40 GbE Transport

IEEE HSSG Meeting, San Francisco — July 16-19, 2007

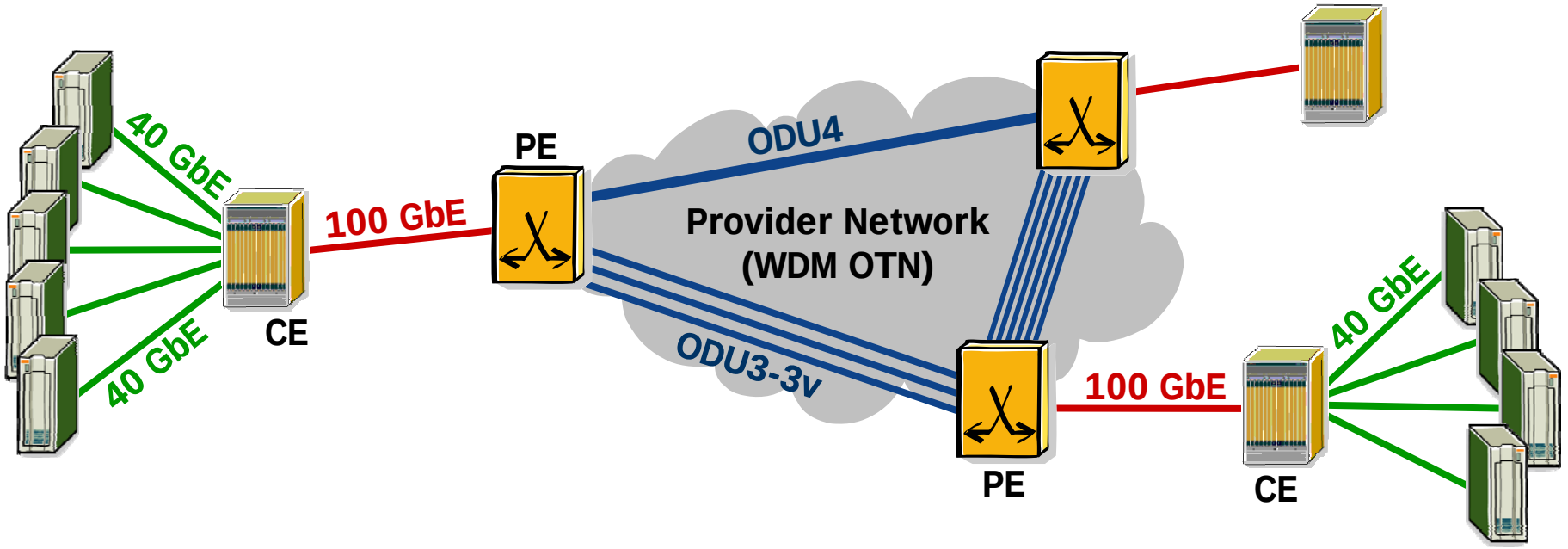
Marcus Duelk
Data Optical Networks Research, Bell Labs / Alcatel-Lucent
duelk@alcatel-lucent.com

