A Tutorial on ITU-T G.709 Optical Transport Networks (OTN)

White Paper

by Steve Gorshe
Principal Engineer

Issue No. 2: June 2011

PMC-Sierra, Inc.

Legal Information

Copyright

PMC-2081250, Issue 2

Disclaimer

Trademarks

Patents

Abstract

The SONET/SDH network that has grown to be the backbone of most of the modern telecommunications network was originally designed for optical interfaces that used a single wavelength per fiber. As optical component technology has advanced, it has become more economical to transmit multiple SONET/SDH signals over the same fiber using wavelength division multiplexing (WDM) rather than going to a higher rate SONET/SDH signal. Based on experience with the SONET/SDH networks, the ITU-T defined a transport network that was optimized for cost-effective transparent transport of a variety of client signals over WDM networks. The optical transport network (OTN) architecture is specified in ITU-T Rec. G.872 and the frame format and payload mappings are specified in G.709.

This white paper provides a tutorial overview of OTN, with primary emphasis on ITU-T G.709. The white paper also discusses various constraints that influenced the development of G.709, its current status in the network, and some factors that will affect its future.

About PMC

PMC-Sierra®, the premier Internet infrastructure semiconductor solution provider, offers its customers technical and sales support worldwide through a network of offices in North America, Europe, Israel and Asia. PMC-Sierra provides semiconductor solutions for Enterprise and Channel Storage, Wide Area Network Infrastructure, Fiber To The Home, and Laser Printer/Enterprise market segments. The Company is publicly traded on the NASDAQ Stock Market under the PMCS symbol. For more information, visit www.pmc-sierra.com.

About the Author

Steve Gorshe, Ph.D. is a Principal Engineer in PMC-Sierra's Technology Office, working on technology for optical transmission and access systems.

Steve is a Fellow of the IEEE. He has been with PMC-Sierra since 2000, and has 29 years of experience in research and development of telecommunications systems and ICs. His previous work included holding the position of Chief Architect for NEC eLuminant Technologies. He is the current Editor-in-Chief for *IEEE Communications magazine*, and former IEEE Communications Society Director of Magazines. He is Chief Editor and a technical editor for multiple standards for the ATIS COAST Committee (formerly T1X1/OPTXS). He is also technical editor for multiple ITU-T standards, including G.7041 (Generic Framing Procedure - GFP), G.8011.1 (Ethernet Private Line Service), and G.Sup43 (Transport of IEEE 10G Base-R in Optical Transport Networks (OTN)). He is a recipient of the Committee T1 Alvin Lai Outstanding Achievement Award and the ATIS Outstanding Contribution award for his standards work. He has 32 patents issued or pending, over 24 published papers, and is co-author of a telecommunications textbook and of three additional book chapters. Steve received his Ph.D. and MSEE from Oregon State University and BSEE from the University of Idaho.

Revision History

Issue No.	Issue Date	Details of Change
1	July 2009	Document created.
2	June 2011	Document reorganized and updated to include the items added to the 2009 revision of G.709, including GMP, delay measurement, ODUflex, and new mappings for 40GE and 100GE.

Table of Contents

Legal Information 2
Copyright 2
Disclaimer 2
Trademarks 2
Patents 2
Abstract 3
About PMC 3
About the Author 3
Revision History 4
List of Figures 7
List of Tables 8
New in this Issue 9
Preface 10
1 Introduction 11
1.1 Background 12
2 Physical Layer 14
3 WDM Multiplexing Approach and Architecture 15
3.1 Background on WDM Network Technical Considerations 15
3.2 ITU-T WDM Network Architecture 16
3.3 Optical Transport Network Equipment 19
4 Signal Formats and Frame Structure 22
5 Payload Mapping 29
5.1 CBR Client Signal Mappings 32
5.1.1 Legacy CBR Client Mapping Methods 33
5.1.2 CBR Signal Mapping Using GMP 36
5.1.3 ODUflex(CBR) Mappings 37
5.1.4 Mappings into ODU4 37
5.2 GFP and ATM Mapping 38
5.2.1 GFP-F Mapping into an Extended OPU2 39
5.2.2 ODUflex(GFP) Mapping 39
5.3 Ethernet Client Signals 39
5.3.1 Gigabit/s Ethernet (GE) 40
5.3.2 10 Gigabit/s Ethernet (10GE) over 10 Gbit/s OTN 42

5.3.3 10GE over 40 Gbit/s OTN ... 47
5.3.4 40 Gigabit/s Ethernet (40GE) ... 50
5.4 Sub-ODU1 rate clients ... 53
5.5 Storage Area Network (SAN) Clients ... 55
6 Synchronization and Frequency Justification / Rate Adaptation ... 57
6.1 Synchronization ... 57
6.2 Frequency Justification and Rate Adaptation ... 57
6.2.1 Asynchronous Mapping Procedure (AMP) Justification for Mapping and Multiplexing ... 57
6.2.2 Bit-synchronous Mapping Procedure (BMP) Justification for Mapping ... 62
6.2.3 Generic Mapping Procedure (GMP) Justification for Mapping and Multiplexing ... 62
7 OAM&P ... 75
7.1 Types of Overhead Channels ... 75
7.2 Maintenance Signals ... 78
7.3 Tandem Connection Monitoring (TCM) ... 79
7.4 Delay Measurement ... 80
8 Forward Error Correction (FEC) ... 82
9 OTN TDM Multiplexing ... 84
9.1 G.709 Multiplexing Concepts and Capabilities ... 84
9.2 Special Considerations for ODU4 ... 90
9.3 Multiplexing Summary ... 90
9.4 Multiplexing Stage Considerations ... 91
10 Virtual Concatenation ... 92
11 OTN Evolution ... 94
12 Conclusions ... 95
Appendix A – Optical Technology Considerations ... 96
Optical Signal Regeneration ... 97
Optical Switching ... 98
Appendix B – Multi-Lane OTN Interface ... 100
Appendix C – Hitless Adjustment of ODUflex (HAO) ... 103
13 References ... 104
14 Glossary and Abbreviations ... 106
15 Notes ... 109

List of Figures

Figure 1 Converged transport over OTN ... 12
Figure 2 Information flow illustration for an OTN signal ... 17
Figure 3 Illustration of OTN network layers ... 18
Figure 4 Packet Optical Transport Platform (OTP) ... 21
Figure 5 Information containment relationships for the electrical signal portions ... 22
Figure 6 G.709 OTN signal frame and overhead structure ... 24
Figure 7 G.709 OPU signal overhead structure with the asynchronous and bit-synchronous mapping procedures ... 27
Figure 8 G.709 OPU signal overhead structure with the generic mapping procedure ... 28
Figure 9 Mapping of CBR (SONET/SDH) signals into OTN ... 34
Figure 10 Mapping for GFP frames and ATM cells into the OPU ... 38
Figure 11 GFP mappings for extended GFP transport of 10GE signals (formerly G.Sup43 Section 7.3) ... 46
Figure 12 Modified OPU2 for extended GFP transport of 10GE signals (formerly G.Sup43 Section 7.3) ... 47
Figure 13 ODU3e1 frame structure and justification control ... 49
Figure 14 512B/513B block construction ... 51
Figure 15 1024B/1027B block construction ... 52
Figure 16 GFP-T superblock construction for FC1200 transport ... 56
Figure 17 Examples showing the AMP Justification Opportunity byte locations ... 61
Figure 18 OPUk payload area octet numbering illustration ... 64
Figure 19 GMP justification control overhead ... 65
Figure 20 C_m bit inversion patterns to indicate increment and decrement ... 66
Figure 21 FC-100 mapping illustration with GMP C_8 = 13062 ... 68
Figure 22 GE mapping illustration with GMP C_8 = 14407 ... 68
Figure 23 Word numbering illustration for GMP multiplexing – ODUflex tributary using five TS (#1, 2, 8, 10, 15) into an OPU3 ... 70
Figure 24 GMP justification control overhead for increased timing resolution ... 73
Figure 25 Illustration of TCM domains ... 79
Figure 26 OTN multiplexing hierarchy ... 85
Figure 27 OPU structure for multiplexing with shared justification overhead ... 87
Figure 28 Optical Add/Drop Multiplexing Illustration ... 99
Figure 29 OTU3/OTU4 parallel lane interleaving word structure ... 100

List of Tables

Table 1 OTN signal and payload rates ... 23
Table 2 Summary of OTN mapping methods ... 29
Table 3 Payload Type mapping code point for OTN signals ... 31
Table 4 Payload summary information for major OTN client signals ... 32
Table 5 Comparison of PDH, SONET/SDH, and OTN AMP frequency justification ... 58
Table 6 Justification control and opportunity definitions for AMP and BMP CBR mappings ... 58
Table 7 Justification control and opportunity definitions for TDM with AMP ... 59
Table 8 Client signal word sizes when using GMP multiplexing ... 70
Table 9 ODUk.ts rates for ODUflex(GFP) ... 72
Table 10 ODUflex(GFP) source C_m values and the ODUk.ts rates ... 72
Table 11 OAM&P channel definitions ... 76
Table 12 APS/PCC multiframe definition ... 77
Table 13 Number of TS and multiframe lengths for HO OPUk ... 86
Table 14 Summary of OTN Multiplexing methods ... 89
Table 15 Summary of TS use and frequency justification methods for OTN multiplexing ... 90
Table 16 Payload type values for virtually concatenated payloads (vcPT) ... 93
Table 17 Starting group of bytes sent in each lane for the OTU3 lane rotation ... 101
Table 18 Starting group of bytes sent in each logical lane for the OTU4 frame lane rotation ... 102

New in this Issue

The OTN standards have evolved significantly since the first issue of this OTN tutorial white paper. The current version of ITU-T Recommendation G.709 with its amendments and companion standards may be considered a ‘next-generation’ of OTN. The key additions to OTN that are covered in this white paper may be summarized as follows:

Addition or Evolution to OTN	White Paper Section(s)
· New ODU0 signal at approximately 1.25 Gbit/s	4
· New ODU4 signal at approximately 112 Gbit/s	4, 5.1.4, 9.2
· Generic Mapping Procedure (GMP) for the flexible rate justification in mapping and multiplexing new client signals	5, 5.1.2, 6.2.3
· New ODUflex signal types for the flexible mapping of both new CBR clients (ODUflex(CBR)) and packet clients (ODUflex(GFP))	4, 5.1.3, 5.2.2, 6.2.3
· Introduction of 1.25 Gbit/s Tributary Slots	6.2.3, 9.1
· Various new client signal mappings into OTN	5
· New delay measurement capability	7.4

Preface

During the “telecom bubble” era of the late 1990s early 2000s, there were high hopes and speculation that all-optical networks would quickly become prevalent. Many envisioned a relatively simple backbone networks where client signals were optically (wavelength division) multiplexed and switched without the core optical network elements having to do any electrical (and hence client signal dependent) processing of the client signals. In many ways, this appeared to be the ultimate integrated network. In response, the ITU-T Study Group 15 (SG15) developed a series of Optical Transport Network (OTN) standards for wavelength division multiplexed (WDM) networks that covered the physical layer, signal rate and format specification, and equipment functional requirements.

Three factors slowed the initial adoption of OTN. First, carriers had huge capital investments in their existing SONET/SDH networks and lacked money to replace or over-lay them with a new network layer and its associated new network management systems. Second, a number of SONET/SDH-based proprietary WDM solutions had already been developed that, while not ideal, were adequately serving the needs of many carriers. In fact, the ITU-T Rec. G.709 standard discussed in this white paper is very similar to SONET/SDH in many ways. Third, carriers had only recently seen bandwidth demand rise beyond what was offered by the combination of the existing WDM equipment and the large amount of fiber deployed in the backbone networks.

Since the mid 2000s, however, compelling reasons to deploy OTN have emerged worldwide, thus making OTN a fundamental component of carrier RFPs for optical metro network equipment. Initially, the most compelling application for OTN was point-to-point links where the enhanced forward error correction (FEC) capability standardized for OTN allowed longer spans of optical cable, higher data rates, or both. Today, OTN is being demanded by carriers worldwide as not just a point-to-point technology but as an entirely new network layer to transition away from SONET/SDH and enable “video-ready” metro optical networks for high bandwidth service delivery to subscribers over broadband access networks. OTN enables carriers to fully utilize each wavelength by building transparent, scalable and cost-optimized networks where client traffic like video and Ethernet is mapped into OTN at the edge of the transport network. In this model, SONET/SDH becomes another client. Another important application is providing cost-effective wide area network (WAN) connectivity for enterprise Ethernet and storage area network (SAN) signals.

This white paper provides an overview of the OTN standards, with primary focus on ITU-T G.709. This second issue of the white paper includes the new material that was incorporated into the 2009 revision of G.709, its subsequent first two amendments, and a brief description of the new G.7044 standard on hitless resizing of ODUflex(GFP) signals.

1 Introduction

The ITU-T has developed a set of new standards covering the wavelengths and signal formats in order to better support the multiplexing of a substantial number of signals onto a single fiber. These signal format and hierarchy standards cover digital signals and include the OAM&P overhead as part of the signal format. In the context of this white paper, Optical Transport Network (OTN) refers to networks using the ITU-T Rec. G.709 standard for Wavelength Division Multiplexed (WDM) signals.

WDM transport networks based on the ITU-T OTN standards are becoming increasingly important. The reasons carriers are moving toward OTN include:

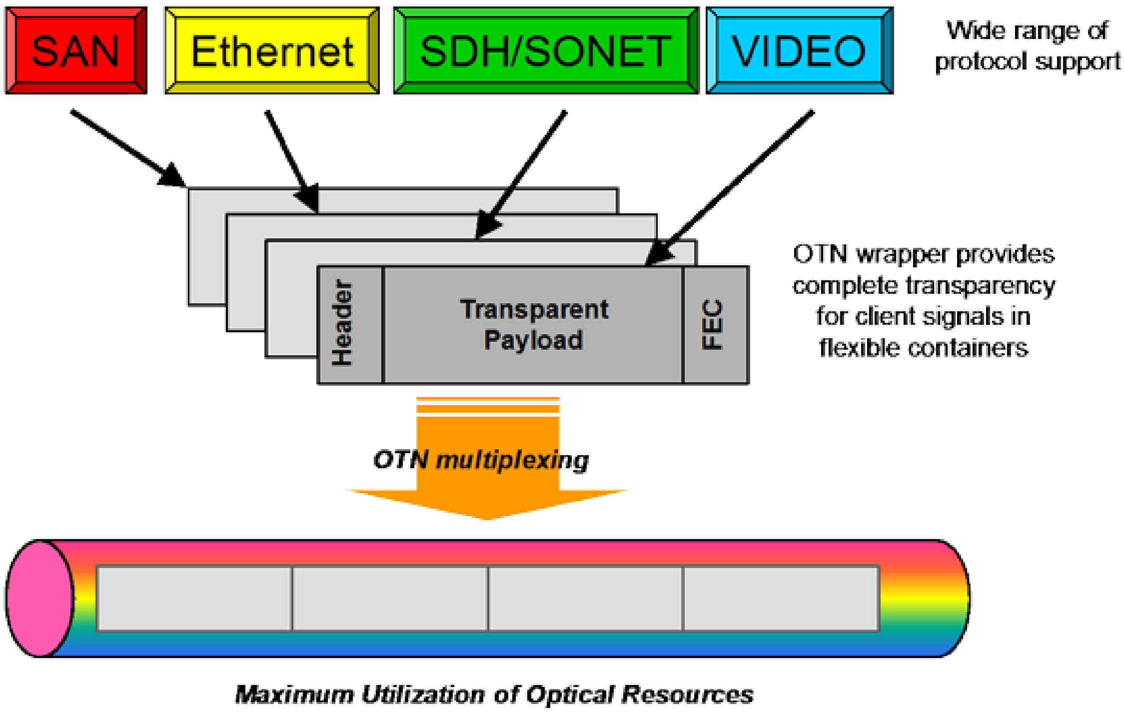
- OTN is a much less complex technology for transport applications than SONET/SDH.
- The OTN signal incorporates overhead optimized for transporting signals over carrier WDM networks.
- The combination of the reduced technology complexity and optimized overhead allows substantial reductions in carrier transport network operations expenses.
- The OTN multiplexing bandwidth granularity is one or two orders of magnitude higher than for SONET/SDH, thus making it more scalable to higher rates.
- OTN now provides a cost effective method for carrying high-speed wide area network (WAN) data clients including Ethernet and storage area network (SAN) protocols.
- OTN provides an integrated mechanism for forward error correction (FEC) that allows greater reach between optical nodes and/or higher bit rates on the same fiber.
- Client signals can be carried over OTN in a transparent manner. This transparency includes native SONET/SDH signals for the “carrier’s carrier” application where the entire client SONET/SDH signal’s overhead must be preserved through the OTN.

In other words, as illustrated in Figure 1, OTN provides an optimum converged transport technology for transparently carrying important legacy and emerging client signals.

This white paper provides a tutorial on G.709 OTN, including the signal hierarchy and formats, client signal mapping and multiplexing methods, OAM&P overhead, and network synchronization considerations. It also includes discussions of background information at the beginning of sections where the reader may find this helpful in understanding the motivations and applications for the different aspects of the OTN standards.

Note: The abbreviations used in this white paper are defined in the glossary section 14.

Figure 1 Converged transport over OTN



1.1 Background

As optical component technology has improved, it has become possible to increase the traffic sent over a fiber by sending multiple signals, each on its own wavelength, rather than increasing the rate of a single signal (e.g., sending 16 OC-48 signals, each on their own wavelength rather than a single OC-768). Such multiplexing is referred to as wavelength-division multiplexing (WDM). When WDM was first discussed, it held the promise of sending each signal in its native format rather than mapping it into the payload of another signal such as SONET/SDH. It is difficult, however, for a network operator to provide operations, administration, maintenance, and provisioning (OAM&P) for each signal if it uses its native signal format, since this would require multiple, client signal dependent management systems. This problem is especially true for analog signals (e.g., TV channels), which have a very different set of channel requirements than digital signals. More will be said on this topic in the discussion of the OTN signal architecture.

One reason for developing a new signal format for WDM signals (instead of just using the existing SONET/SDH signals) was the possibility to add new overhead channels that would give the added functionality required to efficiently perform OAM&P on the WDM network. Another reason for developing a new standard was to provide a means for more powerful forward error correction (FEC) capability. As discussed in PMC-Sierra white paper PMC-2030895 [6], a relatively modest FEC capability was added to SONET/SDH. As signals traverse a multi-hop optical network, however, the signal to noise ratio decreases. Since the carriers hoped to increase the transmission distances and the bit rates per wavelength, the SONET/SDH FEC is not always adequate. Finally, another reason for new standards for transport was to provide a less granular payload envelope for the transport of higher bandwidth individual clients aggregated from access networks. For example, if the Gigabit Ethernet signal eventually becomes the smallest client switched in the network, then providing circuit switching in the transport network at the granularity of STS-1s (51.84 Mbps) does not promote optimal cost and complexity in the network.

2 Physical Layer

A full discussion of lasers, receivers, and the characterization of fiber optic channels is beyond the scope of this white paper. The interested reader can find more detail in books such as [7]]. Appendix A to this white paper introduces some of the basic physical layer concepts so that the reader can appreciate some of the decisions that were made in defining the OTN and its signal formats.

While the optical transport signals (see section 4) were originally specified as serial signals on a single wavelength, in late 2008 the ITU-T also standardized an optional multi-lane interface specification for its 40 and 100 Gbit/s signals. The intention of these multi-lane interfaces is to take advantage of relatively inexpensive Ethernet 40GE and 100GE parallel optical interface modules for applications where a parallel interface was more cost effective than a serial interface. See Appendix B for a description of this parallel interface.

3 WDM Multiplexing Approach and Architecture

After discussing some of the important underlying technical considerations, this section presents the high level view of the WDM architecture adopted by the ITU-T for optical transport networks. The section concludes with an introduction to reconfigurable optical add/drop multiplexers (ROADMs), which are an increasingly important type of WDM network equipment.

3.1 Background on WDM Network Technical Considerations

A number of different approaches had to be examined at the outset of the WDM standardization work, with numerous tradeoffs to be considered. Ideally, any type of native signal could be carried on any of the wavelengths with extensive operations, administration, and maintenance (OAM) capabilities for each signal. This ideal is difficult to achieve in practice, however. Broadly speaking, the approaches fell into two categories: The first is to send the client signal essentially in its native format (with the exception of its normal wavelength) and add OAM capability in some type of separate channel. The second approach is to treat the client signal as a digital payload signal and encapsulate it into a frame structure that includes channel-associated OAM overhead channels.

The approach of assigning each client signal to its own carrier wavelength and carrying it in its native format creates the question of how to create the overhead channel(s). One option is to have the OAM information carried on a separate wavelength. Having the client signal and its associated OAM channel on separate wavelengths has some serious disadvantages, in addition to the obvious drawback of doubling the number of wavelengths used per client. First, the OAM channel won’t necessarily experience the same impairments as the client signal channel. Second, it is possible for provisioning errors to properly connect the OAM signal but not the client signal channel. Another option that received serious consideration was using sub-carrier modulation to create the OAM channel. In this approach, the optical wavelength carrying the high-speed data signal is itself modulated with a low frequency signal that carries the OAM channel and could be removed through low-pass filtering at the termination point. There was some concern that this approach would be too complex, including in its impact on jitter performance.

The approach of carrying the client signals as the payload of a digital frame was referred to as a “digital wrapper” approach.[1] The digital wrapper, which contained the various OAM overhead channels, is conceptually similar to SONET/SDH.

[1] For this reason, the G.709 OTN frame has sometimes been referred to as a “digi-wrapper.”

In the end, a hybrid approach was chosen. The digital wrapper approach was chosen for the basic encapsulation and channel-associated OAM overhead for the client signals. The resulting digital signal is then transmitted a wavelength, or time division multiplexed into a higher rate signal that is transmitted over a wavelength. Additional optical network overhead to support groups of these digital signals is then carried on a separate wavelength.

3.2 ITU-T WDM Network Architecture

The basic signal architecture is illustrated in Figure 2. The client signal is inserted into the frame payload area, which, together with some overhead channels, becomes the Optical Payload Unit (OPU). An OPU is conceptually similar to a SONET/SDH Path. OAM overhead is then added to the OPU to create the Optical Data Unit (ODU), which is functionally analogous to the SONET Line (SDH Multiplex Section). Transport overhead (e.g., frame alignment overhead) is then added to create an Optical Transport Unit (OTU), which is the fully formatted digital signal and functionally analogous to the SONET Section (SDH Regenerator Section). The client signal through OTU layer signal frame relationships are also illustrated in Figure 5. This OTU is then transmitted over a wavelength, which constitutes the Optical Channel (OCh). An Optical Multiplexed Section (OMS) consists of a wavelength division multiplexed group of optical channels, together with a separate wavelength carrying an overhead optical supervisory channel (OSC), that is carried between access points. The Optical Transport Section (of order *n*) consists of an OMS (of order *n*) and an overhead channel (on its own wavelength). The OTS defines the optical parameters associated with the physical interface. The OCh, OMS, and OTS overhead channels provide the means to assess the transmission channel quality, including defect detection, for that layer. The OCh and OTS overhead also provides a means for connectivity verification. The OCh, OMS, and OTS layers are described in ITU-T Rec. G.872, and will not be discussed further here.

Figure 2 Information flow illustration for an OTN signal

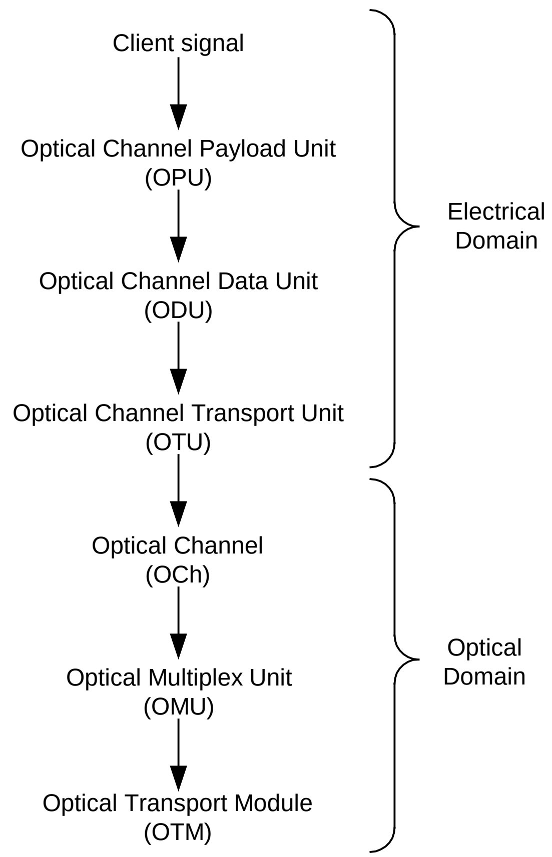
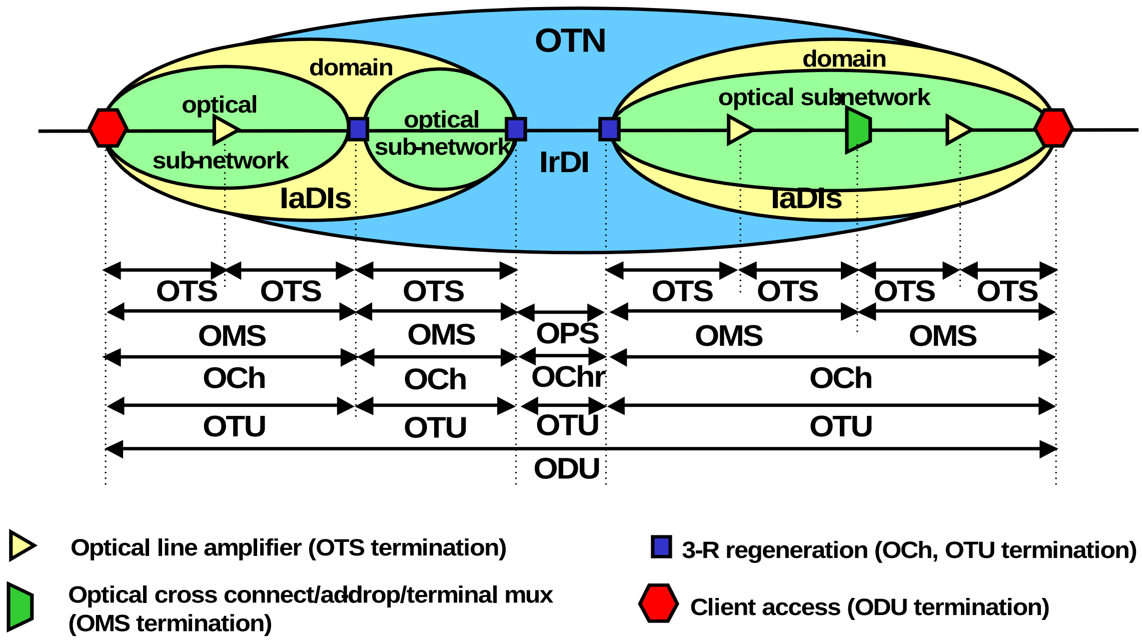


Figure 3 shows an example OTN with the different layers and their relative scope. The IrDI is the inter-domain interface and is specified as having 3R regenerator processing at both sides of the interface. (See Appendix A for an explanation of 3R regenerators.) The IrDI is the interface that is used between different carriers, and can also be useful as the interface between equipment from different vendors within the same carrier's domain. Since the IrDI is the interface for interworking, it was the focus of the initial standard development. The IaDI is the intra-domain interface that is used within a carrier's domain. Since the IaDI is typically between equipment of the same vendor, it can potentially have proprietary features added such as a more powerful FEC.

Figure 3 Illustration of OTN network layers



Once the choice was made to use a digital wrapper approach, the next choice was what client signals should be allowed. Clearly, the digital wrapper approach restricts the clients to being digital signals. Although it would have been ideal to allow both analog and digital clients in the same OTN, the main problem is that analog and digital signals have very different channel requirements. A channel that may be very adequate for a digital signal can have an unacceptably low signal-to-noise ratio or too much distortion for an analog signal. This makes it very difficult, especially administratively, to deploy mixed analog/digital networks in a DWDM environment. The next decision was what types of digital signals to include. Originally, there was a strong desire to carry optical data interfaces such as Gbit/s and 10 Gbit/s Ethernet in addition to SONET/SDH signals. In what appeared to be uncharacteristic shortsightedness, the decision was made to limit the constant bit rate (CBR) clients to the SONET/SDH signals[2]. The assumption was made that other signals could be mapped into SONET/SDH first, with these SONET/SDH signals being mapped into the OTN. This decision not to directly support native Ethernet clients, while potentially simplifying the frame structures, proved to be a significant handicap to early wide-scale deployment of G.709 OTNs[3]. As Ethernet client signals increased in importance, mappings into OTN were eventually added for them. Accommodation of Ethernet client signals is discussed in Section 5.3. In addition to CBR signals, mappings are defined for placing ATM or GFP frames directly into the OPU payload area (i.e., with no SONET/SDH frames). Payload mappings are discussed further in section 5.

3.3 Optical Transport Network Equipment

There are several different types of optical transport network equipment being deployed based on the OTN standards. The most common types include:

- Regenerators,
- OTN terminal equipment,

---

[2] The justification control mechanism described in section 6.2.1 for the original OPU1, OPU2, and OPU3 was limited to SONET/SDH client signals. However, as explained in section 5.1, it is possible to map other CBR signals into the OPUk payload.

[3] The primary reason for not supporting the full 10 Gbit/s payload was the 12.5 Gbit/s bandwidth constraint imposed by undersea cable systems. Supporting the full 10 Gbit/s payload rate would not leave an acceptable amount of overhead bandwidth for FEC. IEEE 802.3, with some initial reluctance, attempted to salvage the situation by defining the 10 Gbit/s Ethernet signal to have a WAN PHY rate of 9.58464 Gbit/s so that it could map directly into a SONET STS-192c (SDH VC-64c) payload envelope rather than the 10.3125 Gbit/s PHY rate used for the LAN. This mapping is described in white paper PMC-2030895. Proposals to carry the full 10 Gbit/s rate LAN PHY information over OTN included increasing the OTN clock rate and using GFP-F encapsulation to provide the frame delineation and eliminate the need for inter-packet Idle characters. No consensus was reached on a single approach, and eventually one of the overclocked solutions and one of the GFP-based solutions were added to G.709. See section 5.3.2 for a full discussion of this topic.

- Optical Add/Drop Multiplexer (OADMs),
- Optical cross connect (OXCs).

