ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Technical Paper

(10/2025)

GSTP-OTN

Evolution of optical transport networks

Technical Paper ITU-T GSTP-OTN

Evolution of optical transport networks

Summary

This Technical Paper provides a comprehensive overview of the optical transport network (OTN), which is the current generation technology used in telecommunication networks for transporting various client signals including Ethernet and legacy protocols. The paper covers multiple aspects of OTN such as architecture, frame structures, operations, protection, synchronization, and management and control mechanisms, reflecting the evolution and standardization efforts by ITU-T since the 1970s.

Keywords

Client mapping, constant bit rate (CBR) clients, delay measurement, electrical interfaces, Ethernet clients, flexible OTN (FlexO), FlexO link security (FlexOsec), forward error correction (FEC), FOIC, justification control (rate adaptation), multiplexing, optical interfaces, optical transport network (OTN), packet clients, performance monitoring, tandem connection monitoring.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Table of Contents

Page

Technical Paper ITU-T GSTP-OTN

Evolution of optical transport networks

0 Preface

The ITU-T, beginning in the mid-1970s, has standardized several generations of transport technologies following the evolution of telecommunication services and transport capabilities. The content of this technical paper is based on the content of the Recommendations issued by ITU-T on optical transport networks.

This Technical Paper is focused on the optical transport network (OTN), which is the current technology generation employed in telecommunication networks. OTN can support other network layers such as Ethernet, but these are only briefly considered as OTN clients. This Preface deals with some general issues related to optical transport networks.

Each clause of this Technical Paper is dedicated to a specific aspect (architecture, multiplexing hierarchy, protection, synchronization, etc.) of OTN.

0.1 Characteristics of optical transport networks

An optical transport network is comprised of a set of optical network elements (wavelength division multiplexers (WDM), optical amplifiers (OA), optical add-drop multiplexers (OADM), optical cross-connects (OXC), etc.) connected by optical fibre links. Optical channels carrying client signals are supported in the network with transport, multiplexing, routing, management, supervision and survivability functions. As shown in Figure 0-1, the optical transport network has the following characteristics:

Client independence: The functions of an OTN can be independent of the client signals. OTN can support the full range of switching technology clients including legacy SDH, Internet Protocol (IP), multiprotocol label switching (MPLS), Ethernet, and so on.

Layered architecture: An OTN is composed of a media layer that supports the transmission of light waves, and a digital layer whose information is modulated into light waves. The digital layer processes the data in the electrical domain and is implemented by equipment which carries out the functions of signal multiplexing, routing, supervision, performance monitoring, surveillance and network survivability. The optical (media) layer provides light paths to the client (digital) layer, with the light paths supported over optical links, optical amplifiers, optical cross connects, and other optical devices.

Wide-range client coverage: The client (digital) layer of an OTN is characterized by its network architecture, multiplexing hierarchy, supervision techniques, network synchronization technique and network survivability procedures. All of these characteristics are specified by the ITU-T in the standardization of the OTN and will be described in this paper.

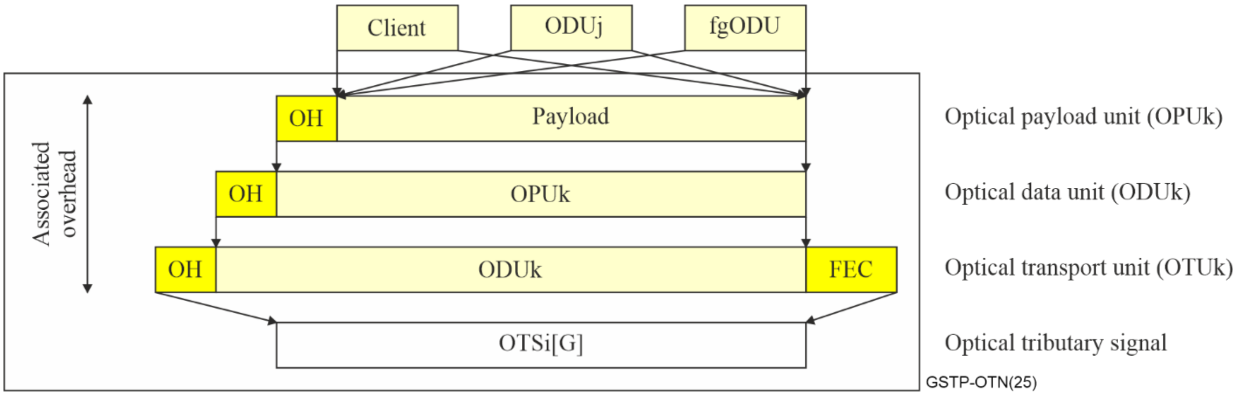


Figure 0-1 – Key components and characteristics of optical transport network

Transparency: It allows an operator to provide a variety of different services using a single infrastructure. The highest degree of network transparency can be obtained with all optical networks, where data is carried from its source to its destination in optical form, without undergoing any optical-to-electrical conversions along the way. The lowest degree of transparency is obtained building a network that handles essentially a single bit rate and protocol. At present, most practical networks are partially transparent, consisting of all-optical sub-networks interconnected by optical-to-electrical-to-optical (OEO) converters. Signals can be regenerated at OEO converters and intelligent control and management functions can be performed on the client (digital) layer once in the electrical domain.

Multiplexing: Most modern optical fibre systems are capable of transmitting at a rate of several hundreds of Gbit/s. To utilize the system capacity fully, many channels can be transmitted simultaneously on the optical fibre by means of multiplexing. This can be accomplished through frequency division multiplexing (FDM) or time division multiplexing (TDM). In the case of TDM, bits associated with different channels are interleaved in the time domain to form a composite bit stream. TDM is generally implemented for digital signals and is commonly used in telecommunication networks to multiplex a large number of client signals (channels) into a single electrical bit stream. In the case of FDM, the channels are spaced apart in the frequency domain. FDM in the optical domain is referred to as wavelength division multiplexing (WDM).

Packet data transport: Packet data signals (e.g., Ethernet) could either be transported via optical wavelength by a specific adaptation of packet frames on optical wavelengths (e.g., [b-IEEE 802.3] PMD) or transported via OTN through frame encapsulation. Generic framing procedure (GFP) is a common method employed to adapt diverse packet protocols at the link layer to be transported over OTN.

Operation: There are several types of optical transport networks which have been deployed over the years in public telecommunication networks as well as in private enterprise networks, including SDH, OTN, Ethernet and MPLS-TP.

0.2 Evolution of optical transport network (OTN) key aspects

Since the 1980s, capabilities were provided by developments in optical component technology, which, in particular, enabled the transport of multiple wavelengths (colours) on a single optical fibre. The OTN technology seamlessly combines multiple networks and services into a common, future-ready infrastructure. As a network, the OTN comprises OTN switching technologies, WDM switching technologies, Ethernet over transport, Management and Control (MC) and Synchronization. The evolution of OTN is described in Figure 0-2.

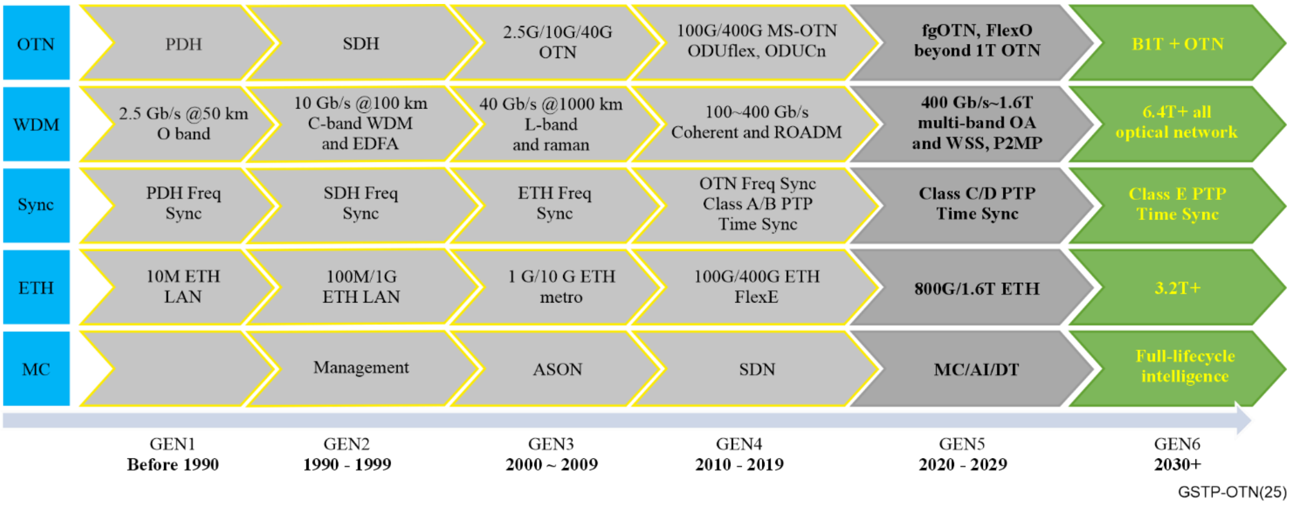


Figure 0-2 – Evolution of optical transport network

0.2.1 OTN switching technologies

After the development of PDH/SDH, the first version of the OTN Recommendations was published in 2001. It defined three types of ODU carrying 2.5G, 10G and 40G STM-N clients. From 2009, OTN evolved to a multi-service optical transport technology, and defined ODU0, ODU4, and ODUflex with GMP to support any bit-rate services including 1G, 10G, 40G and 100G Ethernet. From 2016, OTN further evolved towards the beyond 100G with ODUCn and FlexO to satisfy larger bandwidth transmission requirements including higher rate services e.g., 200G, 400G and 800G Ethernet. Since 2020, OTN evolved towards two directions. One is the fine grain optical transport network (fgOTN), which was defined with fgODUflex and 10M tributary slots for many sub1G clients. Another is beyond 1T OTN, towards an ultra-high rate to meet ultra-large bandwidth requirements such as backbone bandwidth upgrade and data centre interconnection.

The main characteristics of the OTN can be summarized as follows:

– OTN has been designed not only for high data transmission rates, but also for various client rates. The OTN line rates are OTU0, OTU1, OTU2, OTU3, OTU4, and OTUCn/FlexO-n. The OTUCn/FlexO-n has a rate of n × 100 Gbit/s on a single wavelength of a fibre. Moreover, it could multiplex tributaries of lower bit rate ODUs, even sub1G fgODU.
– OTN has been designed to employ forward error correction (FEC) as an overhead. At very high bit rates and over very long distances, noise is significant and becomes a problem when ensuring low bit error rates. FEC is critical to achieving these low bit error rates [b-ITU Handbook].
– OTN provides overhead for monitoring connections end-to-end and over various link segments. This overhead includes signal identification, error measurement and alarm information.
– OTN enables operations, administration and management of connections that are transparent to its clients such as IP, MPLS and Ethernet.
– OTN has asynchronous mapping of client signals into OTN frames where the clock that generates the frames can be a simple free-running oscillator. To account for any mismatch between the clocks of the OTN frames and the client signals, the OTN payload floats within the frame. Using simple free-running oscillators can simplify implementation and reduce costs. OTN also has synchronous mapping where the clock to generate the OTN frames is derived from the client signal.

0.2.2 Optical layer technologies

The optical layer technologies evolve from the following dimensions:

– Wavelength & Spectrum: While pre-telecommunication lightwave systems operated around 850 nm in the early years, the telecommunication systems operate around 1310 nm (O band), 1550 nm (C band) and 1600 nm (L band) respectively.

– Transport distance: transportation distances increased from 10 km in 1970 to >1000 km in 2000.

– Line side rate: It basically expands four times every ten years, starting from 2.5 Gbit/s and up to 400~1 600 Gbit/s. Note that for rates below 40 Gbit/s, intensity modulation direct detection (IMDD) is used for modulation/demodulation, while for rates higher than 100 Gbit/s, coherent technology is used.

– Switching technology: Since 1990 wavelength-division multiplexing (WDM) has increased fibre capacity beyond single wavelength systems. Fixed optical add/drop multiplexer (FOADM) has been developed for 2nd and 3rd generations and then reconfigurable optical add/drop multiplexer (ROADM) become mature since the 4th generation.

– Amplifier: Erbium-doped fibre amplifier (EDFA) and Raman amplifiers were developed in the 2nd and 3rd generations respectively, augmented with multi-band amplifiers in the 5th generation.

– Fibre: Throughout generations 1 to 5, single-mode fibre has been widely deployed. Since generation 5, hollow-core fibre and space-division multiplexing based on multi-core fibre and few-mode fibre have been explored.

0.2.3 Ethernet over transport

The specification of Ethernet by IEEE started in the 1970s as a packet switched data link that connects computers and computer equipment over a single coaxial cable, that is, a bus. Ethernet has since evolved to include a variety of topologies including point-to-point, bus, star and mesh and adapted to a variety of physical communication media, including coaxial cable, twisted pair copper cable, wireless media and optical fibre. It has now a wide range of rates. Traditional rates were 10 Mbit/s and 100 Mbit/s. Typical rates today are 1 Gbit/s, 10 Gbit/s, 40 Gbit/s, 100 Gbit/s and 200 Gbit/s. Recently, 400 Gbit/s and 800 Gibit/s Ethernet are also completing standardization and standardization for 1.6 Tbit/s has started.

Ethernet has a data link layer (PMD) and a physical (PHY) layer. Both layers have been specified in [b-IEEE 802.3] standards. At the PHY layer the IEEE specifies various optical transceivers for short range and long-range reaches, with coherent and intensity modulation direct detection (IMDD) modulation/demodulation techniques.

With the pervasiveness of Ethernet in enterprises, service providers offer Ethernet connectivity across multiple sites. As a consequence, Ethernet is a carrier transport technology within service provider networks themselves and these customers require carrier grade services, including high availability.

0.2.4 Management and control

To support services over OTN, management systems were deployed to create connections between endpoints. Fault management, configurations, accounting, performance monitoring and security, known as FCAPS, were initially used to manage the OTN. Capabilities to automate these were developed as the automatically switched optical network (ASON) and software defined networking (SDN) respectively in later phases.

Both centralized and distributed architectures exist for these capabilities. They share common functions as described in [b-ITU-T G.7701]. The distributed approach evolved first from the early management systems and is ASON. As described in [b-ITU-T G.7703], it uses common functions in [b-ITU-T G.7701] and has been realized in variations of packet signalling and routing. A more centralized approach using SDN concepts is described in [b-ITU-T G.7702]. It too uses functions in [b-ITU-T G.7701] but packages control of resources into SDN controllers that can be hierarchically arranged. Both SDN and ASON approaches can be used to achieve dynamic connection setup in large OTN networks.

0.2.5 Timing and synchronization aspects

Network synchronization has been playing an important role in telecommunications particularly since service providers started to deploy SDH optical systems in the 1990s, as mentioned above. Digital switching equipment required synchronization to avoid slips at ingress elastic stores and guarantee the high quality of circuit switched networks. As a result, a suitable strategy of network synchronization was required, and the master-slave synchronization method was standardized for SDH synchronization networks.