Stephen J. Trowbridge
Alcatel-Lucent

Supporters

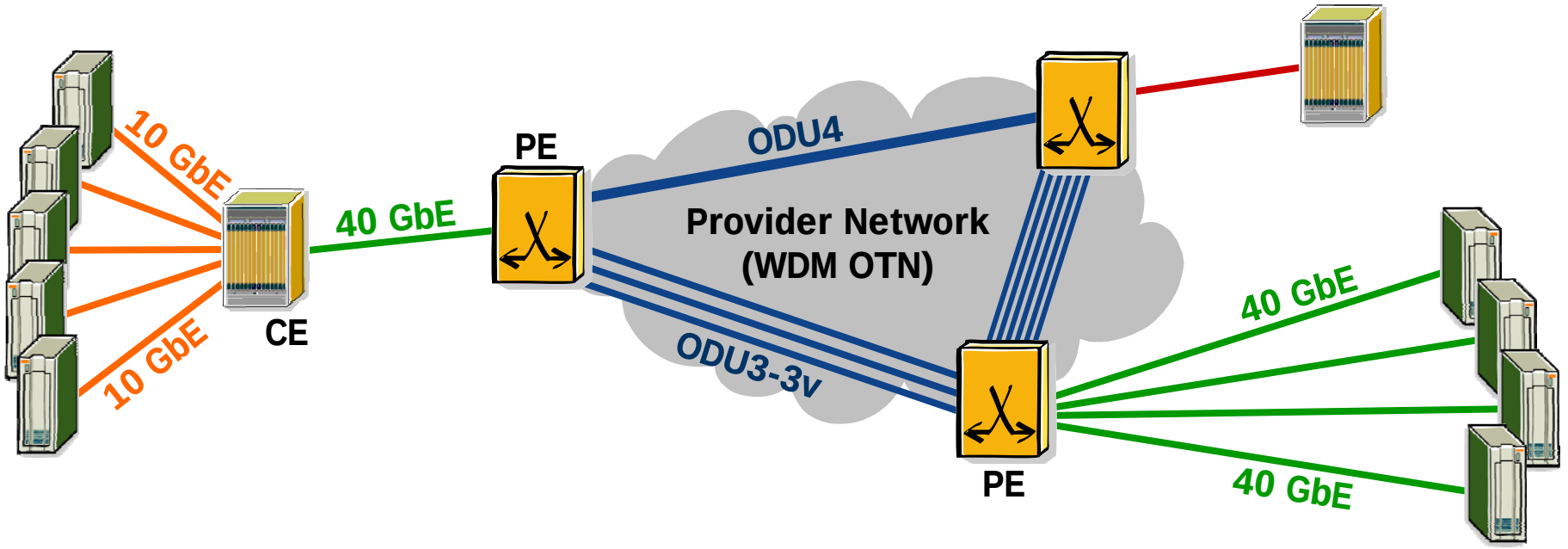
- Pete Anslow, Nortel Networks
- Thomas Fischer, Nokia Siemens Networks
- Ted Woodward, Telcordia
- Matt Traverso, Opnext
- Yann Loussouarn, France Telecom Orange
- Ghani Abbas, Ericsson
- Ralf-Peter Braun, T-Systems
- Med Belhadj, Cortina Systems
- Frank Chang, Vitesse
- Martin Carroll, Verizon
- Neil Peers, ADVA Optical Networking
- Keith Conroy, AMCC

Current View of Future 40 GbE Application Space



- 40 GbE will be only used inside blade servers, or
- 40 GbE will be only used as a server-switch interface, where 'switch' is considered to be customer equipment (CE)
- 100 GbE will be used as uplink switch-switch interface, where one of the switches is provider equipment (PE)
  → network operators will never see a 40 GbE client interface
- Backhauling of 100 GbE clients only

Possible Usage of Future 40 GbE



- 40 GbE server interfaces hooked up to PE switch
- 40 GbE uplink from CE switch to PE switch
  → 40 GbE used as an aggregation interface of 10 GbE (or 1 GbE) access IFs
- 40 GbE uplinks from access network (e.g., DSLAMs)
- Backhauling of 40 GbE clients required

Ethernet Backhauling via OTN

- Optical Transport Network (OTN, ITU-T G.709) is today's standard in WDM transport networks
- Two options for backhauling of client signals:

1. Frame-based mapping

→ remove interpacket gap + preamble + SFD,
map only Ethernet frames via GFP-F (ITU-T G.7041/Y.1303)

→ rate adaptation possible (only information rate is transported),
for example IFS stretch (see ishida_01_0507.pdf)

2. Bit-transparent mapping

→ map Ethernet client signal as CBR signal into payload (every single bit)

→ Ethernet line rate (incl. PCS) has to fit into payload container

Why Customers ask for Bit-Transparent Backhauling ?

1. Preamble or IFG have been used for carrying proprietary non-standard OAM or other type of information
2. L2 encrypted signals are used instead of standard Ethernet signaling, not allowing for frame-based mapping into transport containers
3. "Don't touch my bits" – 802.3 is a LAN technology in which it is not clearly defined which bits belong to a client or to a server and which of them need to be carried across a 3rd party (provider) network
4. "Synchronous Ethernet"
5. Error monitoring of client signals, e.g., remote defect indication

In a perfect world there probably wouldn't be a requirement for bit-transparent backhauling and frame-based mapping would be sufficient !

Overclocked OTU

ITU-T has recently (11/2006) supplemented transport of 10GBase-R LAN PHY in optical transport networks (OTN) → G.Sup43

- Supplement 43 is not a normative standard, it just describes overclocked mapping
- Overclocked 10G OTU (OTU2e) signal must be clocked at a nominal bit rate of 11.0957 Gb/s, as opposed to the standard OTU2 nominal bit rate of 10.7092 Gb/s
- The OTU2e signal is formed by wrapping a signal with the clock tolerance of the underlying Ethernet signal (±100 ppm) rather than that of a standard OTU2 signal (±20 ppm), standard methods for control of jitter and wander according to ITU-T G.8251 do not apply
- OTU2e is not networkable in the ODU layer network
- OTU2e breaks the OTN multiplexing hierarchy, i.e. mapping of 10G ODU2e into overclocked 40G ODU3e can not be accomplished using the same multiplexing mechanisms as mapping ODU2 into ODU3
- Overclocked OTUs are usable only for isolated point-to-point links
- Overclocked OTUs not compatible with existing standard-compliant OTN equipment
- Multiple, incompatible approaches to 10G Ethernet transport over ODU2-like signals (frame based GFP-F, 10GBase-W via STM-64, two flavors of overclocking, AMCC, Intel mappings, etc.)