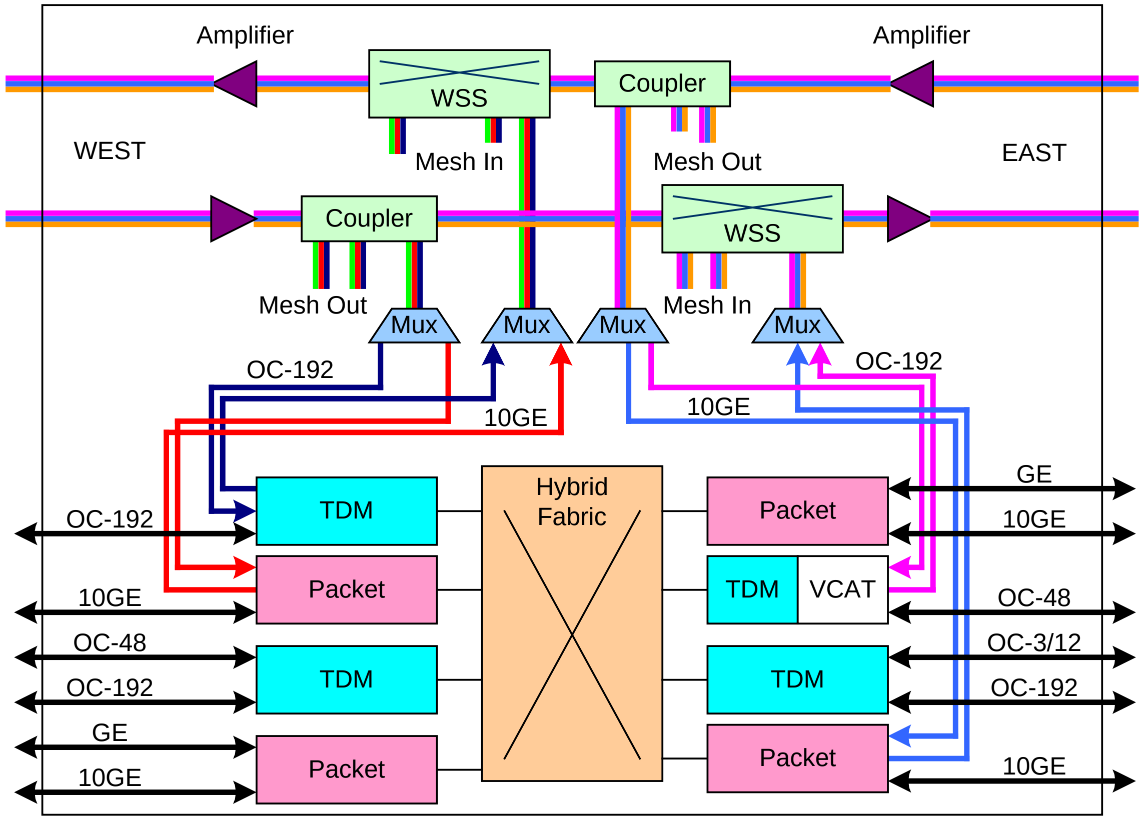
OTN terminal equipment is used for point-to-point connections through WDM networks, mapping the client signals into OPUs, sometimes multiplexing multiple signals in the electrical domain, and finally performing mapping/multiplexing in the optical domain. OADMs, OXCs, and some types of regenerators primarily process the OTN signals in optical domain. See Appendix A for more discussion on these three types of equipment.

In recent years, reconfigurable OADMs (ROADMs) have become popular. The key building blocks of today's ROADM node can be categorized into three primary functions:

- Wavelength add/drop filters or switches – This is generically referred to as a wavelength fabric and operates only in the optical domain. However, it can be implemented with a number of different technologies, including wavelength blockers and Wavelength Selective Switches (WSSs). The wavelength fabric multiplexes and demultiplexes all of the individual DWDM wavelengths from the client interfacing cards. The wavelength fabric also provides optical signal protection.
- Dynamic power control and remote monitoring capabilities at the optical layer – Optical amplification with dispersion compensation and gain equalization, dynamic power control and remote monitoring for the presence/absence of optical signals are just a few of the many advancements that have reduced the need for truck rolls for node engineering.
- Optical service channel termination and generation - Traditionally this is in the form of transponders and muxponders.

Next-generation ROADMs, as illustrated in Figure 4, typically augment ROADM functionality with switching fabrics in the electrical domain, either locally in the form of muxponders or centrally in the form of a switch fabric. Often the designation Optical Transport Platform (OTP) or Packet OTP (P-OTP) are used when a central switch fabric is present. The electrical domain switching can be TDM (SONET/SDH or OTN), packet switching, or both. The motivation for this next generation ROADM or OTP is to allow for much greater flexibility when adding and dropping client signals within the signals carried over the wavelength rather than just adding or dropping the entire wavelength. This finer, sub-wavelength granularity add/drop allows aggregation or grooming for more efficient use of the wavelengths. It also allows more flexible network topologies.

Figure 4 Packet Optical Transport Platform (OTP)



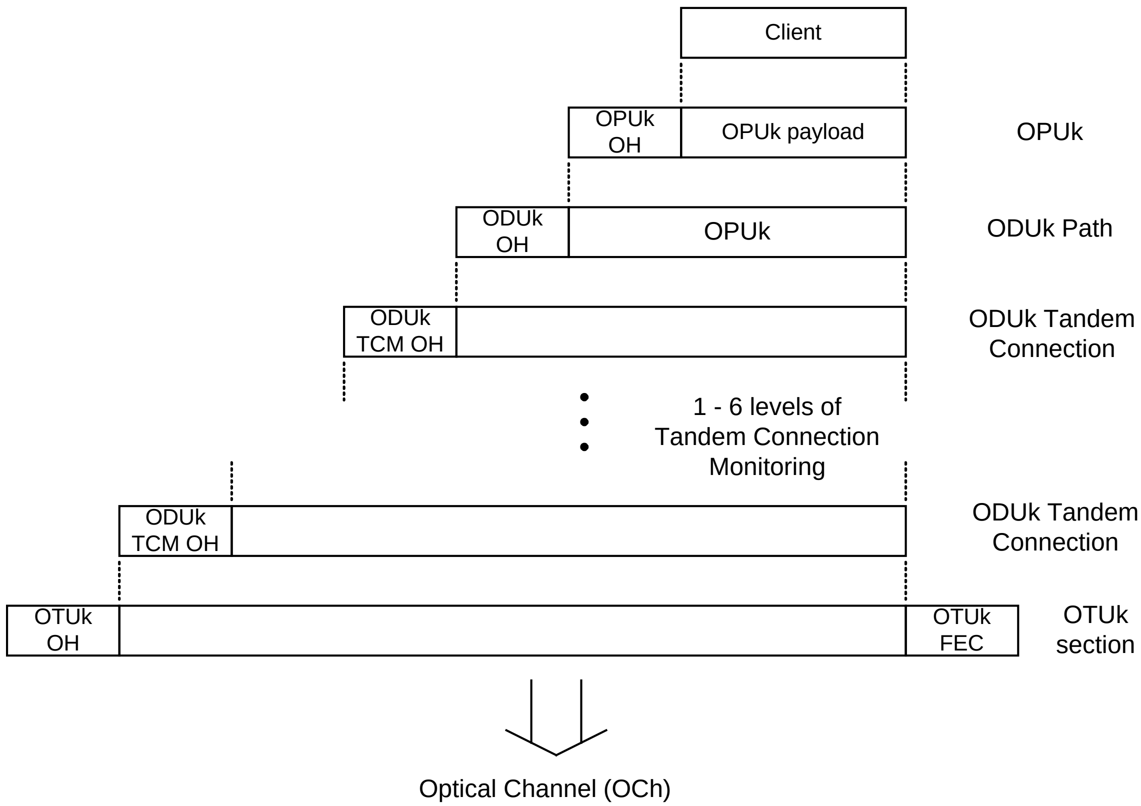
From this discussion, it is clear that the lines are blurring between ROADMs and Multi-Service Provisioning Platforms (MSPPs).

For further discussion on ROADM, OTP, and P-OTP technologies and architectures, please see PMC-Sierra's "ROADMs and the Evolution of the Metro Optical Core" and "Enabling OTN Convergence – Solutions for the New Packet Optical Transport Network" white papers. [16], [21]

4 Signal Formats and Frame Structure

This section describes the signal format for the digital portion of the OTN signal. The containment relationships of the client, OPU, ODU, and OTU layers and their overhead are shown in Figure 5. Figure 5 also illustrates the existence of multiple levels of Tandem Connection Monitoring (TCM), which will be described below. It can also be seen that the FEC is added at the OTU level, which is the last step before the optical transmission of the signal.

Figure 5 Information containment relationships for the electrical signal portions



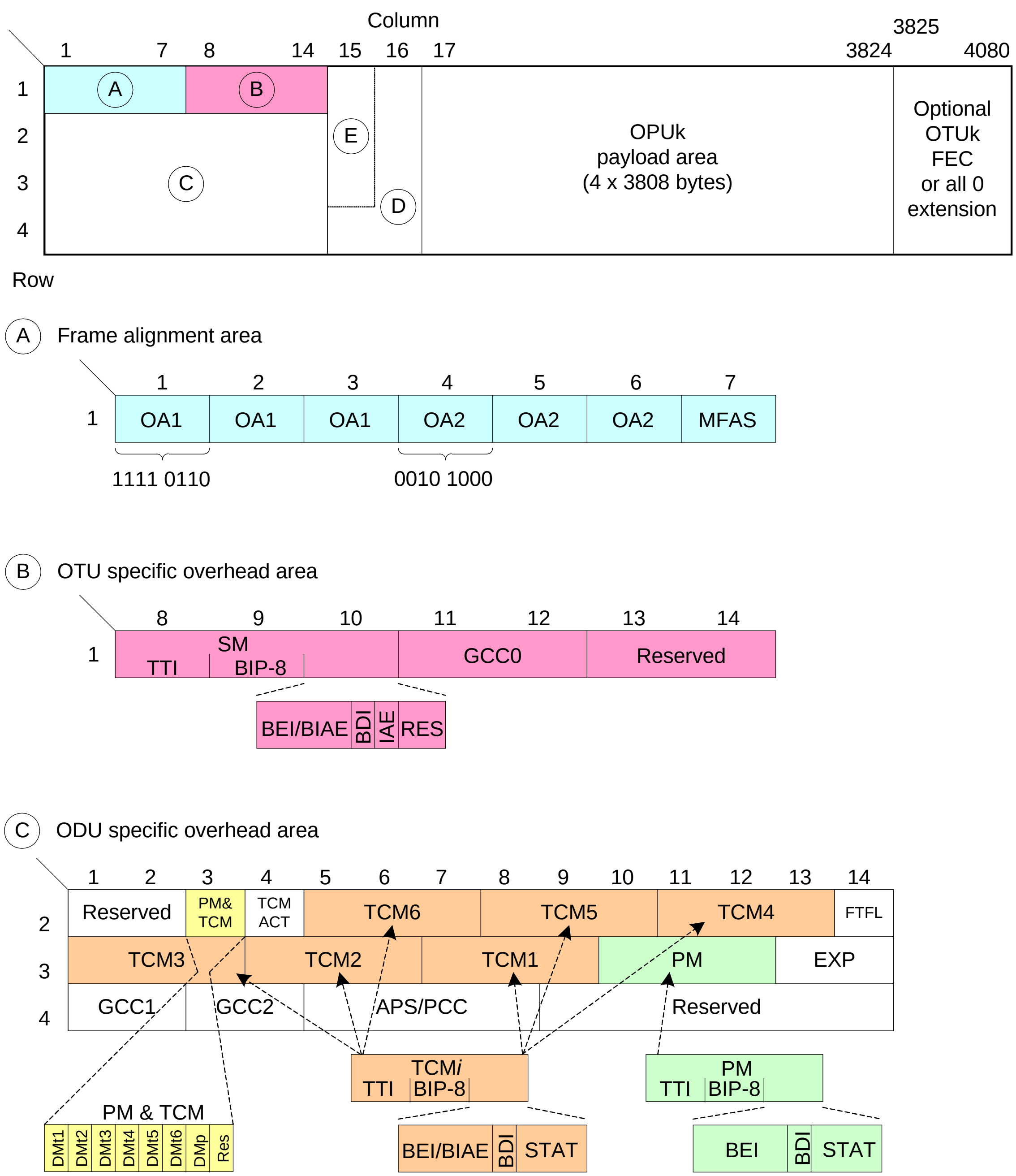
There are four currently defined OTU rates and five fixed OPU/ODU rates. An OPU, ODU, or OTU of a particular rate is referred to as an OPUk, ODUk, or OTUk with k = 0, 1, 2, 3, or 4. The respective signal and payload rates are shown in Table 1. In addition to the fixed ODUk/OPUk rates, flexible ODUflex signals have been defined that have rates adapted to the client signal or service. ODUflex signals are defined in 5.1.3 and 5.2.1.

Table 1 OTN signal and payload rates

k	OTUk signal rate	OPUk payload area rate	OTUk/ODUk/OPUk frame period
0	Not applicable	238/239 × 1 244 160 kbit/s = 1 238 954 kbit/s	98.354 μs
1	255/238 ´ 2 488 320 kbit/s = 2 666 057 kbit/s	2 488 320 kbit/s	48.971 μs
2	255/237 ´ 9 953 280 kbit/s = 10 709 225 kbit/s	238/237 ´ 9 953 280 kbit/s = 9 995 277 kbit/s	12.191 μs
3	255/236 ´ 39 813 120 kbit/s = 43 018 414 kbit/s	238/236 ´ 39 813 120 kbit/s = 40 150 519 kbit/s	3.035 μs
4	255/227 × 99 532 800 kbit/s = 111 809 974 kbit/s	238/227 × 99 532 800 kbit/s = 104 355 975 kbit/s	1.168 μs
flex (CBR)	Not applicable	client bit rate (> 2.488 Gbit/s)	
flex (GFP)	Not applicable	carrier provisioned packet service rate	
Notes: All rates are ±20 ppm for k = 0-4, and ±100 ppm for ODUflex, maximum. All ODUk rates are 239/238 × (OPUk payload rate) For efficient bandwidth use, the ODUflex(GFP) rate is nominally *n* × (Tributary Slot rate). See Table 10 for details regarding ODUflex(GFP) rates.			

The OPU, ODU, and OTU frame structure is shown in Figure 6, Figure 7 and Figure 8, including the overhead for each level. The ODU frame is structured as four rows by 3824 columns, regardless of the signal rate. As the signal rate rises, the frame period shrinks, but the frame size remains constant. The OPU payload area consists of columns 17-3824 for all four rows.

Figure 6 G.709 OTN signal frame and overhead structure



The OTU consists of the ODU, the OTU overhead, and the FEC, if used. The OTU overhead is shown as the A and B areas in Figure 6. The A field contains the frame alignment pattern and the multiframe alignment signal (MFAS). The MFAS field is a binary counter that shows the phase of the current frame within the 256-frame multiframe. Those fields in Figure 6, Figure 7, and Figure 8 and that are defined as spreading across the multiframe (e.g., the PSI and virtual concatenation overhead) use the MFAS to determine the meaning of the byte during that frame. The B area of Figure 6 provides GCC and section monitoring (SM) information for the OTU. The SM overhead includes the trail trace identifier (TTI, similar to SONET Section trace and SDH Regenerator Section trace for connectivity fault detection), a BIP-8 for error detection, and backward error indication (BEI). Similar to the SONET/SDH REI, the BEI is sent by the OTU sink to the OTU source as a (binary) count of the number of errors detected by the previous BIP-8. The backward defect indicator (BDI) is used by the sink to inform the source that it is seeing a signal failure (similar to SONET/SDH RDI) In addition, the SM overhead for OTU includes an incoming alignment error (IAE) indicator. The IAE indicates that a frame alignment error was detected on the incoming signal, with the BIAE informing the source that an IAE was seen. The IAE and BIAE are used to disable the error counting in their respective directions during frame alignment loss conditions.

Note that the final step before transmitting the OTU on the optical channel is to scramble it in order to assure adequate transition density for reliable receiver clock recovery. The scrambling is performed on all OTU frame bits, including the FEC bytes, but excluding the framing bytes. A frame-synchronized scrambler is used with polynomial $x^{16}+x^{12}+x^{3}+x+1$ that is reset to all 1s on the MSB of the MFAS byte.

The ODU consists of the OPU and the ODU overhead, which is functionally similar to the SONET Line (SDH Multiplex Section) overhead. The ODU overhead is area C in Figure 6. It contains the overhead for path performance monitoring (PM), fault type and fault location (FTFL), two generic communications channels (GCC), an automatic protection switching and protection communications channel (APS/PCC), six levels of tandem connection monitoring (TCM), and a set of bytes reserved for experimental purposes. A delay measurement function has also been added to the ODU overhesad to perform round trip delay measurement at the Path and each TCM level. (See 7.4.) The PM and TCM overhead consists of trail trace identifier (TTI, similar to SONET Path trace and SDH Trail trace for connectivity fault detection), a BIP-8 for error detection, status information (to indicate whether this is a normal signal or a maintenance signal), and a BEI. The BEI, which is similar to the SONET/SDH REI, is sent by the ODU sink to the ODU source as a (binary) count of the number of errors detected by the previous BIP-8. The TCM overhead also contains a BIAE and a BDI.

One import aspect of the ODU overhead definition is the coverage of the Path and TCM BIP-8 bytes. Each of these bytes covers only the OPU portion of the frame (columns 15-3824, rows 1-4). In other words, they do not cover the ODU overhead itself. This definition allows the insertion or termination of overhead such as the TCM overhead at intermediate nodes along the ODU path without requiring a re-computation or compensation of the existing BIP-8 values.[4]

The overhead information for the OPU is contained in the D and E areas of Figure 6, and is shown in more detail in Figure 7 and Figure 8. The OPU overhead, which is primarily concerned with the mapping and demapping of the client signal, covers the OPU from the point at which the client signal is mapped into the OPU until it is extracted at the OPU termination point. It includes some functions that are similar to the SONET/SDH Path overhead. For all applications, the Payload Structure Indicator (PSI) byte in row 4, column 15 contains indicators for the payload type (PT) and Multiplex Structure Identifier (MSI). As indicated in Figure 7 and Figure 8, the usage of the other OPU overhead bytes depends on factors including the type of client signal being carried in the OPU and the mapping method being used.

In applications where frequency justification is required in order to adapt between the client signal rate and the OPU payload rate, this frequency justification information is located in the OPU overhead area. The definition of the Justification Control (JC) bytes in the OPU overhead depends on the type of mapping procedure being used. The details of the JC bytes are in Figure 7 are shown in Table 6 and Table 7, and the details of the JCn bytes in Figure 8 are shown in Figure 19.

For applications using the Hitless Adjustment of ODUflex(GFP) (HAO) client signals the E area of the OPU overhead contains the HAO overhead. See section 5.2.2 for a description of ODUflex(GFP) signals and Appendix C for an introduction to HAO.

For applications in which no frequency justification is required, the JC bytes are typically unused. The one exception to this is the 10G-BASE-R mapping discussed in section 5.3.2, in which all of the D and E area except the PSI byte are used to form an extended payload area in which to carry client data.

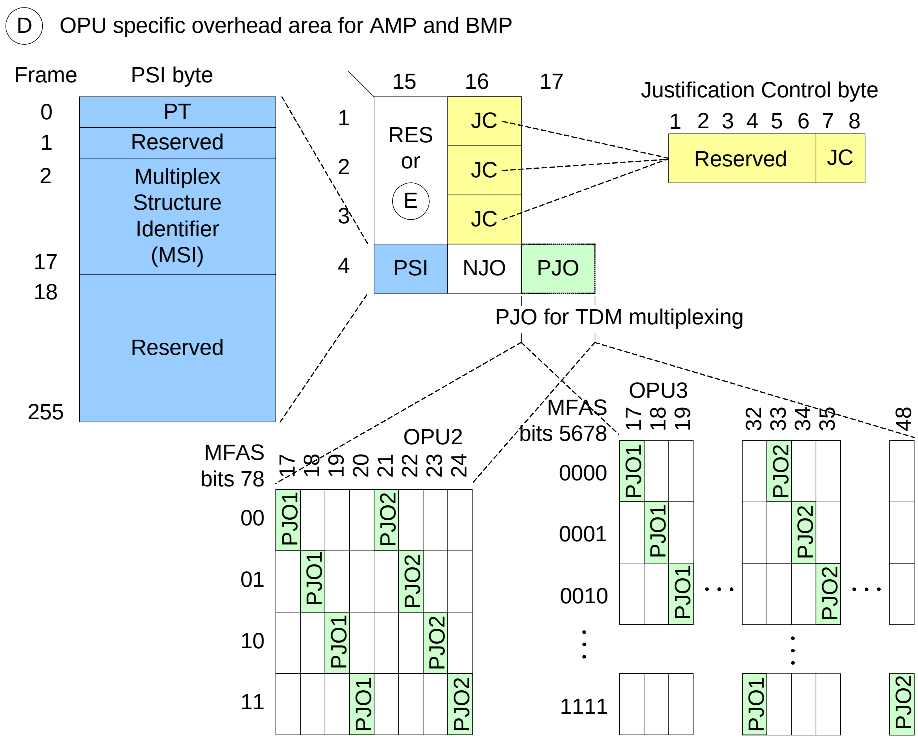
Virtual concatenation is supported when the asynchronous mapping procedure is used. When virtual concatenation is enabled, its overhead is located in the E area of Figure 6, as illustrated in Figure 7. In applications where the E area is not used for virtual concatenation or generic mapping procedure overhead, these three bytes are reserved.

Unlike SONET/SDH Paths, however, the OPU relies on the next lower level (ODU) for end-to-end error detection.

---

[4] Since TCM was added to SONET/SDH after the initial version of these standards, the only option was to add it in an overhead location covered by the Path BIP-8. Consequently, any information insertion/termination of the TCM overhead required that the Path BIP-8 be compensation such that its end-to-end error reporting was correct.

Figure 7 G.709 OPU signal overhead structure with the asynchronous and bit-synchronous mapping procedures



Note: For simplicity, only some representative PJO locations are illustrated here.

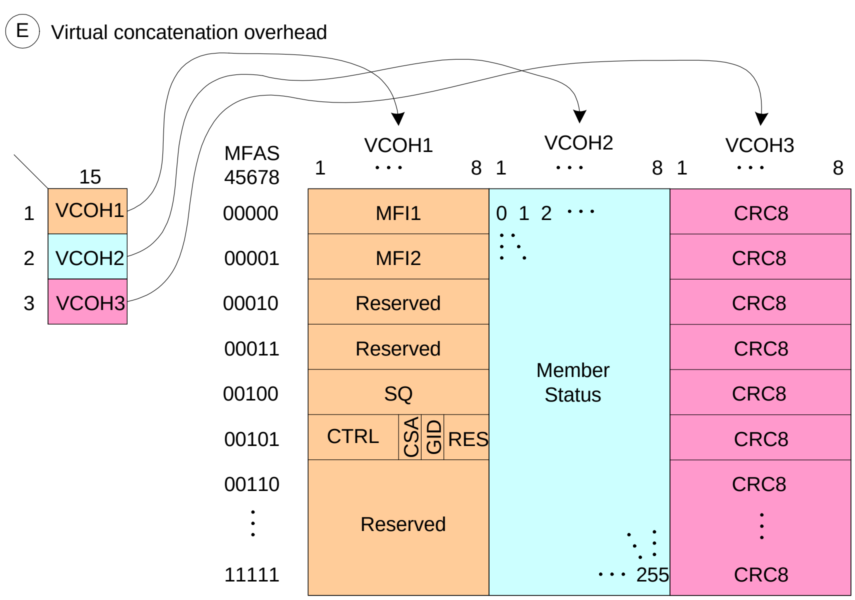
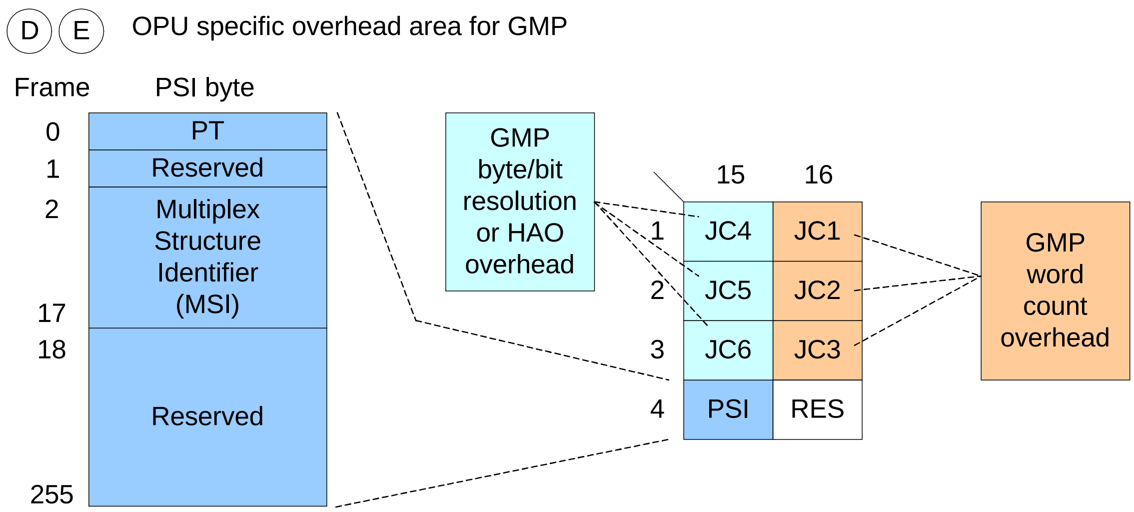


Figure 8 G.709 OPU signal overhead structure with the generic mapping procedure



5 Payload Mapping

In a broad sense, there are five different methods defined in G.709 to map a client signal into OTN. The mapping methods, which support both constant bit rate (CBR) client signals and cell/packet-based signals, are summarized in Table 2. The payload type (PT) overhead values for the defined mappings are shown in Table 3, and the mapping information for major OTN client signals is summarized in Table 4. The rate adaptation mechanisms for the different mappings are described in section 6.2. As explained in section 9, the mechanisms for multiplexing lower rate ODU signals into higher rate ODU signals are provided by extensions to two of these mapping methods. The sub-sections of this section are organized by a combination of mapping type and major client type.

Table 2 Summary of OTN mapping methods

<table>
<tr><th colspan="2">Mapping Method</th><th>Application</th></tr>
<tr><td rowspan="4">CBR</td><td>Asynchronous Mapping Procedure (AMP)</td><td>Mapping SDH client signals into OTN</td></tr>
<tr><td>Bit-synchronous Mapping Procedure (BMP)</td><td>Alternative method for mapping SONET/SDH client signals into OTN
Mapping CBR clients into ODUflex(CBR)</td></tr>
<tr><td>Generic Mapping Procedure (GMP)</td><td>Mapping non-SONET/SDH CBR clients into ODUk (k = 0, 1, 2, 3, 4) and SDH STM-1 and STM-4 clients into ODU0</td></tr>
<tr><td>Timing Transparent Transcoding (TTT) using a combination of GFP and GMP*</td><td>Transcoding the native client signal into a lower bit rate CBR stream in order to increase bandwidth efficiency, and maintain client character and timing information transparency</td></tr>
<tr><td>Packet</td><td>GFP frames or ATM cells into an OPU payload container</td><td>Mapping packet clients into ODUk (k = 0, 1, 2, 3, 4) with GFP (Generic Framing Procedure) encapsulation
Mapping packet clients into an ODUflex(GFP) with GFP encapsulation
Mapping ATM (Asynchronous Transfer Mode) cells into OTN</td></tr>
<tr><td colspan="3">* For the specific case of the FC-1200 client, TTT uses BMP rather than GMP</td></tr>
</table>

The terminology Low Order ODU (LO ODU) and High Order ODU (HO ODU) were introduced to distinguish the function being served by the ODU in that application. Essentially, an ODUk that directly becomes an OTUk is referred to as a HO ODU. When an ODUk is being discussed in terms of carrying a client signal, it is referred to as a LO ODUk. For example, client signals are mapped into the OPU of a LO ODU. A HO ODU will often have multiple LO ODU signals multiplexed into it. The same ODU can be both LO and HO, depending on the functional context. For example, a client signal is mapped into a LO ODUk, but if that ODUk is transmitted directly within an OTUk, then the ODUk is also a HO ODU. Intermediate carrier applications create a similar situation. The HO ODUk transmitted by the end carrier may have multiple LO ODU signals multiplexed into it. When the intermediate carrier multiplexes this ODUk into a higher rate HO ODUj for transparent transport, the intermediate carrier is treating the ODUk as a LO ODU.

Table 3 Payload Type mapping code point for OTN signals

Hex code	Interpretation
01	Experimental mapping (Note 1)
02	Asynchronous CBR mapping
03	Bit synchronous CBR mapping,
04	ATM mapping
05	GFP mapping
06	Virtual Concatenated signal
07	1000BASE-X into ODU0 mapping
08	FC-1200 into ODU2e mapping
09	GFP mapping into Extended OPU2 payload (Note 2)
0A	STM-1 mapping into ODU0
0B	STM-4 mapping into ODU0
0C	FC-100 mapping into ODU0
0D	FC-200 mapping into ODU1
0E	FC-400 mapping into ODUflex
0F	FC-800 mapping into ODUflex
10	Bit stream with octet timing mapping,
11	Bit stream without octet timing mapping,
12	IB SDR mapping into ODUflex
13	IB DDR mapping into ODUflex
14	IB QDR mapping into ODUflex
15	SDI mapping into ODU0 (Note 4)
16	1.5G SDI mapping into ODU1 (Note 4)
17	3G SDI mapping into ODUflex (Note 4)
18	ESCON mapping into ODU0 (Note 4)
19	DVB_ASI mapping into ODU0 (Note 4)
20	ODU multiplex structure
21	OPU2, OPU3 1.25 Gbit/s tributary slot multiplex structure (Note 3)
55	Only present in ODUk maintenance signals
66	Only present in ODUk maintenance signals
80-8F	Reserved codes for proprietary use
FD	NULL test signal mapping
FE	PRBS test signal mapping
FF	Only present in ODUk maintenance signals

NOTES:

Experimental mappings can be proprietary to vendors or network providers. If one of these mappings/activities is standardized by the ITU-T and assigned a code point, that new code point is used instead of the 01 code point.

G.Sup43 had recommended the payload type of code 87 for this mapping since it was experimental at that time.

Equipment capable of using 1.25 Gbit/ tributary slots will initially send PT=21. In order to maintain backward compatibility with legacy equipment that only supports 2.5 Gbit/s tributary slots, equipment that supports 1.25 Gbit/s tributary slots must be revert to using 2.5 Gbit/s tributary slots if the equipment at the other end is sending PT=20.