In the 2000s, packet technologies such as Ethernet and MPLS began to appear in telecommunication networks. With Ethernet as a major telecommunication service, SyncE (sync over Ethernet) technology was specified and widely deployed.

Since OTN technology was specified in 2000s, carrying constant bit rate (CBR) clients required OTN to transparently recovery the CBR client clock in order to successfully transmit the CBR client data. OTN mapping/demapping, multiplexing and demultiplexing clock were specified for some CBR clients, e.g., SDH and Fiber Channel.

After 2010, PTP clock technology-based on [b-IEEE 1588] was standardized to meet synchronization for wireless base stations. From the initial frequency synchronization for 3G/4G FDD to the later time synchronization for 4G/5G TDD, many ITU-T Recommendations were specified, including PTP over Ethernet, OTN and IP networks. As requirements for higher accuracy and more robust synchronization appear, new synchronization technologies are expected.

1 Transport network architecture

1.1 Introduction

The primary purpose of a transport network is to transfer user information from a sender at one location to a receiver at another location. The transport network must also transfer various kinds of network control information such as signalling and operations and maintenance information that are required to operate the transport network.

The objectives of the network architecture (functional modelling) specifications are to support the description of the generic characteristics of transport networks in a way that is independent of how the functions are instantiated. A common language and symbols are used in the specification of transport and management functionalities, which are essential for network design and management.

[b-ITU-T G.800] and [b-ITU-T G.807] provide the functional architecture of transport networks and media in a technology-independent way. These generic functional architectures of transport networks are the basis for a harmonized set of functional architecture Recommendations for specific layer network technologies that use circuit-switching or packet-switching technology (e.g., [b-ITU-T G.872], [b-ITU-T G.8010], [b-ITU-T G.8310]). These technology-specific architecture Recommendations are used as the basis for a corresponding set of Recommendations for interfaces, management, and equipment specifications.

The first generic functional architecture Recommendation was [b-ITU-T G.805], the first version of which was approved in 1996. This describes connection-oriented networks and was used as the basis for [b-ITU-T G.803] (SDH) and [b-ITU-T G.872] (OTN) architecture Recommendations. In connection-oriented networks a connection is set up prior to the transfer of information across the network.

To expand beyond connection-oriented networks, [b-ITU-T G.800] was developed. This Recommendation provides a common framework to describe both connection oriented and connectionless networks. In connectionless networks, datagrams (or individual frames) are transferred through the network without any prior negotiation of routes or allocation of resources. For connection-oriented networks, the descriptions provided in [b-ITU-T G.800] are fully compatible with the descriptions derived from [b-ITU-T G.805], although some of the terminology has been modified.

[b-ITU-T G.807] was developed to document the architecture of the media layer. The media layer is fundamentally different from the digital layer networks in that it cannot be directly monitored and managed. It supports optical signals that contain modulated digital layer information.

1.2 OTN architecture

Optical transport networks are comprised of both digital and optical functions. The functionality of the optical transport networks (OTN) is described from a network level viewpoint in [b-ITU-T G.872] using the generic principles defined in [b-ITU-T G.800] and [b-ITU-T G.807]. [b-ITU-T G.872] describes the specific aspects concerning the optical transport network layered structure, characteristic information, client/server layer associations, network topology and layer network functionality, while [b-ITU-T G.807] describes the optical media network functions and their management.

In accordance with [b-ITU-T G.800], the optical transport network is decomposed into independent transport layer networks where each layer network can be separately partitioned in a way that reflects the internal structure of that layer network. Layer networks often have a client-server relationship.

In the following functional description, optical signals (OTSi) are supported in network media channels that are related to frequency and wavelength grids in other ITU-T Recommendations.

Deployment of a pure optical network is not practical for a variety of reasons, including the need to digitally multiplex signals to higher bit rates to maximize spectral efficiency, the need to regenerate optical signals to support long-haul transmission, and the need to monitor the performance of the network to ensure service requirements are being satisfied.

As such, [b-ITU-T G.709] defines a hierarchy of digital layer networks. Client signals are mapped into a path layer that provides monitoring from the point the client enters the OTN to the point where it leaves the OTN; this path layer is called the optical data unit (ODU). An ODU may be aggregated (along with other ODUs) into a higher-order ODU or may be mapped into a section layer called Optical Transmission Unit (OTU). For rates up to 100 Gbit/s, the OTU is directly modulated onto an Optical Tributary Signal (OTSi) for transmission over the media layer. For rates beyond 100G, the server layer is further mapped into an FEC frame and potentially inverse multiplexed; the technology supporting those functions is called "FlexOTN (FlexO)". The FlexO frame also supports direct mapping of Ethernet signals to facilitate optimized transport of Ethernet signals over an optical layer.

The high-level structure of the OTN is provided in Figure 1-1.

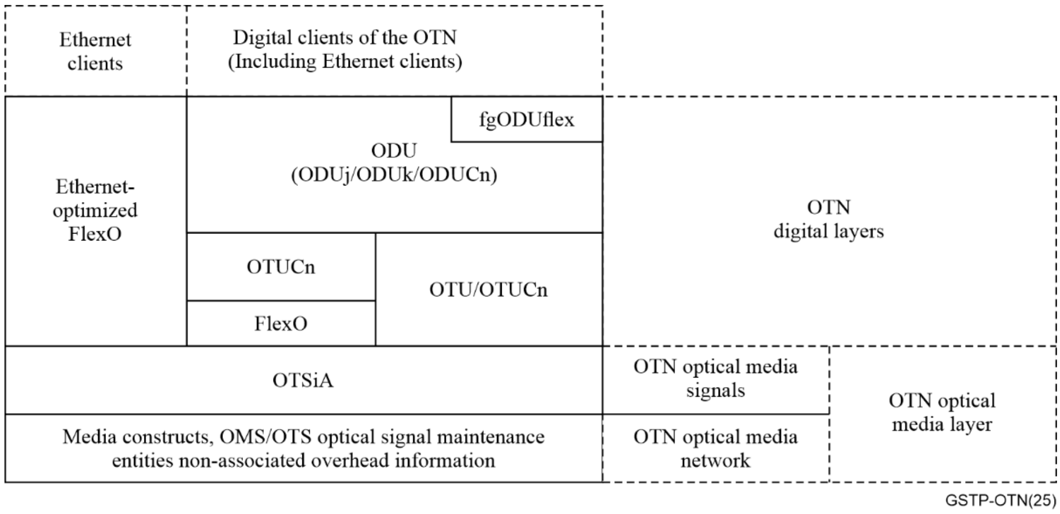


Figure 1-1 – Overview of the OTN (based on Figure 6-1 of [b-ITU-T G.872])

1.3 Digital layers of OTN

1.3.1 Optical channel data unit (ODU) layer network

The ODU layer network provides functionality for end-to-end networking of digital path signals that transparently convey client information. The client of an ODU is either a non-OTN signal or a tributary slot (TS) structure that enables a higher-rate ODU to carry a multiplex of lower-rate ODUs. As in any connection-oriented layer network, the topological components of the ODU layer network are subnetworks and links. The links are supported by an OTU trail or a server ODU trail. Since the resources that support these topological components support a heterogeneous assembly of ODUs at different bit rates, the ODU layer is modelled as a single layer network that is independent of bit rate. To provide end-to-end networking the following capabilities are included in the layer network:

– ODU connection rearrangement for flexible network routing;
– ODU overhead processes for ensuring the integrity of the transfer of the ODU-adapted information;
– ODU operations, administrations, and maintenance functions for enabling network-level operations and management functions, such as connection provisioning, quality of service parameter exchange and network survivability.

The structure of the ODU and the associated functions are described in clause 2.2.

1.3.2 Optical channel transport unit (OTU) layer network

The OTU layer network provides functionality for the transport of ODU client signals through an OTU trail between 3R points of the OTN. It transparently conveys a single ODU client signal. The capabilities of this layer network include:

– OTU overhead processes and conditioning for the transport over optical channels for ensuring the integrity of the OTU-adapted information;
– OTU operations, administrations, and maintenance functions for enabling section level operations and management functions, such as OTU survivability.

The structure of the OTU and the associated functions are described in clause 2.2.

1.3.3 Optical tributary signal (OTSi)

OTU connections are supported by one or more OTSi in the optical media layer. As discussed in clause 1.4, there are additional digital layers used to provide non-associated monitoring of the optical layer.

1.3.4 Application scenarios

There are different types of FlexO interfaces defined in the ITU-T G.709.x series of Recommendations. Figure 1-2 provides an overview of the applications. Short-reach interfaces are often referred to as client or grey interfaces, and they can be used as handoffs between different administrative domains (e.g., different carriers, different equipment vendors, metro vs core). These typically target Ethernet-type of optical modules and can be used directly from routers or from transport equipment. Long-reach interfaces are often referred to as line or DWDM interfaces, and they can be used for multi-span metro applications over longer distances. Long-reach interfaces can be single-vendor bookended or multi-vendor interoperable.

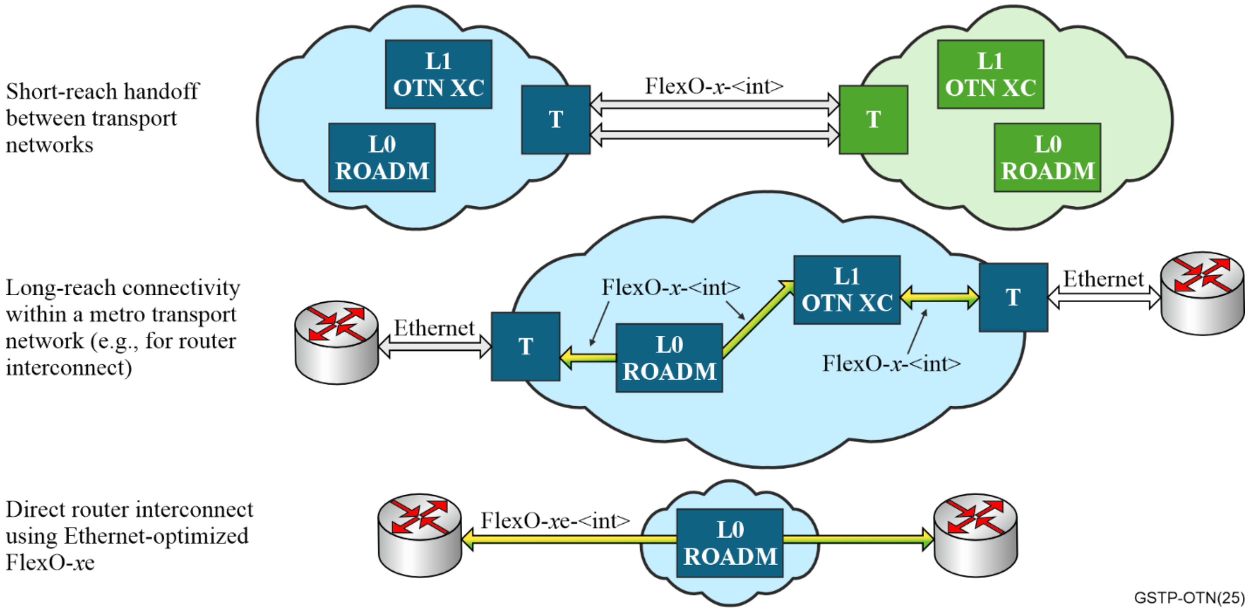


Figure 1-2 – FlexO – Application examples

1.4 Optical media layer

The network architecture described so far concerns the transfer of information in digital layer networks. The information transferred by a layer network is called the characteristic information (CI) of that network. Examples of CI are the Ethernet MAC frame and OTN ODUk frame. To transfer CI over optical media, the CI is modulated onto one or more optical signals. [b-ITU-T G.807] describes the generic functional architecture of the optical media layer from the viewpoint of the propagation of optical signals and the non-associated digital overhead, in the context of a transport network. It extends some of the symbols and constructs of [b-ITU-T G.800] / [b-ITU-T G.805], is independent of the client CI carried in the optical signals, and is generic to optical technologies such as modulation.

[b-ITU-T G.807] provides a modelling construct to modulate and demodulate information into and from a signal. This is illustrated in Figure 1-4 where the output of OTU adaptation is input to a transmit/receive function which includes modulation/demodulation, denoted by the symbol .

The output of a modulator is an optical tributary signal, or OTSi, that is defined in [b-ITU-T G.959.1]. The received OTSi is input to a demodulator where the digital information stream is recovered.

The optical media layer is defined as the set of constructs that support and constrain the propagation of OTSi. Figure 1-3 illustrates the boundary of an optical media layer and is shown below. A key difference between the optical media layer and the digital layers previously discussed is that the media cannot directly be monitored; only the signals that the media carries can be monitored.

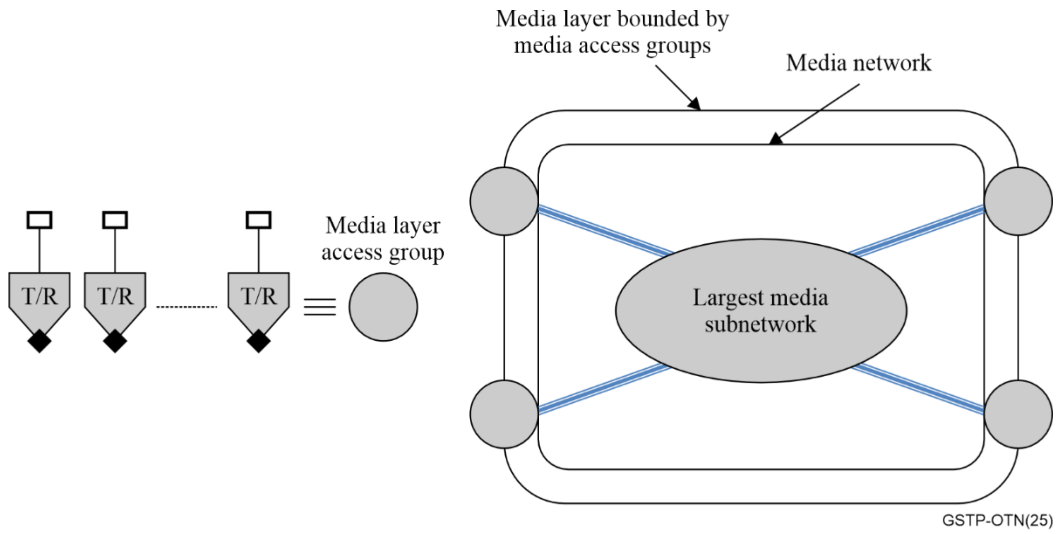


Figure 1-3 – Overview of optical media layer (based on Figure 6-1 of [b-ITU-T G.807])

The optical media layer is bounded by sets of modulators/demodulators (labelled "T/R") which originate and terminate the OTSi that are supported by the media. Within the media subnet are other constructs that model functions that operate on OTSi. These functions include filters, couplers, optical parameter monitors, amplifiers, and frequency mixers.