OTU overclocking is not an acceptable networking solution !

Difference between 40 GbE and 100 GbE (Provider Perspective)

40 Gigabit Ethernet

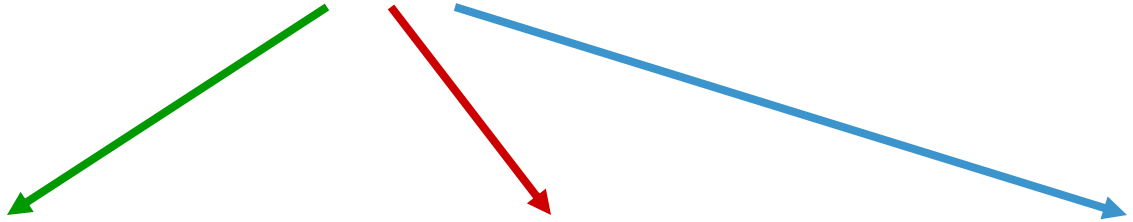
- 40G transport networks have been on the market for six years now
- 40G (OTU3) transport equipment is shipping today in growing volumes
- By 2010 there will be large installed OTU3 WDM infrastructure in most provider networks
- Bit-transparent backhauling of 40 GbE clients will be required with existing OPU3 payload rates !

100 Gigabit Ethernet

- No existing 100G transport network
- ITU-T SG15 decided in June 2007 to extend G.709 to the next higher rate
- ITU-T SG15 is monitoring the HSSG/TF activities and will define an ODU4 rate that is sufficiently large for bit-transparent backhauling of 100 GbE

Options for Defining OTN-Compatible 40 Gigabit Ethernet

- Existing 40G OTN payload rate (OPU3) = 40.150519 Gb/s
- Clock tolerance for OTN = ±20 ppm
  → min. OPU3 rate = 40.149716 Gb/s



Option 1: Reuse 64B/66B PCS but use a lower MAC/PLS rate such that the PCS-encoded 40GbE signal (incl. +100ppm clock tolerance) fits into min. OPU3 container

→ max. 40G MAC/PLS rate = 38.929165 Gb/s

Option 2: Use a MAC/PLS rate of 40.000 Gb/s in combination with a different PCS such that the PCS-encoded 40GbE signal (incl. +100ppm clock tolerance) fits into min. OPU3 container

→ max. PCS coding overhead = 0.36425%

Option 3: Use a lower MAC/PLS rate and a different PCS such that the PCS-encoded 40GbE signal (incl. +100ppm clock tolerance) fits into min. OPU3 container