These payload types were provisionally accepted at the time this white paper was released, with formal acceptance expected in late 2011.

Table 4 Payload summary information for major OTN client signals

Client Signal	OPUk	Justification Method	Comment
1GE	OPU0	GMP	Uses GFP-T
CBR clients: client ≤1.238 Gbit/s	OPU0	GMP	Includes STM-1, STM-4, FC-100
CBR clients: 1.238 Gbit/s < clients ≤ 2.488 Gbit/s	OPU1	GMP	Includes FC-200
STS-48	OPU1	AMP or BMP	
STS-192	OPU2	AMP or BMP	
10GE LAN PHY	Extended OPU2	GFP Idles	
10GE LAN PHY	OPU2e	BMP	
Fiber Channel FC-1200	OPU2e	BMP	Uses TTT/GFP-T
STS-768	OPU3	AMP or BMP	
40GE	OPU3	GMP	Transcoded
100GE	OPU4	GMP	
CBR clients: client > 2.488 Gbit/s	OPUflex(CBR)	BMP	Includes FC-400, FC-800, IB-SDR, IB-DDR, IB-QDR
Packet clients	OPUk, k=0,1,2,3,4	GFP Idles	GFP-F
Packet client streams	OPUflex(GFP)	GFP Idles	GFP-F

Key	Legacy	New Standard	Extension to Legacy

5.1 CBR Client Signal Mappings

The different methods for mapping CBR signals into OTN are summarized in Table 2. The first two methods were defined in the original version of G.709 for carrying SDH/SONET client signals or signals at those rates. In addition, although not listed in Table 2, the original version of G.709 supported a non-specific client stream mapping, which could be used for a CBR stream at the rate of (i.e., synchronous to) the OPU payload area. These AMP, BMP, and non-specific client stream mappings are described first in this section as legacy mapping methods. The other two mapping methods were added to allow flexible support for the mapping of client signals into OPU0 and OPU4, and also for client signals with rates or frequency tolerances that did not fit well with OPU1, OPU2 or OPU3 payload rates. These more recent CBR client mapping methods are described in the second part of this section. Note: Only a single CBR mapping method is defined for each client signal into a particular OPUk. In this manner, legacy mappings are preserved rather than creating new options for them.

5.1.1 Legacy CBR Client Mapping Methods

The initial set of CBR client signal mappings defined for G.709 OTN were SDH STM-16, STM-64, and STM-256 (SONET STS-48, STS-192, and STS-768), which are referred to as CBR2G5, CBR10G, and CBR40G, respectively. The CBR2G5, CBR10G, and CBR40G are in turn respectively mapped into the OPU1, OPU2, and OPU3. The OPUk payload area structures associated with these mappings are shown in Figure 9 where D indicates a payload data byte and FS indicates a fixed stuff byte. The mapping methods for these signals are:

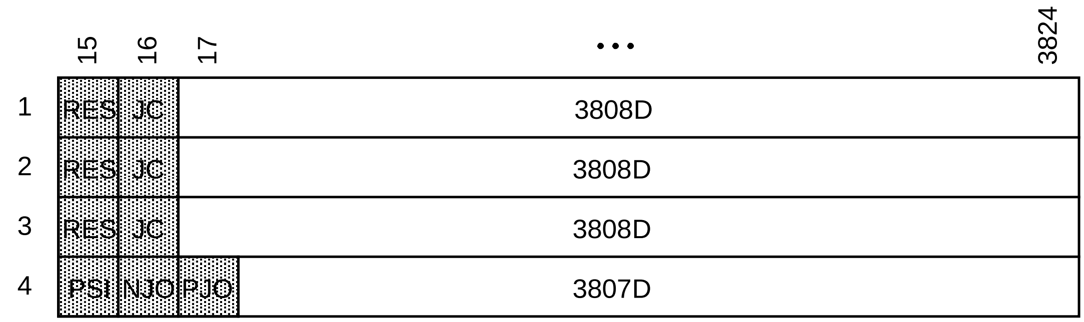
- **Asynchronous mapping procedure (AMP)**
- **Bit synchronous mapping procedure (BMP)**

As shown in Figure 9, the ODU2 and ODU3 signal payload areas include additional fixed stuff columns that are used when SDH signals are mapped into them. The purpose of the fixed stuff columns is to increase the OPU rate enough to accommodate multiplexing lower rate ODUs into it with the fixed stuff columns removed[5].[6]

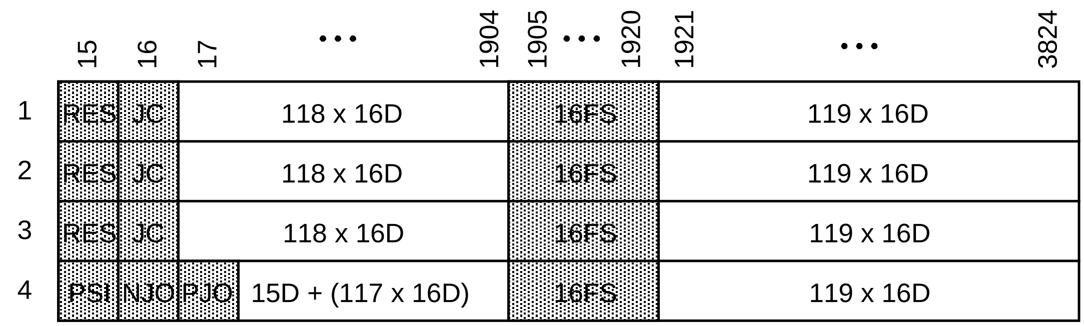
[5] Alternatively, one can say that an OPUk rates was defined in order to accommodate multiplexing the LO ODU signals into its payload container. The fixed stuff bytes are required in order to reduce the OPUk container rate to be close enough to the nominal STM-*n* rate that the AMP periodic justification opportunities can handle the remaining frequency offsets.

[6] For example the rate of an ODU1 signal is (239/238)(STM-16 rate), with the 239/238 factor accounting for the ODU1 and OPU1 overhead. In order to carry four ODU1 signals, the OPU2 payload rate must be at least (4)(239/238)(STM-16 rate) = (239/238)(STM-64 rate). The actual ODU2 rate is (239/237)(STM-64 rate) = (239/238)(238/237)(STM-64 rate). The 239/238 factor here accounts for the ODU2 and OPU2 overhead, and the 238/237 factor accounts for the additional OPU2 payload bandwidth required to carry the ODU1 overhead of the four ODU1 signals. When an STM-64 is mapped into the OPU2, the 16 fixed stuff columns fill the bandwidth corresponding to this 238/237 factor. Similarly, the OPU3 includes enough fixed stuff columns to accommodate carrying 16 ODU1 signals.

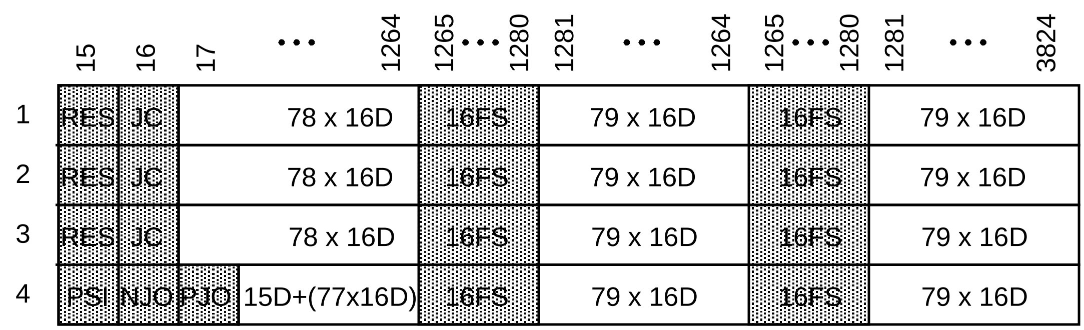
Figure 9 Mapping of CBR (SONET/SDH) signals into OTN



a) CBR2G5 mapping into OPU1



b) CBR10G mapping into OPU2



c) CBR40G mapping into OPU3

Legend:		
NJO = Negative Justification Opportunity byte	JC = Justification Control byte	D = payload Data byte
PJO = Positive Justification Opportunity byte	FS = Fixed Stuff Byte	RES = Reserved byte

Asynchronous mapping procedure (AMP)

With AMP, the OPU clock is generated locally. The adaptation between the OPUk payload rate and the client signal rate is performed through the use of the justification control (JC) bytes and their associated Negative Justification Opportunity (NJO) and Positive Justification Opportunity (PJO) bytes, as described in section 6.2.1.

The primary drawback to AMP is that it can only accommodate a limited range of client signal rates and frequency tolerances. Different client signal rates could be accommodated by defining an appropriate number of fixed stuff bytes for each type of client signal, however this needs to be defined for each new client signal and the column widths will typically not be a regular pattern across the frame[7]. The AMP technique can accommodate a client frequency range of up to ±45 ppm. This range is more than adequate for SDH signals, which typically have a frequency range of ±4.6 ppm, with ±20 ppm for their AIS signals. Additional NJO and/or PJO bytes would be required to accommodate wider client frequency tolerances such as the ±100ppm that is typical for Ethernet and other common data signals. The GMP method described in sections 5.1.2 and 6.2.3 was developed to provide a simpler mechanism to handle arbitrary client signal rates and clock tolerances in a consistent manner without unique per-client signal definitions.

Bit-synchronous mapping procedure (BMP)

With BMP the OPU clock is derived from the client signal clock (e.g., CBR10G signal). Because the OPU is frequency and phase locked to the client signal, there is no need for frequency justification. The JC bytes contain fixed values, the NJO contains a justification byte, and the PJO contains a data byte. Note that with BMP, the client signal data's relationship to the OPU payload contain can be bit-aligned or byte aligned. For most clients, including SONET/SDH and ODUflex(CBR), byte alignment is used.

BMP may be used for mapping SDH STM-16, STM-64, or STM-256 signals into ODU1, ODU2, or ODU3, respectively. When an STM-16 is bit-synchronously mapped into an OPU1 payload container, the STM-16 rate is used directly for OPU1 container rate. The resulting ODU1 rate is:

$$\text{ODU1 rate} = (239/238)(\text{STM-16client rate})$$

When an STM-64 is bit-synchronously mapped into an OPU2, the resulting ODU rate takes into account the presence of the OPU2 Fixed Stuff columns that are used for the mapping. As a result, the ODU2 rate is:

[7] The legacy fixed stuff columns had a regular spacing between groups of columns, and each group of columns had the same width that it maintained in all rows. It is typically not possible to define such a regular fixed stuff structure for new client signals. For example, the groups of fixed columns may need to have different widths, and one of the fixed column groups will typically require a different width in different rows.

ODU2 rate = (239/238)(STM-64 client rate + fixed column rate) =

= (239/237)(STM-64 client rate)

The BMP STM-256 into ODU3 mapping similarly takes the Fixed Stuff columns into account.

As discussed in section 5.1.3, BMP is also used for all CBR client signal mappings into an ODUflex(CBR). (Note that the ODU2e is a special case of ODUflex(CBR) that is used for carrying 10GBASE-R Ethernet clients, and retains the same fixed stuff columns that are used for STM-64 into ODU2.) As is explained later in this section, the ODUflex(CBR) signals are then multiplexed into a HO ODU using GMP for rate adaptation.

Non-specific client stream mapping

In addition to the CBR2G5, CBR10G, and CBR40G, G.709 also allows for mapping a non-specific client bit stream into the OPU. In this mapping, a client signal (or set of client signals) is encapsulated into a CBR stream at the rate of (i.e., synchronous to) the OPU payload area. Any rate adaptation must be performed within the CBR bit stream as part of the process that creates it.

5.1.2 CBR Signal Mapping Using GMP

As explained above, AMP is limited in its ability to conveniently accommodate new client signals with rates that are significantly different than the OPUk payload rates, and handle client signals with a clock tolerance significantly greater than ±20 ppm. The adoption of ODU0 and ODU4 created the opportunity to define both signals with a more flexible mapping method. This method, GMP, was also chosen for use when mapping new client signals into OPU1, OPU2, and OPU3 signals.

The concept behind GMP is that the JC bytes of each frame are used to communicate the number of payload words that will be mapped into the OPUk payload area during the next frame. The transmitter and receiver use modulo arithmetic based on this count value to determine the location of payload and stuff words within the payload area of the frame. See section 6.2.3 for a complete description of the GMP process, including some specific mapping examples. The word size is an integer number of bytes. For the purposes of mapping, the word size is equal to the maximum number of Tributary Slots (TS) that could be supported by that OPUk in a multiplexed payload. Specifically, ODU0, ODU1, ODU2, ODU3 and ODU4 word sizes of 1, 2, 8, 32 and 80 bytes, respectively.

The flexibility of GMP comes from it being able to accommodate a CBR client with a rate that needs between one and 15232 bytes per OPU frame. It also inherently allows great flexibility in accommodating different client frequency clock tolerances; although in practice most CBR client signal clocks are no worse than ±100ppm. In the case of ODU0, this allows mappings including the 155.52 Mbit/s STS-3/STM-1 and, as discussed in 5.3.1. Gigabit/s Ethernet (GE).

5.1.3 ODUflex(CBR) Mappings

CBR signals with a rate lower than OPU0 or OPU1 are mapped directly into an OPU0 or OPU1, respectively, using GMP. Similarly when the client signal rate is only somewhat lower than of an OPU2, OPU3 or OPU4 it is mapped directly into that signal. However, many important CBR client signals do not fit well into an OPUk. Examples of such clients include Fiber Channel signals with rates over 2 Gbit/s. The ODUflex(CBR) is a flexible rate ODU was created to simplify mappings for arbitrary-rate CBR client signals.

The ODUflex(CBR) is created by wrapping the client signal with an OTN frame. In other words, a BMP process is used that simply adds OTN overhead to the client signal stream to create a signal with a rate that is exactly (239/238)(client rate). The ODU rate is thus directly derived from the client signal rate. The ODUflex(CBR) signals are then multiplexed into the tributary slots of an OPU2, OPU3 or OPU4.

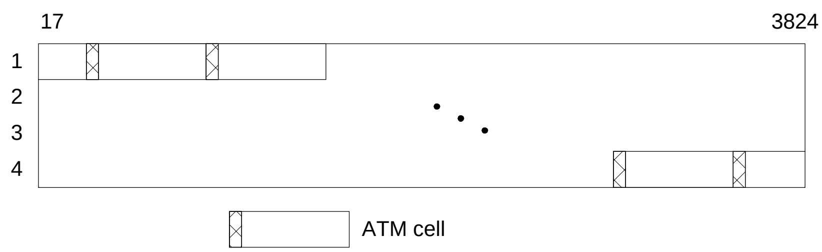
5.1.4 Mappings into ODU4

The ODU4 signal was defined somewhat differently than the ODU1, ODU2 and ODU3, which were optimized for carrying SDH payloads. In order to avoid the compatibility issues associated with 10GE and the OTN, the ITU-T defined the ODU4 such that it is optimized for carrying the 100GE signal rather than following its traditional approach of defining it as four times the capacity of the next lower signal (i.e., 160 Gbit/s = 4 x ODU3). The assumption was that 100GE would become the most important signal at that data rate, and the hope was that the OTU4 will become the preferred method for serial transmission of 100GE whenever FEC is required. IEEE initially focusing on shorter reach multi-fiber or multi-wavelength physical layer options that do not require FEC. The OTU4 rate of around 112 Gbit/s also allowed using optical technology that was substantially less complex and expensive than would have been required for a 160 Gbit/s interface. The drawback was that the resulting OPU4 is not a clean integer multiple of lower rate ODUk signals. An OPU4 structure was chosen that consists of 80 tributary slots with a rate of approximately 1.25 Gbit/s per tributary slot. (See section 9.1 for a description of the tributary slots and their rates.) GMP was chosen as the justification mechanism for all non-GFP mappings into the OPU4.

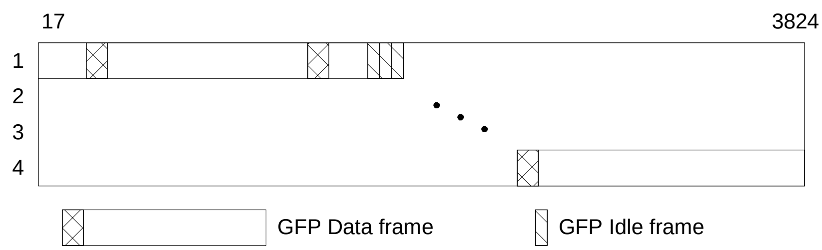
5.2 GFP and ATM Mapping

Direct mappings for frame-mapped GFP (GFP-F) frames and ATM cells into the OPU payload area are also defined[8]. In these mappings, a continuous stream of GFP-F frames or ATM cells are mapped in an octet-aligned manner into the whole OPU payload area with no OPU fixed stuff columns. The mapping is illustrated in Figure 10. The delineation of the GFP frame or ATM cell boundaries is performed using the header information of these protocols. Rate adaptation between the client and the OPU payload rates is performed by sending GFP Idle frames or ATM Idle cells when there is no data to send.

Figure 10 Mapping for GFP frames and ATM cells into the OPU



a) ATM cell mapping into OPUk payload area



b) GFP frame mapping into OPUk payload area

---

[8] Transparent GFP (GFP-T) frames could also be accommodated through this mapping. However, most of the mappings that use GFP-T create a CBR stream of GFP-T frames that is then handed as a CBR client signal for the purposes of rate adaptation. See section 5.3.1.

5.2.1 GFP-F Mapping into an Extended OPU2

A special case GFP-F mapping exists for carrying 10GBASE-R Ethernet client signals. The GFP mapping process was extended to accommodate transparency for the Ethernet frame preamble and Ordered Set information, and some of the OPU2 overhead bytes are used to create an extended payload area. See section 5.3.2 for the details of this mapping.

5.2.2 ODUflex(GFP) Mapping

The ODUflex(GFP) mapping is used to provide a fixed-rate channel for carrying packet-oriented client data, with flexibility in choosing the channel size. This mapping is essentially a hybrid of the ODUflex(CBR) and GFP-F mapping methods described above. As with all other GFP-F mappings, it encapsulates client data frames with GFP-F, maps the GFP-F frames into the OPU (i.e., the OPUflex here), and fills any unused OPU bandwidth with GFP Idle frames. There are some special considerations with ODUflex(GFP), however, that merit treating it separately in this section.

As discussed in section 5.1.3, the ODUflex(CBR) rate is derived directly from the client signal rate. In contrast, the ODUflex(GFP) rate is predetermined for efficient use of the OPU payload area. Although the approaches to deriving the rates are different, both types of ODUflex are multiplexed into the HO OPU by GMP, and are handled in the same manner within the network. Consequently, there is no need for intermediate switching nodes to know which type of ODUflex they are switching or multiplexing.

The ODUflex(GFP) rate is determined as follows. As discussed in section 9.1, the smallest bandwidth granularity within an OPU is the 1.25Gbit/s tributary slot (TS). The network provider chooses a service rate for this mapping such that the resulting ODUflex(GFP) uses the bandwidth available in the *N* tributary slots. In other words, the service rate for clients mapped into ODUflex(GFP) is nominally *N* x 1.25Gbit/s. As discussed in section 9.1, the available bandwidth in these tributary slots must take into account the potential variations in TS rate that the ODUflex(GFP) could potentially encounter within the network. See Section 6.2.3, including Table 10 for the exact rates and the explanation of how they are determined.

5.3 Ethernet Client Signals

Ethernet has become very important for both enterprise customer WAN interfaces and as an emerging telecom network infrastructure technology. Consequently, Ethernet signals are becoming increasingly important client signals for OTN. For that reason, the Ethernet signal mappings into OTN are discussed in their own section of this white paper.

Ethernet was identified as a potential OTN client signal during the initial OTN standards development. However, the decision was made at that time to not directly support Ethernet mappings into the OTN[9]. It appeared that there were acceptable alternatives to map Ethernet signals into SONET/SDH signals, which could then be mapped into OTN. In retrospect, this was an unfortunate decision. The lack of standard Ethernet client mappings delayed the widespread use of OTN for Ethernet clients and complicated both OTN equipment and networks through the introduction of multiple non-standard mappings. This situation was resolved in late 2008 for 1 Gigabit/s Ethernet (GE) clients with specification of a mapping for GE into the new ODU0 that was optimized for carrying GE clients. Transport over OTN was addressed as an integral part of the development of the recent 40 Gigabit/s (40GE) and 100 Gigabit/s (100GE) standards. IEEE 802.3 agreed to specify these interfaces in a manner that would be ‘friendly’ to OTN, and the ITU-T defined the mappings for 40GE and 100GE into OTN.

The GE mapping into ODU0 is described in 5.3.1, and the 40GE and 100GE mappings are discussed in 5.3.4 and 5.1.4, respectively. Since it was not possible to define a single 10 Gigabit/s Ethernet mapping that fulfilled the requirements of every carrier, a compromise was reached to standardize multiple mappings. The 10GE mappings are discussed in 5.3.2.

5.3.1 Gigabit/s Ethernet (GE)

Gigabit/s Ethernet (GE) client signals are becoming increasingly important in telecommunications networks. The emerging applications for GE include:

- GE as a UNI for enterprise customers;
- GE as an interface to broadband access equipment (e.g., IP-DSLAM, PON OLT, and wireless base station); and
- GE interconnections for using Ethernet as a metro network switching technology.

GE client signals were initially carried by using one of the standard GFP-F or GFP-T mappings into SONET/SDH STS-48/STM-16 signal and mapping the SONET/SDH signal into an ODU1. The main drawback to this method is that it requires maintaining a complex SONET/SDH TDM layer in the network just for the transport of the packet-oriented Ethernet clients. Ideally, the transport could be greatly simplified by eliminating the SONET/SDH layer for these packet client signals. Some equipment vendors and silicon vendors, such as PMC-Sierra, have developed methods to combine two GE signals in an ODU1 without using a SONET/SDH layer. However, these methods are only effective for point-to-point links with the same vendor’s equipment on each end. Extending the capabilities of these methods to support switching capability within the OTN would add considerable cost and complexity to the equipment and the network.

---

[9] As noted above, bandwidth limitations of undersea carrier systems were a major factor in this decision. At the time, carriers were more concerned about universal OTN deployment than with carrying native 10GE LAN signals. They wanted their land and undersea links to be compatible and part of the same OTN.

A substantial number of carriers requested a standard method for GE transport over the OTN that was efficient (i.e., supported two GE clients within an ODU1 bandwidth), supported switching within the OTN electrical domain (e.g., did not require a SONET/SDH layer), maintained maximum compatibility with the existing OTN, provided transparency down to the character and timing levels of the GE client signal, and allowed simple and economical equipment and network implementations. In order to meet this carrier demand, in 2008 the ITU-T standardized an ODU0 structure optimized for GE clients that meets these requirements. The GE into ODU0 mapping is described as follows.

The 1.25 Gbit/s line rate of the 8B/10B-encoded GE signal exceeds the OPU0 payload capacity. Consequently, transparent mode of GFP (GFP-T) is used to adapt the GE signal into the OPU0 such that character-level transparency of the GE signal is maintained. [17], [18] While GFP-T has a built-in method for rate adaptation between the GE client and the transport payload container rates, a more general rate adaptation mechanism was chosen that can also be used for non-GE clients. This rate adaptation mechanism allows timing transparency for the client signal and is applicable to any CBR client signal with a bandwidth less than the OPU0 payload bandwidth. The mapping method for GE into OPU0 can be summarized as follows:

1. Adapt the incoming GE signal into GFP-T:
   - Transcode the incoming GE 8B/10B characters into 64B/65B code blocks,
   - group eight 64B/65B blocks into a 67 byte superblock, and
   - map one superblock into a GFP frame, with no 65B_PAD or GFP Idles.
2. Map the resulting CBR stream of GFP data frames into the OPU0 using GMP for rate adaptation.

Each GFP frame consists of one 67-byte superblock and 8 total bytes of GFP frame overhead (i.e., the Core and Type headers). As a result, the data rate of the CBR stream of GFP data frames is:

$$\left(\frac{1.25Gbit}{s}\right)\left(\frac{8b}{10b}\right)\left(\frac{(67+8)bytes}{64bytes}\right)=\left(\frac{1.25Gbit}{s}\right)\left(\frac{15}{16}\right) \quad [1]$$

This fixed ratio between the GE client signal rate and the GFP stream rate preserves the client signal timing information. Consequently, this method is referred to as a Timing Transparent Transcoding (TTT) mapping. The use of a TTT mapping for GE allows full transparent support for G.8261 Synchronous Ethernet (SyncE) GE signals.

5.3.2 10 Gigabit/s Ethernet (10GE) over 10 Gbit/s OTN

Transporting 10GE client signals has become one of the largest opportunities for OTN, and has been one of its most contentious problems. As noted above, the maximum rate of the OTU2 signal was established based on the limitations of undersea cable links. Unfortunately, this rate was less than the native rate of the 10GE LAN signal (10.3125 Gbit/s). Two standard methods were initially developed for 10GE transport over OTN; however, neither method adequately satisfied all requirements of all carriers. Consequently, additional methods were developed to address specific carrier applications. Eventually, the most popular non-standard methods were documented in ITU-T supplementary document G.Sup43[10], and two of these methods were subsequently moved into the G.709 standard at the end of 2008. This section describes the different 10GE into 10 Gbit/s OTN mapping methods, in roughly the chronological order in which they were standardized.[11] The most popular mappings are the statistical approach, one of the overclocked approaches, and the method using an extended GFP with a modified OTN frame format. The next section (5.3.3) describes the methods for multiplexing 10GE signals into 40 Gbit/s OTN signals.

Some background is required to understand what led to the variety of 10GE mappings into OTN. From an IEEE 802.3 perspective, the Ethernet information consists of the Ethernet MAC frames. The preamble and inter-frame gap (IFG) characters are not regarded as part of the Ethernet information. The preamble was originally used to allow receivers to synchronization to a new data frame in networks that used a shared transmission medium. Since only full-duplex operation is specified for 10GE, the preamble is unnecessary. However, it was included in the 10GE signal specification for the sake of commonality with previous lower-rate Ethernet interfaces. Since the preamble information is ignored by a 10GE receiver, some equipment vendors “borrowed” some of the preamble bytes in order to create a proprietary OAM channel. Because the initial Ethernet mapping into GFP-F only included the Ethernet MAC frame bytes, this preamble-based OAM channel would not be carried across a GFP-F link.[12] In addition, some carriers have defined a proprietary OAM channel that uses Ethernet Ordered Sets as part of the IFG in place of Idle characters. As a result, some carriers demanded effective bit/character transparency for the entire Ethernet PHY signal. Other carriers, however, insisted that the mapping must not change the bit rate or frame structure of the OTU2. Unfortunately, these requirements are mutually exclusive. A further complication was the desire for full bit-transparency in order to carry Synchronous Ethernet (G.8261) physical layer timing information across the OTN link.

---

[10] Both the standard and non-standard mappings are described in G.Sup43.

[11] Note that in addition to the options discussed here, another alternative for carrying the 10GE LAN signal over OTN is in a container of five concatenated ODU1 signals. This is an inefficient mapping that is not popular with carriers.

[12] At this point in time, the use of the preamble for OAM information is largely historic. In practice, Ethernet OAM is carried in Ethernet OAM frames defined by the IEEE and ITU-T for carrier applications.

10GE WAN (10GBASE-W and ITU-T G.709 Derivative)

This method was the first standard for 10GE transport over OTN. Carriers sent representatives to IEEE 802.3 during the development of the 10GE standard with the hope of defining a signal that could be readily transported over the SONET/SDH-based WAN as well as the LAN. Since 802.3 was primarily concerned with LAN signals, they preferred to adopt 10 Gbit/s in keeping with their tradition of increasing their MAC data rates by a factor of 10. The resulting compromise was separate 10GE LAN and WAN signal rates and formats. The 10GE LAN signal has a MAC data rate of 10 Gbit/s. The 64B/66B block code was chosen as the typical line code for serial transmission, resulting in a 10.3125 Gbit/s PHY signal rate. The signal format is essentially the same as all Ethernet LAN signals in that it is a stream of Ethernet frames that begin with a preamble, and with a minimum number of IFG characters between the Ethernet data frames. The 10GE WAN signal was defined to have an outer TDM frame with the same rate and format as a SONET STS-192c (SDH VC-64c), with a minimum amount of the SONET/SDH overhead active. The 64B/66B characters were mapped directly into the payload envelope/container. This 9.58464 Gbit/s signal is also known as the 10GE WAN-PHY. The hope was that this signal could then be carried either directly over a SONET/SDH network, or mapped into an ODU2 for transport over OTN.

There were two primary problems with the 10GE WAN signal. The first concerns the signal's clock accuracy requirements. Ethernet signals have typically specified ±100ppm clock accuracy for the transmitted signals. SONET/SDH, however, requires ±4.6ppm clock accuracy. Lower accuracy clocks can lead to excessive pointer adjustments in the network that trigger network alarms. As a compromise, the IEEE ultimately agreed to specify a ±20ppm clock accuracy, which corresponds to the "SONET minimum clock" accuracy. This is still unacceptable for reliable data transport in carrier networks, so equipment vendors typically implement the interface with a ±4.6ppm clock. ITU-T specified its own version of the WAN interface that is essentially identical to the IEEE definition except for specification of a ±4.6ppm clock with tighter jitter and wander requirements.

The second problem is that for a variety of reasons, not all technical, 10GE LAN port units were typically significantly less expensive than WAN ports. Consequently, enterprise customers prefer using LAN interfaces for their connection to the carriers rather than the WAN interface.

Statistical Approaches Using GFP-F

In practice, it is rare for an Ethernet link to operate at its full rate for a sustained period. Consequently, it should be possible to map the 10GE LAN signal into an OPU2 by discarding the IFG characters between data frames. The standard approach to this type of mapping was to encapsulate the Ethernet frames with GFP and map the GFP frames into the OPU2. If required, some buffering could be used to handle peak rate bursts.[13]

While this approach is acceptable for most applications, unfortunately it is not acceptable to carriers that want to pass OAM information in the IFG portion of the signal. There is also a concern that user data frames can be lost due to congestion if the period at which the user transmits at the full 10GE rate lasts longer than the mapping buffers can handle. Carriers want to be able to offer premium services with guaranteed, deterministic performance.