An OTSi propagates in the media within a **frequency slot**, which is a contiguous frequency range, characterized by a nominal central frequency and slot width, that does not overlap with any other frequency slot. It is associated with a **media channel** which represents topology. A media channel guides zero or more OTSi between its ports and "may be formed by a serial concatenation of multiple media channels, each with its own frequency slot". Figure 7-2 of [b-ITU-T G.807] illustrates the OTSi being generated, combined in a coupler to one fibre, and how they occupy frequency slots. It is replicated in Figure 1-4 below.

The figure also illustrates **media ports** which represent the ends of a media channel. They may also represent the S_S and R_S reference points defined in [b-ITU T G.698.1].

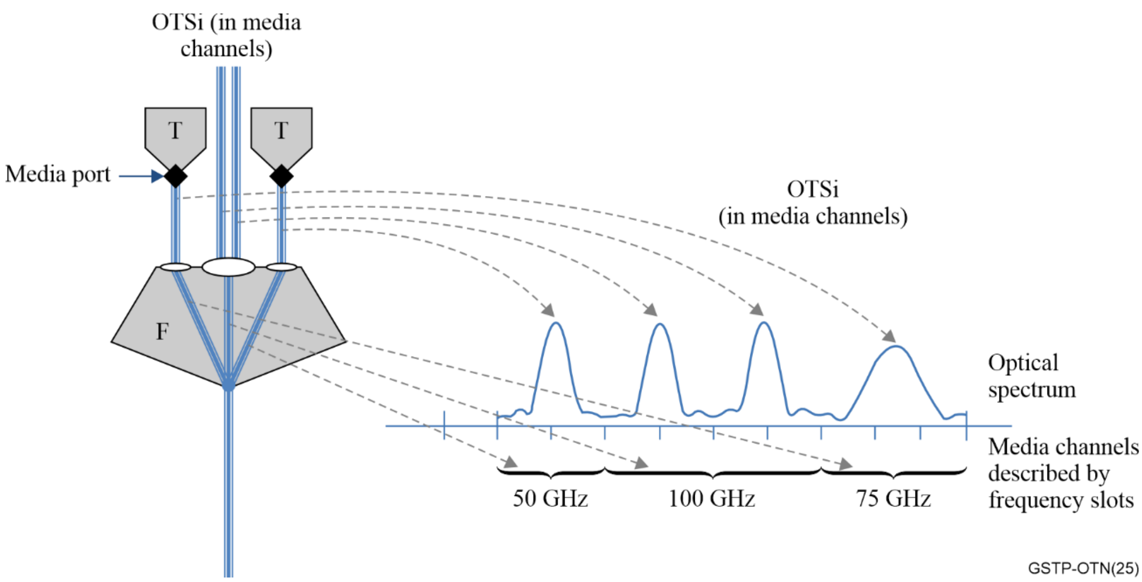


Figure 1-4 – Relationship between signals and media channels (based on Figure 7-2 of [b-ITU-T G.807])

A **network media channel** (NMC) is a type of media channel that is formed by the serial concatenation of all media channels between the media port of a modulator and the media port of a demodulator. It must be present to allow the OTSi to be propagated from the modulator to the demodulator, cannot be divided into smaller media channels, and cannot be concatenated with another media channel.

Figure 1-5 (see Figure 10-6 of [b-ITU-T G.807]) illustrates an OTSi carried in an NMC. The NMC may exist prior to the OTSi. Some types of clients are supported in the media using multiple OTSi and the management abstraction representing these are known as an **optical tributary signal group** (OTSiG). The management abstraction representing the set of NMCs these OTSi use are known as an **NMC group** (NMCG). In the figure, client CI is adapted to media-layer-adapted information, or **M-AI** which is input to a modulator. The media layer access point, or **M-AP**, is analogous to the ITU-T G.800 access point and is the reference point "at the boundary of the media layer that represents the binding between the modulator/demodulator and adaptation functions".

Management of media-layer networks is necessarily external because the media itself cannot be monitored. Media supports and guides optical signals, the propagation of which is used to accomplish information transfer. [b-ITU-T G.807] defines two topological constructs, the OTS media channel group (MCG) and OMS MCG, which are used for management control purposes.

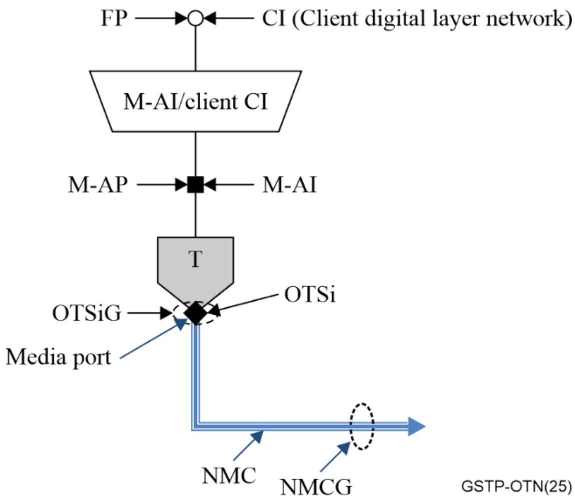
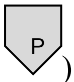


Figure 1-5 – Mapping a client digital stream to an OTSiG that contains one OTSi (based on Figure 10-6 of [b-ITU-T G.807])

The OTS MCG represents the topological relationship between the output media port of one amplifier and the input media port of the next amplifier. Non-associated overhead (OTS-O) may be provided for the OTS MCG; the combination of the OTS MCG and OTS-O are referred to as the OTS media channel assembly (MCA).

The OMS MCG represents the topological relationship between the media port on a filter or coupler where a set of media channels are aggregated and the media port on a filter or coupler where one or more media channels is added to or removed from that aggregation. Non-associated overhead (OMS-O) may be provided for the OMS MCG; the combination of the OMS MCG and OMS-O are referred to as the OMS MCA.

Media networks do not have any monitoring capabilities of their own, so the state of a media channel is inferred by monitoring the OTSi that are expected to be present in that media channel and via the use of the non-associated overhead. Monitoring of OTSi is achieved by attaching an optical parameter monitor (**OPM-x**) to the media channel. The output of the OPM-x is digital information related to what is being monitored (e.g., power). OPM-x operates on the optical wave and not on any client information in the OTSi. In fact, there may not be any client information modulated on the OTSi.

Monitoring of OTS and OMS MCGs is accomplished by an optical signal maintenance entity (OSME) for the set of the OTSi that traverse a MCG. When an OPM-pwr (denoted as 

) function is attached at each end of an MCG the OSME and its reference points are created. The OPM-pwr measures the power of any optical signals (OTSi and optical noise) that is present in the MCG at the OSME reference point. This is applied to both OTS and OMS MCGs and is illustrated in Figure 1-6.

The non-associated overhead provides information about what OTSi are expected to be present in an OSME and allows additional monitoring (e.g., detection of misconnections).

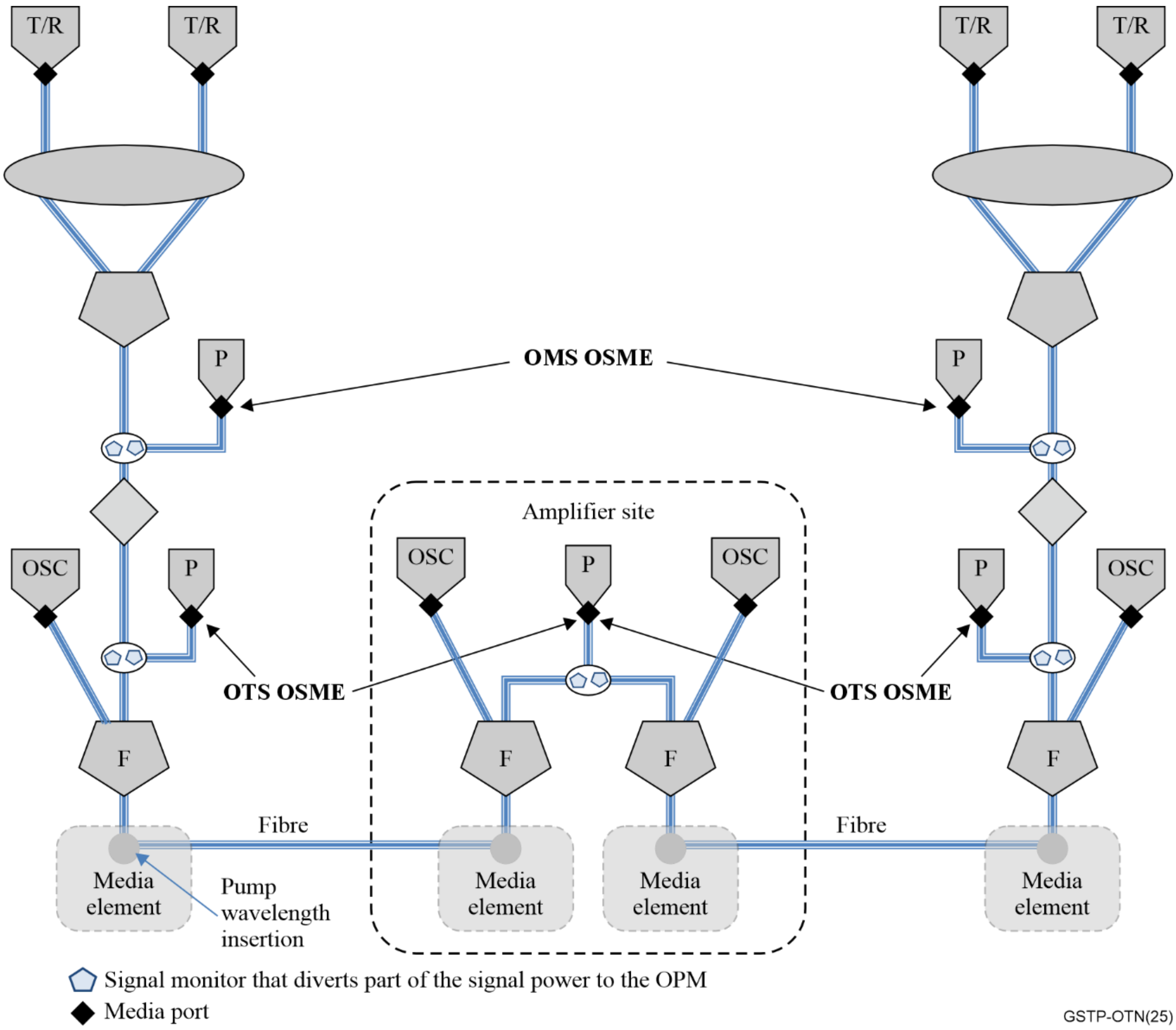


Figure 1-6 – Optical signal maintenance entities with discrete amplifiers (based on Figure 9-2 of [b-ITU-T G.807])

2 Frame structures, mapping and multiplexing

2.1 Introduction

Today's optical transport network technologies are focused on efficient and reliable transport of information that is digitized for transport. Whether it is voice, video, or raw data of some other type (e.g., connections between data centres), the digital content of those signals must be transparently transported, monitored, and protected. To avoid requiring different transport equipment for every different type of signal that might be carried, all types of digital information are adapted to fit into standardized containers and carried over the same equipment. In this way, only the adaptation function is unique to each payload or client signal type.

The digital transport signal consists of a repeating, fixed-length frame structure that provides a "container" field for carrying the digital client payload and additional bytes of data called overhead information. The overhead is used to identify the payload type and structure, to carry control information used for encapsulating and extracting the data, for performance monitoring of the frame (e.g., error detection), detection of network faults and degradations, and for carrying control/management messages. The frames are structured in such a way that the transmitting and receiving equipment can identify the start and end of each frame and extract the overhead and payload reliably. As illustrated below in this clause, the frame structures are defined in terms of rows and columns of bytes. Each frame is transmitted one byte at a time for each row, one row at a time, until all rows in the frame are sent, followed immediately by the next frame. This clause describes the optical transport network (OTN), which is the most commonly used optical transport protocol defined in ITU-T Recommendations.

The payload area of an OTN signal can either be entirely dedicated to a single client signal or carrying a combination of multiple lower-rate client signals. The OTN signal carrying multiple clients can be treated as a single entity, which can significantly reduce network management complexity.

OTN uses a time division multiplexing (TDM) method for carrying multiple client signals in the same OTN payload area. TDM works by dividing the payload area into regularly spaced intervals (e.g., bytes) that enable time-sharing the payload area among the client signals on a round-robin basis. The sink node can easily recover the multiple client data streams since the location of the data from each client is identified by the location within the frame. OTN frame error detection overhead ensures that all the clients received in an error-free OTN frame have not experienced any errors during that frame period, which eliminates the necessity to independently monitor each of the individual lower speed client signals over that portion of the network.

The regular spacing significantly reduces the complexity of mitigating jitter for each client signal[1]. (see clause 5.)

Figure 2-1 illustrates a basic OTN network where user network elements (NEs) interface with OTN NEs using either their native client interfaces or OTN interfaces. Typical OTN CPEs include but are not limited to Ethernet switches, IP routers, and SDH equipment. OTN NEs may include add drop multiplexers (ADMs), digital cross connects, Reconfigurable optical add drop multiplexers (ROADMs), and integrated packet optical transport systems (POTSs).

[1] The most common alternative to TDM is packet multiplexing, which is used by Ethernet. With packet multiplexing, information from a client is accumulated to form the payload area of a packet (typically between 64 and 1500 bytes). There is no requirement that the packets all be the same length, which means that there is a large spacing and typically irregular spacing between packets associated with a given client. Consequently, handling client jitter is much more complicated with packet multiplexing than with TDM.

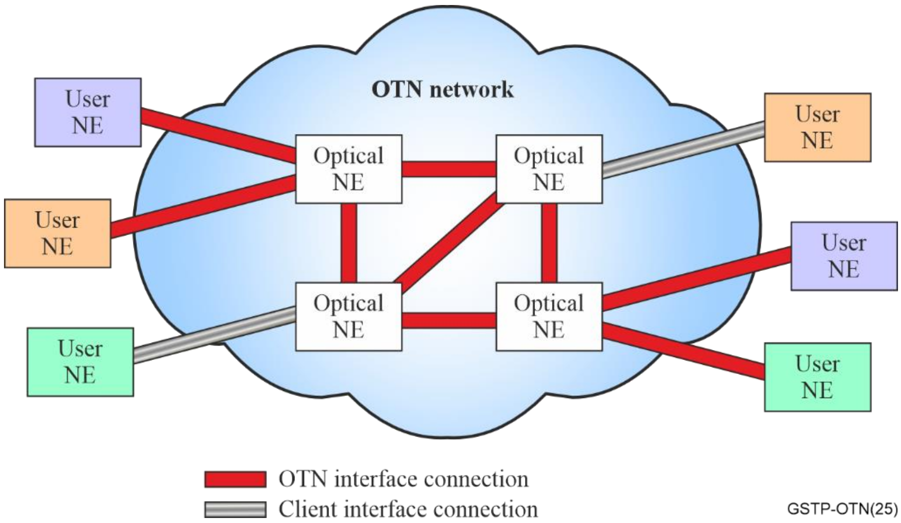


Figure 2-1 – Client signals multiplexed and transport using optical network elements

Ethernet has become one of the most important clients of optical transport systems. Consequently, the OTN standard and related standards have evolved to enable interoperable and efficient transport of Ethernet over OTN transport networks for Ethernet rates of 1 Gbit/s through multiples of 100 Gbit/s, as explained below in this clause.