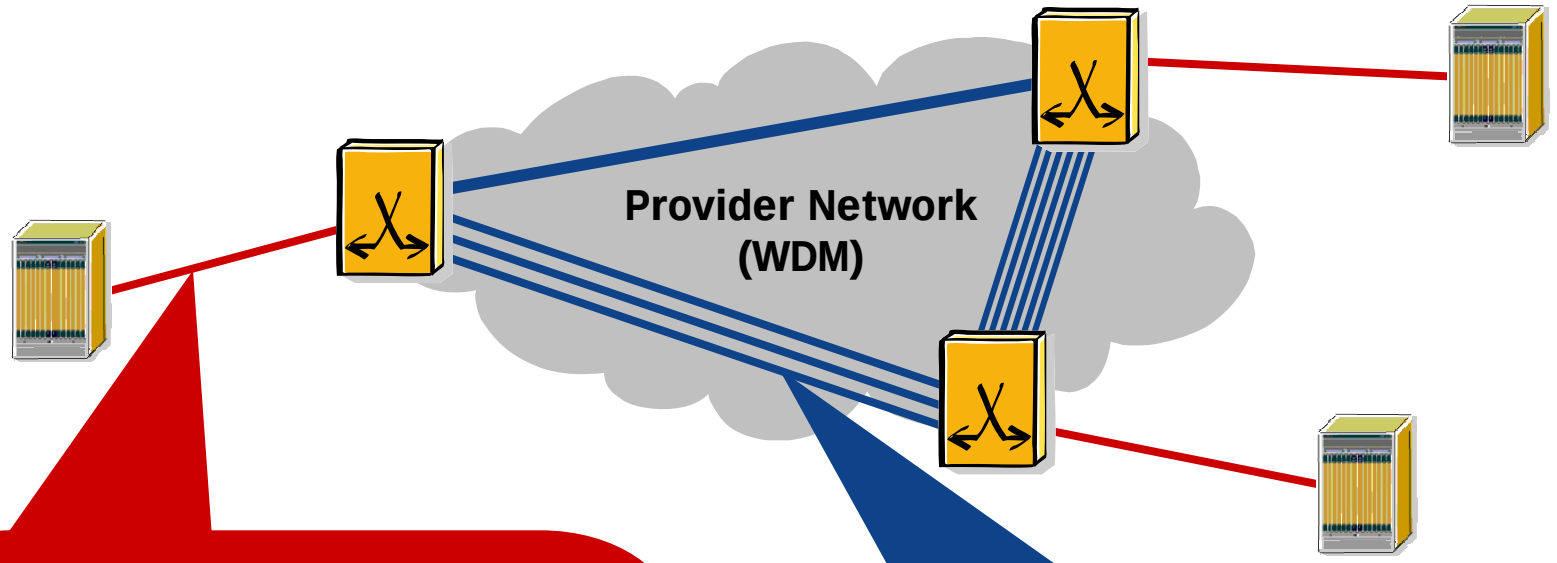
PCS Options for OTN Compatibility

Block Code	Block Length	Overhead	Comments
64B/66B	8 bytes	3.125%	Used in 10GBase-R, suitable PCS code for MAC/PLS rate ≤ 38.929 Gb/s
64B/65B	8 bytes	1.5625%	
128B/129B	16 bytes	0.78125%	
256B/257B	32 bytes	0.390625%	
512B/513B	64 bytes	0.1953125%	Suitable PCS code for MAC/PLS rate = 40.000 Gb/s

see trowbridge_01_0707.pdf for more details !

Benefits of Making 40GbE OTN-Compatible

Mapping 40GbE Client Interfaces into 40G Transport Interfaces



Client Interface

(UNI = User-Network Interface)

- → 40G Ethernet (IEEE)
- → Parallel PHY = multiple wavelengths (e.g., 4x10G) or multiple fibers (e.g., OM3 MMF)
- → short reach (100m to a few km)
- → single-channel point-to-point links
- → cost dominated by CAPEX

Transport Interface

(NNI = Network-to-Network Interface)