"Overclocked" OTN

Since typical OTN signals do not go through undersea cable links, a simple alternative is to simply increase the rate of the OTN signal so that it can accommodate the 10GE LAN signal in the OPU. There are two versions of this approach. One version increases the rate of the ODU2 signal. Consequently, it is become known as the ODU2e (extended ODU2) signal. As shown in Figure 9, the OPU2 contains fixed stuff bytes for CBR10G mappings. This version was originally described in G.Sup43 section 7.1. The ODU2e was moved into the G.709 standard at the end of 2008, with the associated OTU2e description remaining in G.Sup43. The other overclocked approach eliminates the fixed stuff columns in order to minimize the increase to the signal rate. Since the elimination of the fixed stuff columns gives the same OPU structure as the OPU1, this approach is referred to as ODU1e. The ODU1e is described in G.Sup43 section 7.2, and will not be moved into G.709. Both ODU2e and ODU1e approaches essentially wrap an OTN frame around the Ethernet client signal. Consequently, both inherit the Ethernet ±100ppm clock tolerance and jitter/wander characteristics. Although it is not designated as such, the ODU2e is effectively the first defined ODUflex(CBR) signal.

Issues associated with multiplexing ODU2e (and ODU1e) into ODU3 are a significant drawback to this approach. These issues and their solutions are discussed in 5.3.3.

---

[13] The need for peak-rate buffering depends on multiple factors. If the maximum length of the Ethernet frames is limited to the 802.3 restriction of 2000 bytes and the preamble is discarded, the Ethernet frame information can fit within an OPU2. The use of 9600 byte jumbo frames and retaining the preamble information causes the client signal information to exceed the OPU2 capacity at the peak 10GE rate.

Extended GFP with Modified OPU2

Another approach uses GFP-F to effectively obtain a character transparency for 10GE that is similar to what GFP-T provides for GE. This approach was originally described in G.Sup43 section 7.3, and was moved into the G.709 and G.7041 standards at the end of 2008. In order to preserve any information encoded in the preamble bytes, it uses a different Ethernet frame mapping that includes the preamble bytes when the Ethernet frame is mapped into a GFP-F frame. This modified mapping into GFP-F, which uses a 0x13 User Payload Indicator (UPI), is illustrated in Figure 11.a. The /S/ character at the beginning of the preamble is replaced by a 0x55 pattern[14].

In order to preserve any Ordered Set information between Ethernet frames, the four bytes of each Ordered Set are mapped into a separate GFP-F frame, with a modification to the first byte (the /O/ character). This mapping, which uses the 0x14 UPI code, is illustrated in Figure 11.b. The result is that the relevant information from the 10GE physical layer can be reconstructed for the egress signal at the GFP sink.[15] Note that since this mapping operates on the physical layer signal, there is no MAC frame processing. For example, no error checking is performed on the Ethernet frame.

In spite of the removal of the inter-frame Idle characters, the resulting GFP stream does not quite fit within an OPU2. In order to provide the additional bandwidth, this mapping also modifies the OPU2 structure. As shown in Figure 12, seven of the eight OPU payload specific overhead bytes are used to carry payload.[16]

The key advantage to this approach is that the resulting ODU2 signal retains the same rate as all other ODU2 signals. Consequently, it does not require a separate OTU2 specification and is completely compatible with the OTN multiplexing hierarchy. These characteristics are very important for carriers that have already deployed a significant amount of OTN networks or want to maintain an efficient mix of Ethernet and SONET/SDH client transport within the same OTN.

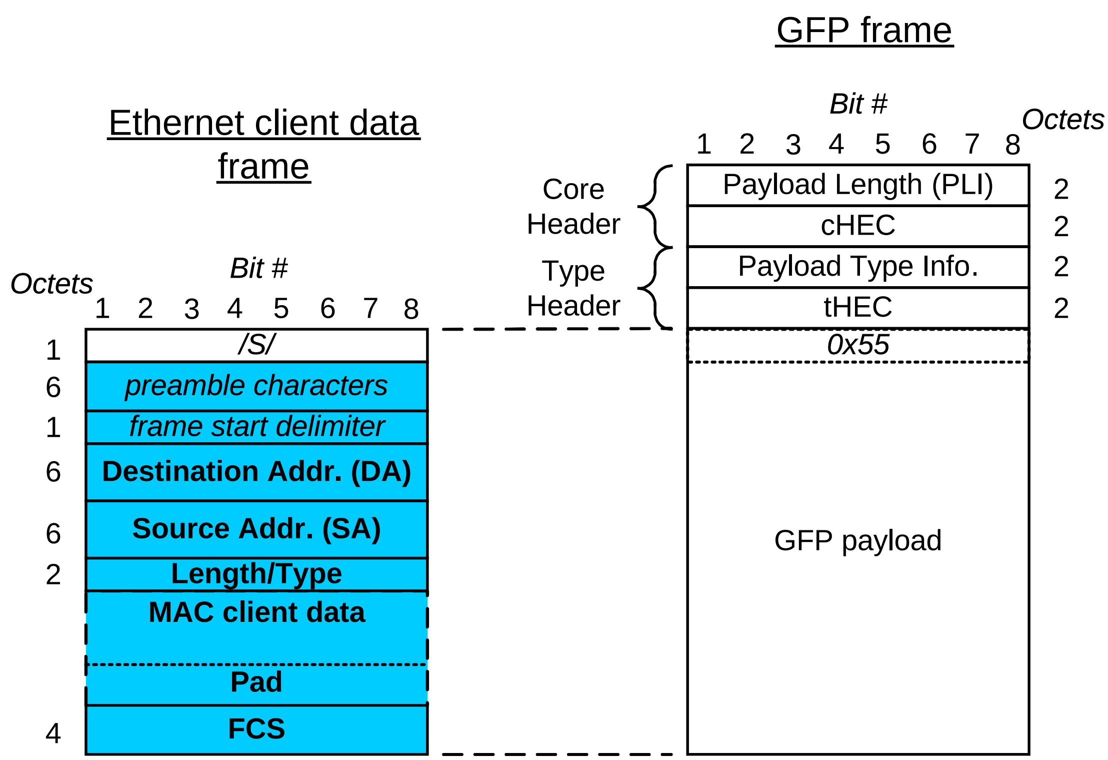
---

[14] Since the /S/ character is redundant information here, it would have been more bandwidth efficient to omit it from the GFP frame. It was apparently replaced with the 0x55 character for convenience of implementation.

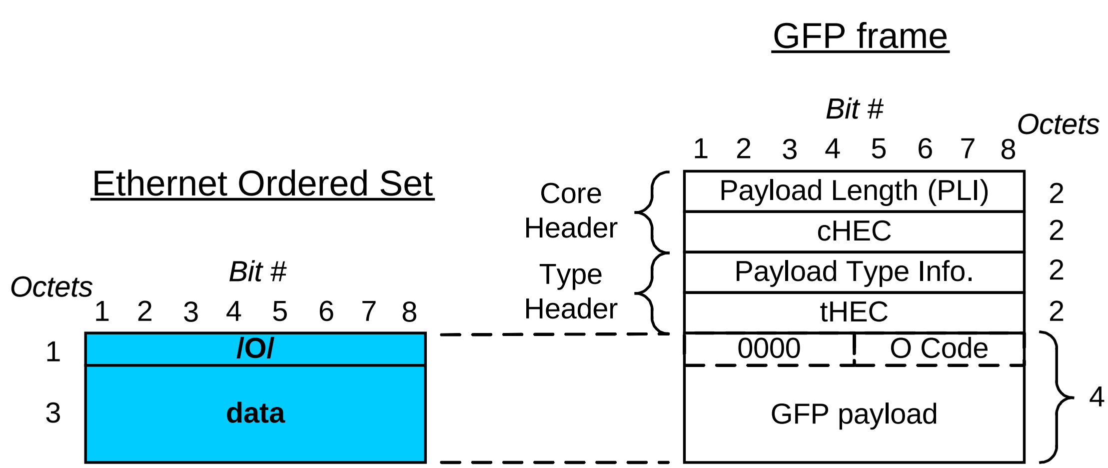
[15] The GFP frame overhead adds eight bytes to the Ordered Set. This bandwidth expansion is typically not an issue since Ordered Sets occur between data frames and there are typically enough Ethernet inter-frame Idle characters to make up for the additional GFP frame overhead bandwidth. The only cases where Ordered Sets are sent consecutively (back-to-back) are when they indicate a local or remote link failure. Due to the GFP mapping bandwidth expansion, only about one of every three of these Ordered Sets will be encapsulated in a GFP frame and transmitted. This is acceptable since these fault indication Ordered Sets may be treated like Idle characters in that they can be removed or inserted for rate adaptation.

[16] While it was argued that using these OPU overhead bytes was a “payload specific” use, using the overhead for data is a layer violation that creates potential compatibility problems with other framers handling ODU2 frames.

Figure 11 GFP mappings for extended GFP transport of 10GE signals (formerly G.Sup43 Section 7.3)

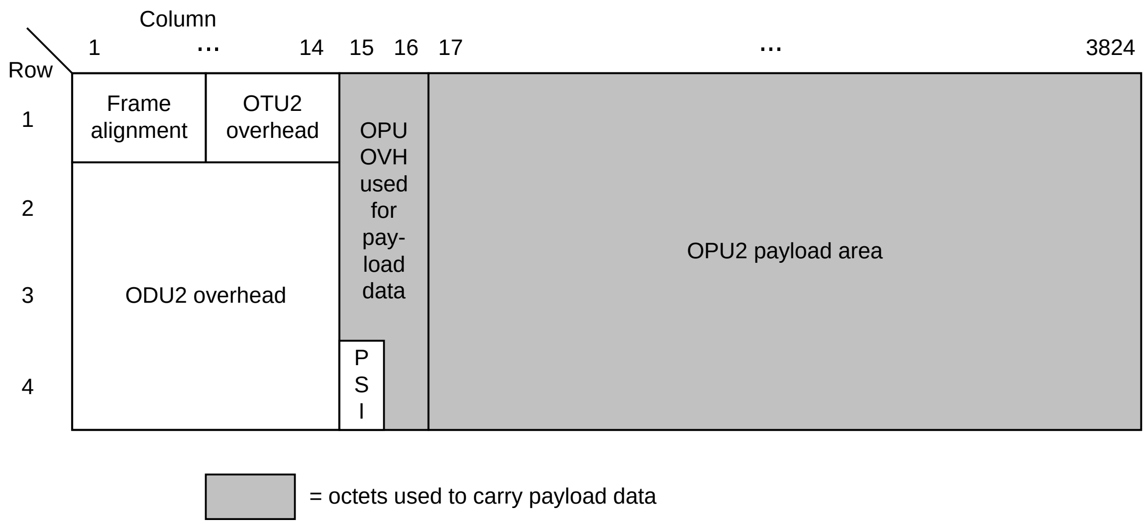


a) Modified Ethernet data frame mapping



b) Ordered Set mapping

Figure 12 Modified OPU2 for extended GFP transport of 10GE signals (formerly G.Sup43 Section 7.3)



5.3.3 10GE over 40 Gbit/s OTN

As can be seen in Table 1, the OPU3 payload container has a bandwidth that exceeds 40 Gbit/s. Hence, it would be possible to fit four 10GE signals into the OPU3 if either the 64B/66B line code was transcoded into a more efficient line code, or the Ethernet frames were mapped into GFP. For applications that require character or timing transparency, however, carriers preferred to simply multiplex the ODU2e signal into the 40 Gbit/s OTN signal. A central issue for ODU2e is that the ODU3 rate is too low to carry four ODU2e signals. Using an overclocked ODU3 (referred to as an ODU3e) introduces several other issues. First, the existing justification mechanism lacks the frequency accommodation range to multiplex both ODU2 and ODU2e signals into an ODU3e. Hence, carriers would be forced into the undesirable situation of requiring separate OTNs for Ethernet clients and SONET/SDH clients. Further, even if the ODU3e was limited to ODU2e clients, the ±100ppm clock range of the ODU2e is beyond the capability of the existing OTN justification mechanism.

As with 10GE transport over 10Gbit/s OTN signals, no single method for carrying ODU2e signals over 40Gbit/s OTN satisfies every carrier’s requirements. The ITU-T addressed multiplexing ODU2e into 40Gbit/s OTN signals in multiple ways.

The standard mapping for an ODU2e into an ODU3 signal is to multiplex it into nine 1.25G tributary slots of the ODU3[17]. GMP is used for the rate adaptation, which allows the mapping to easily accommodate the ±100ppm clock range that the ODU2e signal inherits from the 10GE client it carries. This mapping allows an ODU3 to carry up to three ODU2e signals, and to carry a mixture of ODU2 and ODU2e clients. While this mapping lacks some bandwidth efficiency, it is favored by several carriers with substantial OTN deployments. These carriers prefer using non-overclocked approaches as their typical method for 10GE client transport, and see applications requiring full bit transparent transport as rare enough that the mapping inefficiency is not important.

To address the needs of carriers that wanted more bandwidth efficient ODU2e transport, the ITU-T added two overclocked ODU3 descriptions to G.Sup43 at the end of 2008. These options are referred to as ODU3e1 and ODU3e2.

ODU3e1

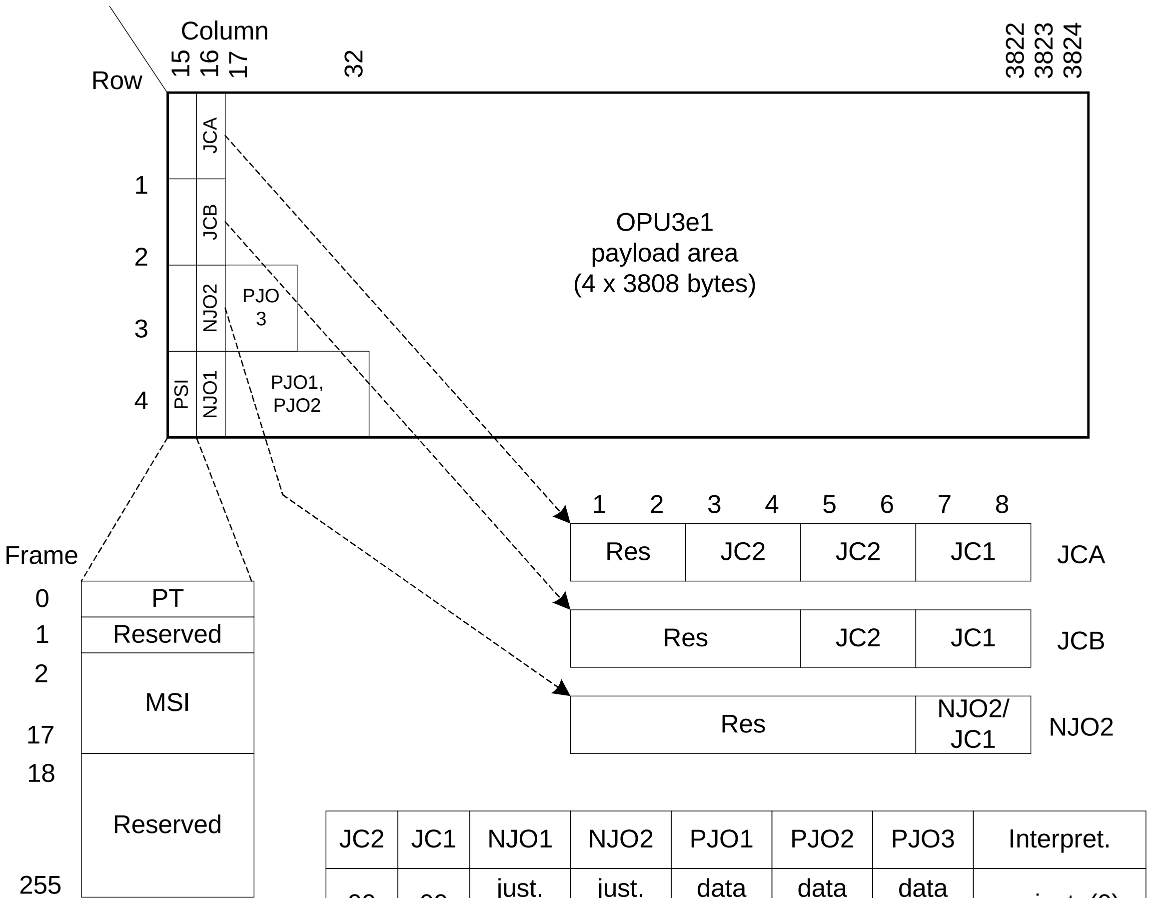
The ODU3e1 is designed to carry four ODU2e signals as its only client. The frame format and multiplexing techniques for the ODU3e1 are similar to the ODU3 with the following exception. In order to accommodate the ±100ppm clock tolerance of the ODU2e clients, the ODU3e1 justification mechanism has been extended with an additional PJO and NJO byte and a corresponding JC-byte format to use them. The modified frame structure and JC byte definitions are illustrated in Figure 13.

The ODU3e1 bit rate is (ODU2e rate)(4)(239/238) = 41.774364407 Gbit/s ±20ppm.

The carriers that prefer ODU3e1 are those that have made extensive use of Ethernet as a network technology and use 10GE over ODU2e in much of their transport networks. Since ODU2e is their primary 10 Gbit/s OTN signal, they need an efficient means to carry four ODU2e signals over their 40 Gbit/s OTN links and are not as concerned about also carrying ODU2 signals. NTT has already deployed ODU3e1 in its network.

[17] See section 9.1 for a description of the 1.25G tributary slots.

Figure 13 ODU3e1 frame structure and justification control



JC2	JC1	NJO1	NJO2	PJO1	PJO2	PJO3	Interpret.
00	00	just. byte	just. byte	data byte	data byte	data byte	no just. (0)
00	01	data byte	just. byte	data byte	data byte	data byte	Neg. just. (-1)
00	10	just. byte	just. byte	just. byte	just. byte	data byte	double Pos. just. (+2)
00	11	just. byte	just. byte	just. byte	data byte	data byte	Pos. just. (+1)
01	-	data byte	data byte	data byte	data byte	data byte	double Neg. just. (-2)
11	-	just. byte	just. byte	just. byte	just. byte	just. byte	triple Pos. just. (+3)

ODU3e2

The ODU3e2 bit rate is (239/255)(243/217)(16)(2.488320Gbit/s) ≈ 41.78596856 Gbit/s ±20ppm with a corresponding OPU3e2 rate of 41.611131871 Gbit/s. The 3808 columns of the OPU3e2 payload area are divided into 32 tributary slots. The method for multiplexing client signals into these tributary slots is the same GMP that is specified for multiplexing client signals into ODU4.

While G.Sup43 only addresses transport of 10GE clients, carriers who prefer ODU3e2 see it a potential universal 40 Gbit/s OTN signal. Since it uses the same mapping as ODU4, it is capable of carrying any lower-rate client. These lower rate clients include ODU3, ODU2e, ODU2, ODU1, and ODU0. For example, the ODU3e2 can carry four ODU2e clients, a combination of *n* ODU2e and 4-*n* ODU2 clients, and various other arbitrary combinations of lower rate ODUk signals with a total bandwidth less than the OPU3e2 capacity. These carriers typically do not have a large embedded OTN network and want to avoid the multiple network issues by deploying ODU3e2 as their only 40 Gbit/s OTN signal.

5.3.4 40 Gigabit/s Ethernet (40GE)

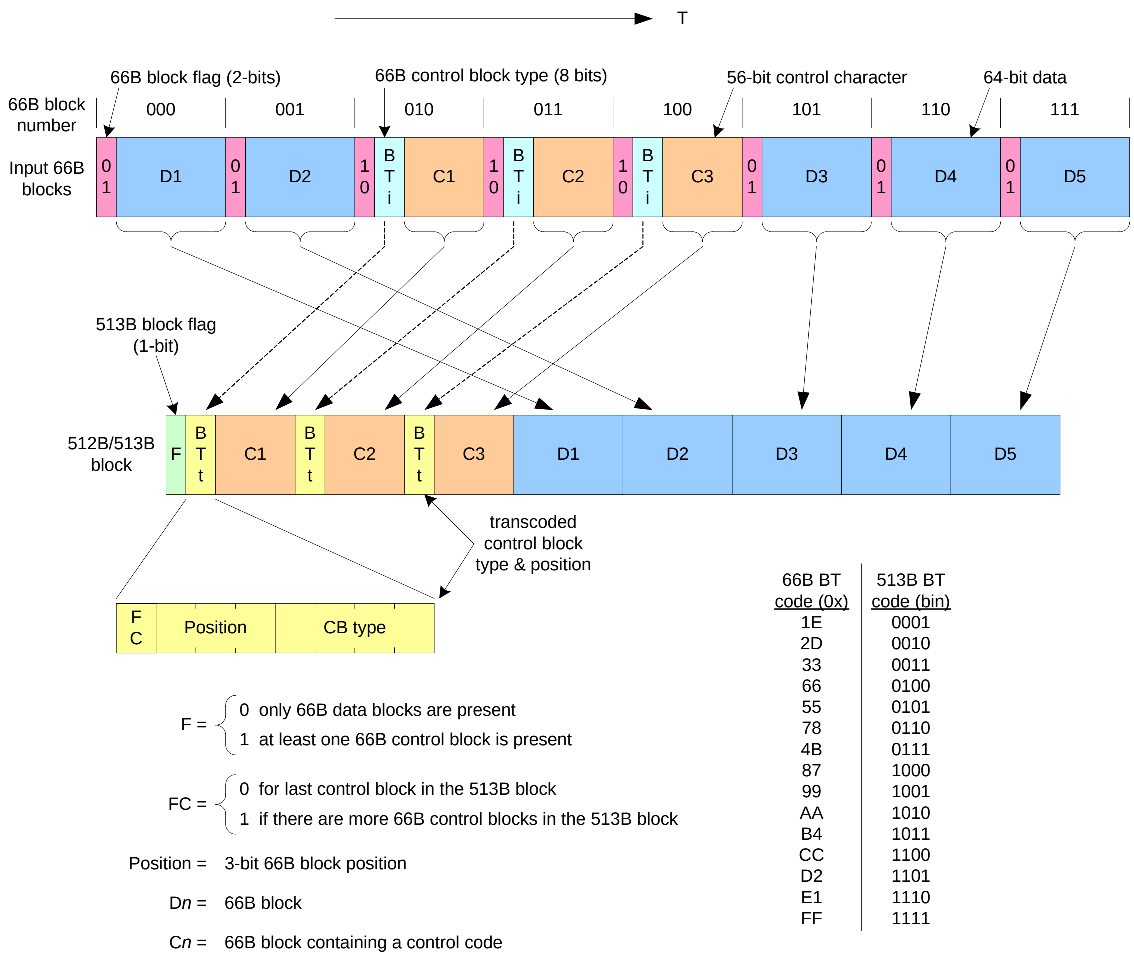
The IEEE 802.3ba working group created an Ethernet interface standard with a MAC data rate of 40 Gbit/s. There was close liaison between ITU-T SG15 and IEEE 802.3ba in order to avoid the type of incompatibility issues that occurred with 10GE and OTN. Fortunately, since the OPU3 payload rate (40.150519 Gbit/s) is greater than 40 Gbit/s, there were more options for finding a solution that achieved full character-level and timing transparency without using an overclocked ODU3. The key points of the resulting 40GE LAN signal and its mapping into OTN may be summarized as follows:

- The 40GE LAN interface uses the same 64B/66B line coding as 10GE, which results in a 41.25 Gbit/s serial rate.
- For transport over ODU3, the 64B/66B blocks are transcoded into a more efficient 1024B/1027B block code, which results in a client signal rate of 40.117088 Gbit/s.
- The 40.117088 Gbit/s stream is mapped into a standard-rate OPU3 using GMP.

The 1024B/1027B block code is constructed as a concatenation of two 512B/513B block codes, with an additional synchronization bit added as a parity check over the flag bits of the two 512B/513B blocks. The 512B/513B block construction is illustrated in Figure 14, and the concatenation to create the 1024B/1027B block is illustrated in Figure 15.

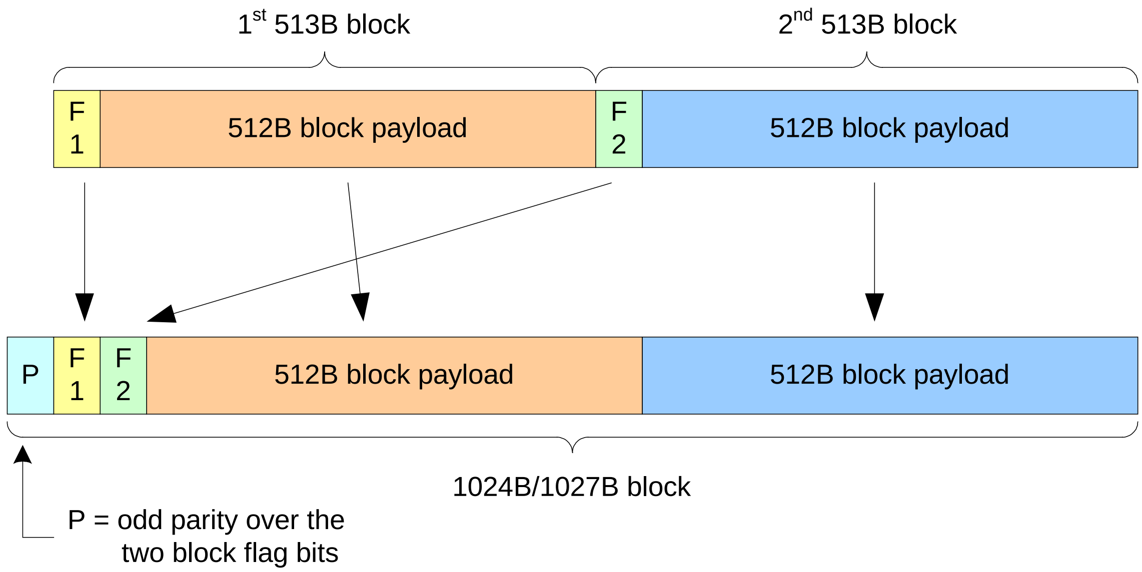
The 512B/513B block code is an extension of the transcoding technique used to create the 64B/65B block code of GFP-T. The 513B block flag bit indicates whether any 66B control characters are present in the block. The control characters are moved to the beginning of the block, with each preceded by fields that indicate the control character’s original location within the input stream, a compact encoding of the control character type, and a flag continuation (FC) bit to indicate whether this is the last control character in that 513B block.

Figure 14 512B/513B block construction



In general, the bandwidth efficiency of block codes is increased by reducing the amount of redundant information contained within the codes, which in turn increases their potential vulnerability to undetectable bit errors. Various techniques to detect corrupted 512B/513B codes are described in G.709 Appendix VIII. An error that corrupts the 513B flag bit can lead to the corruption of a substantial amount of data, which can be difficult to detect. The OPU3 bandwidth is not sufficient to use a pair of flag bits per block, as is done with 64B/66B. Consequently, the structure of Figure 15 was adopted. Two 513B blocks are concatenated such that their flag bits are moved to the beginning of concatenated pair with an odd parity bit added to cover the flag bits.

Figure 15 1024B/1027B block construction



The initial IEEE 802.3 definition for the 40GE interface is a parallel interface using four parallel lanes at 10 Gbit/s. Each lane uses 64B/66B characters, and includes a periodic special 66B control block. This special control block serves a lane alignment marker, to identify the lane, and is also used by the receiver for timing de-skew between the lanes. All four lane alignment marker control block types use the same control block type and are distinguished by the content of the data portion of the block. For transmission over ODU3, the characters from the four lanes are reassembled into their proper order in a serial stream. The lane alignment marker control blocks are preserved in their proper locations by mapping them into the 1024B/1027B blocks in the same manner as other 66B control blocks. The ODU3 receiver can then directly de-interleave the characters into a parallel interface with the lane markers again in their proper positions.

The lane alignment markers require additional process for error communications. Each lane alignment marker carries a BIP-8 calculated over data between lane alignment markers. The BIP-8 values are calculated by the Ethernet source PCS layer after the self-synchronous scrambling is applied to the 64B/66B block payloads. Since transcoding into 512B/513B block codes requires descrambling the 64B/66B blocks, the lane alignment marker BIP-8 values lose their meaning. In order to preserve the end-to-end error detection capability of the BIP-8 codes, compensation processes are added at both the OTN mapper and demapper.

In order to preserve error detection between the Ethernet source and the OTN mapper, the mapper calculates the expected received value for each BIP-8 on the unscrambled data, and XORs it with the actual received BIP-8 in order to generate an 8-bit error mask. In other words, an error mask bit contains a ‘1’ whenever an incoming error is detected in that received BIP-8 bit location. This error mask is then inserted into the transcoded value of the lane alignment marker. An additional error check is also inserted in order to detect any errors that occur over the OTN link. For this purpose, the mapper calculates a new BIP-8 over the unscrambled 64B/66B blocks and original lane alignment marker and inserts this BIP-8 into the transcoded lane alignment value. This BIP-8 is referred to as the OTN BIP-8[18] since it covers the OTN portion of the link. Hence, the transcoded lane alignment marker contains both the error mask for the incoming signal and the OTN BIP-8.

The OTN demapper regenerates the lane alignment markers when it re-creates the 40GE signal. For error control, the demapper calculates the expected OTN BIP-8 values for the stream of 64B/66B blocks immediately after transdecoding (i.e., prior to scrambling). Comparing the received OTN BIP-8 with this calculated value creates a second error mask with a ‘1’ indicating an error on the OTN portion of the link in that bit position. When the 64B/66B block stream is scrambled for transmission, a new BIP-8 is calculated for the reconstructed lane alignment markers. This new BIP-8 is then compensated by XORing it with both the error mask calculated by the receiver, and the error mask created and transmitted by the mapper.[19]

5.4 Sub-ODU1 rate clients

Three methods exist to carry client signals with rates significantly lower than the OPU1. One is a standard technique, one is in the process of being standardized, and one is very useful even though it is not specified in a standard. Each has is its own target application, although efficient use of the OPU1 is a common goal. These methods are described here.

Intermediate mapping into another transport technology

The client signals can be first mapped into SONET/SDH. The SONET/SDH multiplexing can then be used to combine the clients into an STS-48/STM-16 that is mapped into the ODU1. The advantages to this approach are that standardized mappings exist for nearly all clients into SONET/SDH and SONET/SDH is a full switched layer network. The drawback of this approach is that it requires SONET/SDH as an intermediate layer when carriers are looking to reduce the number of layers in their transport networks.

---

[18] Note that in spite of its name, the OTN BIP-8 here has no relation to the various BIP-8 codes in the OTN overhead.

[19] The existence of the two error masks provides the potential for sectionalized error isolation on the 40GE connection. However, the carriers chose to only have end-to-end error monitoring standardized for both 40GE and 100GE client signals.