Clauses 2.2 and 2.3 deal with the frame structures, mapping and multiplexing aspects specific to OTN.

The ITU-T Recommendations dedicated to these subjects are shown in clause 2.6.

2.2 Specific aspects of OTN – OTN signal structure and rates

This clause describes the framing structures, overhead, and mapping/multiplexing specific to OTN. More details on this subject can be found in [b-ITU-T G.709].

The basic OTN signal structure is illustrated in Figure 2-2. It can be summarized as follows:

– An optical payload unit level k (OPUk) contains OPUk overhead and one or more clients that may be lower-rate ODUk signals.

– An optical data unit level k (ODUk) signal contains an OPUk and ODUk overhead.

– In addition to the ODUk Path overhead, the ODUk overhead area provides for up to six levels of tandem connection monitoring (TCM) overhead. The TCM overhead functions are essentially the same as ODUk Path overhead and allow performance monitoring of different segments of the ODUk path.

– For transport signal rates of ≤ 100 Gbit/s, an optical transport unit level k (OTUk) signal contains an ODUk, OTUk overhead and OTUk forward error correction (FEC).

– For transport signal rates of ≥ 100 Gbit/s, an n × 100 Gbit/s modular ODUCn/OTUCn contains one or more ODUk and ODUCn/OTUCn overhead and is carried over a flexible OTN (FlexO) structure that contains FlexO overhead and FEC[2].

– An optical tributary signal assembly (OTSiA) signal, which contains a digital signal (OTUk or FlexO) and OTSiA overhead, the latter of which is carried in a 'non-associated' manner over the optical supervisory channel. "Non associated" means that OTSiA Overheads are not contained in the OTSiA frames.

[2] The "C" in the OTUC/ODUC/OPUC names comes from C, which is the Roman numeral for 100.

The "k" subscript indicates the rate of the OTUk, ODUk and OPUk signals within the OTN hierarchy. (See Table 2-1.) OTN initially had three OTUk signals defined in [b-ITU-T G.709], denoted as OTU1, OTU2 and OTU3, with OTU4 and OTU0 added subsequently:

– OTU1 rate has been optimized to transport SDH STM-16 client signal;
– OTU2 rate is approximately four times the rate of an OTU1 (4.017x) and has been optimized to transport SDH STM-64 client signal and a multiplex of lower rate ODUk signals;
– OTU3 rate is approximately four times the rate of an OTU2 (4.017x) and has been optimized to transport SDH STM-256 client signal and a multiplex of lower rate ODUk signals;
– OTU4 rate is approximately two and a half times the rate of an OTU3 (2.599x) and has been optimized to transport one 100GBASE-R Ethernet client, to transport up to ten 10GBASE-R clients that have each been mapped into an ODU2e and to transport a multiplex of lower rate ODUk signals;
– OTU0 rate has been optimized to transport the 1GBASE Ethernet client signal and to allow multiplexing two ODU0 clients into the ODU1.

The ODU0, ODU2e and ODUflex signals were added to OTN around the same time as ODU4. The ODU0 is optimized for carrying a 1 Gbit/s 1000BASE-R Ethernet (GbE) client signal and to transport two ODU0 signals in one ODU1. The corresponding OTU0 can use GbE optical modules. The ODUflex provides a flexible way to carry constant bit rate (CBR) and packet-oriented client signals of any rate in the OPU of a higher rate ODU frame. The ODU2e, which is optimized to transport a 10 Gbit/s 10GBASE-R Ethernet (10GbE) client signal was essentially the first ODUflex signal to be defined.

The OTU25 and OTU50 were added later in order to provide an OTN signal into which a 25GBASE or 50GBASE Ethernet client signal can be efficiently mapped for short reach (e.g., UNI) applications where it may be possible to reuse the corresponding Ethernet optical modules. Note that neither provides a space in the OTU frame for FEC.

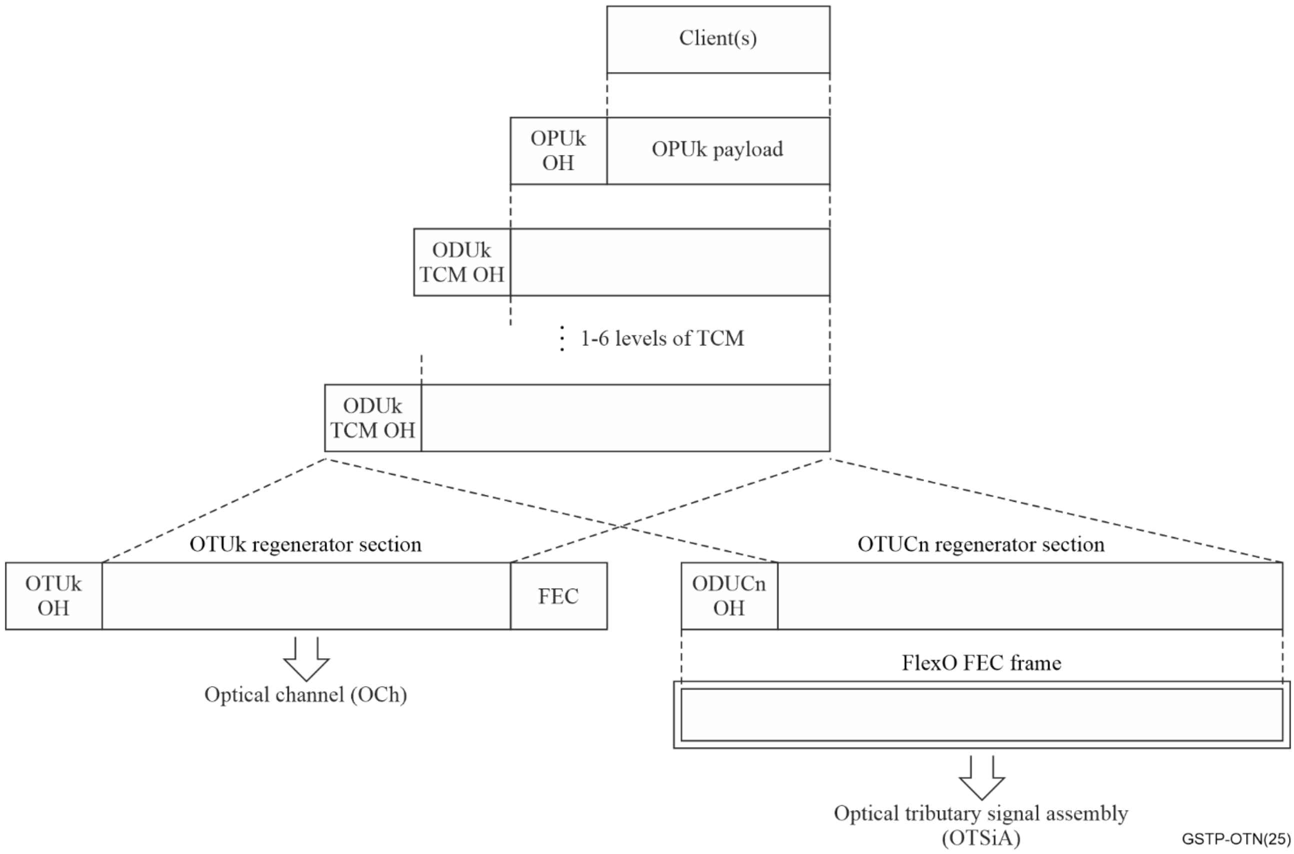


Figure 2-2 – OTN information containment relationships

The OTUk, ODUk and OPUk signals have a synchronous relationship, while the OPUk and client signals may either have a synchronous relationship, or an asynchronous relationship (see clause 2.3.4).

As illustrated in Figure 2-3, all OTUk signals have the same frame structure, i.e., same number of rows and columns, and different bit rates. The rate differences of 4.017, 4.017 and 2.599 times instead of exactly 4, 4 and 2.5 times between the successive OTUk signals is due to the capability to multiplex lower rate ODUk signals into the higher rate ODUk signal. The difference accounts for the bandwidth used by the overhead of the lower rate ODUk into which the client signal is mapped. The resulting frame periods are different for each OTUk signal, the OTU2 frame period is approximately one quarter of of the OTU1 frame period, the OTU3 frame period is approximately one quarter of of the OTU2 frame period and the OTU4 frame period is approximately two fifths of the OTU3 frame period. Table 2-1 describes the OTN hierarchical bit rates.

Table 2-1 – OTN hierarchical bit rates OTUk/ODUk in [b-ITU-T G.709]

level [b-k]	Optical Transport Unit [b-OTUk]	Hierarchical bit rate (kbit/s)	Optical Data Unit [b-ODUk]	Hierarchical bit rate (kbit/s)	Period (µs)	Tolerance (ppm)
0	OTU0 (Note 1)	255/239/2 × 'STM-16' (≈ 1 327 451)	ODU0	1/2 × 'STM-16' (≈1 244 160)	98.354	±20
1	OTU1	255/238 × 'STM-16' (≈ 2 666 057)	ODU1	239/238 × 'STM-16' (≈ 2 498 775)	48.971	±20
2	OTU2	255/237 × 'STM-64' (≈ 10 709 225)	ODU2	239/237 × 'STM-64' (≈ 10 037 273)	12.191	±20

Table 2-1 – OTN hierarchical bit rates OTUk/ODUk in [b-ITU-T G.709]

level [b-k]	Optical Transport Unit [b-OTUk]	Hierarchical bit rate (kbit/s)	Optical Data Unit [b-ODUk]	Hierarchical bit rate (kbit/s)	Period (µs)	Tolerance (ppm)
3	OTU3	255/236 × 'STM-256' (≈ 43 018 413)	ODU3	239/236 × 'STM-256' (≈ 40 319 218)	3.035	±20
4	OTU4	255/227 × 'STM-640' (≈ 111 809 973) (Note 3)	ODU4	239/227 × 'STM-640' (≈ 104 794 445)	1.168	±20
Cn	OTUCn	n × 239/226 × 'STM-640' (≈ n ×105 258 138) (Note 2)	ODUCn	n × 239/226 × 'STM-640' (≈ n ×105 258 138)	1.163	±20
2e	N/A	N/A	ODU2e	239/237 × '10GBASE-R' (≈ 10 399 525)	11.767	±100
Flex	N/A	N/A	ODUflex (CBR)	239/238 × CBR client bit rate	(Note 3)	max of ±100
Flex	N/A	N/A	ODUflex (GFP)	A value in the range 1 244 160 to largest OPUk payload bit rate (Note 5)	(Note 4)	±20
Flex	N/A	N/A	ODUflex (IMP)	239/238 OPUflex(IMP) payload signal rate (Note 6)		±100
25	OTU25	61677/58112 × 24 883 200	ODU25	61677/58112 × 24 883 200	4.633	±20
50	OTU50	61677/58112 × 49 766 400	ODU50	61677/58112 × 49 766 400	2.317	±20

NOTE 1 – For OTUk (k = 0, 1, 2, 3 and 4), the OTUk rate is 255/239 times the ODUk rate. This fraction presents the total number of bits in the frame to the number of non-FEC bits.

NOTE 2 – The OTUCn does not include an FEC overhead area.

NOTE 3 – Period is 121856 / client_signal_bit_rate.

NOTE 4 – Period is 122368 / ODUflex_bit_rate.

NOTE 5 – 'STM-640' represents a bit rate of forty times the 'STM-16' bit rate. Note however that an STM-640 signal is not defined in SDH.

NOTE 6 – For IMP mapped FlexE client signals (ODUflex(IMP, s), the rate is s × 239/238 × 5 156 250 kbit/s for s = 2, 8, n × 5

2.3 OTN frame structure and overhead

2.3.1 Basic OTN frame structure

The basic OTN frame structure is the OTUk frame (Figure 2-3). The OTUC uses the same frame structure as the OTUk except that it does not include an FEC overhead area. As explained below in clause 2.4, the OTUC uses the Flexible OTN (FlexO) structure for its FEC. The OTUk frame is visually represented in four rows of 4080 bytes per row. For OTUk, one byte is referred to as a column, and the OTUk frame is said to have 4080 columns. The first 16 columns of each row of the OTUk frame contain OTN overhead information. The last 256 columns of each row of the OTUk frame may contain optional OTUk forward error correction (FEC) information. The overhead is grouped into four entities, called Alignment (7 bytes for frame and multiframe alignment), OTUk overhead (7 bytes), ODUk overhead (42 bytes) and OPUk overhead (8 bytes) overhead. The ODU (ODUk/ODUC) overhead is functionally part of the ODU frame, which consists of 4 rows by 3824 columns including the 42 ODU overhead bytes and the OPU frame. The OPU overhead is functionally part of the OPU frame. The OPU frame consists of 4 rows by 3810 columns (15 to 3824) including the 8 OPU overhead bytes and 15232 OPUk Payload bytes.

The alignment overhead is attached to the OTUk/OPUC frame and is also attached to an ODUj frame to create the extended-ODUj frame (Figure 2-5). This latter frame furthermore contains 7 bytes of reserved overhead. The OTUk, ODUk and OPUk frame structures are independent of the value of 'k' and applies for all values of k (i.e., 0, 1, 2, 2e, 3, 4, flex). The OTUCn frame is based on n OTUC instances, each of which has the same OTU frame structure except that some of the overhead fields are only active in the first OTUC1 instance.

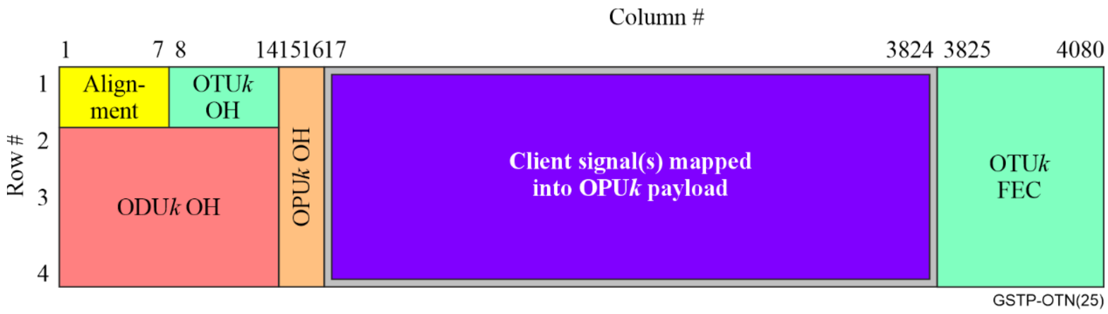


Figure 2-3 – Basic OTN frame structure

The OPUk originates at the point where a client signal is encapsulated into an OTN signal and client mapping agnostic and optional client mapping specific OPUk overhead (Figure 2-4) is attached to it. It is terminated at the point where the client signal is extracted, so the OPUk OH spans from end to end of the OTN network. The OPUk may also originate at the point where one or more ODUj (j < k) signals are multiplexed into an OTN signal and multiplexing agnostic specific payload overhead (OPUk OH, see Figure 2-5) is attached to it. It is terminated at the point where the multiplexed ODUj signals are extracted, so the OPUk OH spans from multiplexing end to demultiplexing end of the OTN network. The OPUk OH is carried as the first two columns of the OPUk, which are located in columns 15 and 16 of the ODUk and OTUk frames.