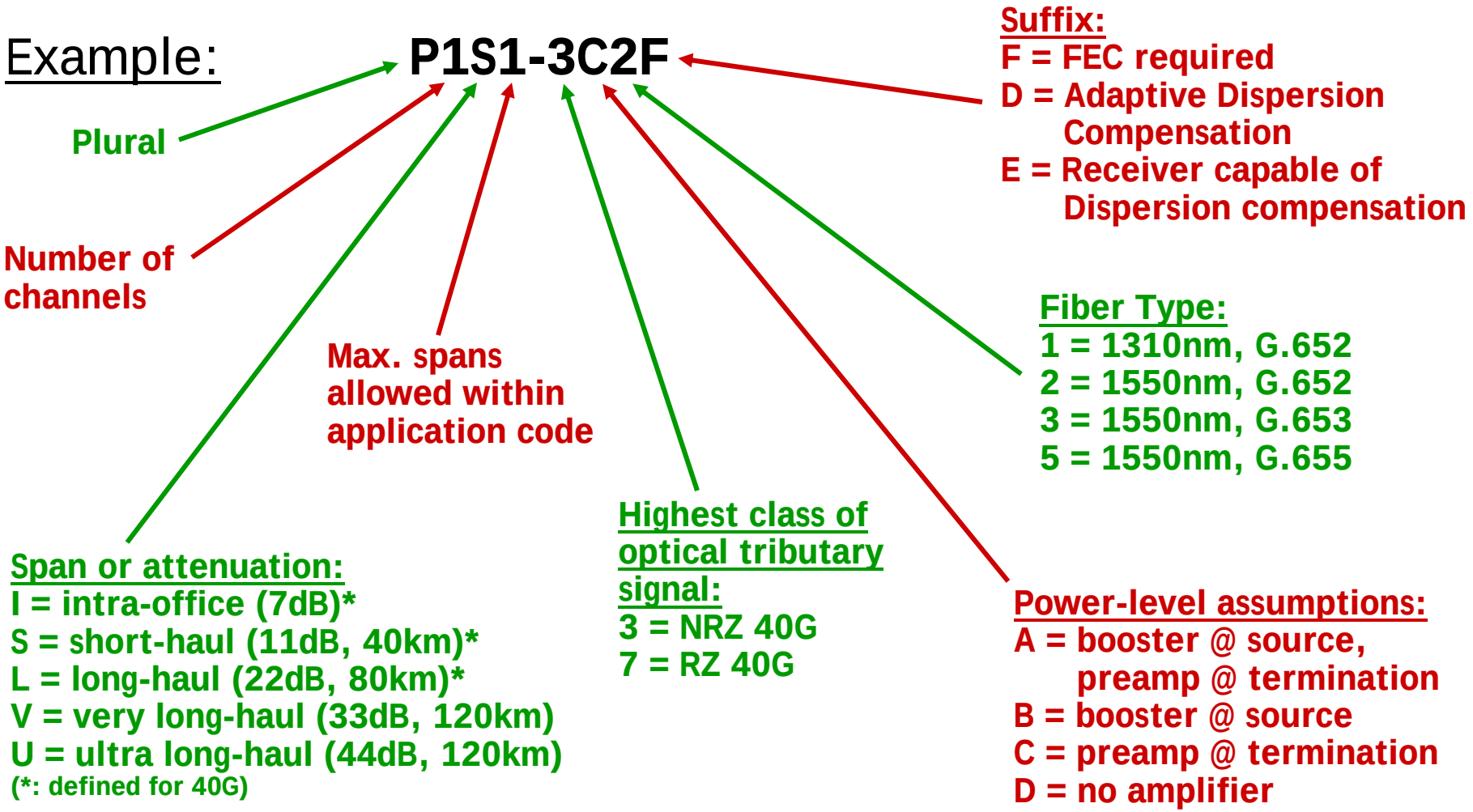
- → OTU3 (ITU-T)
- → Serial PHY = single wavelength and single fiber
- → long reach (100-3,000km)
- → multi-channel WDM systems
- → cost dominated by total capacity per link and by OPEX

Which PMDs for 40 GbE do you get "for free" (no new standards required!) if it fits standard ODU3?

Rec.	Title	Content
G.709	Interfaces for the Optical Transport Network	Frame Format for single or multi-channel systems with optional FEC
G.693	Optical interfaces for intra-office systems	Physical layer specifications for OTU3 signals, VSR interfaces for 40G (single-channel client interface)
G.694.1	Spectral grids for WDM applications: DWDM frequency grid	
G.696.1	Longitudinally compatible intra-domain DWDM applications	Multiple amplified spans of various reaches
G.698.2	Amplified multichannel DWDM applications with single channel optical interfaces (Next version will include 40G interfaces)	Metro networking with ROADMs and OAs
G.959.1	Optical transport network physical layer interfaces (2008 version will include new application codes for 40G)**	Single and multi-channel physical interfaces for various single-span reaches

**** For low-cost 40G applications using EMLs without the need for optical dispersion compensation, e.g., P1I1-3D1 (10-km reach), 1S1-3D1F (20-km reach with FEC) or 1L1-3C1F (40-km reach with FEC)**

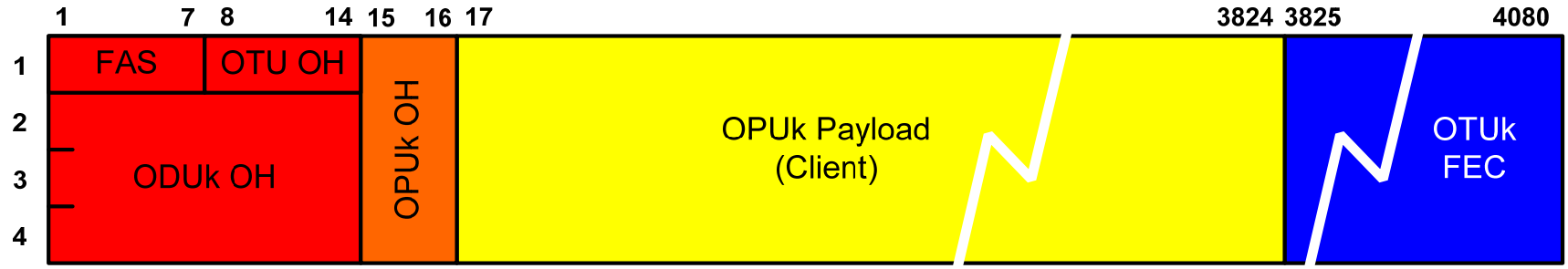
G.959.1 Application Codes



Example: P1S1-3C2F is single-channel 40G NRZ at 1550nm, 11dB span loss, target reach 40km over G.652 fiber, preamp at termination, w/ FEC