Mapping into ODU0

The motivation for the ODU0 was to provide a switchable OTN signal with lower bandwidth than ODU1. There was universal agreement among the carriers that GE is the most important sub-ODU1 rate client. While there were other sub-ODU1 rate clients such as storage area network clients and TOH-transparent transport of STS-3/12 (STM-1/4), they are relatively uncommon in the network. Consequently, it was more important to optimize the ODU0 for simple transport of GE clients than to be bandwidth efficient for the other sub-ODU1 rate clients.

GMP was specified first for ODU0 (see section6.2.3) in order to allow it to support any CBR client stream with a bandwidth lower than the OPU0. While GE was the only client initially defined for the ODU0, mappings for several other clients have been added. See Table 3 for the current client mapping list. .

Mapping into sub-ODU1 tributary slots

For point-to-point applications, it is possible to define tributary slots into which various sub-ODU1 clients can be multiplexed. This approach is restricted to point-to-point applications since the tributary slots lack the frequency justification and OAM overhead required for switching. It also typically requires having the same vendor's equipment on each end. However, this approach can have significant value in access grooming applications where a variety of lower speed clients are combined to make efficient use of OTN at the metro edge.

For example, PMC-Sierra has implemented a 155 Mbit/s channel structure that can transparently carry a variety of arbitrary sub-ODU1 client signals in an ODU1 on a point-to-point basis, including up to 16 STS-3/STM-1 signals. These channels can also be concatenated to transparently carry STS-12/STM-4 and up to two GE signals. PMC-Sierra has demonstrated that asynchronous sub-ODU1 clients, including GE, can be multiplexed together by taking advantage of standard GFP extensions in combination with advanced rate adaptation techniques implemented at the silicon level. Of course, combining wander-insensitive clients such as GE with wander-sensitive clients such as STM-4 into the same ODU1 requires additional considerations. PMC-Sierra has demonstrated that a mix of any sub-ODU1 clients is possible using its OTN Phase Signaling Algorithm (OPSA™) and OTN Payload Tributary Mapping (OPTM™) technology.

There are multiple advantages to the PMC-Sierra method. Since the 155Mbit/s channels are fully transparent to an STS-3/STM-1 signal, including its transport overhead, it allows a carrier to connect to SONET/SDH-based access or enterprise network equipment with a relatively simple OTN signal rather than providing full SONET/SDH functionality in the access network. The signals from multiple enterprise customers and other access equipment can then be multiplexed into the OTN signal in order to make efficient use of the access network, where bandwidth is typically limited. Since these access applications are effectively point-to-point, there is no need to add the significant cost and complexity of per-tributary slot path overhead that some vendors have implemented.

5.5 Storage Area Network (SAN) Clients

SAN clients are less common than Ethernet clients, however they have become increasingly important. Concern about incidents such as natural disasters and terrorist attacks has motivated many enterprises to use remotely-located storage sites to protect their corporate data and computing infrastructure. In many cases, this remote data protection is mandated by the government. Three methods have been defined for carrying SAN clients over OTN without an intermediate mapping into SONET/SDH. These methods are summarized as follows. For simplicity, this discussion ignores SAN protocol issues such as “spoofing” to accommodate longer links.

Transport with GFP

The SAN clients can be mapped into either GFP-F or GFP-T. The resulting GFP streams are then carried over OTN as described in section 5.2.

Transport as CBR signals

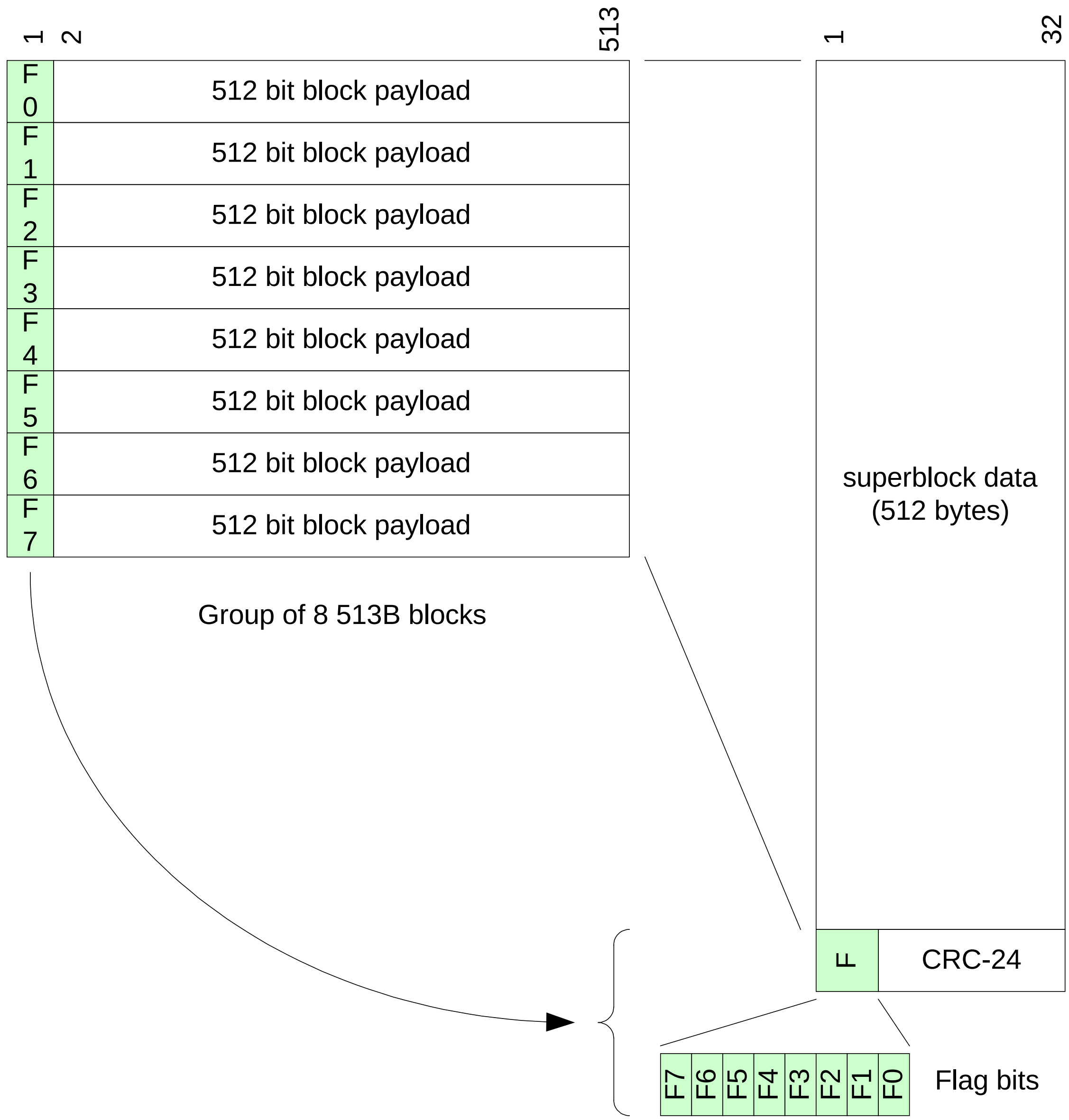
CBR mappings can be used for SAN clients as described in sections 5.1.2 and 5.1.3. These mappings carry the SAN client within an ODU0, ODU1, or ODUflex(CBR) signal, depending on the client data rate.

FC1200

The 10 Gbit/s Fiber Channel interface (FC1200) is a special case. Although it uses the same 64B/66B line code as 10GE, its interface rate is higher than 10GE and hence cannot be directly carried over ODU2 or ODU2e. To accommodate FC1200, a new GFP-T mapping was defined that is conceptually the same as the GFP-T mapping for lower rate SAN clients. The FC1200 64B/66B line codes are first transcoded into 512B/513B block codes. This is the same 512B/513B block code described in section 5.3.4 for use with 40GE. As illustrated in Figure 16, eight 513B blocks are grouped into a 16-block (64-octet) superblock that includes a CRC for error protection.[20] A group of 17 superblocks is carried in each GFP-T frame, with no GFP Idle frames between the GFP-T frames. The resulting signal has the same bit rate as 10GE (10.3125 Gbit/s). This signal is then transported through the OTN using an ODU2e.

[20] A CRC-24 is used with the generator polynomial $G(x) = x^{24} + x^{21} + x^{20} + x^{17} + x^{15} + x^{11} + x^9 + x^8 + x^6 + x^5 + x + 1$.

Figure 16 GFP-T superblock construction for FC1200 transport



6 Synchronization and Frequency Justification / Rate Adaptation

6.1 Synchronization

One of the key decisions for the OTN was that it would not be required to transport network synchronization as part of the OTN signal. Since client signals such as SONET/SDH can transport this synchronization, there was no compelling reason to add the extra complexity and stringent clock requirements to the OTN signals. The only constraint was that the OTN justification scheme for mapping SONET/SDH clients had to guarantee that these clients could be carried without causing them to violate the ITU-T Rec. G.825 jitter and wander specifications.[21]

6.2 Frequency Justification and Rate Adaptation

As discussed above in section 5.1, G.709 originally only defined two frequency justification mechanisms for adapting the rate of client signals into OTN. These mechanisms are the asynchronous mapping procedure (AMP) and the bit-synchronous mapping procedure (BMP). In order to accommodate the flexible mapping of new client signals, a new generic mapping procedure (GMP) has been defined. AMP and GMP are also used for multiplexing lower rate ODU signals into higher rate OPU payloads areas. These three justification mechanisms are described in this section.

6.2.1 Asynchronous Mapping Procedure (AMP) Justification for Mapping and Multiplexing

Frequency justification in OTN is required for some of the CBR mapping techniques and for TDM multiplexing. As indicated in Table 5, the AMP justification technique is a hybrid of the techniques used for asynchronous/PDH networks and SONET/SDH. Similar to the PDH networks, the justification is based on an asynchronous technique with justification control fields rather than the pointer-based approach of SONET/SDH. Like SONET/SDH, however, it provides for both positive and negative byte-wise adjustments rather than the bit-oriented positive adjustments of PDH.

[21] The jitter and wander requirements for OTN network interfaces are specified in ITU-T G.8251.

Table 5 Comparison of PDH, SONET/SDH, and OTN AMP frequency justification

Hierarchy	Technique	Adjustment increment
PDH	Positive justification (stuff)	Single bit
SONET / SDH	Positive/negative/zero (pnz) justification (via pointers)	Single byte for SONET VTs and STS-1 (SDH VC-1/2/3). *N* bytes for SONET STS-*N*c, 3 bytes for SDH VC-4, and 3*N* bytes for SDH VC-4-*N*c.
OTN	Positive/negative/zero justification	Single byte

The justification overhead (JOH) in the OTN is the Justification Control (JC), Negative Justification Opportunity (NJO), and Positive Justification Opportunity (PJO) bytes. As illustrated in Figure 7, these bytes are part of the OPUk overhead. The NJO provides a location for inserting an additional data byte if the client signal is delivering data at a faster rate than the OPUk payload area can accommodate. The PJO provides a stuff opportunity if the client signal is delivering data a lower rate than the OPUk payload area can accommodate. The NJO is thus analogous to the SONET/SDH H3 byte and the PJO to the SONET/SDH positive stuff opportunity byte. The demapper ignores the contents of the NJO or PJO bytes whenever they carry a justification byte. Bits 7 and 8 of JC are used to indicate the contents of the NJO and PJO, somewhat analogous to the SONET/SDH H1 and H2 or the PDH C-bits. The mapper assigns the same value to each of the three JC bytes in an OPUk frame so that the demapper can perform a two-of-three majority vote for error correction.

AMP Mapping Frequency Justification

Table 6 shows the definitions of JC, NJO, and PJO for the AMP and BMP CBR mappings. Here, for the mapping application, PJO is a single byte in the OPUk payload area. As noted in section 5.1, justification is only required for asynchronous mapping since the OPUk clock is generated independently of the client signal clock. Since the bit-synchronous mapping uses an OPUk clock derived from the client signal, it can use fixed assignments for NJO and PJO.

Table 6 Justification control and opportunity definitions for AMP and BMP CBR mappings

<table>
<tr><td rowspan="2">JC [78]</td><td colspan="2">Generation by AMP mapper</td><td colspan="2">Generation by BMP mapper</td><td colspan="3">Interpretation by a demapper</td></tr>
<tr><td>NJO</td><td>PJO</td><td>NJO</td><td>PJO</td><td>NJO</td><td>PJO</td><td></td></tr>
<tr><td>00</td><td>justification byte</td><td>data byte</td><td>justification byte</td><td>data byte</td><td>justification byte</td><td>data byte</td><td>(0)</td></tr>
<tr><td>01</td><td>data byte</td><td>data byte</td><td colspan="2" rowspan="3">not generated</td><td>data byte</td><td>data byte</td><td>(-1)</td></tr>
<tr><td>10</td><td colspan="2">not generated</td><td>justification byte</td><td>data byte</td><td>(0)</td></tr>
<tr><td>11</td><td>justification byte</td><td>justification byte</td><td>justification byte</td><td>justification byte</td><td>(+1)</td></tr>
<tr><td colspan="8">NOTE – Since the mapper never generates the JC [78] = 10, the interpretation by the demapper is based on the assumption that an error has corrupted these bits.</td></tr>
</table>

AMP Multiplexing Frequency Justification

The AMP justification multiplexing frame format differs somewhat from the mapping frame format of Figure 7. First, in addition to an NJO, the ODTUjk includes two PJO bytes rather than a single PJO. This structure easily accommodates the ±20 ppm of the LO ODU signals. The JOH interpretation for the use of the NJO, PJO1, and PJO2 is shown in Table 7.

Table 7 Justification control and opportunity definitions for TDM with AMP

JC [78]	NJO	PJO1	PJO2	Interpretation by the demapper
00	justification byte	data byte	data byte	no justification (0)
01	data byte	data byte	data byte	negative justification (-1)
10	justification byte	justification byte	justification byte	double positive justification (+2)
11	justification byte	justification byte	data byte	positive justification (+1)

The second frame modification is to accommodate the frequency justification for multiple LO ODU tributaries. Each LO ODU tributary that is being multiplexed into the HO OPU requires its own justification. Since there is only a single set of JC bytes and NJO byte in each OPUk frame, they must be shared among the tributary ODUs. This sharing is done based on the frame number within the multiframe. Figure 17 illustrates this sharing of the OPUk JOH overhead. As can be seen, the number of justification opportunities for client signal in each HO OPU multiframe is equal to the number of TS that the client signal uses.

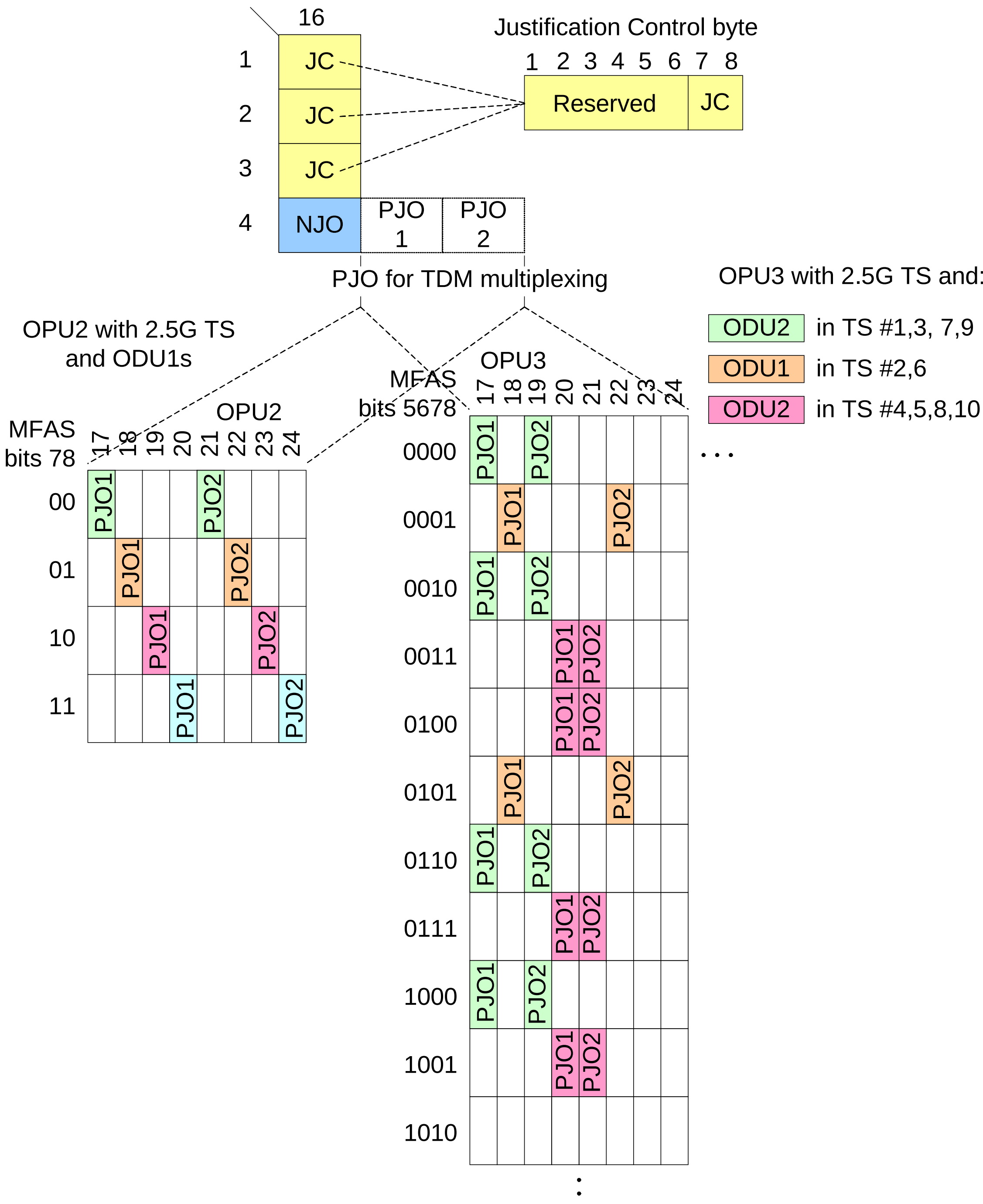
When a client signal is mapped into an OPUk, the PJO is always located in the fourth row of the first payload column (column 17), as illustrated in Figure 7. The PJO locations continue to use the fourth row of the HO ODU in the multiplexing structure. However, when signals are multiplexed into an OPUk with AMP, the PJO locations must be located within a TS column associated with that client. For that reason, the first PJO (PJO1) is located in the lowest numbered column used by that client signal and the second PJO (PJO2) is located in the next column used by that client. Further, the NJO, PJO1 and PJO2 bytes for a client signal appear during each frame in the HO ODU multiframe where the frame number is equal to the number of a TS used by that client. The PJO1 and PDO2 locations are illustrated in the examples of Figure 17.

Figure 17 illustrates an example of a LO ODU2 that uses TS # 4, 5, 8 and 10 when it is multiplexed into an OPU3. Its NJO and PJO bytes appear in frames 4, 5, 8 and 10 of the 16-frame ODU3 multiframe, as indicated by the MFAS LSB values 3, 4, 7 and 9, respectively. In each of these four frames, PJO1 is located in column 20, since that is the first OPU3 payload column associated with the lowest number TS used by that ODU2[22]. The PJO2 is located in column 21 since that is the first payload column associated with the next lowest TS used by that client. Similarly, if an ODU2 is multiplexed into eight 1.25G TS of an OPU3, its JOH, NJO, and PJO bytes appear in the eight frames of the 32-frame ODU3 multiframe corresponding to the TS that it uses.

It is important to note that while the SONET JOH bytes are located within the transport overhead, the OTN JOH bytes are located within the OPUk overhead, which is analogous to the SONET Path overhead. This JOH location choice has an important implication: Retiming an OTN signal requires demultiplexing back to the client signal.

[22] Since column 17 is the first payload column, the first column occupied by TS #*y* is *y* + 16.

Figure 17 Examples showing the AMP Justification Opportunity byte locations



6.2.2 Bit-synchronous Mapping Procedure (BMP) Justification for Mapping

As described above in section 5.1.1, there is no frequency justification for client signals using BMP because the ODU rate is derived directly from the client signal rate. As shown in Table 6, the NJO is always a stuff byte, and the PJO always contains data for BMP mappings. The JC bits are set to a value of 00, consistent with that NJO and PJO usage.

6.2.3 Generic Mapping Procedure (GMP) Justification for Mapping and Multiplexing

GMP was developed in order to provide a consistent asynchronous mapping method to accommodate new CBR client signals having arbitrary rates and frequency tolerances. The mapping is achieved in a straightforward manner without having to define a unique set of fixed stuff bytes or additional NJO/PJO bytes (e.g., to accommodate ±100pm client signal rates) for each new client. GMP was developed together with ODU0 and ODU4 and is the only CBR client mapping rate adaptation method used with these signals. GMP is also used for most non-SDH CBR client mappings into ODU1, ODU2, and ODU3.

For simplicity, the GMP method will first be described for the mapping case. The distinctions associated with GMP multiplexing are described in the next part of this section. The section concludes with a description of an additional fine-grained phase/frequency information capability that is required with many GMP applications.

GMP for Mapping

The GMP justification method works as follows for mapping. A count value, referred to as C_m,[23] is sent in the JC octets of frame *i* (see Figure 8 and Figure 19) to indicate the number of client signal payload words that will be transmitted in the OPUk payload area during frame *i*+1. The stuff words are distributed throughout the OPU payload container in a manner that the receiver can derive directly from the received count value.

[23] The terminology "C_m" indicates that the data and stuff word size is *m*-bits, and hence the corresponding count increment is *m*-bits. Specifically, $m = (8)(\#TS)$, where #TS is the number of 1.25 Gbit/s TS used by that client.

For the purposes of GMP, the OPU payload words are numbered from 1-15232/*M* for ODUk, k = 0-3, and 1-15200/*M* for ODU4, where *M* is the number of 1.25 Gbit/s TS that OPU accommodates[24]. Specifically, for mapping into OPUk, *M* = 1 for OPU0, *M* = 2 for OPU1, *M* = 8 for OPU2, *M* = 32 for OPU3, and *M* = 80 for OPU4. The numbering is illustrated in Figure 18 for the ODU0 and ODU1. For example, for k = 0-3, in the first row of the frame, column 17 is the beginning of the first payload word location, column 17+M is the beginning of the second payload word location, and etc. with the last byte of word location 3808/M appearing in column 3824. Column 17 in the second row is the beginning of payload word location (3808/M)+1, and so forth, with column 3824 of the fourth row being the last byte of payload word 15232/M. The structure is similar for the ODU4, except that, as shown in Figure 27 in section 9.1, the last 8 columns of the OPU4 contain fixed stuff. Consequently, the last word location of the multiframe ends in column 3816 rather than 3824. The byte location designated as NJO for the AMP mappings (e.g., in Figure 2.xx) is not defined for GMP and never carries data. Similarly, the byte location(s) designated as PJO for the AMP mappings is regarded as part of the payload container for GMP.

The method for determining the data and stuff word locations is based on modulo arithmetic. In modulo arithmetic, the modulo remainder of *X* divided by *Y*, which is expressed as (*X*)mod*Y*, is the integer remainder of *X* when it is divided by *Y*. For example, (49)mod13 = 10, since 49 = (3)(13) + 10. The number of payload words in the OPUk frame is referred to as P_{server}, since the OPUk is the server layer channel for the client signal. Let *n* be the payload word location number and let C_m be the count of the number of *m*-bit words to be transmitted in the next frame The contents of word *n* in frame *i*+1 is determined by:

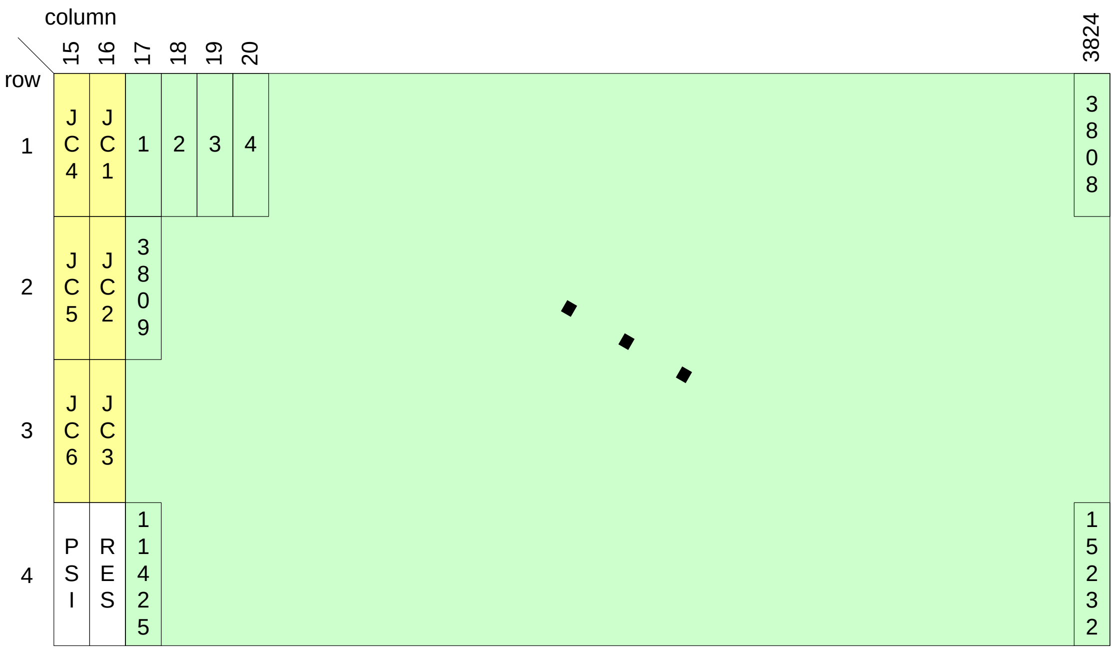
$$\text{Word } n = \begin{cases} \text{data} & \text{for: } (n)(C_m) \bmod P_{server} < C_m \\ \text{stuff} & \text{for: } (n)(C_m) \bmod P_{server} \geq C_m \end{cases} \qquad [2]$$

$$P_{server} = 15232/M \text{ for } k = 0\text{-}3$$
$$P_{server} = 15200/M \text{ for } k = 4$$

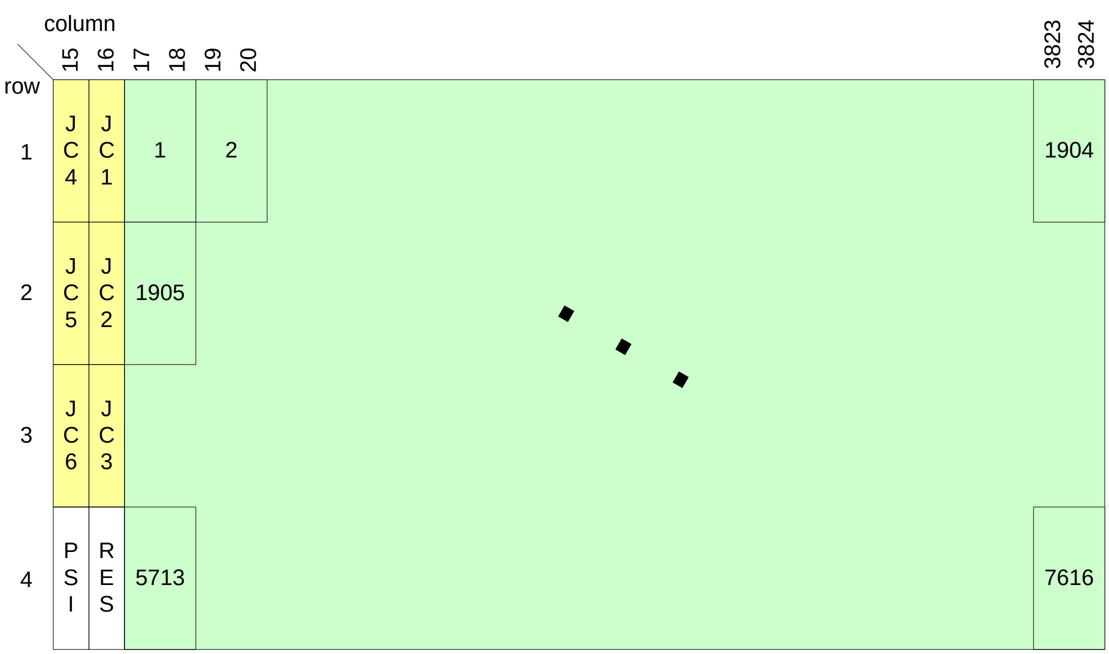
---

[24] As data rates increase, there are significant advantages to using wider data paths at lower clock rates within integrated circuits. In order to better facilitate wider data paths, an *M*-byte word size is used with GMP.

Figure 18 OPUk payload area octet numbering illustration



a) OPU0 example



b) OPU1 example

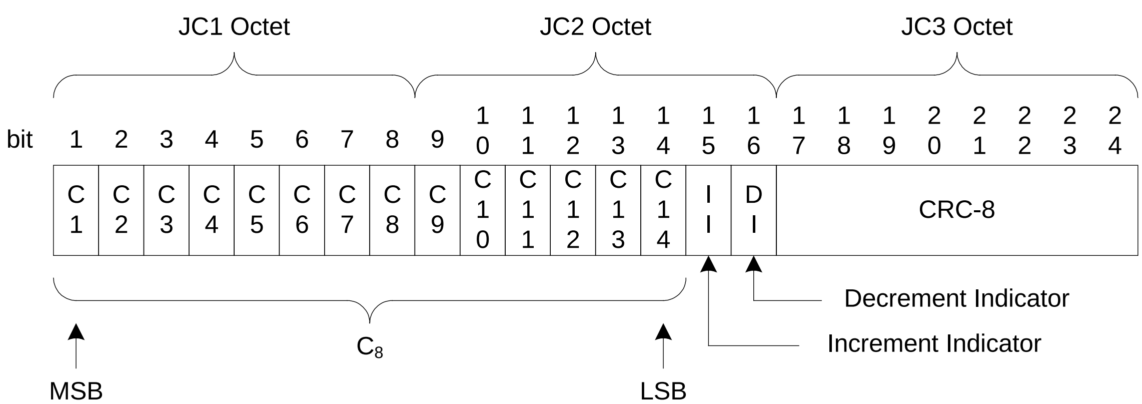
The result is evenly spaced groupings of payload bytes and all-zero stuff bytes. The average number of payload bytes per frame is determined by the ratio of the encoded client signal rate to the payload container rate:

$$C_{m\ average} = (P_{server})(\text{client stream rate / OPUk payload container rate}) \quad [3]$$

Another way to describe this relationship is that remainder of $(n)(C_m) \bmod P_{server}$ is incremented by an amount equal to $(P_{server} - C_m)$ each time the payload octet number is incremented by one. When this $(n)(P_{server} - C_m) \bmod P_{server}$ remainder reaches a value that is less than previous remainder (i.e., less than C_m when the remainder is taken modulo P_{server} a stuff word is inserted. For example, consider a client that use will use 11424 bytes in the next frame, which is exactly (7/8)(15232). The remainder value increments by 1904 for each increase in the value of *n*, which results in a remainder of zero for every eighth octet. This type of modulo count is easily implemented in hardware.