The ODUk originates and terminates at the points where an OPUk originates and terminates, so the ODUk OH spans either from end to end of the OTN network, or from multiplexing end to demultiplexing end of the OTN network. The ODUk OH is attached to the OPUk (see Figures 2-4 and 2-5). A subset of the ODUk OH is intended to provide monitoring and operational support between intermediate points along the ODUk connection (e.g., for tandem connection monitoring (TCM)), so this ODUk OH spans between mid-points of the OTN network. The ODUk OH is carried as the first 14 columns of the ODUk, which are located in columns 1 to 14 of rows 1 to 3 of the OTUk frames.

The OTUk is originated and terminated at every point where the signal goes through an electrical-optical or optical-electrical conversion or at each point where the signal is electrically regenerated (optical-electrical-optical conversion). The OTUk OH, OTUk FEC and Alignment overhead therefore spans from electrical-optical to optical-electrical conversion points. At electrical regeneration points OTUk OH, OTUk FEC and Alignment overhead is terminated and re-originated. OTUk OH, OTUk FEC and Alignment overhead is attached to the ODUk (see Figures 2-4 and 2-5). The OTUk OH is carried in row 1, columns 8 to 14. The Alignment overhead is carried in row 1, columns 1 to 7. The OTUk FEC is carried in columns 3825 to 4080. The OTUk signal is converted into an optical baseband signal, called the OTSiA.

A frame-synchronous scrambler is used over the OTUk frame such that, with the exception of OA1 and OA2 alignment fields, every bit of the OTN frame, including the FEC fields, is scrambled. Consequently, the frame synchronous is reset to 0xFFFF on the MSB of the MFAS byte and runs continuously at the OTUk line rate over the remainder of the frame. The scrambler prevents long sequences of "1" or "0" patterns, which helps ensure a sufficient number and spacing of bit transitions in the OTUk signal to aid signal-timing recovery at the receiver. Further, it helps to ensure that the receiver framer has correctly determined the start of the OTN frame when it sees the 6-byte OA1 + OA2 pattern.

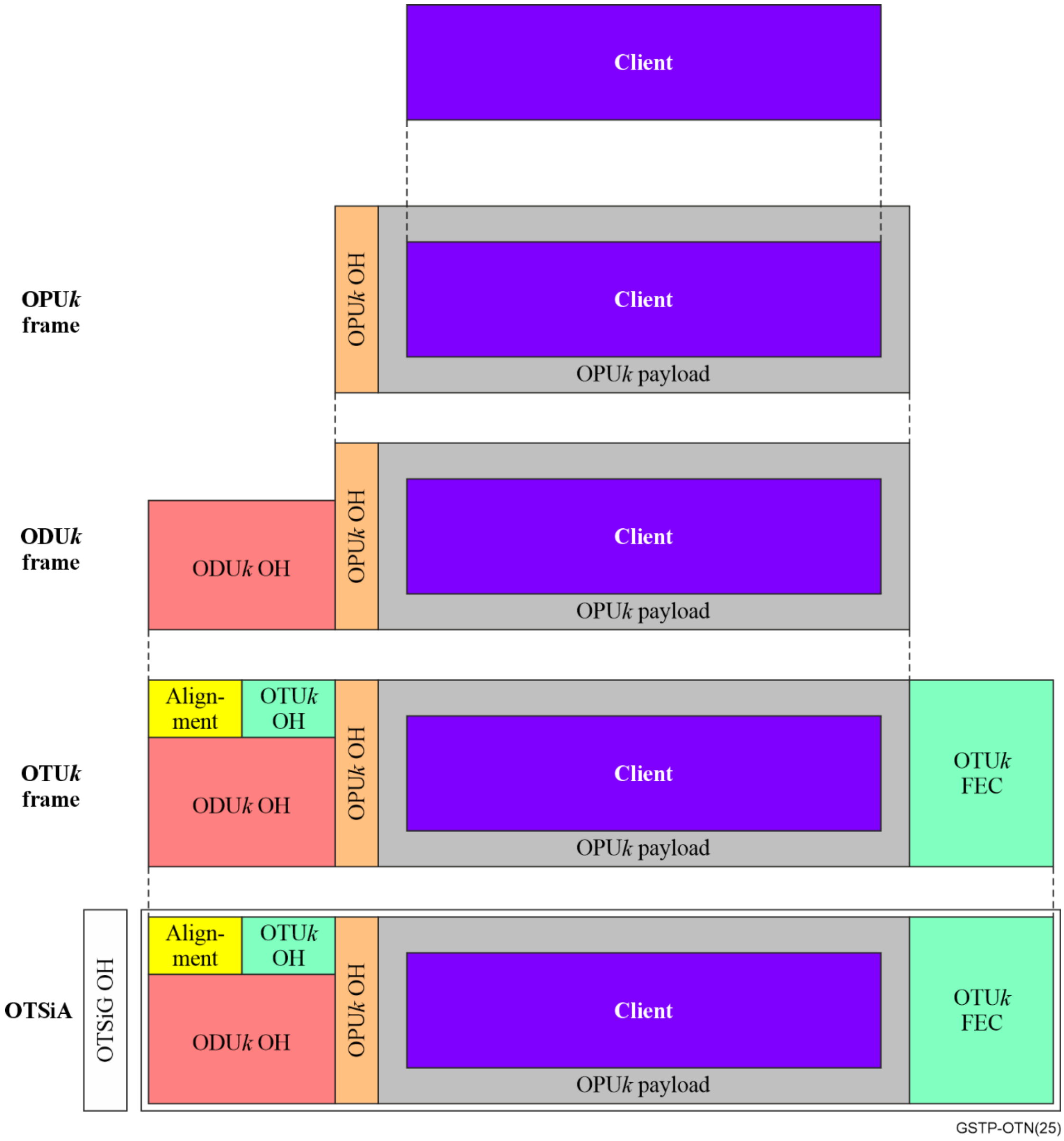


Figure 2-4 – OTUk, ODUk and OPUk frame structures when carrying non-OTN client signal(s)

The OTUk and ODUk frame structures may be replaced by some specific maintenance signals to indicate to equipment downstream of the insertion point that there is a specific maintenance condition, e.g., absence of ODU signal, upstream fault condition, upstream administrative locking condition. The portions of the OTN frame that are replaced by the maintenance signal pattern and the pattern for each maintenance signal type are illustrated in Figure 2-6.

During an incoming OTUk signal failure, the OTUk signal is replaced by an OTUk Alarm Indication Signal (OTUk-AIS). The ODUk frame structure is replaced when a signal failure condition occurs in the connection transporting the ODUk signal. During such signal failure condition, the incoming ODUk signal is replaced by an ODUk alarm indication signal (ODUk-AIS).

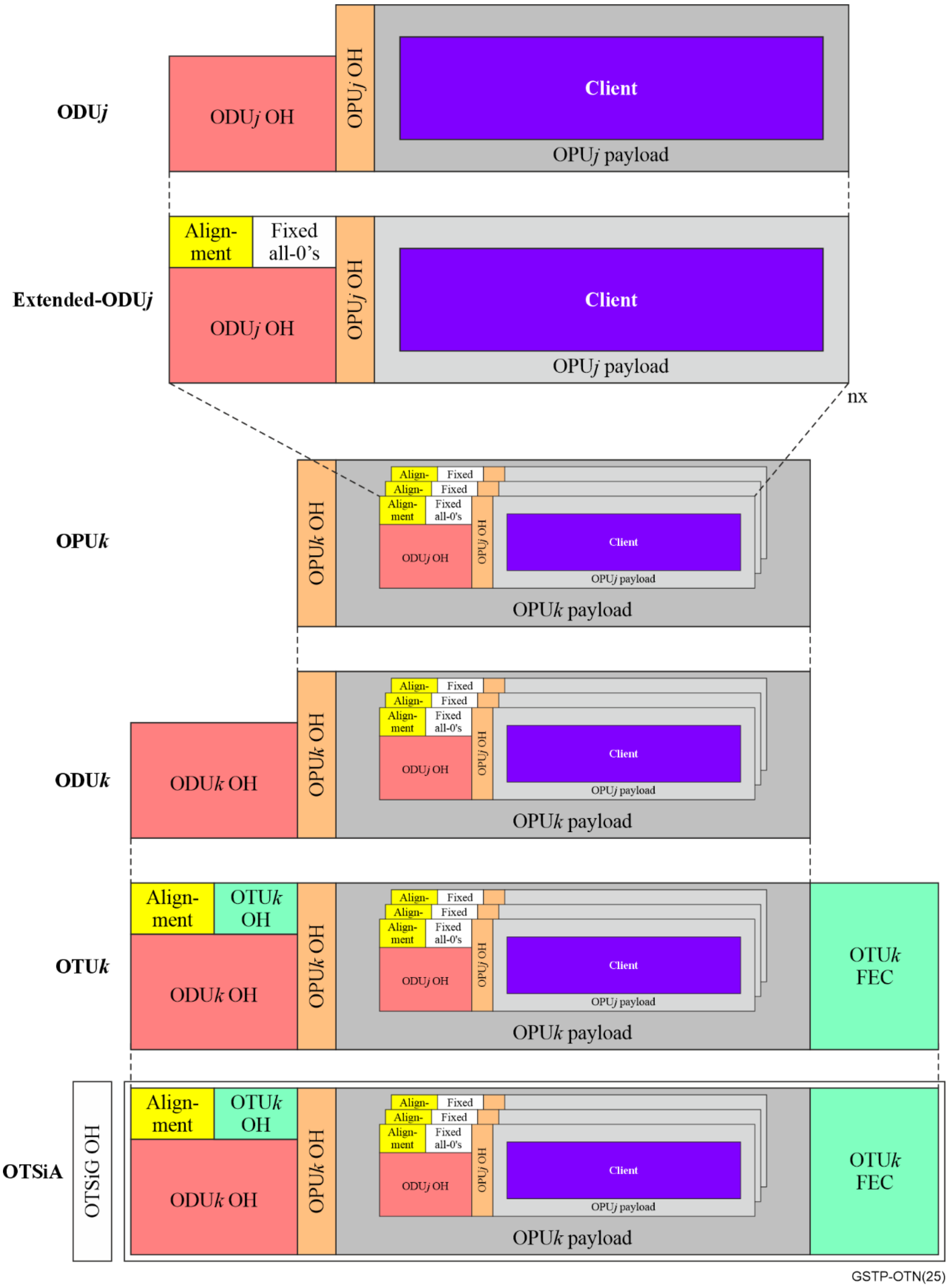


Figure 2-5 – OTUk, ODUk and OPUk frame structures when carrying a multiplex of ODUk signals

The ODUk signal is also replaced when a line or tributary output port is not connected with an ODUk source. During such an open-connection condition, the ODUk signal that is sent out is an ODUk open connection indication signal (ODUk-OCI). The ODUk frame structure is replaced when a line or tributary input or output port is administratively locked. During such an administrative-lock condition, the incoming ODUk signal is replaced by an ODUk locked indication signal (ODUk-LCK).

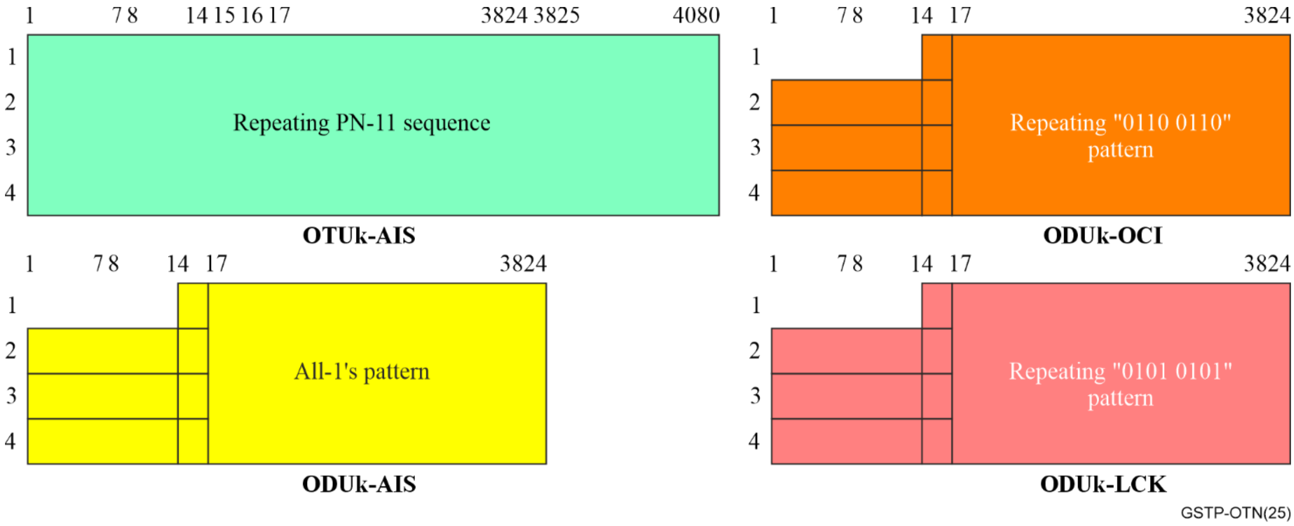


Figure 2-6 – OTUk and ODUk maintenance signal formats

2.3.2 OTUk, ODUk, OPUk overhead

An overview of the names and location of each of the OTUk, ODUk and OPUk OH bytes is shown in Figure 2-7, with additional details shown in Figures 2-8, 2-9 and 2-10. The names and defined functions of each of these OTUk, ODUk and OPUk OH bytes are shown in Table 2-2. More information is available on each one in clause 4 and [b-ITU-T G.709].