Alcatel·Lucent

G.709 OTU3 Frame Format



OTU = OCh Transport Unit
ODU = OCh Data Unit
OPU = OCh Payload Unit
(OCh = Optical Channel)

- Optional use of forward error correction (FEC) → RS(255/239)
  - Fixed all-0s stuff bytes are to be used when no FEC is used
- OTU-3 line rate = 43.018413559 Gb/s ± 20 ppm
- OPU-3 payload rate = 40.150519322 Gb/s ± 20 ppm
  - Min. OPU3 rate = 40.149716312 Gb/s
  - Max. OPU3 rate = 40.151322332 Gb/s

Summary

- Compatibility of 40 GbE with existing already-installed and standard-compliant transport networks is important, especially by 2009 or beyond
- Compatibility is not necessarily a burden but will offer many advantages
  - Many existing 40G transport PMDs
  - More 40G low-cost transport solutions will emerge
  - Compatibility helps reducing the cost of networks
- True compatibility means that the PCS-encoded 40 GbE line rate (incl. +100 ppm clock tolerance) will fit into existing OPU3 payload rate (incl. -20 ppm clock tolerance)
  - Lower MAC rate (e.g., 38.000 Gb/s) with 64B/66B PCS
  - 40.000 Gb/s MAC/PLS rate with other PCS, e.g., 512B/513B
  - Will allow for bit-transparent mapping (required by certain customers)
- Only one 40G MAC/PLS rate is acceptable
  - No repetition of LAN PHY/WAN PHY disaster
  - Only one MAC rate would ensure true transport compatibility
  - No one wants two different MAC rates

Appendix – More on OTN

OTN Networking

Early visions of "All-Optical" Transport (mid/late 1990's)

- Any signal (analog/digital), from anywhere to anywhere (transparency), anytime

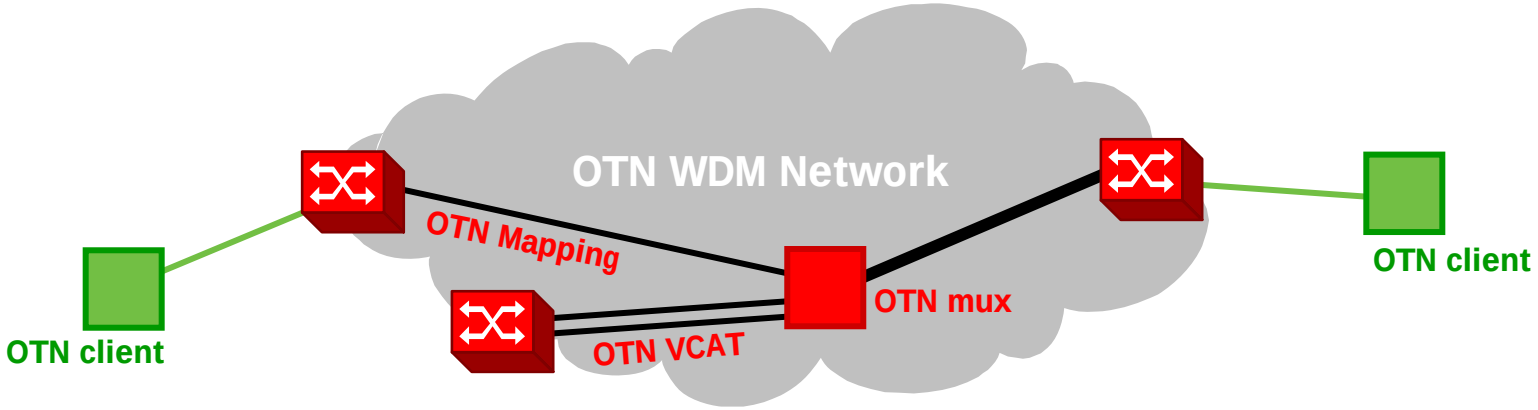
The reality of Optical Transport

- Some OEO conversion is needed to mitigate noise, and to maintain network
- Use SDH/SONET transponders in WDM networks

What is Optical Transport in the light of OTN ?