The OPUk justification control (JC) octet format is illustrated in Figure 19. The average value of C_m will rarely be an integer. Consequently, C_m must occasionally be adjusted from frame to frame. Since a mismatch between the source and sink C_m value would cause significant data corruption, it is critical to communicate C_m and its adjustments in a very robust manner.

Figure 19 GMP justification control overhead



Robust count communication is achieved through two mechanisms.[25] The first is a count increment or decrement indication based on inverting a subset of the C_m bits. The second mechanism is a CRC-8 error detection and correction code over the three-octet JC field.[26]

[25] The robust JC octet format adopted for GMP was originally developed and proposed by PMC-Sierra.

[26] The CRC-8 generator polynomial, $G(x) = x^8 + x^3 + x^2 + 1$, was chosen such that it could provide an efficient, low-latency implementation with parallel logic.

The bit inversion mechanism is somewhat similar to the SDH/SONET pointer adjustment method, but with three important differences. While the SDH/SONET pointer adjustment is limited to ±1, the C_m can be adjusted by ±1 and ±2. The source signals the sign and magnitude of the adjustment by transmitting the C_m with different subsets of its bits inverted. These bit inversion patterns, shown in Figure 20, were chosen to have a per-octet Hamming distance of at least four between every pattern. Another difference from SDH/SONET pointers is that the JC field includes explicit Increment Indicator (II) and Decrement Indicator (DI) bits. The final major difference between the C_m encoding and the SDH/SONET pointers is that the JC fields are protected by a CRC-8 error check code. The CRC-8 allows per-frame changes to the C_m of any magnitude by eliminating the need for the persistency checking used with SDH/SONET pointers.

Figure 20 C_m bit inversion patterns to indicate increment and decrement

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	II	DI	Δ
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	+1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	-1
0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	+2
1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	-2
Binary Number														1	1	>±2

The CRC-8 is capable of detecting any 8-bit burst error, and hence can protect against the corruption of any single JC octet.[27] [28] Consequently, the combination of using the count value inversion patterns (including the II and DI) and the CRC-8 allow the receiver to correctly interpret the received C_m value in the presence of any error pattern affecting a single JC octet. The combination of the Increment and Decrement Indicators and the CRC allow communicating an entirely new C_m, in any situation in which it is necessary. This type of change will typically only occur upon initialization, or upon entering or exiting a client signal fault condition.

[27] Due to the spacing between JC bytes, it is assumed that an error burst will affect no more than a single JC octet per frame.

[28] The CRC-8 polynomial was also chosen to have adequate Hamming distance for single error correction. However, since it is difficult to know a priori which links would be characterized by random errors rather than burst errors, G.709 only specifies using the burst error capability of the JC encoding.

Initial source-sink C_m synchronization or recovery of from corruption of the sink's expected C_m value can be achieved within two frames, even in the presence of continuous increment and decrement actions.[29] [30]

One example of a mapping with GMP is the 1 Gbit/s Fiber Channel (FC-100) client signal. The nominal bit rate of the FC-100 signal is 1.0625 Gbit/s ±100ppm. The average GMP count value to carry the FC-100 in an OPU0 is 13062.6 bytes, based on the nominal rates of both the client and server signals[31]. Consequently, the ODU0 transmitter will alternate between sending 13062 and 13063 bytes per frame in order to achieve this average rate. Figure 21 illustrates the GMP data and stuff placement for the FC-100 client with a C_8 GMP count value of 13062.

Another example is the 1000BASE-R Ethernet (GE) client signal. This client uses the timing transparent transcoding mapping discussed in section 40. Since the 1.25 Gbit/s line rate of this client slightly exceeds the payload capacity of the OPU0, the GE client signal is first mapped into GFP-T in order to take advantage of its more efficient line code. The GFP-T stream is constructed such that its rate is directly derived from the GE client signal rate[32]. This GFP-T stream is then mapped into the OPU0 with GMP. When both the GE and OPU0 run at their nominal rates, the GFP-T stream supplies 14407.311 bytes/ frame, which would result in the mapper varying between sending 14407 and 14408 bytes/ frame. Figure 22 illustrates the data and stuff byte locations for this mapping with $C_8 = 14407$.

---

[29] Achieving synchronization at the receiver requires receiving error free JC octets.

[30] Another criterion for choosing the count value inversion patterns was the requirement to achieve fast synchronization at the receiver with a state machine of minimum complexity. If there is no increment or decrement operation, the receiver can directly accept the received Cm value and achieve synchronization with a single received set of JC octets. While the II and DI indicate whether an increment or decrement is occurring, the receiver cannot determine the magnitude of the change since it does not have a correct base Cm value with which to compare the inversion patterns. However, the increment or decrement operation will change the next transmitter C*m* value in a predictable manner. Consequently there are a small number of valid combinations of the two count value LSBs (C13 and C14), II, and DI that could be received in the next frame. As a result, the combination of C13, C14, II and DI in two consecutive frames uniquely identifies the type and magnitude of the increment/decrement operation in the second frame, which allows the receiver to directly determine the transmitter's base C_m value by reversing the inversion pattern.

[31] The nominal ODU0 frame period is (1/1.24416 Gbit/s)(3824 x 4 x 8 bits) = 98.354 µs. In this period of time, the FC-100 signal at its nominal 1.0625 Gbit/s rate delivers (1.0625 Gbit/s)(byte/8-bits)(98.354 µs) = 13062.6 bytes.

[32] Exactly 64 GbE characters are mapped into each GFP-T frame. In other words, each GFP-T frame contains one superblock of 64B/65B block codes.

Figure 21 FC-100 mapping illustration with GMP C_8 = 13062

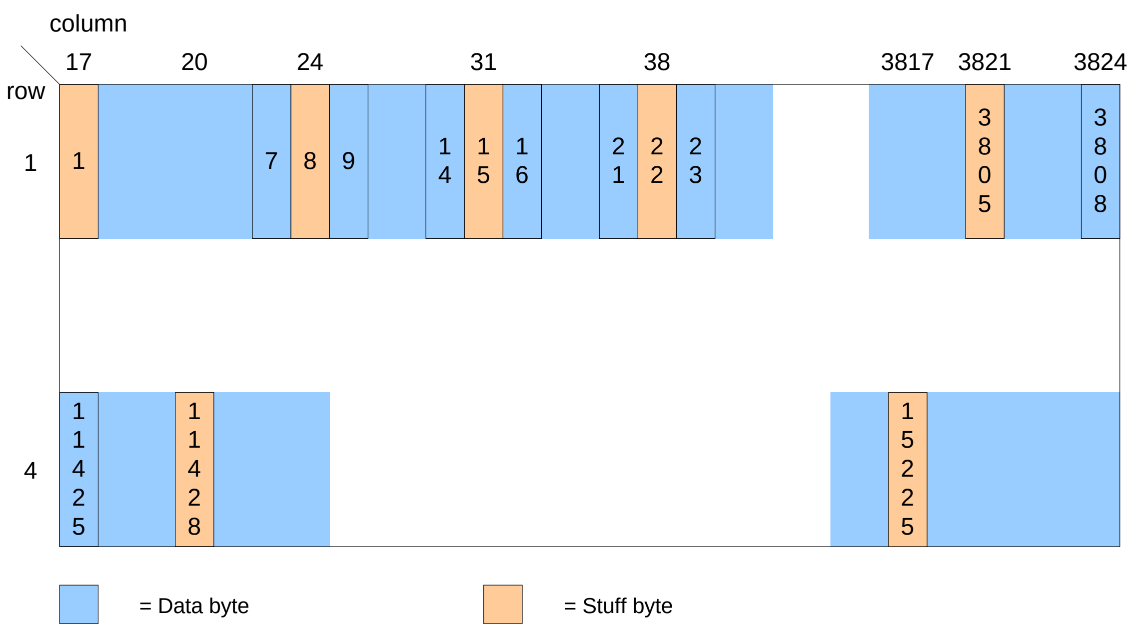
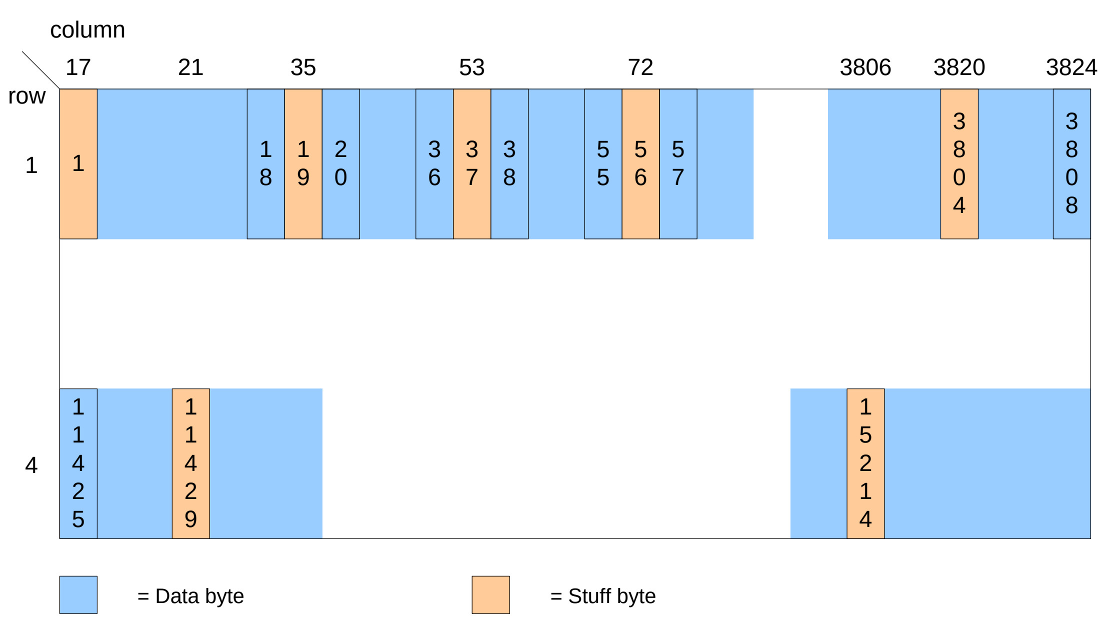


Figure 22 GE mapping illustration with GMP C_8 = 14407



GMP for Multiplexing

The GMP method used for multiplexing is similar to the method used for mapping, although with some important differences that are described below.

As with AMP, LO ODU clients multiplexed into a HO OPU are first mapped into an ODTU that must time share the JOH in the HO OPU overhead. The ODTU for a LO ODU mapped with GMP is referred to as an ODTUk.ts, where ‘ts’ is the number of TS occupied by that LO ODU in the HO OPUk. Since there can be an arbitrary spacing between the different TS used by the client, it was simpler to have the GMP count refer to the amount of client data sent in the next multiframe rather than the amount of client data in the frames between JOH opportunities for that client.

The OPUk JC overhead for GMP is time shared in a manner similar to AMP. With both AMP and GMP, the JC bytes for frame *n* of the multiframe pertain to T.S. *n*. With AMP, the JC bytes are active in each frame in which they pertain to that client. With GMP the JC bytes are only required once per multiframe. They are active in the frame corresponding to the highest number TS used by that client. In summary, the GMP count field in the HO ODU JOH of the frame associated with the highest number TS used by that client LO ODU signal indicates the number of data words that will be sent for that client signal during the next HO OPU multiframe.

As noted above, the use of M-byte words allows better use of parallel data paths within devices, and in some situations can eliminate the need for barrel-shifters to achieve proper data bus alignment. The values for M used by the different LO ODU client signals are shown in Table 8.

Since a client signal using M of the N possible T.S. of an OPU will use M-byte words, we have:

(15232 bytes/frame)*(N frames/multiframe)(M/N Trib. Slots / client) / (M bytes/word) [4]

= 15232 words / multiframe[33]

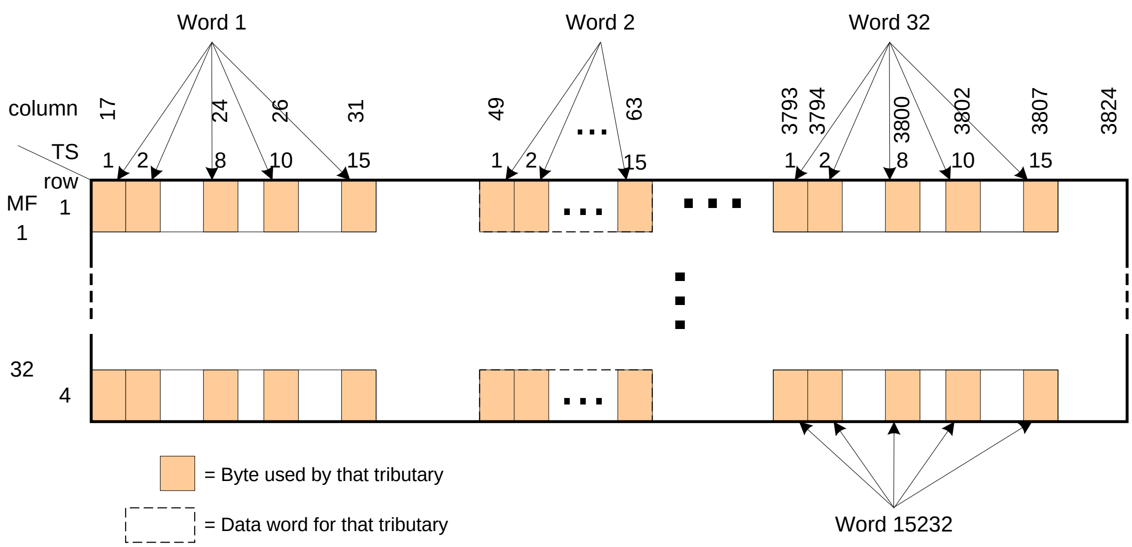
Consequently, even though the GMP count value pertains to a multiframe rather than a frame, it can still use exactly the same GMP count encoding as for the mapping case. Figure 23 illustrates the word numbering and locations for the example of an ODUflex using five TS that is multiplexed into an OPU3.

[33] The OPU4 has 15200 bytes rather than 15232.

Table 8 Client signal word sizes when using GMP multiplexing

Client LO ODU	# of HO OPUk TS used by the client (= # bytes in data word)
ODU0	1
ODU1	2
ODU2	8
ODU3	32
ODUflex(CBR)	>2
ODUflex(GFP)	*x*
NOTE – Since clients with a rate lower than OPU1 are mapped directly into an OPU0 or OPU1, the ODUflex(CBR) will always occupy at least three TS. An ODUflex(GFP) can occupy any number of TS.	

Figure 23 Word numbering illustration for GMP multiplexing – ODUflex tributary using five TS (#1, 2, 8, 10, 15) into an OPU3



Multiplexing ODUflex(GFP) clients with GMP presents a special case. Since a LO ODU client signal can only be multiplexed into an integer number of TS in the HO ODU, the ODUflex(GFP) rate was defined to use the available bandwidth in an integer number of TS. Otherwise, any fraction of a TS bandwidth that is unused by the ODUflex(GFP) client would be lost, since it cannot be used by any other client. However, the rate of a TS is not a fixed value in the network. As previously noted, the TS rate increases when moving up the ODU hierarchy. This could cause a potential problem for signals that are carried by different rate HO OPU signals within the network. For example, if a client is initially multiplexed into seven TS of an OPU4, but is then carried by an HO OPU2 at some point in the network, a signal that fills the seven OPU4 TS will not fit within seven OPU2 TS. That would potentially force the OPU2 to use an inefficient eight TS to carry this signal[34]. Clock frequency differences can create this same problem even for signals that remain at the same HO ODU hierarchical rate through the network.

In addition to this link rate issue, there are other reasons for a more precise specification of the ODUflex(GFP) rates. For example, specifying the rates helps to ensure multi-vendor interoperability. A consistent ODUflex(GFP) rate definition across all links is also a requirement for the Hitless Adjustment of ODUflex(GFP) protocol (HAO) that is described in Appendix C.

In February 2011, SG15 agreed to specify a set of nominal rates for all possible ODUflex(GFP) sizes. These rates were chosen to take into account the TS rate of the smallest HO ODU into which that ODUflex(GFP) signal can fit. For example, since ODUflex(GFP) signals occupying 1-8 TS can fit within an ODU2, the ODU2 TS rate was used to determine their nominal ODUflex(GFP) signal rates. Further, the ODUflex(GFP) rates were chosen such that the worst case combinations of the ODUflex(GFP) rate, ±100ppm, and the HO ODU TS rate, ±20ppm still leaves adequate margin to accommodate jitter or wander[35]. In other words, the rate of an ODUflex(GFP) signal that occupies *N* TS was effectively defined as:

$$(N)(\text{nominal TS rate of the smallest applicable HO OPUk})(\text{Scaling factor}) \quad [5]$$

where the scaling factor takes into account the clock tolerances as described above. G.709 refers to the scaled TS rate as the “ODUk.ts” rate for this application. The resulting ODUk.ts rate specifications are shown in Table 9.

---

[34] Even if the original routing for the ODUflex(GFP) signal in this example only used ODU4 links, it is possible that it could need to be routed over an ODU2 link for fault restoration. In order to simplify the network management, there should be no restrictions on signal routing.

[35] Specifically, the nominal ODUflex(GFP) signal rate occupying *N* TS was chosen to be 186ppm less than *N* times the TS rate of the smallest HO OPUk into which it could fit.

Fortunately, GMP provides a simple mechanism to implement the scaling factors for each HO ODUk. The ODUflex(GFP) source node can use a constant C_m value that takes into account the combination of the number of TS used by the ODUflex(GFP) signal, the LO and HO ODU clock tolerances, and its hierarchical OTUk rate (i.e., k = 2, 3, or 4). The resulting C_m values and their associated ODUk.ts rates are shown in Table 10. These ODUk.ts rates, with their ±20ppm tolerance, fall within the range specified in Table 9. Note that as long as the resulting clock rate conforms to the Table 9 requirements, it is also possible for the source node to use an internal system clock to generate the ODUflex(GFP) signal rather than implementing the clock as a direct choice of C_m for the transmitted HO ODUk.

Table 9 ODUk.ts rates for ODUflex(GFP)

k	ODUk.ts (Gbit/s)
2	1.249177230 ±100ppm
3	1.254470354 ±100ppm
4	1.301467133 ±100ppm

Table 10 ODUflex(GFP) source C_m values and the ODUk.ts rates

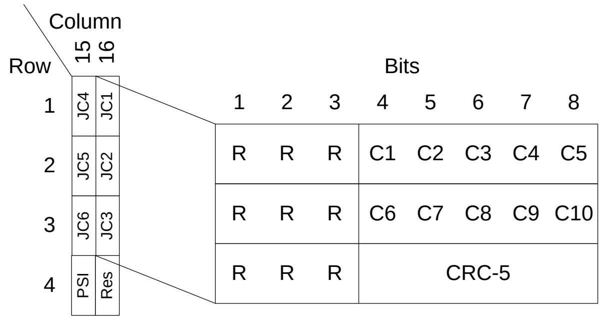
ODUflex(GFP) rate (Number of TS)	HO OPUk rate at the ODUflex(GFP) source	C_m value at the ODUflex(GFP) source	Resulting bit rate per ODUk.ts (Gbit/s)
1-8	OPU2	15230	1.249245570 ±20ppm
	OPU3	15165	1.249184746 ±20ppm
	OPU4	14587	1.249212687 ±20ppm
9-32	OPU3	15230	1.254538983 ±20ppm
	OPU4	14649	1.254522291 ±20ppm
32-80	OPU4	15498	1.301537974 ±20ppm

Fine-Grain Phase/Frequency Information with GMP

The timing resolution when mapping a signal into OTN is 8 bits, and the timing resolution when multiplexing with GMP is M*8 bits. It is difficult to satisfy the jitter and wander requirements of some client signals if they are mapped with just byte-level resolution. Examples include STS-3 and STS-12 (STM-1 and STM-4) signals mapped into an ODU0. The M-byte word resolution with GMP multiplexing also potentially increase the jitter and wander challenges when desynchronizing a client signal. For these reasons, an additional capability was added with GMP to provide the receiver with finer resolution phase/frequency information.

The additional timing information is carried in the J4-J6 bytes, which are located in the first three rows of column 15. The encoding of J4-J6 is illustrated in Figure 24.

Figure 24 GMP justification control overhead for increased timing resolution



The method chosen for communicating the finer resolution timing information was to encode it in terms of phase relative to the data words. The GMP mapper (multiplexer) encoder makes a decision once per frame (multiframe) regarding how many bytes (M-byte words) of data it will transmit during the next frame (multiframe). There will typically be some number of additional bits (bytes) remaining in the transmitter's buffers that cannot be transmitted since they constitute a fraction of a byte (M-byte word). This fractional value is encoded in the C1-C10 field of the J4-J6 bytes as a binary number representing the count of the remainder number of bits (bytes). This phase count value is referred to as the CnD, which means the Count of the Difference in the number of bits (bytes) that could be transmitted as a whole byte (M-byte word) and the bits (bytes) remaining untransmitted at the mapper (multiplexer). Note that in practice, transmitter can use an estimated filtered CnD value rather than a strict measurement of the ingress buffer fill remainder. This filtering removes the effects of incoming jitter and wander.

At the GMP receiver, the CnD information helps the desynchronizer PLL control by providing a more accurate picture of the client frequency. Specifically, when the GMP byte (M-byte word) count value (J1-J3) is constant, the rate of change of the CnD value indicates the frequency offset between the client signal and the OTN channel. Another way to view the CnD is that it provides the receiver with an accurate indication to anticipate upcoming changes in the GMP byte (M-byte word) count value, and hence avoids byte (M-byte word) sized steps in the PLL control.

The concept of finer-grained frequency/phase communication can be illustrated with the following GMP mapper example. Consider a client signal that delivers (Z bytes + Y bits)/frame to the GMP mapper per ODU frame, with Y < 8. The mapper will initially send Z bytes of data in the next frame and communicate the Y value in the CnD. This remainder number of bits will accumulate at the mapper over successive frames until the mapper has Z+1 bytes to send and increases its GMP count value accordingly for the next frame. As a specific example, let Z=15100 bytes and Y=5 bits. If we begin with an empty ingress buffer, the transmitter will send the following sequence of GMP count (C_m) and CnD values: C_m=Z and CnD=5; C_m=Z+1 and CnD=2; C_m=Z and CnD=7; C_m=Z+1 and CnD=4; C_m=Z+1 and CnD=1; C_m=Z and CnD=6; C_m=Z+1 and CnD=3; C_m=Z+1 bytes and CnD=0; etc. The receiver anticipates the jump from Z to Z+1 or Z+1 to Z bytes by observing the CnD count changes and can hence smooth the transition in its demapper phase-locked loop (PLL) filter. Consequently, the jitter and wander of the demapped signal can be significantly reduced with a simpler desynchronizer than would otherwise be required.

A phase offset encoding approach was chosen for CnD rather than a frequency offset encoding approach. A frequency offset approach encodes difference between the number of bits to be transmitted in the next frame (i.e., [word size]x[C_m]) and the average number of bits the client signal delivers during a frame period. As shown, above, the phase encoding is a running sum of the number of bits that can’t be transmitted in the next frame. Although both approaches are mathematically equivalent the phase encoding approach is more robust to transmission errors in CnD. A transmission error with the frequency offset approach would result in a transient frequency error at the receiver. Transmission errors with the phase offset approach are more readily filtered and at worst result in a transient phase error. Hence, the phase offset approach is more robust to preventing short-term frequency deviations due to transmission errors.

7 OAM&P

The key to saving network operational costs is having an effective OAM&P capability built into the signal format. The lack of this capability has been one of the reasons that Ethernet has been slow to take hold as a carrier technology.[36] The OTN OAM&P overhead was built on the experience gained from the SONET/SDH overhead. While many client signals have their own native OAM&P capabilities, they are typically not comprehensive or powerful enough to provide carrier-grade functionality. Consequently, carriers prefer to map client signals into OTN for transport. Mapping all client signals into OTN also provides a the possibility of a common network management framework and set of systems across their entire network.

7.1 Types of Overhead Channels

The different OAM&P overhead channels and their functions are summarized in Table 11. Most of these overhead functions (e.g., BIP-8 and TTI) have been discussed in the context of SONET in another PMC-Sierra white paper [6]. The OTN BDI, BEI, GCC, and OA are functionally equivalent to the SONET/SDH RDI, REI, DCC and A1/A2 overhead channels, respectively. The functions that are unique to G.709 OTN are the following:

[36] Carrier-type OAM&P capability is being added to Ethernet through activities in IEEE 802 and ITU-T SG13. Any Ethernet OAM&P, however, must travel in-band as an Ethernet frame in the same channel as the client data frames. This means that Ethernet OAM&P frames consume client signal bandwidth and require all NEs that make use of this OAM&P information to be capable of removing and inserting the OAM&P frames from the client data stream.

Table 11 OAM&P channel definitions

OAM&P channel	Used in	Function
APS / PCC	ODU	Automatic Protection Switching / Protection Communications Channel – Similar to the SONET/SDH K1 and K2 bytes, but with potential for additional capability. The APS/PCC byte is time-shared across the multiframe to create channels for the control of sub-network connection protection at the ODUk Path and each TCM level.
BDI	OTU, ODU PM, each TCM	Backward Defect Indication – Sent from the overhead sink to the source to indicate that a defect has been detected in the forward direction. (Similar to SONET/SDH RDI.)
BEI	OTU, ODU PM, each TCM	Backward Error Indication – A binary count of the number of BIP-8 bits indicating errors, sent from the overhead sink to the source. (Similar to SONET/SDH REI.)
BIAE	OTU, each TCM	Backward Incoming Alignment Error – Indication sent from the overhead sink to the source that it received an IAE.
BIP-8	OTU, ODU PM, each TCM	8-bit Bit Interleaved Parity- Used in the OTU SM, ODU PM, and each level of TCM overhead.
FTFL	ODU	Fault Type and Fault Location – A 256 byte message with the first 128 bytes applying to the forward direction and the last 128 to the backward direction.
GCC	OTU, ODU	General Communications Channel – Similar to the SONET/SDH DCC. One available in the OTU overhead and 2 in the ODU overhead. GCC1 and GCC2 in the ODU are clear channels whose format is not specified in G.709.
IAE	OTU	Incoming Alignment Error – Indication sent downstream to inform the receiving NEs that a framing alignment error (e.g., a slip) was detected on the incoming signal. Primarily used to suppress BIP error counting.
MFAS	OTU	Multiframe Alignment Signal – Binary counter used to establish the 256-frame multiframe that is used for the time-shared overhead channels that spread their content over the course of a multiframe.
OA	OTU	Optical Alignment – Frame alignment signal for the OTU. OA1 = 1111 0110 and. OA2 = 0010 1000
TCM ACT	ODU	Tandem Connection Monitoring Active: Indication that TCM is being used on the ODUk.
TTI	OTU, ODU PM, each TCM	Trail Trace Identifier – Similar to the Trace identifiers used in SONET/SDH. Used to check that the ODU being received is the one expected.

- **FTFL** – OTN networks can potentially be much more complex than SONET/SDH networks due to the mixture of TDM and WDM technologies. For this reason, it is very advantageous to have better fault type and fault location indication capability. As noted in Table 11, the FTFL information is spread across the 256-byte multiframe with the first 128 bytes pertaining to the forward direction and the last 128 bytes to the reverse direction. The first byte of the 128-byte frame is the fault indication field, the next 9 bytes are an operator identifier field (country and carrier codes), and the remaining bytes 118 bytes are an operator-specific field. No fault (00_H), signal fail (01_H), and signal degrade (02_H) are the only currently defined fault types.
- **IAE** and **BIAE** – The IAE gives a specific indication that a frame alignment error was detected on an incoming signal. This indication allows the receiving NE to suppress invalid BIP-8 errors caused by a frame slip, rather than by actual bit errors and to distinguish between a loss of signal and a loss of frame alignment when AIS is received. In SONET/SDH, only AIS is available for both types of failures. The IAE can be used by the receiver to inhibit counting BIP-8 errors until proper frame alignment is achieved. BAIE is the reverse IAE indication. Reporting BAIE in each of the TCM channels gives greatly enhanced fault location capability compared to an end-to-end indication like the BDI or SONET/SDH REI.
- **APS/PCC** –SONET/SDH use the K1 and K2 bytes for a Line (MS-Section) protection channel and reserves K3 and K4 for Trail protection. OTN, however shares a common protection channel to allow subnetwork connection protection at the level of the ODU and each TCM level. As shown in Table 12, the protection channel is time multiplexed across the signal multiframe. See PMC-Sierra white paper [19] for more discussion of subnetwork connection protection.