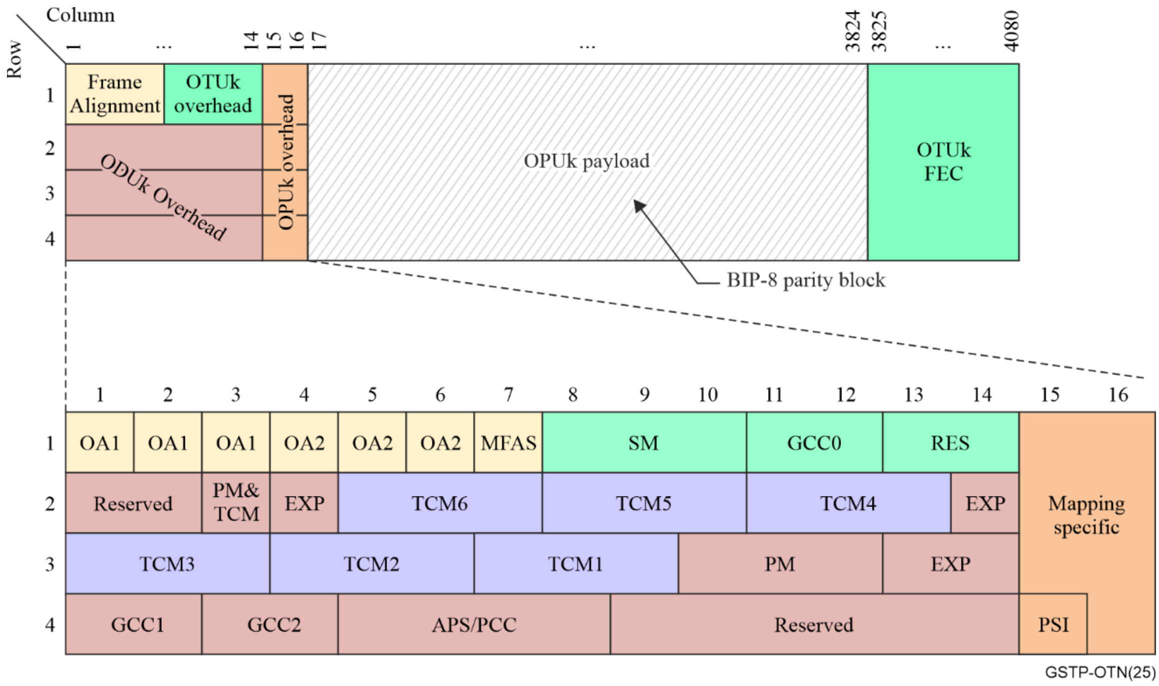


Figure 2-7 – OTU/ODU/OPU overhead overview

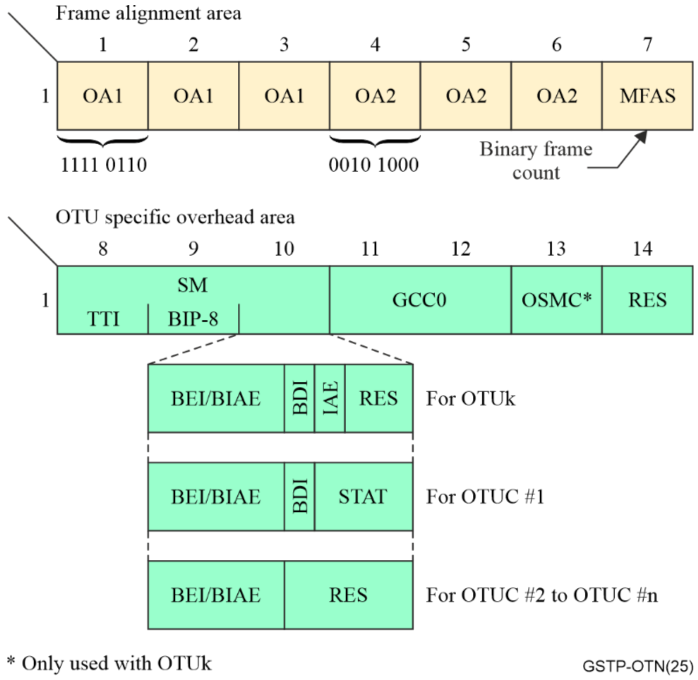


Figure 2-8 – OTU and frame alignment overhead

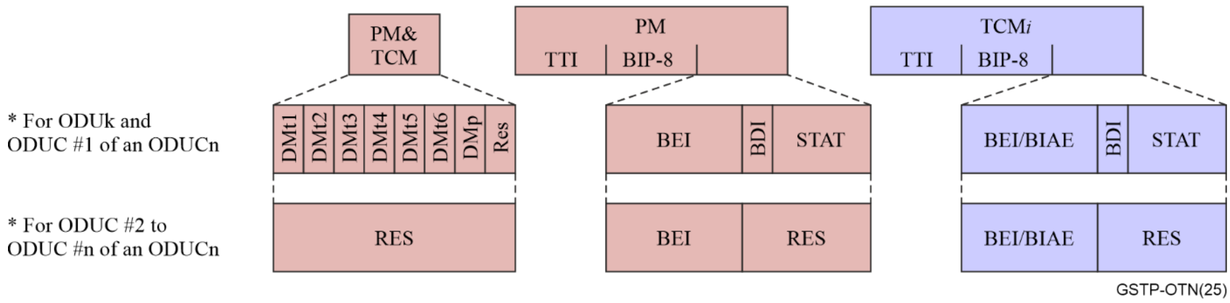


Figure 2-9 – ODU overhead details on Path and TC monitoring overhead

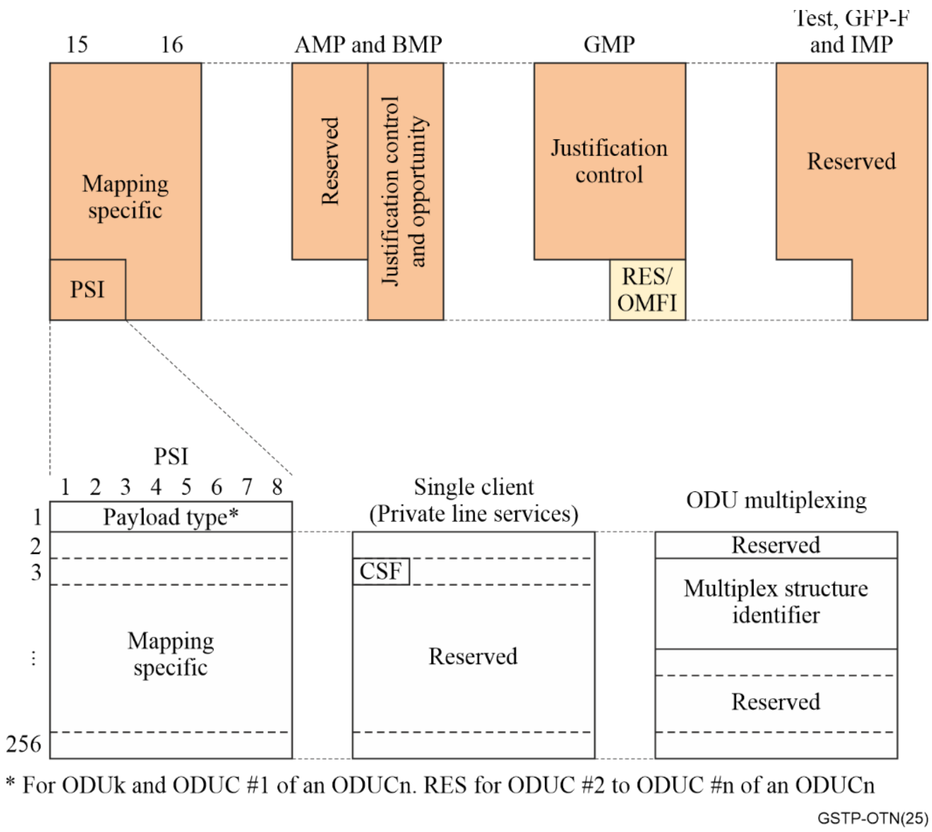


Figure 2-10 – OPU overhead details

Table 2-2 – Types of OTN overhead

OH Type	Location(s)	Comments
Alignment		
FAS	OTUk, OTUC	Frame alignment signal: Used for finding OTUk/ODUk and OTUC frame alignment
MFAS	OTUk, OTUC, FlexO	Multi-frame alignment signal: Indicates the start and location within a multi-frame, which allows either time-sharing of overhead fields for different information in each frame (e.g., justification overhead) or to provide a longer OH field length (e.g., for TTI)
OMFI	OPU4, fgOTN	OPU multi-frame indicator: Used for time-sharing the OPU OH fields across a multi-frame
AM	FlexO	Alignment marker: Used for finding FlexO frame alignment
Performance monitoring		
xM	OTUk, ODUk Path, ODUC, TCM	Performance monitoring with "x" = S for Section, P for Path, TC for Tandem Connection
TTI	OTUk SM, ODUk PM, ODUC SM, FlexO EOH, TCM	Trail trace identifier: Source and destination access point identifier to enable layer end points to verify connectivity

Table 2-2 – Types of OTN overhead

OH Type	Location(s)	Comments
BIP-8	OTUk SM, ODUk PM, ODUC SM, ODUC PM, TCM	BIP-8: Bit interleaved parity 8 code using even parity for detecting errors at that layer
BDI	OTUk SM, ODUk PM, ODUC PM, TCM	Backward defect indication: Used to convey in the upstream direction the signal-failure status detected by the connection monitoring sink at that layer.
BEI	OTUk SM, ODUk PM, ODUC PM, TCM	Backward error indication: Used to convey in the upstream direction the count of interleaved-bit blocks that have been detected in error by the corresponding connection monitoring sink at that layer.
IAE	OTUk SM	Incoming alignment error: Used to forward to the downstream OTUk section monitoring sink that the OTUk monitoring source has detected an alignment error in the incoming signal.
BIAE	OTUk SM, TCM	Backward incoming alignment error: Used to convey in the upstream direction an incoming alignment error (IAE) condition that is detected in the corresponding connection monitoring sink at that layer.
STAT	OTUk SM, ODUk PM, ODUC SM, TCM, FlexO, FlexO Regen	Status: Used to indicate the connection or PHY status, including remote fault conditions and the presence of an ODUk maintenance signal (e.g., AIS, OCI, LCK). Includes FlexO RSTAT for regeneration applications.
DMp	ODUk PM, TCM	Delay measurement – Used to convey the start of the ODUk and TCM path/network connection-level delay measurement test.
GCCx	OTUk, ODUk	General communication channel: GCC0 is located in the OTUk OH and used as a message communication channel between OTUk termination points. Both GCC1 and GCC2 are located in the ODUk OH and are used as two message communication channels between ODUk termination and/or intermediate points.
FCCx	FlexO	FlexO communications channel: FCC0 is similar to GCC0 and FCC1 is similar to GCC1.
OSMC	OTUk, FlexO BOH	OTN synchronization message channel
APS	OTUk, ODUk, TCM	Automatic protection switching channel: Used to communicate the status and synchronize the configuration of the bridge and selector processes at the endpoints of the protection span. Only linear APS is specified.
PCC	ODUk, ODUCn	Protection communication channel: Used to convey additional protection related configuration information between the endpoints of the protection span via a message-based channel. Up to eight levels of nested protection coordination signals are supported. The bytes in a given frame are assigned to a dedicated connection monitoring level (path, tandem connection i, section) depending on the value of MFAS.
RES	OTUk, ODUk, OTUC, FlexO	Bits reserved for future international standardization. Network operators and vendors must not use those bytes for proprietary applications.

Table 2-2 – Types of OTN overhead

OH Type	Location(s)	Comments
EXP	ODUk	Experimental: Used to allow a vendor and/or a network operator within their own network to support a proprietary application, which requires additional ODUk overhead.
Client and mapping OH		
PSI	OPUk	Payload structure indicator: Indicates the structure of the payload area, including PT, OPUk MSI, and CSF or CDI
JC	OPUk, OPUC, FlexO BOH	Justification control overhead
PT	OPUk, OPUC, FlexO BOH	Payload type
MSI	OPUk, FlexO BOH	Multiplex structure indicator: Provides information related to TS occupancy
CSF	OPUk	Client signal fail
TPID	OPUk, FlexO BOH	Tributary port ID:

2.3.3 OTN client signal mapping and multiplexing

As explained above, each OTN client is mapped into the OPU of its own ODU signal. In order to increase the usage efficiency of each optical media channel (wavelength), multiple lower-rate ODU signals can be multiplexed together into the OPU of a higher-rate ODUk. A generic description of OTN multiplexing is illustrated in Figure 2-11. This clause describes both the OTN mapping and multiplexing methods, and the associated rate adaptation techniques.

When an OPUk is dedicated to a single client (i.e., client mapping), the client occupies the entire OPUk payload area except for "stuff" bytes. The stuff bytes provide padding for rate adaptation and may use fixed (dedicated) or dynamic locations, as described below. In the case of the OPU2 and OPU3, the OPUk rate was chosen for a corresponding client signal (e.g., an SDH client). The fixed stuff columns provide the additional bandwidth for multiplexing applications that is required in order to accommodate the overhead of the ODU signal into which the client was mapped. See the asynchronous mapping procedure (AMP) section for additional uses of fixed stuff bytes. Multiple client signals are multiplexed into an OPUk by the two-step process of first mapping each client into its own ODU and then multiplexing the lower rate ODU signals into the higher rate OPUk. Multiplexing requires dividing the OPU payload area into separate regions called tributary slots (TS). A given client can be carried in one or more TS, depending on the bandwidth (channel capacity) required for carrying that client. For OPU1, OPU2 and OPU3, the OPUk byte columns are assigned to the different TS on a round robin basis.

The original nominal TS rate was 2.5 Gbit/s for OPU2 and OPU3. Subsequently, the addition of the 1.25 Gbit/s ODU0 and the desire for greater multiplexing flexibility into the OPU4 led to adopting a nominal TS rate of 1.25 Gbit/s. OPU2 and OPU3 use both TS rates, as specified by their payload type overhead, while OPU1 and OPU4 only use the 1.25 Gbit/s TS. When the ODUC was defined for Section layer rates of $\geq$ 100 Gbit/s, it was more practical to define them to use a nominal 5 Gbit/s rate.

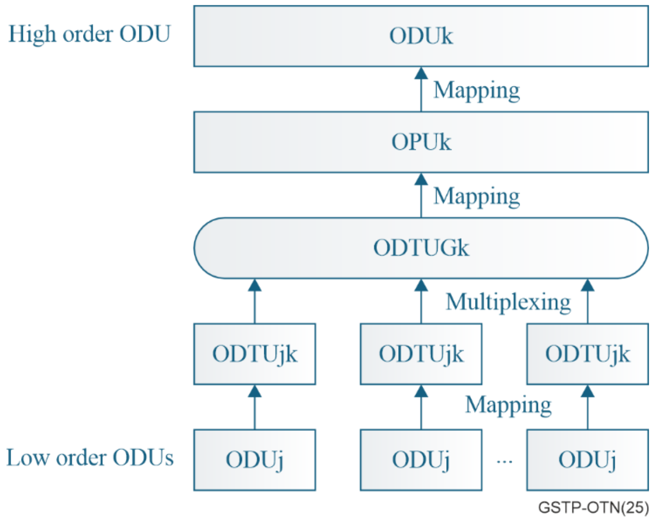


Figure 2-11 – OTN multiplexing

The introduction of the 1.25G TS made the new ODUflex attractive. Constant bit rate (CBR) clients can be mapped into an ODUflex(CBR) and packet-oriented clients can be mapped into an ODUflex(GFP) or ODUflex(IMP). The resulting ODUflex occupies an appropriate number of TS in the higher-rate OPUk into which it is multiplexed. Mapping clients into an OPUflex is described below.

Whenever a client signal is mapped into the payload container of a server layer signal, some mechanism is required to adapt between the respective information rates of the client signal and the server container. The process of adapting the client signal to the payload container rate is referred to as synchronization, and the process of recovering the rate of the client signal when the receiver extracts it from the payload container is referred to as desynchronization. OTN supports mapping and transporting both constant bit rate (CBR) clients and packet-oriented clients, as described in the following sections.