- Define new granularity of transport/switching containers → Optical Channel (OCh)
- Manage frequency-slots (OCh's) instead of time-slots (e.g., STS-1's)

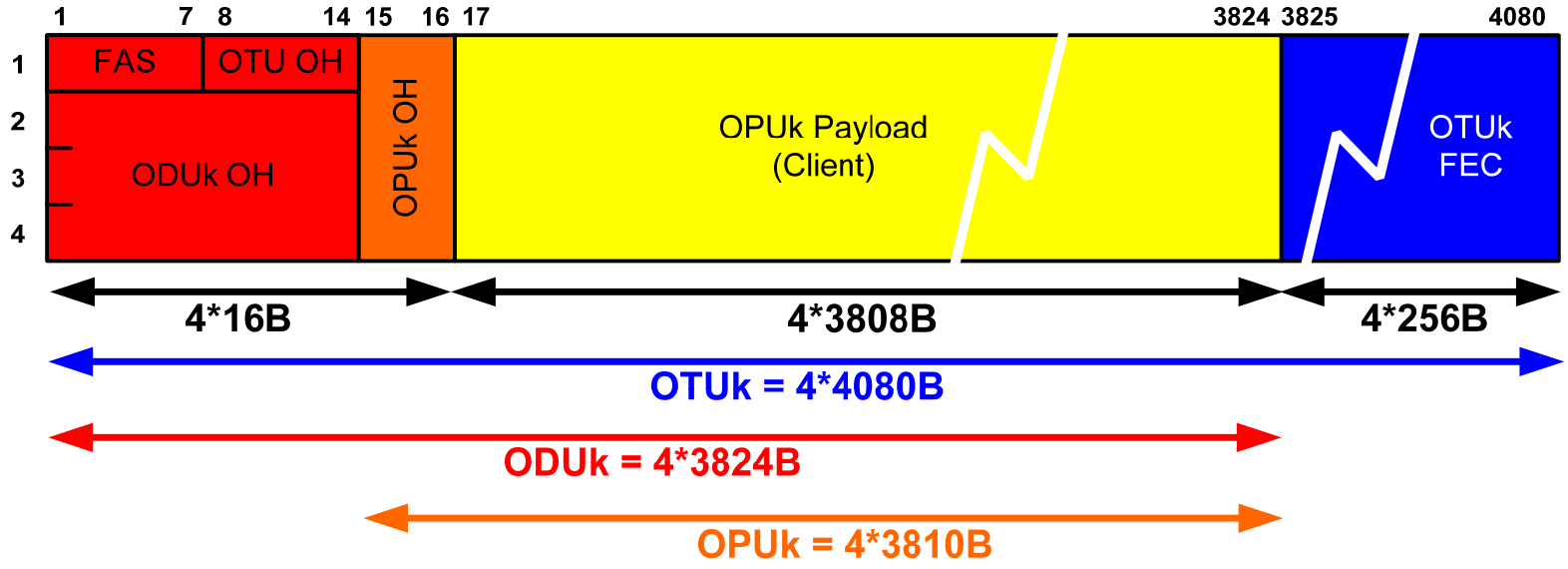
OTN Standards



Optical Transport Network (OTN) ITU-T standards

- architecture and network interfaces (G.872)
- multiplexing hierarchy and mapping, frames and formats (G.709)
- functional blocks (G.798)
- linear protection switching (G.873)
- management aspects (G.874)
- jitter & wander (G.8251)
- http://www.itu.int/ITU-T/2001-2004/com15/otn/transport.html

OTN Framing (G.709)



- OTU = OCh Transport Unit
- ODU = OCh Data Unit
- OPU = OCh Payload Unit

FAS = Frame Alignment Signal

- Frame length is constant, independent of line rate → no synchronous network !
- The use of FEC is optional (for OTU1, OTU2, OTU3)

OTN Rates

OTUk	OTUk rate (line rate)	ODUk rates	OPUk rates	Payload rates	OTUk Frame rate
OTU1	2.666 Gb/s	2.498 Gb/s	2.490 Gb/s	2.488 Gb/s	20.4 kHz
OTU2	10.709 Gb/s	10.037 Gb/s	10.003 Gb/s	9.995 Gb/s	82.0 kHz
OTU2e overclk**	11.096 Gb/s	10.400 Gb/s	10.361 Gb/s	10.3125 Gb/s	
OTU3	43.018 Gb/s	40.319 Gb/s	40.171 Gb/s	40.151 Gb/s	329.5 kHz

FEC overhead = 6.27% → FEC = RS(255,239), enhanced/ultra FEC not defined !

