-7
    - 850 nm version of the 10GBASE-LX4
    - Potential issue for iSCSI
Fibre Channel and SAN 10GFC

- Key issue for 10G SAN - regardless of IO technology - is timing of cost effective 10G Optical XCVRS
  - Will 4-lane 10G optics be more cost effective than 1-lane 850nm 10G optics?
    - SAN can not withstand expensive XVRS
  - Meanwhile, 2G optics rule SAN while 4G copper enters in-box, loop application

- 4Gb FC is non-fabric, copper only, mostly CMOS, non-”SAN”, in-the-box disc storage “loop” migration and does not address same usage as 1, 2, & 10 Gb FC out-of-box SAN “fabric”
Fibre Channel Speed

*FC800 for intrabox applications, i.e., disk drive (Copper)
Will 10Gb Be “SANable”? When?

Core SAN Market

Adoption

Length

SAN “Sweet Spot”

SAN Sweet Spot
for any IO Technology

- <300 meters
  mostly <100 meters
- <$500/GBIC (max!)
  mostly <$100 (10G will bear some premium, but not much)
- For 2004, translates to:
  4-lane optics (FC only)
  and/or
  850nm shortwave
SAN Optical Transceiver Migration

Integrated 4-lane CMOS 10G Copper (FC and IB Only), 2002 <$20/port
4G FC Disks 1-lane CMOS Copper, 2003 <$10/port – No plans for 4G Optical xvr!
Resistance Is Futile

Ethernet Everywhere!

Shared Media
ADSL Rules
Limit Broadband
No VoIP
Save SONET/SDH
Modems Forever
Preserve our Copper