Table 12 APS/PCC multiframe definition

MFAS bit 678	Level to which the APS/PCC applies
000	ODUk Path
001	ODUk TCM1
010	ODUk TCM2
011	ODUk TCM3
100	ODUk TCM4
101	ODUk TCM5
110	ODUk TCM6
111	ODUk SNC/I APS

7.2 Maintenance Signals

The maintenance signal set for OTN is somewhat richer than the simple AIS of SONET/SDH and PDH, which is a reflection of the added wrinkles of the TDM/WDM mixture. These maintenance signals are summarized as follows:[37]

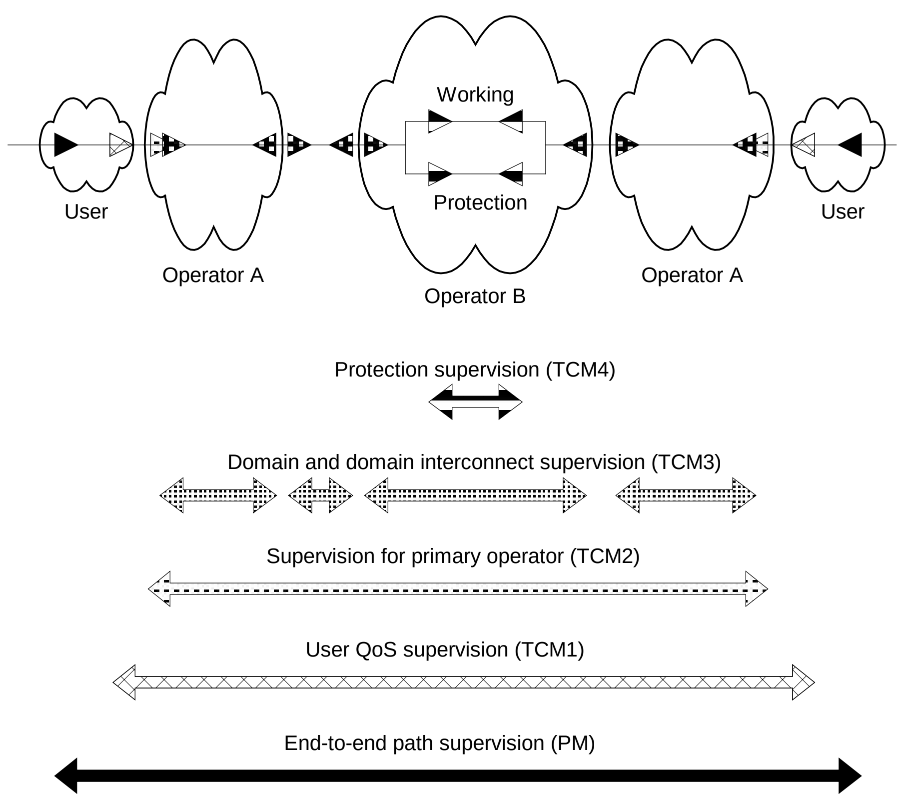
- **OCI** – Open Connection Indication. OCI is provided at the OCh and ODU levels. The OCI indicates to the OCh or ODU termination point that, due to management command, no upstream signal is connected to their corresponding source. This allows the termination point to distinguish the intentional absence of a signal from an absence due to a fault condition. The ODUk-OCI is carried in the entire ODU payload and overhead, and is detected by monitoring the STAT field of the PM byte and each active TCM byte in order to allow monitoring each of these points.
- **AIS** – The AIS signal is sent at the OTU and ODU levels in response to upstream failures. ODUk-AIS is an all-ones pattern in the OPUk (payload and overhead) and ODUk overhead except for the FTFL byte. OTUk-AIS fills the entire OTUk frame with the “generic AIS” pattern, which is defined as the pseudo random sequence generated from the $1+x^9+x^{11}$ polynomial (PN-11, as defined in ITU-T Rec. O.150). The PN-11 gives better signal transition characteristics than an all-ones signal.
- **LCK** – Lock condition. The ODUk-LCK is a downstream indicator that the upstream signal is “locked” and no signal is passing through. It is carried in the entire ODU payload and overhead, and is detected by monitoring the STAT field of the PM byte and each active TCM byte in order to allow monitoring each of these points.

[37] It should be noted that other maintenance signals have been defined for use in the optical domain. PMI (Payload Missing Indication) applies to the OTS and OMS layers and indicates the absence of an optical signal. FDI-O and FDI-P provide a Forward Defect Indication for the payload (server layer) and overhead layers, respectively at the OMS and OCh levels. At this time, neither of these overhead signals has been implemented in commercial equipment.

7.3 Tandem Connection Monitoring (TCM)

As discussed in PMC-Sierra white paper [6], TCM allows the insertion and removal of performance monitoring overhead at intermediate points in the network that correspond to some administrative boundary. It is done such that their insertion and removal do not destroy the performance monitoring overhead that traverses that region. The OTN TCM application is illustrated in Figure 25. The PM byte is used for end-to-end path performance monitoring. Here TCM1 is used by the user to monitor the physical layer connection QoS. TCM2 is used by the primary network operator (Operator A) to monitor the connection from ingress to egress of its network, including the link through Operator B. TCM3 is used by each operator to monitor the connection through its own subnetwork and for the connection between the operator domains. Operator B uses TCM4 to monitor the connection through the protected facility. The TCM overhead fields, as illustrated in Figure 6, include bit error detection, signal trace identifiers, status information, and backward error and defect information associated with that TCM segment.

Figure 25 Illustration of TCM domains



Recall that the SONET/SDH TCM BIP is part of the field covered by the Line BIP-8 (B2) and hence B2 had to be compensated whenever the TCM BIP was modified (i.e., on insertion or removal). This situation resulted from TCM being added after the SONET/SDH signal was defined. In the OTN, however, each TCM only covers the OPUk payload area, and hence no compensation is required.

7.4 Delay Measurement

Some services carried over OTN are sensitive to delay. Examples of such service applications include SAN signal transport and financial transaction communications such as for stock or commodity trading. Delay is important for SAN transport applications because the SAN signal throughput is a function of the delay in communicating a bandwidth grant from the sink to the source. The financial transaction application was brought about by computer aided trading. Studies have shown that a having a latency advantage of just several milliseconds over competitors can result in a substantial increase in profits from timely trades. While network engineers can calculate the delay for a circuit through their network, the users want a guarantee that the delay has been confirmed by measurement, and that it has not changed due to events like network protection switching. Consequently, an integrated delay measurement capability was added to OTN in 2009 at the request of many carriers.

The concept behind the delay measurement is to send test information that is looped back at a far-end NE in order to determine the round-trip delay to the sending NE. The test information is located in the ODU overhead, and is associated with the Path Monitoring and TCM sub-layers. There were several advantages to using Path and TCM sub-layers for the measurement, including:

- The impact to adding this capability to devices is potentially reduced since a type of information loopback already exists in the PM and TCM bytes (for BEI and BDI).
- It provides management consistency by performing the function at the Management End Points (MEPs). Since the MEP is aware of whether a fault exists on the connection, it knows whether to ignore the measurement test result.
- A TCM trail will often already exist or can be readily provisioned between the two NEs where the measurement is desired..
- Allowing delay measurement tied to the Path and TCM overhead allows for simultaneous, independent measurements at the Path and multiple TCM levels.

Latency measurement does not require a complex signaling protocol. Consequently, the latency overhead was defined to use a single bit for the PM and each TCM level in the ODU overhead byte of row 2, column 3. (See Figure 6.) One node is provisioned to be the source of the delay measurement pattern, and the node at the other end of the path or tandem connection is provisioned to be the loopback node for that measurement signal. For simplicity, the latency signal pattern is simply to toggle the bit.

Specifically, the source initiates the latency measurement by toggling the appropriate PM/TCM delay measurement bit and initiating an OTN frame counter. The loopback node for that signal transmits the received delay measurement bit value back to the source in the same overhead bit location. The source node measures the round-trip delay as the number of OTN frame periods between when it transmits the toggled bit and when it receives that bit value from the loopback node.

There are three contributors to uncertainty in the delay measurement. The first comes from there being no fixed relationship between the boundaries of a node’s received and transmitted frames. A bit error at the toggle point can also contribute a frame worth of uncertainty in detecting the toggle at the source node’s receiver. The loopback implementation can also contribute uncertainty. In other words, the total uncertainty (i.e., the measurement resolution) cannot be better than two ODU frame periods plus the loopback implementation latency. However, this is not a significant problem since all ODU signals have a frame period less than 100µs and the accuracy requirement for the delay measurement is ≤500µs. As shown in Table 1, the frame periods of the ODU0, ODU1, ODU2, ODU3, and ODU4 are approximately 98, 48, 12, 3, and 1µs, respectively. Consequently, there is typically considerable loopback implementation flexibility. For example, this accuracy resolution can potentially allow for a firmware-based loopback implementation. Also, the latency measurement will typically have much better resolution than 500µs. Even with the ODU0, a resolution under 250µs is feasible.

8 Forward Error Correction (FEC)

One of the primary benefits to the G.709 OTN is that it provides for a stronger FEC code than the one available with SONET/SDH. This is especially important to allow improved bit error rate and link reliability in ROADM systems. As shown in Figure 6, a four-row by 256-column area at the end of the OTUk frame is reserved for FEC. The FEC code specified in G.709 is a Reed-Solomon RS(255,239). The RS(255,239) symbol size is 8-bits, and the code is implemented as a byte-interleave of 16 separate RS(255,239) codes. The advantage to interleaving is that it allows the resulting interleaved codes to detect/correct burst errors of a length up to the interleaving depth (i.e., 16 bytes here). The Hamming distance of the RS(255,239) code is 17, which allows each code to correct up to 8 symbols (for the error correcting mode) or detect up to 16 symbol errors (in error detection mode). When used for error correction, this leads to a raw coding gain of 6.2 dB for systems with an operating BER of 10^{-15}. This coding gain can be used to allow higher rates over existing facilities, longer span lengths, higher numbers of DWDM channels, or relaxed parameters for the system optical components.

If FEC is not implemented, the FEC field contains all zeros.

In the case of an IaDI, the signal remains within a single domain where there will often be longer spans than would exist for the typical IrDI interface. As a result, there has been some desire for an even stronger FEC option for these IaDI applications. Since such applications will typically have the same equipment vendor’s equipment on each end of the link, there is no compelling need to standardize this FEC. Vendors, then, are free to develop their own FEC to gain a competitive advantage. Different implementations place the FEC bytes in different locations. Some use the frame structure of the OTU frame in Figure 6, but use different algorithms for generating the FEC bytes and correcting errors. Others dispense with the byte order in Figure 6 entirely, and place the FEC bytes in locations convenient for their proprietary FEC codec, intersperse the FEC bytes throughout the frame. Various types of strong FEC codes have been proposed and/or used, and several of these are listed and defined in ITU-T Rec. G.975.1. The G.975.1 strong FEC codes were designed primarily for OTU2. However, some are also being used for OTU3 and OTU4. As with all FEC codes, there are tradeoffs between the error correction/detection capability of the code and the decoder complexity, the encoding and decoding latency, and the transmission bandwidth required to carry the FEC overhead.

As signal rates increase, the need for FEC increases. Links that were capable of OTU2 transmission with no FEC or the RS(255,239) FEC may require a strong FEC for OTU3 transmission. Consequently, different vendors have developed FEC codes that are stronger than the G.975.1 codes. One such code is the PMC-Sierra “Swizzle” FEC described in [20]. These ultra-strong FEC codes use a hard decoder to achieve coding gain performance significantly greater than 9dB. For example, when using the 6.7% FEC overhead of the standard OTUk IrDI frame, the Swizzle achieves a 9.45dB gain, correcting an incoming signal with a BER of 4.8×10^{-3} to an output 10^{-15} BER.

Note that ultra-long-haul applications can use even stronger FECs. These applications can use FEC embedded in the optical modules for a bookended solution. The additional coding gain can be achieved by either adding more than the 7% FEC overhead of the IrDI (e.g., closer to 20%), or use soft decoding techniques, or a combination of both. **

9 OTN TDM Multiplexing

9.1 G.709 Multiplexing Concepts and Capabilities

As noted above, G.709 supports time division multiplexing (TDM) into the OPU payload areas of an OPUk (k = 1,2,3,4).[38] The OTN TDM hierarchy is shown in Figure 26

In contrast to mapping, where the client is mapped directly into an OPUk, a LO ODU is multiplexed into a HO OPU by first asynchronously mapping it into an intermediate structure called an Optical Channel Data Tributary Unit (ODTU). The overhead of the ODTU is the information required for timing justification. The HO OPU payload is divided into Tributary Slots (TS), with each LO ODU occupying an integer number of TS. For ODUk, k = 1,2,3, each TS occupies an integer number of full columns within the OPUk. As will be explained below, this structure is modified slightly for ODU4. The ODTU is byte synchronously mapped into a set of HO OPUk TS.

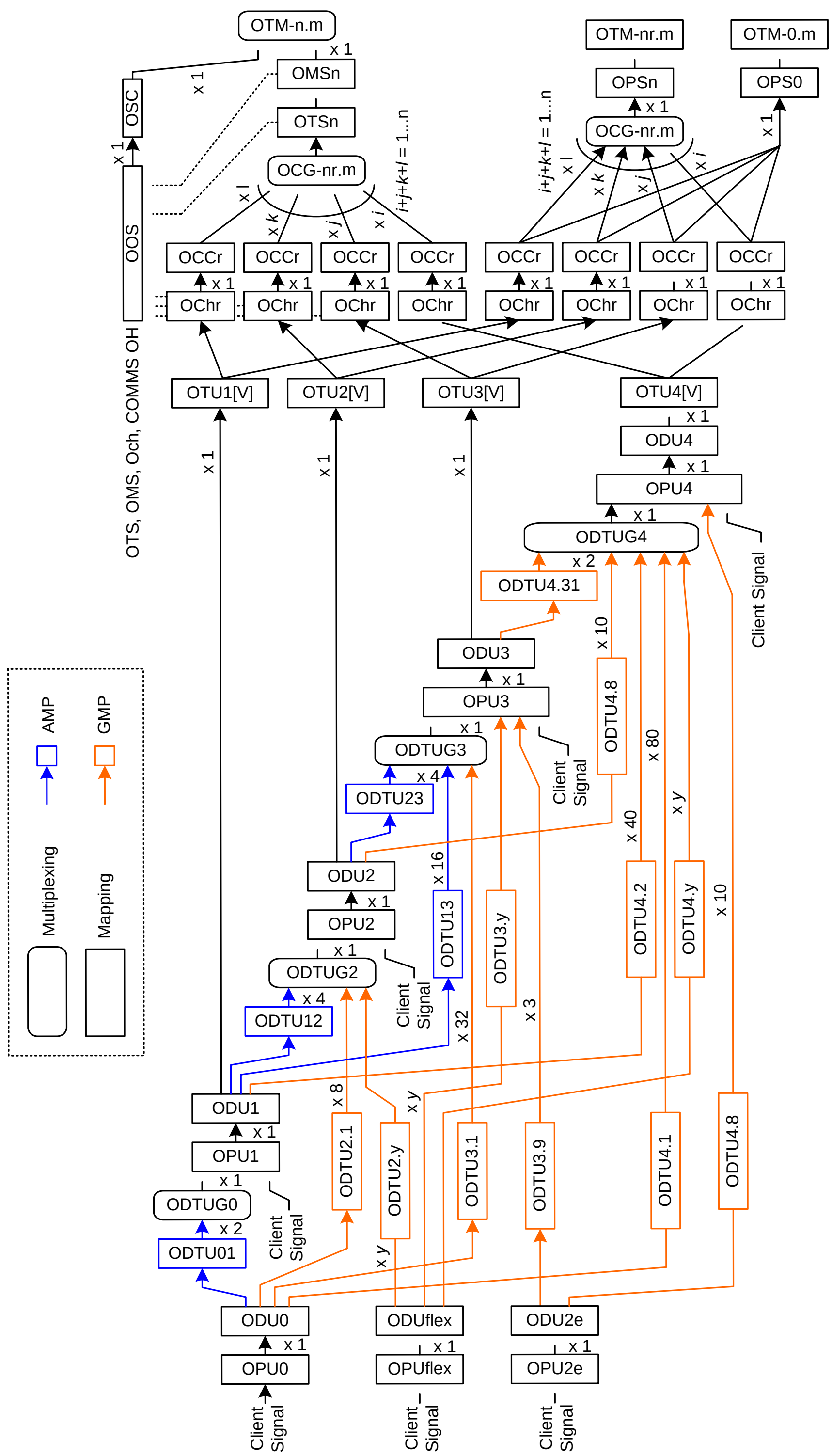
The initial G.709 TDM structure defined the TS in the ODU2 and ODU3 to have a bandwidth of approximately 2.5 Gbit/s (i.e., a rate adequate to carry a LO ODU1 client). The TS at this rate are referred to a 2.5G TS.

The introduction of the ODU0 in late 2008 required defining a TS of approximately 1.25 Gbit/s (referred to as the 1.25G TS) to efficiently carry LO ODU0 clients. An additional motivation for defining a 1.25G TS was that the smaller bandwidth granularity provided greater bandwidth efficiency for multiplexing non-SDH and future client signals into a HO ODUk. Consequently, the 1.25G TS was added as an option for the ODU1 (which previously had not supported multiplexing), the ODU2, and the ODU3. Since the ODU4 signal was being defined at about the same time as the ODU0, it was specified to only support the 1.25G TS. Note that since the tributary multiframe length for each ODUk is equal to the number of TS it can support, multiframe lengths for the ODU2 and ODU3 double with 1.25G TS. See Table 15.

---

[38] Note that originally, there was no intention to provide TDM within the OTN frame. The thinking was that the OTN signal should be kept as simple as possible, and that any TDM multiplexing could be performed within the SONET/SDH client signal prior to the OTN. In the end, however, TDM multiplexing was added. Among the driving applications was that of the “carrier’s carrier.” Carrying each of the original carrier signals on its own wavelength was regarded as too inefficient when these are OC-48s (STM-16s). Multiplexing the original signals into a higher rate SONET/SDH signal, however, requires terminating the incoming SONET Section and Line (SDH RS and MS) overhead, including the Section Data Communications Channel (SDCC). Some proprietary schemes accomplished Section and Line overhead preservation by shifting these bytes to some otherwise unused transport overhead byte locations, but mapping the SDCC is not trivial since the SDCC source clocks will be different than the multiplexed signal SDCC clock. As a result, SDCC packet store and forward buffering may be required. The solution was TDM multiplexing within the OTN.

Figure 26 OTN multiplexing hierarchy



As shown in Table 3, a payload type of PT=20 is used to indicate that the OPU is structured for clients multiplexed into 2.5G TS, and PT=21 is used to indicate an OPU structured for multiplexing into1.25G TS.[39]

Recall from section 6.2 that when multiple LO ODU signals are multiplexed into a HO OPU, the JC1-JC3 JOH bytes in the HO OPU overhead that control the rate adaptation must be shared between the client LO OPU signals. The sharing is implemented through a multiframe structure in which each ODTU has a specific frame or frames in which the JOH bytes in the HO OPU overhead belong to it. In summary, the ODTU is conceptually the portion of the HO OPU that contains both the LO ODU and the JOH that is associated with that LO ODU. The number of frames in a multiframe is equal to the number of TS supported by that OPU. See Table 13. The multiframe structure and JC byte sharing are illustrated in Figure 27. Note that the legacy mappings using the 2.5G TS remain essentially unchanged with the 1.25G TS. They simply use twice as many TS. The actual justification operation, and definitions and operation of the JC bytes, are the same as for legacy multiplexing into 2.5G TS.

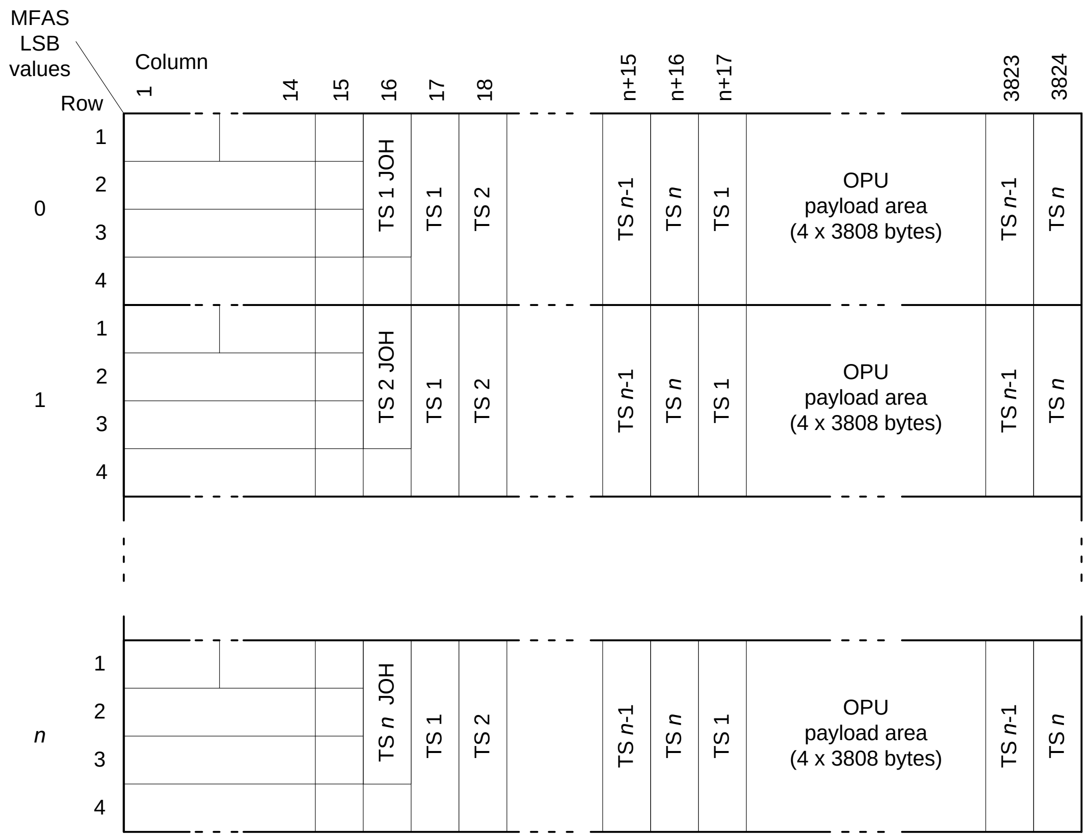
Table 13 Number of TS and multiframe lengths for HO OPUk

HO OPU Type	TS Size	Number of TS
ODU1	1.25G	2
ODU2	2.5G	4
ODU2	1.25G	8
ODU3	2.5G	16
ODU3	1.25G	32
ODU4	1.25G	80
Note: The number of TS = number of frames per multiframe		

The two frequency justification methods used for OTN multiplexing were described in section 6.2. Table 14 summarizes the applications for which they are used.

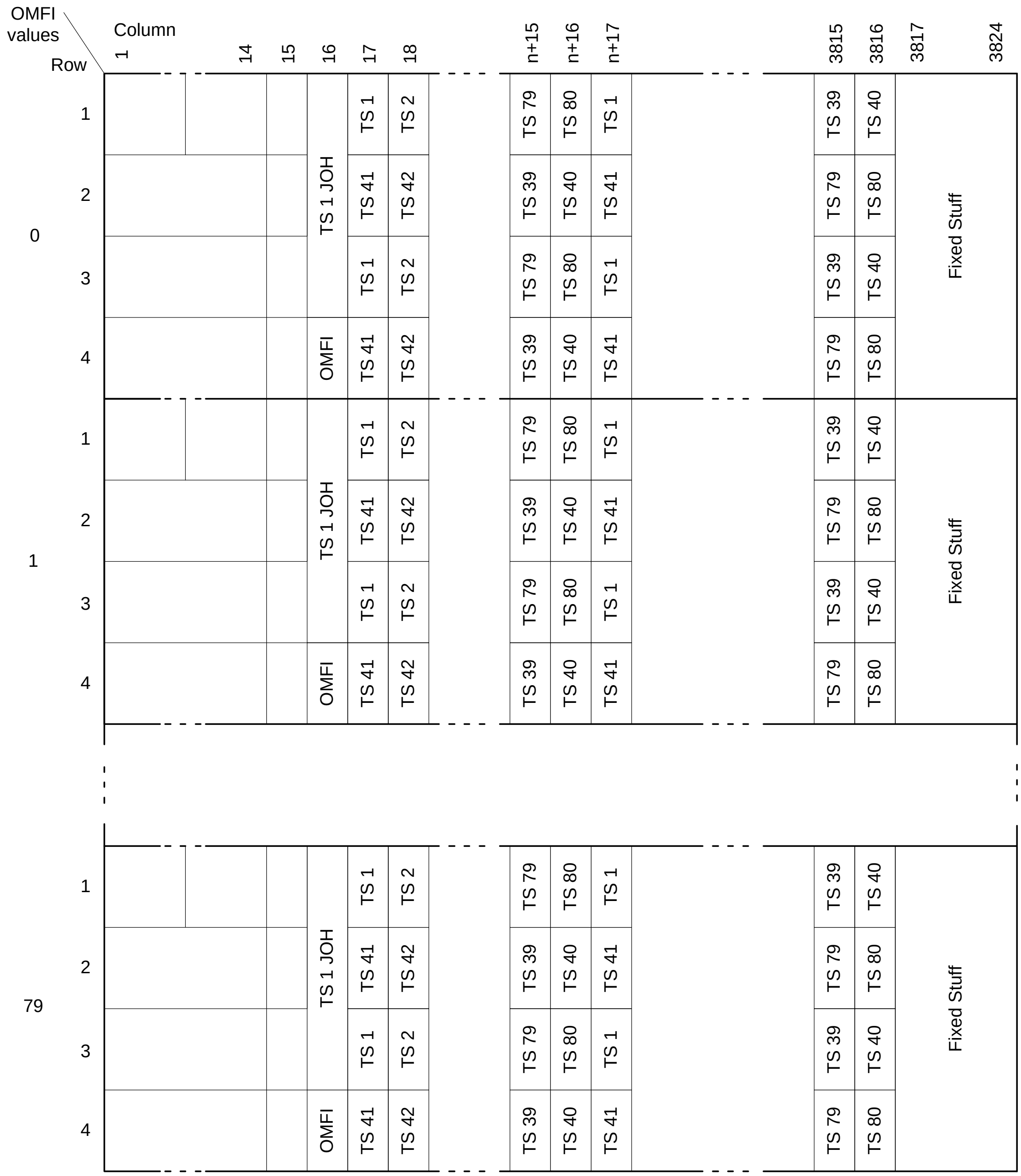
[39] Note that auto-detection of TS granularity is possible on an OTUk link. When both NEs are configured for multiplexed payloads, they initially transmit the smallest granularity they support. In other words, an NE sends PT=21 if it supports the 1.25G TS, and PT=20 otherwise. If both NEs receive PT=21, they will both use the 1.25G TS. If an NE sending PT=21 receives PT=20, it will switch to sending PT=20 and both will use the 2.5G TS.

Figure 27 OPU structure for multiplexing with shared justification overhead



a) Payload and justification location structure for multiplexing into ODUk, k = 1, 2, 3

Figure 26 – *continued*



b) Payload and justification location structure for multiplexing into ODU4

Table 14 Summary of OTN Multiplexing methods

Multiplexing Method	Application
Asynchronous Mapping Procedure (AMP)	Multiplexing LO ODU signals into 2.5 Gbit/s tributary slots Multiplexing some LO ODU signals into 1.25G tributary slots (ODU0 into ODU1, ODU1 into ODU2, ODU1 into ODU3, and ODU2 into ODU3)
Generic Mapping Procedure (GMP)	Multiplexing LO ODU signals into 1.25 Gbit/s tributary slots (except ODU0 into ODU1, ODU1 into ODU2, ODU1 into ODU3, and ODU2 into ODU3)
Note 1 – The 2.5 Gbit/s and 1.25 Gbit/s tributary slot rates are approximate rate values used for convenience of notation. The actual size of the tributary slot is slightly different for each HO ODU signal rate (see section 2.2.1). Note 2 – Since GMP only supports positive frequency justification, AMP is required when the relative bandwidth of the LO ODU client and the HO OPU TS requires the ability to perform negative frequency justification. The cases for which this is true are elaborated above in this table.	

The multiplex structure identifier (MSI) overhead is carried in frames 2-17 of the PSI byte (see Figure 7 and Figure 8). The first two bits (MSBs) of each MSI byte indicate whether the ODU is an ODU1 (00), ODU2 (01), ODU3 (10), or ODU0 (11). The six LSBs of each MSI byte contain the number of the tributary port associated with each OPU TS. The first MSI byte (i.e., the byte carried in frame 2 of the PSI multiframe) contains the tributary port number that is mapped into TS 1, the second MSI byte contains the tributary port number associated with TS 2, etc. While it is possible to use a fixed mapping between tributary ports and tributary slots, the MSI can thus be used to increase the flexibility of the assignments. A different payload type indicator is used when the OPUk is structured for 1.25 Gbit/s (ODU0-capable) tributary slots.[40] [41]

Fixed stuff columns are added to the AMP OTDU structure when they are required to adjust the nominal payload area rate for the nominal tributary rates. For example, in the case of ODU1 to ODU3 multiplexing, an additional fixed stuff column (Columns 1889-1904) is also required.

[40] Although OTN was specified as only allowing a single digital multiplexing stage, it may be necessary in the interim to multiplex ODU0s into ODU1s in order to maintain backward compatibility with ODU crossconnects that support only 2.5 Gbit/s time slots.

[41] There have been proposals to define an intermediate mode of operation in which ODU1 and ODU2 client signals would be restricted to the legacy 2.5 Gbit/s time slot boundaries, with ODU0s effectively occupying a half 2.5 Gbit/s time slot. The motivation for this approach is to allow upgrades to line cards that would terminate the ODU0 clients, while allowing the NE switch fabric to continue to support switching 2.5 Gbit/s time slots. From the ODU0 perspective, there is no difference between this and full 1.25 Gbit/s time slot structure.

9.2 Special Considerations for ODU4

The ODU4 is a special case since the number of columns in the payload area is not an integer multiple of the number of TS. Since 3808-(47)(80) = 48, too much bandwidth would be wasted if these 48 columns were all left as fixed stuff. Consequently, the column sharing approach illustrated in Figure 27.b was used in order to reduce the number of fixed stuff columns to eight. The byte in the fourth row of column 16 is used as the OPU4 multiframe indicator (OMFI) instead of using the OTU4 MFAS. The OMFI implements a modulo 80 count, with 0000 0000 indicating frame 1 and 0100 1111 indicating frame 80.

9.3 Multiplexing Summary

The LO ODU to HO ODU multiplexing parameters are summarized in Table 15.

Table 15 Summary of TS use and frequency justification methods for OTN multiplexing

LO ODU	**HO ODUk [Number of T.S. used (Justification Method)]**						
	2 ½ G T.S. (PT=20)		1 ¼ G T.S. (PT=21)				
	OPU2	OPU3	OPU1	OPU2	OPU3	OPU3e2	OPU4
ODU0	-	-	1 (AMP)	1 (GMP)	1 (GMP)	1 (GMP)	1 (GMP)
ODU1	1 (AMP)	1 (AMP)	-	2 (AMP)	2 (AMP)	2 (GMP)	2 (GMP)
ODU2	-	4 (AMP)	-	-	8 (AMP)	8 (GMP)	8 (GMP)
ODU2e	-	-	-	-	9 (GMP)	8 (GMP)	8 (GMP)
ODU3	-	-	-	-	-	32 (GMP)	31 (GMP)
ODU3e2	-	-	-	-	-	-	32 (GMP)
ODU4	-	-	-	-	-	-	-
ODUflex (CBR)	-	-	-	*n* (GMP) (Note 1)	*n* (GMP) (Note 1)	*n* (GMP) (Note 1)	*n* (GMP) (Note 1)
ODUflex (GFP) (Note 2)	-	-	-	*n* (GMP) $n \leq 8$	*n* (GMP) $n \leq 32$	*n* (GMP) $n \leq 32$	*n* (GMP) $n \leq 80$

Note 1 – The number of T.S required is determined as:

$$n = \left\lceil (\max [\mathit{CBR\_bitrate}]) / (\min [\mathit{OPUk\_TS\_bitrate}]) \right\rceil$$

Note 2 – For ODUflex(GFP), the number of T.S. is recommended to be an integer multiple of the 1.25G T.S. rate for smallest OPUk into which that ODUflex can be mapped.