2.3.3.1 Mapping and multiplexing for CBR clients

In the simplest mapping method, referred to as bit-synchronous mapping (BMP), the OPU clock rate is derived directly from the client signal clock. Specifically, the ODU rate resulting from BMP is the client rate multiplied by the ratio of the number of ODU columns (3824) divided by the number of available payload columns (i.e., 3824 minus overhead and fixed stuff columns). In most cases the resulting LO ODU rate is simply equal to (239/238)×(the client signal rate). Because the OPU is frequency and phase locked to the client signal, there is no need for dynamic frequency justification and the justification control (JC) bytes in the OPU OH contain fixed values. See Table 2-2. Note that for most clients, byte alignment is used rather than bit alignment. BMP is used for mapping CBR clients into the payload area of the OPUflex and also for SDH client signal mapping into ODU1, ODU2 and ODU3.

When the client and server signal clocks are asynchronous from each other (i.e., independent), the two signals require an asynchronous mapping process. This situation is typical for most clients mapped into OTN and is most practical for multiplexing since there can be multiple lower-rate ODU signals with their own clock sources. If the container rate is any faster than the client signal rate, a mechanism is required to fill the unused container bandwidth in a manner that can be recognized and removed at the receiver. Filling unused payload container bandwidth is referred to as either stuffing or positive justification. Neither the server nor client clocks will be ideal but will have a frequency tolerance range that must be taken into account. The need for positive justification or stuffing can arise even if the nominal client signal and payload container rates are the same, but the server clock is running in the faster end of its tolerance and the client signal clock is running in the slower end of its tolerance. Further, if the nominal client signal and payload container rates are close, additional payload area (bandwidth) will periodically be required when the server clock is at the slower end of its range and the client clock is at the faster end of its range. Filling normally unused bits with payload data is referred to as negative justification.

Periodic justification opportunities are provided in the server signal in order to accommodate the client signal and server container rate differences. The mapper uses justification control (JC) overhead in the OPU overhead to inform the receiver when and how the justification opportunities are being used.[3] See Figure 2-10 for the justification control overhead location. The required amount of positive and negative justification bandwidth is determined by the worst-case difference between the client signal rate and the server container rate with each operating at the opposite extremes of its respective frequency tolerance range. The number of justification bytes specified in a server container is determined by the combination of this required justification bandwidth and the period of time between justification opportunities.

Describing asynchronous mapping in terms of high-level conceptual implementation, the incoming CBR client signal is written into a FIFO (First-In-First-Out) buffer at the client signal rate and is read from the buffer at the server payload container rate. The justification mechanism provides the method for occasionally sending additional client data to prevent buffer overflow (negative justification) or less client data to prevent buffer underflow (positive justification). The demapper also uses a FIFO buffer when extracting the client signal from the payload container. The average fill rate of this buffer is based on the client signal rate. The demapper determines this client signal rate using a phase locked loop (PLL) that is controlled by the average fill level of this buffer. Since the positive and negative justifications change the amount of data written into the demapper buffer at that point in time, the resulting buffer fill level variations must be filtered to determine the client signal's long-term average buffer fill level. This filtering, which is required to minimize jitter and wander in the egress client signal rate, is also part of the desynchronizer process.

There are two types of OTN asynchronous mapping procedures for CBR clients. The method that was part of the original version of ITU-T G.709 became known by the general title of Asynchronous Mapping Procedure (AMP). Subsequently, a more flexible method, referred to as the Generic Mapping Procedure (GMP), was added. Both are described below.

In the case of client mapping, the JC bytes are updated once per OTN frame. The AMP and GMP CBR client mapping methods can also be used for multiplexing. However, since each client signal needs its own justification overhead, as explained below, multiplexing requires that the JC bytes are time shared on a per-frame basis across a multiframe. Specifically, since the number of TS in an OPUk is a power of 2, the appropriate subset of the MFAS LSBs are used to indicate the TS to which the current JC bytes apply[4]. As illustrated below in Figure 2-12, the JC overhead for TS #*m* is present when the MFAS LSB values are *m*-1.

2.3.3.1.1 AMP (asynchronous mapping procedure)

The ODU1, ODU2, and ODU3 signal rates were chosen to be optimized for carrying SDH client signals. Generically, signals with the SDH STM-16, STM-64 and STM-256 rates are referred to in [b-ITU-T G.709] as CBR2G5, CBR10G, and CBR40G clients, respectively. The OPUk rates were defined to be close enough to these client rates that the occasional justification only requires adding a single addition payload-area byte or removing a payload area byte by assigning it to be a stuff

---

[3] Some mapping procedures use dedicated "fixed" stuff bandwidth when the container rate is significantly higher than the client signal rate. The presence of the fixed stuff bits/bytes reduces the resulting nominal container rate to be close enough to the client signal rate that the periodic justification opportunities can accommodate any remaining rate offset.

[4] Since the OPU4 has 80 TS, it uses a dedicated OPU multiframe indicator (OMFI) in its OPU4 overhead to indicate which TS is currently using the JC bytes rather than using the MFAS.

(padding) byte. As illustrated in Figure 2-12 and Table 2-3, AMP was designed for this function. The source transmits each of the JC bytes in a frame with identical JC bit values. The sink uses a majority vote of the three received JC values to determine the correct JC bit values in the event of an error corrupting one of the JC bytes in that frame.

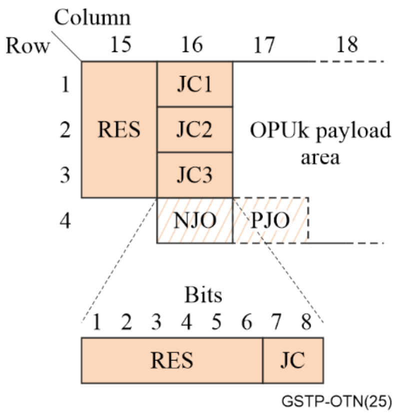


Figure 2-12 – JC byte definitions for AMP mapping

The JC bytes are used in the same manner for client multiplexing with two exceptions. First, the client and container rate differences can require providing an additional positive justification overhead (PJO2) byte. Second, since the client information appears in the TS to which it is assigned, the PJO bytes must be located the bytes of those TS rather than in column 17 as is used for mapping. See Figure 2-13. Table 2-3 shows how the JC bytes are used.

Table 2-3 – Justification control and opportunity definitions for AMP CBR client mappings

JC [b-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)

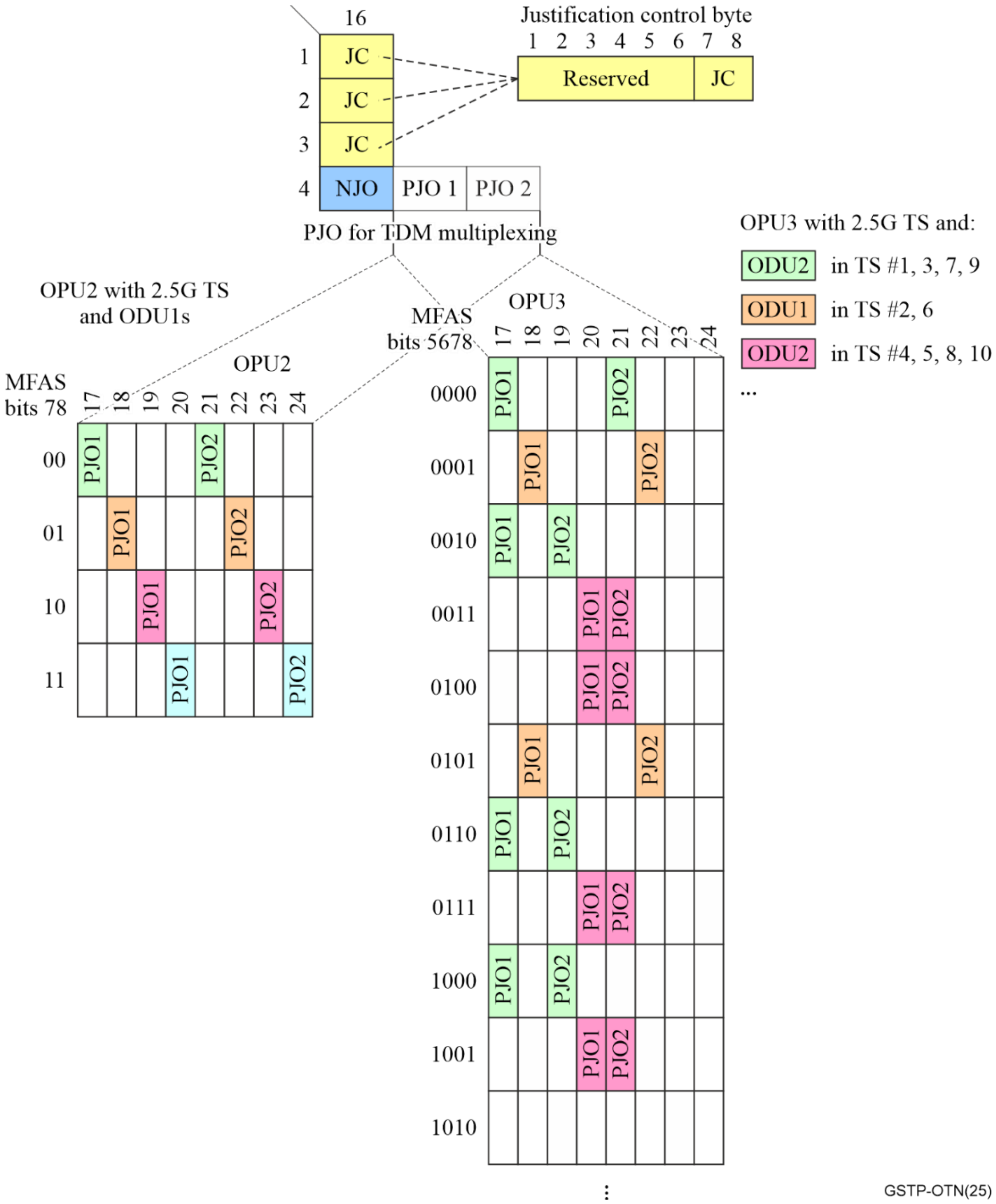


Figure 2-13 – JC byte illustration for AMP multiplexing

2.3.3.1.2 GMP (generic mapping procedure)

GMP was developed in order to provide a consistent asynchronous mapping and multiplexing 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 for each new client. GMP was developed together with ODU0 and ODU4 and is the only mapping rate adaptation method used for OPU0 and OPU4 as well as OPUC and multiplexing into FlexO. GMP is also used for most mappings into OPU1, OPU2, and OPU3 when their payload is structured with 1.25 Gbit/s tributary slots.

Since GMP applies to both mapping and multiplexing, the following description is written from a mapping perspective with the corresponding multiplexing term or concept following in parentheses.

The GMP mapping (multiplexing) performs rate justification by communicating the number of OPU payload words that will contain client data during the next ODU frame (multiframe). The count value is referred to as C_m (count of the integer number of m-bit words to be sent). For mapping, this C_m count value information is carried in the JC bytes of the current frame. For multiplexing, the C_m count information is carried in the JC bytes of the frame within the multiframe that corresponds to the highest number of TS used by that client. The m-bit stuff words are distributed throughout the OPU payload container using a modulo arithmetic algorithm such that the receiver can derive the value directly from the received C_m count value.

The modulo arithmetic algorithm guarantees that stuff words are distributed evenly across the payload area of the frame (multiframe), which results in much less client signal jitter seen at the sink node PLL than with AMP. The robustness of the transmitted C_m overhead is ensured through having the source signal ±1, and ±2 count increment/decrement operations through inverting subsets of the C_m bits that were chosen to maximize Hamming distance, combined with a CRC error check code over the GMP overhead bit fields. See Annex D in [b-ITU-T G.709] for a full description of GMP operation.

The flexibility of GMP comes from it being able to accommodate mapping a CBR client with a rate that needs between one and just under the number of word locations in the OPUk 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.

As data rates increased, there were 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. The word sizes are summarized in Table 2-4.

Table 2-4 – GMP word size and server channel capacities

	Word size M	**Base word granularity**	**P_{server}**	**Comment**
OPUk mapping	maximum number of TS that could be supported by that OPUk	8-bits (1-byte)	15232 for k = 0-3 15200 for k = 4	ODU0, ODU1, ODU2, ODU3 and ODU4 have word sizes of M = 1, 2, 8, 32 and 80 bytes, respectively.
OPUk multiplexing	(number of TS) × (base granularity)	8-bits (1-byte)	15232 for k = 1-3 15200 for k = 4	1) Max number of words in the server payload area 2) M for Number of TS occupied by that client 3) P_{server} for case of a single TS
OPUC multiplexing		128-bits (16-bytes)	952	
FlexO-n	n_i × 256	256 bits	10260	For OTUCn$_i$ clients (∑n_i ≤ n)
FlexO-ne	y × 257	257 bits	10220	y00GBASE Ethernets (∑y ≤ n)

One example of a mapping with GMP is carrying the 1 Gbit/s fibre channel (FC-100) client signal over an OPU0. The nominal bit rate of the FC-100 signal is 1.0625 Gbit/s ±100 ppm. The average GMP count value to carry the FC-100 in an OPU0 is 13062.6 bytes, based on the nominal rates of both signals[5]. Consequently, the transmitter will alternate between sending 13062 and 13063 bytes per frame in order to achieve this average rate. Figure 2-14 illustrates the GMP data and stuff placement for the FC-100 client with a C_8 GMP count value of 13062.