ODUk portion = OTU frame w/o FEC trailer = 93.73%

OPUk portion = Optical Channel payload frame = 93.38%

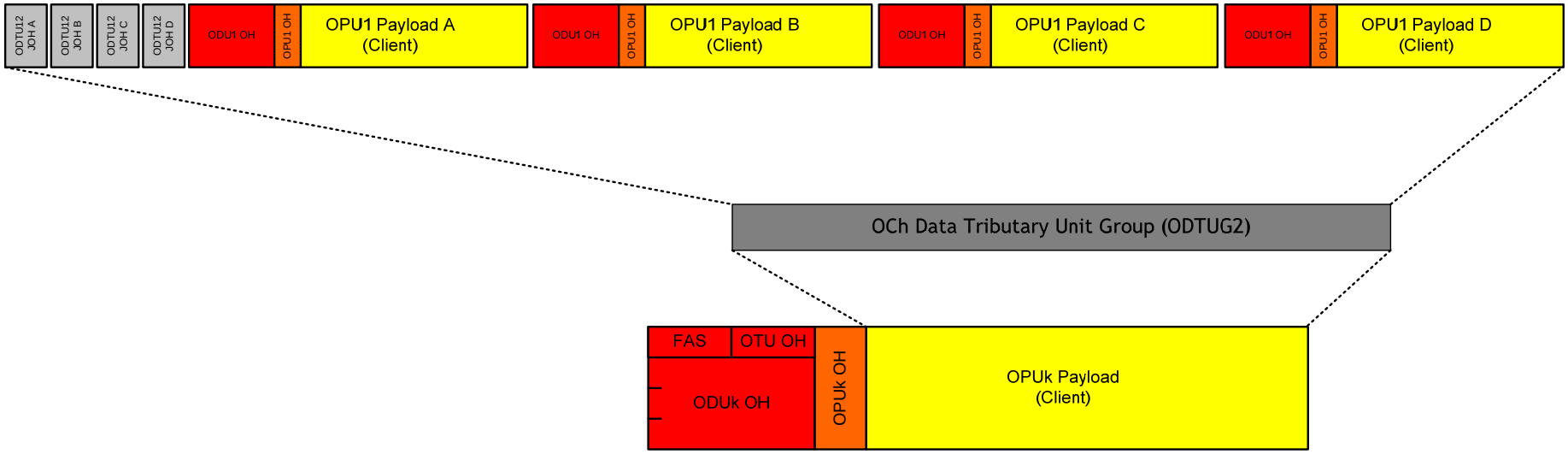
Payload portion = OPUk frame w/o OH = 93.33%

OTNk frame length = 130,560 bits (independent of bit rate)

OTUk clock tolerances = ±20 ppm

**** overclocked OTU2e, according to ITU-T G.Sup43**

OTN Multiplexing



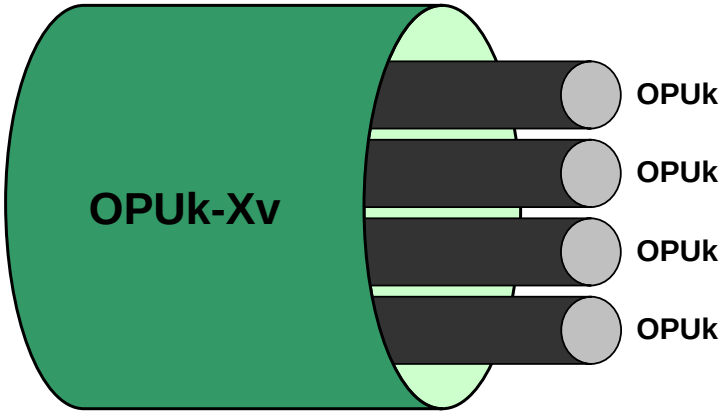
Multiplexing into 10G OTU-2

- Map up to 4 ODU1's into one ODTUG2 using time-division multiplexing
- Map ODTUG2 into OPU2

Multiplexing into 40G OTU-3

- Map a mixture of j ($j \leq 4$) ODU2's and 16-4j ODU1's into one ODTUG3
- Map ODTUG3 into OPU3

Virtual Concatenation (VCAT, G.7041)



- Well understood and widely deployed (SONET/SDH) technology to create larger logical transport "pipes"
  - High-order (HO) VCAT with STS1-Xv (SONET) or VC4-Xv (SDH) → 51.84 Mb/s granularity
  - Low-order (LO) VCAT down to E1 (SDH) or T1 (SONET) level → 2.0/1.5 Mb/s granularity
- In-service hitless link capacity adjustment to add/remove members to increase/decrease the virtual pipe size (ITU-T G.7042, LCAS)
- **OTN VCAT**
  - Same technology concept as from SONET/SDH
  - OPUk-Xv with X=1..256, only same k within one VCG
  - Max. service rate today = OPU3-256v = 10.28 Tb/s !