9.4 Multiplexing Stage Considerations

One of the important concepts that OTN inherited from SONET/SDH is that the number of multiplexing stages should be minimized. This avoids the complexity of the PDH hierarchy in which extracting the lowest rate client signal required demultiplexing at every stage of the hierarchy. Initially, OTN was defined to use only a single stage of multiplexing in the electrical domain. In other words, a client ODU could be multiplexed directly into a server OPU of any higher rate. The introduction of ODU0 and ODUflex changed that situation for at least the near and mid-term. In order to carry an ODU0 or ODUflex through a legacy network (i.e., one that only supported ODU1, ODU2 and ODU3), they needed to first be multiplexed into a legacy signal. In other words, ODU0 needed to be mapped into an ODU1 for transport over a legacy ODU2 path, or into an ODU1 or ODU2 for transport over a legacy ODU3 path. Similarly, an ODUflex must be mapped into an ODU2 for transport over a legacy ODU3 path.

Consequently, 2-stage multiplexing is required for this specific interim set of cases. The principle of limited multiplexing stages remains, however, and no other cases have been identified that require multiple multiplexing stages.

It is important to understand the multiplexing stage restriction applies to a network domain and the equipment within that domain. A carrier providing transparent transport for a client ODU from another carrier (or for a different domain of its own network) will typically multiplex this client into a higher rate OTN signal. The resulting OTN signal itself can thus contain an arbitrary number of multiplexing stages. However, from the perspective of each network domain, only one or two multiplexing stages have been performed.

10 Virtual Concatenation

Virtual concatenation (VCAT) is specified in G.709 for OPUk channels to create payload containers with rates that are more efficient for carrying a particular client signal than using the next higher rate ODUk signal. For example, Fiber Channel FC-400 is more efficiently carried by combining two OPU1s rather than using an entire OPU2. The ODUflex signals were added to G.709 in order to avoid some of the implementation and network management complexities associated with VCAT. VCAT uses the Link Capacity Adjustment Scheme (LCAS) to vary the size of the VCAT channel. ODUflex(GFP) uses the G.7044 Hitless Resizing of OFUfles (HAO) protocol to accomplish a similar function, but in a much more powerful manner.[42] See the appendix (0) for an overview of HAO.

In concept, OPUk virtual concatenation works the same as described in [6] for SONET/SDH and PDH signals. The group of OPUks is launched with the same frame and multiframe phase. Each OPUk is allowed to take a different, independent path through the network with the receiver using the multiframe information (including the LCAS overhead MFI) to perform the compensation for the differential delay between the members. A virtually concatenated channel is referred to as an OPUk-*X*v, where *X* is the number of OPUks that are concatenated. Each OPUk is placed into its own ODUk with the X ODUks being referred to as an ODUk-*X*v. The virtually concatenated OPUk has no fixed stuff columns, giving a payload capacity of $X * 238/(239-k) * 4^{(k-1)} * 2\,488\,320$ kbit/s ± 20 ppm.

The virtual concatenation and LCAS overhead location and structure is shown in Figure 6. The individual fields in the overhead are the same as those discussed previously for SONET/SDH and PDH. The actual structure differs in that there is a member status (MST) byte and a CRC-8 included for each of the other bytes. Allowing for up to 255 members seems somewhat optimistic about the future progress of optical components since the OPU1 is about 2.5 Gbit/s.

The payload mappings into an OPUk-Xv are identified in the PSI byte with the values shown in Table 16. The first byte of the PSI byte frame (PT) specifies that this OPU is part of a virtually concatenated group. The second byte (frame 1, shown as Reserved in Figure 6) is used as the virtual concatenation payload type indicator (vcPT), with the values as defined in Table 16. The mapping techniques into virtually concatenated channel are essentially the same as discussed above for non-concatenated channels.

[42] While LCAS hitlessly varies the size of the transmission channel, HAO hitlessly varies both the size of the transmission channel and the signal (i.e., the ODUflex(GFP) signal) that carries the GFP payload.

Table 16 Payload type values for virtually concatenated payloads (vcPT)

Hex code	Interpretation
01	Experimental mapping (NOTE 1)
02	asynchronous CBR mapping
03	bit synchronous CBR mapping
04	ATM mapping,
05	GFP mapping
10	bit stream with octet timing mapping
11	bit stream without octet timing mapping
55	Present only in ODUk maintenance signals
66	Present only in ODUk maintenance signals
80 - 8F	reserved codes for proprietary use (NOTE 2)
FD	NULL test signal mapping
FE	PRBS test signal mapping
FF	Present only in ODUk maintenance signals
NOTE 1 – Experimental mappings can be proprietary to vendors or network providers. If one of these mappings/activities is standardized by the ITU-T and assigned a code point, that new code point is used instead of the 01 code point. NOTE 2 – Proprietary mappings are similar to experimental mappings. If the mapping subsequently becomes standardized, the new code point is used. ITU-T has rules associates with this specified in G.806.	

11 OTN Evolution

In the years since the original OTN standards were developed, the carriers' views on OTN application requirements have evolved. The motivations behind this evolution include the following:

- Transport of additional LAN and SAN signals over the WAN;
- Efficient direct mapping of lower rate clients into the OTN[43];
- Desire to eliminate network layers in order to reduce operations expenses; and.
- New higher speed Ethernet interfaces being developed by IEEE 802.3.

Network Layer Reduction

Each network layer that a carrier must manage creates a substantial amount of operating expense. In some carriers, an entire organization exists for managing each network layer. Since operations constitute 30-40% of carriers' annual expenses, they want to minimize the number of different layers in their networks whenever possible. OTN contains most of the features of SONET/SDH with respect to higher order transport signals. With some additional enhancements, carriers can potentially use the OTN for all transport functions and use SONET/SDH as a client rather than a separate transport network layer. Client signals would be mapped directly into OTN. The specification of how to map the various clients directly into OTN is the main enhancement required.

[43] Of course, this direct mapping only applies to signals over 100 Mbit/s. Lower rate signals such as DS1/E1 would be multiplexed into a higher-rate container (e.g., a SONET/SDH signal or higher rate Ethernet signal) separately outside the OTN.

12 Conclusions

The ITU-T OTN hierarchy has many merits and advantages for DWDM systems and for optical transport networks as a whole. The G.709 OTN frame provides some very significant OAM&P capabilities, especially in the nesting of multiple TCM overhead channels. The recent addition of 1.25 Gbit/s Tributary Slot granularity, ODU0 and ODUflex, and GMP has greatly increased bandwidth efficiency of the OTN and its flexibility for carrying non-SONET/SDH clients. The initial focus of G.709 OTN systems was in core long haul networks. As bandwidth growth continues unabated, largely driven by application such as video, peer-to-peer networking, and SAN extension, OTN is becoming a key requirement for the metro, access, and long haul networks. OTN addresses the need for reducing carrier operating expenses, takes advantage of the decreasing cost of optical components, provides a standards-based hierarchy to address the growing demand for bandwidth throughout the network, and addresses the requirement to achieve higher rates over existing facilities. The recent additions to optimize OTN for Ethernet transport create opportunities to reduce network complexity by eliminating SONET/SDH as an intermediate layer. Many carriers are planning to move to networks that use OTN as the transport technology and a combination of Ethernet or MPLS, and OTN for their switching layers. All network equipment vendors are being asked to answer the call for OTN through integration of this technology into their micro-OTP, ROADM, OTP, and P-OTP portfolios. Even Carrier Ethernet Switches and Routers (CESRs) are adding OTN wrapping for data clients in order to seamlessly interconnect with the new transport network, reducing transponders and providing for end-to-end OAM.

PMC-Sierra is a leader in OTN solutions for micro-OTP, ROADM, OTP, P-OTP, and CESR line cards, beginning with the innovative HyPHY and META families.[44] PMC-Sierra's OTN solutions enable carriers to accelerate their network transition to OTN, delivering cost effective grooming and switching of services end-to-end in the new Packet Optical Transport Network. For more detailed information on the HyPHY and META product families and the complete portfolio of OTN solutions from PMC-Sierra, please visit the PMC-Sierra Wireline Infrastructure solutions web page at http://www.pmc-sierra.com/

[44] PMC-Sierra played a leading role as one of the most active participants in the evolution of the ITU-T OTN standards. PMC-Sierra was author or co-author on many of the key standards contributions and also served in key editor roles.

Appendix A – Optical Technology Considerations

This appendix provides a basic introduction to some of the optical domain concepts so that the reader can appreciate some of the decisions that were made in defining the OTN and its signal formats.

Glass fibers are not fully transparent to light, and in fact typically have three wavelength windows where the light attenuation is lowest. The first is in the 820-900 nm wavelength region, which is easily generated by inexpensive GaAlAs lasers, but does not allow single-mode transmission[45]. The next window is 1280-1350 nm, which has substantially lower loss than the 820-900 nm region and supports single mode transmission. The third window is the 1528-1561 nm region, which has the lowest attenuation, but also requires lasers that are more expensive than those for the other two regions.[46]

An ideal unmodulated laser would output light with a single wavelength, which would appear as a single line on a spectral graph. In practice, however, physical realities mean that a laser outputs light with some spread around its central wavelength. Once modulated, the laser signal develops sidebands, and occupies a bandwidth determined by the data rate and modulation scheme. Since this spread looks like a fatter line on a spectral graph, a laser's wavelength spread is often referred to as its line-width. Clearly, the line-width of the lasers determines how many lasers' signals can be combined in a wavelength window with WDM. If the wavelengths of two lasers overlap, they will interfere with each other at their respective receivers. The width of the sidebands (literally, the bandwidth) determines the minimum spacing between laser frequencies in a wavelength-division multiplexing system. Additional guardbands must be added to account for nonidealities in the laser and other optical components. Many DWDM systems use 50GHz, which causes substantial difficulties in transmission of 40Gbps and 100Gbps signals.

Although the lasers for the 1555 nm region are more expensive than those for the 1310 nm region, the good news is that it is also possible to manufacture these lasers with relatively narrow line widths that are compatible with the erbium-doped fiber amplifier (EDFA) optical amplifiers discussed below. Multiple-quantum well (MQW) lasers with distributed feedback (DFB) can achieve line widths of a few hundreds of kHz (a few millionths of a nm).

---

[45] Light propagates through a fiber by the process of total internal refraction in which the light going through the core of the fiber is refracted back into the core when it hits the cladding that surrounds the core. (The cladding has a lower index of refraction than the core.) In multi-mode transmission, the light "bounces" through the fiber as it encounters the cladding in such a manner that the portion of a light pulse that encounter the fewest bounces has a shorter path than the one that has the most bounces. The result is a time spreading of the pulse at the receiver that limits possible spacing between pulses (i.e., the possible data rate for a digital signal). In single mode transmission, only the light that goes directly through the core is able to propagate, thus minimizing any pulse spreading.

[46] This section will follow a common practice of referring to the regions by their center wavelength, i.e., 1310 and 1555 nm.

In Recommendation G.694.1 and G.694.2, the ITU-T has defined "grids" of wavelengths that can be used for WDM. These grids specify the wavelengths that lasers can use, but does not specify which of these may or should be used within a WDM system.[47] In dense WDM (DWDM), the wavelengths in a WDM system are close together, with 50 or 100 GHz spacing. In coarse WDM (CWDM)[48], the wavelengths are much farther apart, which lowers cost at the expense of fiber capacity. As wavelengths closer together on the grid are used, crosstalk can become a problem. The primary way to minimize or eliminate cross talk is to require the optical signals to have a sufficiently narrow line width and low enough drift from their center wavelength that their modulated signal does not overlap with adjacent channels. Another phenomenon that creates crosstalk is called four-wave mixing, , which arises due to slight non-linearities in the fiber. In four-wave mixing, several optical signals constructively interfere ('beat') with each other to create quite large fields inside the fiber, to which the glass responds in a slightly nonlinear way. The result is the generation of interfering signals at various "beat frequencies". In the case of three or more evenly spaced signal frequencies, some of these "beat frequencies" align with the original signals, causing crosstalk.. These crosstalk and mixing problems are familiar to people who are experienced in frequency division multiplexed (FDM) systems, since wavelength modulation is essentially a form of frequency modulation. Bit rate, fiber type, and fiber length are also factors in determining how many channels are possible in a DWDM system. DWDM systems for metro networks will use up to 40 wavelengths (100 GHz spacing), while DWDM for core networks commonly use up to 80 wavelengths (50 GHz spacing).

Optical Signal Regeneration

There are three aspects to the regeneration of optical signals (the three "Rs"):

- Re-amplification of the optical signal
- Re-shaping of the optical pulses
- Re-timing/re-synchronization

An all-optical amplifier is a 1R amplifier. A 2R regenerator both amplifies and re-shapes the optical pulses. A 3R regenerator also performs clock recovery on the incoming signal and re-times the outgoing signal in order to remove jitter on the pulses. Currently, both 2R and 3R regenerators convert the optical signal to an electrical signal (OE conversion) and create a new optical pulse (EO conversion) after amplification (and re-timing for 3R regenerators).

---

[47] The wavelengths in the grid are called the C-band and are on evenly spaced wavelengths of 0.39 nm (50 GHz) starting at 1528.77 nm. The grid was originally assembled largely as a collection of the wavelengths supported by the various laser vendors. In addition to the C-band, DWDM systems can use the L-band (1561-1620 nm). The S-band (1280-1350 nm) is typically used for single wavelength rather than DWDM applications.

[48] A common, extreme example of CWDM is to use just two wavelengths, one at 1310 and the other at 1555 nm.

Clearly, a 3R regenerator needs to be aware of the client signal (e.g., SONET/SDH), at a minimum having the ability to perform clock recovery at that client signal rate, but possibly requiring the ability to frame on the signal (e.g., to function as a SONET STE / SDH RS terminating NE).

An alternative is to amplify the signal in the optical domain (i.e., a 1R amplifier), thus removing any requirements on the regenerator concerning the rate of the client signal. This is accomplished through optical amplifiers, of which the EDFA is the most common. In an EDFA, the signal passes through a section of fiber that has been doped with erbium. A strong signal from a pump laser is coupled into this fiber segment. The 980 (or 1480) nm wavelength energy of the pump laser excites the erbium atoms, and the presence of 1555 nm signals causes the erbium atoms to transfer their energy to the 1555 nm signal through stimulated emission. Gains of over 20 dB are possible. Other types of amplifiers, such as the Raman amplifier, can amplify a much wider range of wavelengths (1300-1600+ nm). A full discussion of optical amplifier technology is beyond the scope of this white paper.

Carriers prefer to use 1R regenerators whenever it is practical in order to preserve maximum signal transparency and minimize the amount of costly electrical domain processing in the network.

Optical Switching

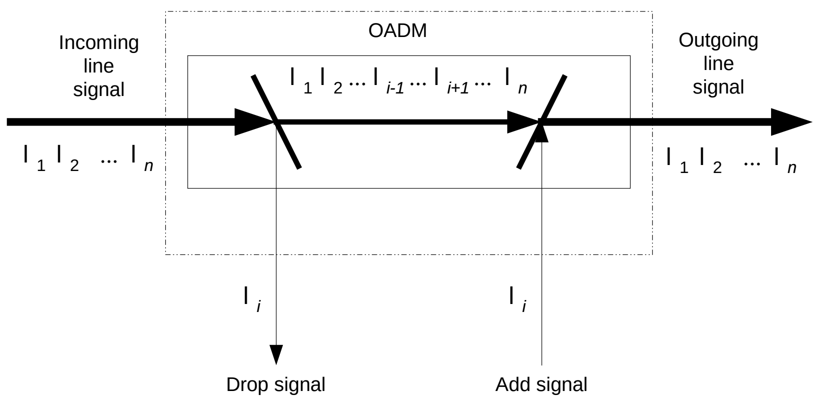
A key motivation for developing all optical networks is that if signals are kept in the optical domain, the network equipment can be agnostic to the client payload signals and eliminate the circuitry required for conversions between electrical and optical domains. For example, switching SONET signals requires STS-1 electronic cross connect fabrics and the OEO functions to convert the signal between the optical and electrical domains. Switching in the optical domain has the promise of lower equipment and provisioning costs at the expense of large granularity in the switched signals. As an additional capability, OTN also supports the use of use cross-connects that switch the client signal in the electrical domain whenever it is desirable to groom the signals that are placed on the wavelengths. Such switches are referred to as hybrid switches.

A number of different technologies exist for switching in the optical domain, including solid-state devices (e.g., directional couplers that are combined to form multi-stage switch fabrics), free-space techniques (e.g., waveguide grating routers), and micro-electrical-mechanical switches (MEMS). MEMS technology allows the construction of an array of mirrors in silicon where each mirror’s reflection angle can be controlled by an electrical signal. The optical input signals to the MEMS array can be steered to the appropriate output ports. MEMS switch times are in the order of microseconds. For fast switching, $LiNbO_3$ solid-state switches can achieve switch times in the order of nanoseconds. See [6] for additional information on switching component technologies.

Optical equipment with extensive cross-connect capability is known as an optical cross-connect (OXC). Simpler equipment that is capable of adding or dropping wavelengths is known as an optical ADM (OADM). As illustrated in Figure 28, an OADM filters an incoming wavelength(s), removing it from the incoming signal and steering it to a drop port. At the transmitter of the OADM, the signal from the add port is then optically merged back into the outgoing signal. OADMs can either add/drop fixed wavelengths or dynamically select which wavelengths to add/drop.

It should be noted that a typical OXC or OADM implementation would have an EDFA at the input to the NE that acts as a pre-amplifier prior to the cross-connect fabric. The pre-amplifier boosts the signal amplitude to compensate for the attenuation over the fiber, and sets it to an appropriate signal level for the switch fabric. Another EDFA is usually present at the output of the OXC or OADM to amplify the signal for transmission.

Figure 28 Optical Add/Drop Multiplexing Illustration



Appendix B – Multi-Lane OTN Interface

IEEE 802.3 has defined parallel interfaces for 40GBASE-R and 100GBASE-R, and the ITU-T chose to define corresponding OTU3 and OTU4 parallel interfaces. These OTN interfaces could then be used in applications that may benefit from using the higher volume, lower cost Ethernet PHY modules. The OTN signal formats have been added to G.708, and in late 2009, the physical layer specifications will be added to G.695 and G.959.1.

An inverse multiplexing method is used across the physical/logical lanes, based on a 16-byte boundary aligned with the OTU3/OTU4 frame. See Figure 29. The 16-byte increments are distributed among the lanes in a round-robin manner. The lane assignments are rotated on the boundary of each OTUk frame such that the starting group of 16-bytes rotates among the lanes.

Figure 29 OTU3/OTU4 parallel lane interleaving word structure

Rows \ Columns (Bytes)	1 ... 16	17 ... 32	17 ... 32	...	4065 ... 4080
1	(1-16, with FAS)	(17-32)	(33-48)	...	(33-48)
2	(4081-4096)	(4097-5012)	(5013-5028)	...	(5013-5028)
3	(9161-9176)	(9177-9192)	(9193-9208)	...	(9193-9208)
4	(12241-12256)	(12257-12272)	(12273-12288)	...	(16305-16320)

OTU3

The OTU3 interface uses four parallel lanes. The two LSBs of the MFAS are used to determine the lane assignment and rotation, as illustrated in Table 17.

The lane rotation causes the FAS to appear periodically on each lane, which allows framing to be recovered for each lane. The lane can be identified by examining the two LSBs of the MFAS, which will have the same values on each time the FAS appears on a given lane. Lane identification is necessary since the optical modules may not preserve their respective positions. The lanes can be deskewed by comparing the 8-bit MFAS of each lane. Deskew can be achieved as long as it doesn't exceed 127 (2^7-1) OTU3 frames ($\approx$385μs).

Table 17 Starting group of bytes sent in each lane for the OTU3 lane rotation

MFAS 7:8	Lane			
	0	1	2	3
0 0	1:16	17:32	33:48	49:64
0 1	49:64	1:16	17:32	33:48
1 0	33:48	49:64	1:16	17:32
1 1	17:32	33:48	49:64	1:16

OTU4

In order to support both four and ten lane physical interfaces, OTU4 uses 20 logical lanes. However, a different lane marking mechanism is required since the MFAS period is not divisible by 20. A lane marker byte is implemented as a virtual MFAS by borrowing the sixth FAS byte (third OA2 byte) for the OTN domain carrying the parallel interface. The counter in this lane marker byte increments per frame from 0 to 239[49]. The logical lane is recovered from this count value modulo 20. See Table 18 for the OTU4 byte distribution across the 20 logical lanes.

The first five FAS bytes provide the pattern for frame recovery on each logical lane, with the lane marker byte allowing the lane identification and deskewing, similar to the OTU3 case. The deskew range can be extended by combining the lane marker and MFAS counts, to give a maximum deskew range of 1912 OTU4 frame periods (≈2.223 ms).

Five of the 20 logical lanes are bit multiplexed onto each of the four physical lanes to form the four lane interface. The sink performs the inverse interleaving to recover the five logical lanes from each physical lane. Note that each logical lane can appear in any bit position.

The 10-lane interface is implemented in a similar manner to the four-lane interface, with 2-bit multiplexing being used per lane. Note that there is no ITU-T physical interface specification for the IEEE 100GBASE-R 10-lane interface.

[49] As seen in Figure 6, the FAS contains six framing bytes. The sixth byte can be borrowed here because G.798 specifies that only the first four FAS bytes are required for frame alignment.

Table 18 Starting group of bytes sent in each logical lane for the OTU4 frame lane rotation

Lane Marker count (decimal value, mod20)	Lane				
	0	**1**	**2**	**…**	**19**
0	1:16	17:32	33:48	…	305:320
1	305:320	1:16	17:32		289:304
…	…	…	…		…
19	17:32	33:48	49:64	…	1:16

Appendix C – Hitless Adjustment of ODUflex (HAO)

Carriers have found that it is typically much more efficient in terms of equipment cost and power consumption to switch packet streams in a circuit switched manner rather than using a packet switch for rates higher than roughly 500 Mbit/s. One drawback to using circuit switching is that it has typically lacked the ability to adjust the channel bandwidth without tearing down and re-building the connection. In contrast, packet switches have full flexibility to change their rates of packet transmission, however it comes at the cost and complexity of processing each packet.

ODUflex(GFP) provides a very flexible capability to size the transport channel to the desired rate, especially for connections between routers. GFP Idles adapt the packet rate to the OPUflex capacity and GMP provides the mechanism for multiplexing the ODUflex into the HO ODU TS. So, GMP provides an inherent capability for adjusting the ODUflex signal rate. Specifically, since GMP can adjust its C_m to occupy any amount of bandwidth within that set of HO OPU TS, it allows for the possibility of growing the ODUflex(GFP) signal into more TS, or shrinking it into fewer TS. Further, it can allow this resizing to be done in a hitless manner.

Consequently, ITU-T defined a protocol for the hitless adjustment of ODUflex(GFP) channels (HAO) in G.7044. While the details of HAO are a topic for another white paper, its concepts are described in this appendix.

The HAO protocol consists of two parts. One part is a protocol that allows changing the number of TS being used on each link. This part is known as the Link Capacity Resizing (LCR) protocol. The other part adjusts the actual ODUflex signal rate to fit within the target number of TS, and is known as the bandwidth resizing (BWR) protocol.

In the case of increasing the ODUflex(GFP) rate, the first step is for the nodes to add one or more TS to each link, and increase their own internal connections to that rate. A LCR handshake is defined to allow the nodes on each end of a link to start using the new number of TS at the same time. Once all the links have been resized, the sink informs the source that entire path is ready, and the source begins increasing the rate of the ODUflex(GFP signal using BWR. HAO terminates when the ODUflex(GFP) signal is operating at the appropriate rate for the new number of TS.

The case of decreasing the ODUflex(GFP) signal rate is similar, except that the ODUflex signal rate has to first be reduced to fit into the target number of TS. Specifically, the first step is for the LCR protocol to set up the path. Then the BWR reduces the rate of the ODUflex, and finally the LCR protocol completes by coordinating the removal of the TS that are being removed.

13 References

[1] ITU-T Recommendation G.709 (2003), *Interfaces for the Optical Transport Network (OTN)*

[2] ITU-T Recommendation G.798 (2004), *Characteristics of optical network hierarchy equipment functional blocks*

[3] ITU-T Recommendation G.872 (2001), *Architecture of optical transport networks*

[4] ITU-T Recommendation G.7044/Y.1347 (2011) *Hitless Adjustment of ODUflex(GFP) (HAO)*

[5] M. Elanti, S. Gorshe, L. Raman, and W. Grover, *Next Generation Transport Networks – Data, Management, and Control Plane Technologies*, Springer, 2005.

[6] PMC-2030895, "A Tutorial on SONET/SDH," PMC-Sierra white paper by S. Gorshe.

[7] S. Kartalopoulos, *Introduction to DWDM Technology*, IEEE Press, Piscataway, HJ, 2000

[8] ITU-T Recommendation G.692 (1998), Optical interfaces for multichannel systems with optical amplifiers

[9] ITU-T Recommendation G.8251 (2001), T*he Control of Jitter and Wander within the Optical Transport Network*

[10] ITU-T Recommendation G.707 (1996), Synchronous Digital Hierarchy Bit Rates.

[11] ITU-T Recommendation G.7042/Y.1305 (2001), "Link capacity adjustment scheme for virtual concatenated signals."

[12] ITU-T Recommendation G.694.1 (2002) *Spectral grids for WDM applications: DWDM frequency grid*

[13] ITU-T Recommendation G.694.2 (2003) *Spectral grids for WDM applications: CWDM frequency grid*

[14] ITU-T Recommendation G.975.1 (2004), Forward error correction for high bit rate DWDM submarine systems

[15] ITU-T Supplement G.Sup43 (2008), Transport of IEEE 10G Base-R in Optical Transport Networks (OTN)

[16] PMC-2061695, "ROADMs and the Evolution of the Metro Optical Core," PMC-Sierra white paper by T. Mashologu and M. Orthodoxou

[17] PMC-2041083, "Generic Framing Procedure (GFP)," PMC-Sierra white paper by S. Gorshe

[18] ITU-T Recommendation G.7041 (2008) Generic Framing Procedure (GFP)

[19] PMC-2050248, "A Tutorial on SONET/SDH Automatic Protection Switching (APS)," PMC-Sierra white paper by S. Gorshe

[20] PMC-2091509, "Overview of the "Swizzle" FEC for 40G and 100G Optical Transmission," PMC-Sierra white paper by Phil Northcott

[21] PMC-2110752, "Enabling OTN Convergence – Solutions for the New Packet Optical Transport Network," PMC-Sierra white paper by Mark Orthodoxou

14 Glossary and Abbreviations

Term	Definition
3R	Re-amplification, Reshaping and Retiming
10GE	10 Gbit/s Ethernet
40GE	40 Gbit/s Ethernet
100GE	100 Gbit/s Ethernet
CBR	Constant Bit Rate
CESR	Carrier Ethernet Switch / Router
CM	Connection Monitoring
CRC-*n*	*n*-bit Cyclic Redundancy Check error detection code
CWDM	Coarse Wavelength Division Multiplexing
DFB	Distributed Feedback laser
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbium-Doped Fiber Amplifier
FEC	Forward Error Correction
GE	Gbit/s Ethernet
GFP	Generic Framing Procedure (ITU-T Rec. G.7041)
GFP-T	Transparent mode of GFP
GMP	Generic Mapping Procedure
HAO	Hitless Adjustment of ODUflex(GFP) signals
IaDI	Intra-Domain Interface
IrDI	Inter-Domain Interface
JC	Justification Control
JOH	Justification Overhead
MFAS	MultiFrame Alignment Signal

MQW	Multiple Quantum Well laser
MSI	Multiplex Structure Identifier
naOH	non-associated overhead
OADM	Optical Add-Drop Multiplexer
OCC	Optical Channel Carrier
OCh	Optical channel with full functionality
OCI	Open Connection Indication
ODU	Optical Channel Data Unit
ODUflex(CBR)	Flexible rate ODU for carrying CBR client signals
ODUflex(GFP)	Flexible rate ODU for carrying packet client signals that use a GFP-F mapping into the OPUflex
ODUk	Optical Channel Data Unit-k
ODTUjk	Optical channel Data Tributary Unit j into k
ODUk-Xv	X virtually concatenated ODUk's
OH	Overhead
OMS	Optical Multiplex Section
ODTU	Optical channel Data Tributary Unit
ODTUG	Optical channel Data Tributary Unit Group
ODTUjk	ODTU for multiplexing a LO ODUj signal into a HO ODUk
ODTUk.ts	ODTU for multiplexing a LO ODU into “ts” Tributary Slots of a HO ODUk
OMU	Optical Multiplex Unit
ONNI	Optical Network Node Interface
OOS	OTM Overhead Signal
OPS	Optical Physical Section
OPU	Optical Channel Payload Unit
OPUk	Optical Channel Payload Unit-k

OPUk-Xv	X virtually concatenated OPUk's
OSC	Optical Supervisory Channel
OTH	Optical Transport Hierarchy
OTM	Optical Transport Module
OTP	Optical Transport Platform
OTS	Optical Transmission Section
OTU	Optical Channel Transport Unit
OTUk	completely standardized Optical Channel Transport Unit-k
OTUkV	functionally standardized Optical Channel Transport Unit-k
OXC	Optical cross-connect equipment
PDH	Plesiochronous Digital Hierarchy
P-OTP	Packet Optical Transport Platform
PSI	Payload Structure Identifier
PT	Payload Type
ROADM	Reconfigurable Optical Add / Drop Multiplexer
TC	Tandem Connection
TCM	Tandem Connection Monitoring
TxTI	Transmitted Trace Identifier
vcPT	virtual concatenated Payload Type
WDM	Wavelength Division Multiplexing

15 Notes

Contacting PMC-Sierra

PMC-Sierra
8555 Baxter Place
Burnaby, BC
Canada V5A 4V7

Tel: +1 (604) 415-6000
Fax: +1 (604) 415-6200

Document Information: document@pmc-sierra.com
Corporate Information: info@pmc-sierra.com
Technical Support: apps@pmc-sierra.com
Web Site: http://www.pmc-sierra.com