5 The nominal ODU0 frame period is (1/1.24416 Gbit/s)(3824 × 4 × 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.

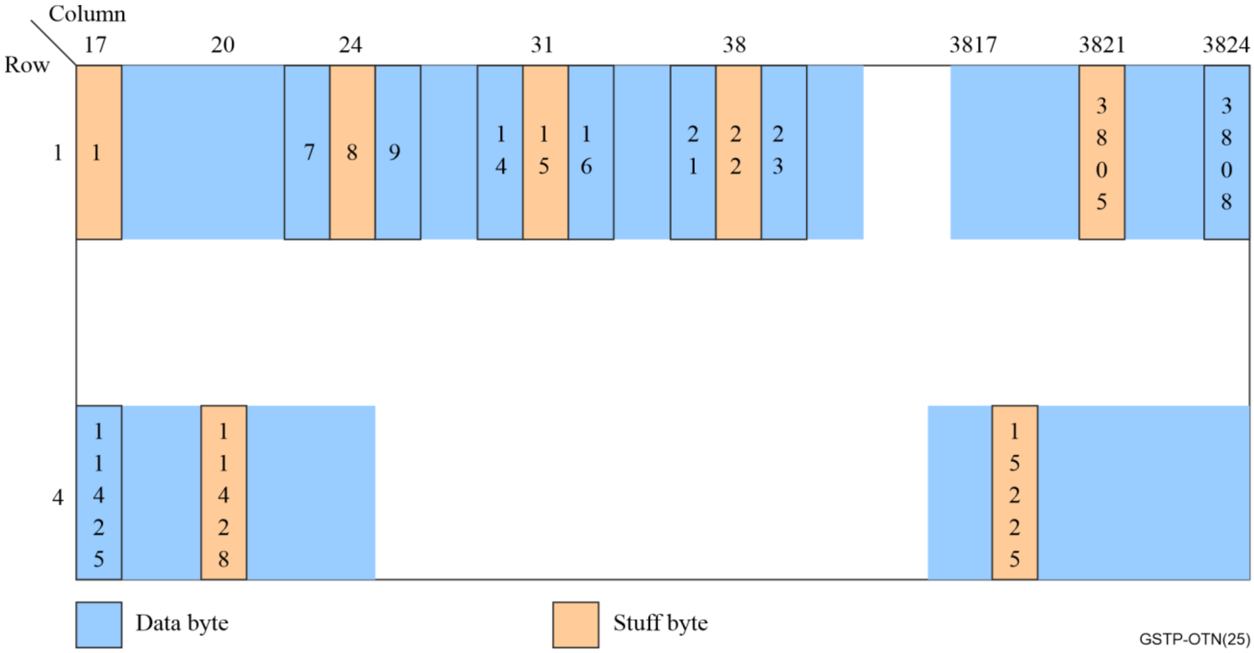


Figure 2-14 – FC-100 mapping illustration with GMP C_8 = 13062

Using GMP results in the average server container rate matching the client signal rate. GMP count value adjustments to the container size introduce instantaneous jitter. In order to simplify the receiver desynchronizer PLL implementation for mitigating this jitter, the GMP overhead also includes overhead for finer resolution related to the phase offset between the client signal and container clocks. The overhead field carrying this finer-grain information is referred to as C_{nD} and is carried in JC4-JC6. The phase encoding is essentially a running sum of the number of bits at the source node that cannot be transmitted in the next frame due to being smaller than the size of a word. Transmission errors with the phase offset approach are more readily filtered and at worst result in a transient phase error.

2.3.3.2 ODTU concept

Rather than being multiplexed directly into the TS of an OPUk, non-OTN client signals are always first mapped into their own ODU and these lower rate ODU signals are multiplexed into the higher rate OPUk. The ODU carrying the client signal is assigned to the minimum number of TS in the higher rate ODUk such that the capacity of the set of TS exceeds the rate of the ODU carrying the client signal. (1 ≤ #TS ≤ OPUk max)[6]

Since there is no requirement for the TS occupied by a client to be contiguous, their locations are typically distributed throughout the OPUk payload area. Consequently, it was convenient to introduce the optical data tributary Unit (ODTU) concept. The ODTU consists of the specific OPUk TS bytes that are occupied by the client ODU along with the JC bytes that pertain to that client. The data portions of the ODTUs (data and stuff words) are byte synchronously placed (byte-interleaved) into their respective TS of the higher rate OPUk with their justification control overhead appearing in the JC bytes of their associated frame within the multiframe. See the example in Figure 2-15. When using AMP and 2.5G TS, the ODTU is called an ODTUjk, where j refers to the OTN hierarchical rate of the client ODU (j = 0, 1, 2) and k refers to the OTN hierarchy rate of the server OPU (k = 1, 2, 3). When using GMP mapping into 1.25G TS, the ODTU is called an ODTUk.TS where TS is the number of server TS occupied by that client.

---

[6] For example, the 25GBASE-R Ethernet rate is 25.78125 Gbit/s. The ODUflex into which is mapped would have a nominal rate of (239/238)×(25.78125G) = 25.88957 Gbit/s. The ODU3 is the lowest rate ODUk that could carry this ODUflex. The OPU3 TS rate is 1.2547 Gbit/s and (25.88957)/(1.2547) = 20.63, so 21 TS must be used in the higher rate OPUk.

Note that while the ODTU is a useful descriptive concept, there is no requirement to construct an intermediate ODTU "signal" as a distinct entity in the multiplexing process.

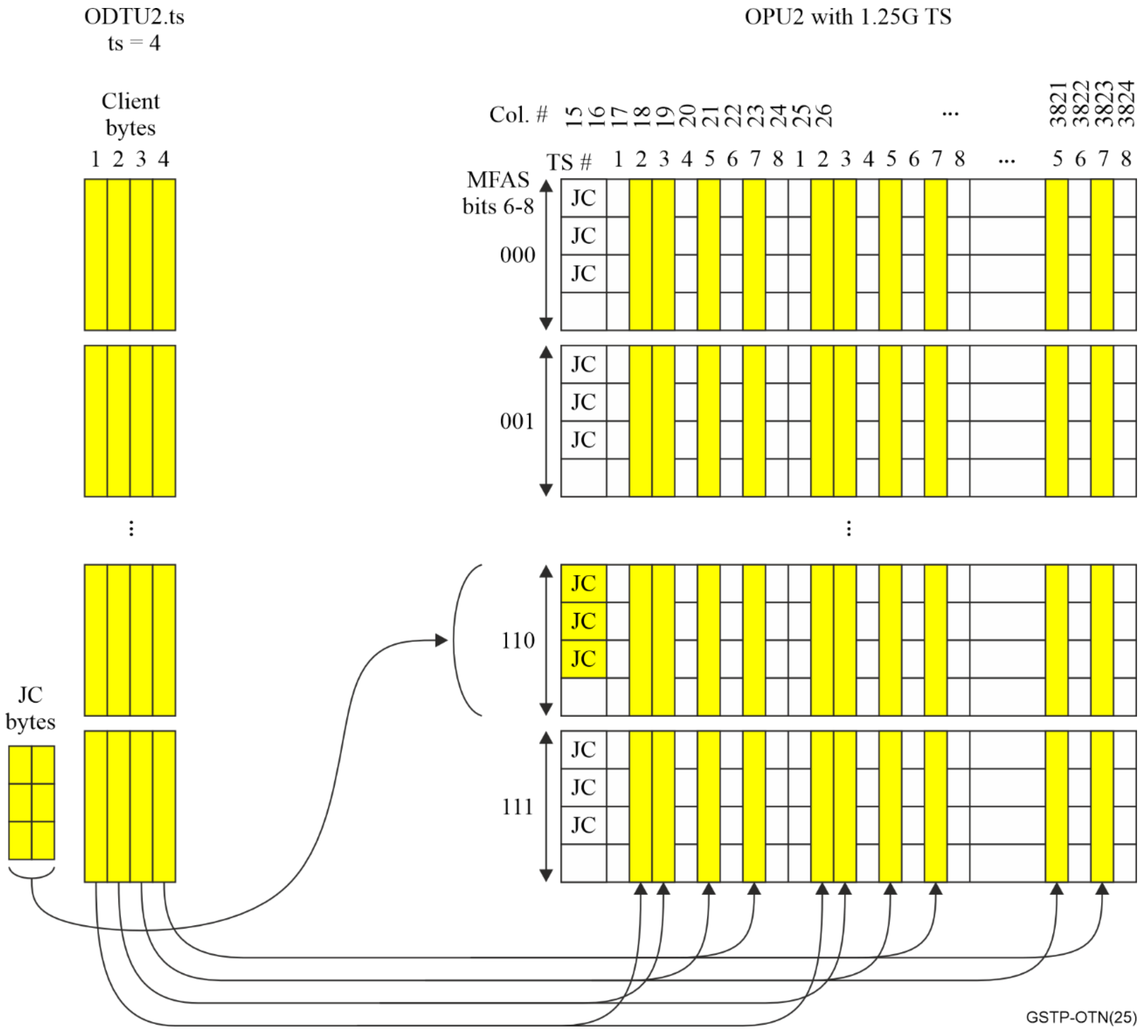


Figure 2-15 – Illustration of the conceptual ODTU for ODTU2.ts where TS # 2, 3, 5 and 7 of the OPU2 are occupied by the client

2.3.4 Mapping and multiplexing for packet clients

At a high level, the rate justification methods associated with carrying packet clients rely on adding padding elements to the encoded client information stream at the source that the sink can recognize and remove rather than relying on the OPUk (k = 0, 1, 2, 3, 4 or flex) JC bytes for the dynamic rate justification. When the packet clients are fist mapped into an ODUflex(GFP) or ODUflex(IMP), the associated ODUflex is multiplexed into the server OPUk in same manner as an ODUflex(CBR). Consequently, ODU multiplexing and switching nodes within the network can remain agnostic to the type of ODUflex they multiplex or switch.

The ODUflex clock rate for ODUflex(GFP) and ODUflex(IMP) can be derived from a local clock at the source node. The rate is chosen such that it is sufficient to carry the highest expected rate of the GFP-F frame stream or 64B/66B encoded signal associated with that client and use the minimum number of OPUk TS. If the clock is derived from the rate of the ODUk into which they are multiplexed, then the relationship between their rates can be accommodated through the choice of either a fixed C_m or using two alternating C_m values for mapping the ODUflex into the server OPUk. In the case of ODUflex(IMP), it is possible to use the clock rate of the 64B/66B block-encoded client stream as the basis for the ODUflex(IMP) rate. In this case, the ODUflex(IMP) rate is simply

(239/238) times the 64B/66B encoded packet client bit rate, which directly preserves the client clock rate.

Regarding client maintenance (replacement) signals during fault conditions, packet clients use Ethernet for their layer 2 encapsulation. Consequently, when a fault condition impacts the client signal, it is replaced by the appropriate fault indication signal for that packet client type.

2.3.4.1 Generic framing procedure (GFP) frame mapping

The OTN mapping for packet-oriented clients with rates < 100 Gbit/s is to use the packet encapsulation method defined in [b-ITU-T G.7041], generic framing procedure (GFP). GFP encapsulation is defined for a variety of client packet streams, including Ethernet, MPLS and video signals. Each client packet arriving over a data interface is encapsulated into the payload bytes of a GFP frame.

GFP frames are mapped in a byte-aligned manner into the OPUk. As shown in Figure 2-16, the space (bandwidth) between GFP client data frames in the OPUk payload area is filled by GFP Idle frames. In other words, GFP Idle frames are used for the rate justification when mapping into an OPUk. GFP frame delineation is achieved using the GFP overhead, as defined in [b-ITU-T G.7041].

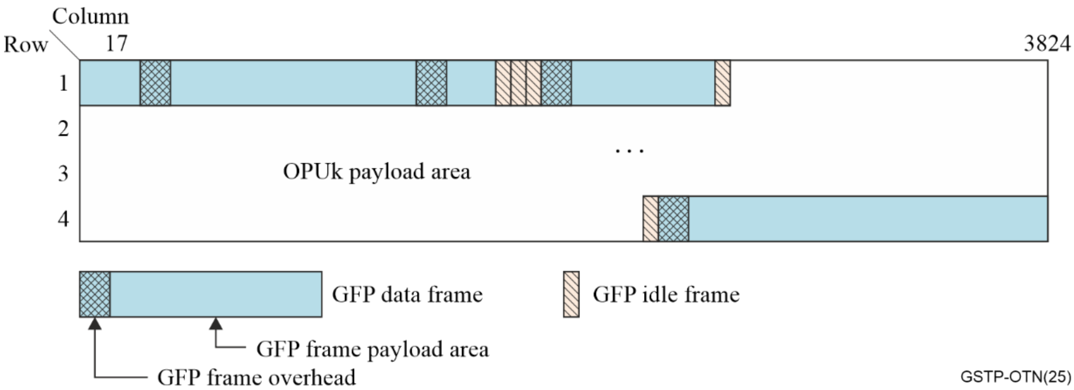


Figure 2-16 – GFP frame mapping into an OPUk

Note that [b-ITU-T G.7041] also defines a transparent GFP mapping (GFP-T) in which the client 8B/10B line code is transcoded for bandwidth efficiency in manner that preserves the 8B/10B code information.

2.3.4.2 ODUflex(GFP) mapping

The ODUflex(GFP) is a CBR signal with a frame rate generated from a local source clock. As with any GFP mapping into an OPUk, GFP Idle frames provide the mechanism for rate adapting the client packet stream into the OPUflex(GFP) payload area.

One of the valuable applications for ODUflex(GFP) is that it allows the use of ODU switching as an economical way to provide pseudo-wire (PW) or Ethernet virtual LAN (VLAN) switching within the OTN domain, thus avoiding the equipment and network complexity of using packet switching.

The use of GFP Idle frames for rate adapting the client data flow into the OPUflex(GFP) and the ability to use GMP to incrementally change the effective rate of the ODUflex(GFP) into the set of server TS allowed the possibility of performing a hitless change of the number of TS used by that client. The resulting HAO (hitless rate adjustment of ODUflex(GFP)) protocol is defined in [b-ITU-T G.7044].

2.3.4.3 Idle mapping procedure (IMP) mapping

As background, when higher rate Ethernet interfaces were defined in [b-IEEE 802.3], it was typical to define them in terms of a stream of 64B/66B blocks. Each 64B/66B block either contained only data from an Ethernet packet or included some type of control information. The blocks with control information include start of packet (/S/), termination of packet (/T/), Ordered sets (/O/) to communicate client status, and idles (/I/) to fill the inter-packet gap (IPG). The 64B/66B block encoding as defined in [b-IEEE 802.3] clause 82.

[b-IEEE 802.3] defines a minimum gap between packets, referred to as the inter-packet gap (IPG), which is filled with idle characters. Idle characters are inserted or removed from the transmitted Ethernet signal in order to adapt the rate of the Ethernet packet information flow to the PHY capacity rate. In the case of an interface using 64B/66G encoding, idle characters are grouped into /I/ idle blocks such that idle insertion/removal is performed on a 64B/66B block basis. (See [b-IEEE 802.3] clause 82 for additional information about this process, including the rules for removing /O/ characters for rate adaptation.)

IMP was specified as the mapping method for carrying Ethernet MAC clients and FlexE clients that have a bit rate ≥ 100 Gbit/s, although it can be used for 64B/66B-encoded clients at any rate. When carrying 64B/66B-encoded Ethernet client signals or FlexE[7] client signals, they are mapped into an ODUflex(IMP). The OPUflex payload area contains the stream of 64B/66B client blocks. Note that in order to reduce the complexity for the OTN sink node to perform 64B/66B block synchronization, the 64B/66B blocks are 2-bit aligned within the OPU payload bytes. In other words, the first bit of the sync header (i.e., block bit 0) must be aligned with an odd numbered bit of the OPU payload byte.

Similar to the GFP frame mapping approach, when using a local clock, the rate adaptation is performed within the client information stream (i.e., within the OPU payload area) rather than using the OPUk justification control overhead. Specifically, the mapper uses the IEEE 802.3 clause 82.2.3.6 rate adaptation process of adding or removing /I/ blocks within the IPG regions of the client block stream to adjust its rate to fill the OPUflex(IMP) payload area capacity. Hence the name "Idle Mapping Procedure" (IMP).

2.3.5 Summary information for OTN client mapping and multiplexing

Table 2-5 provides a brief summary of the type of mapping procedure used for some of the important OTN client signals.

Table 2-5 – OTN mapping method summary

Mapping Method	Application
Bit-synchronous Mapping Procedure (BMP)	• Alternative method for mapping SDH client signals into OTN • Mapping CBR clients into ODUflex(CBR)
Asynchronous Mapping Procedure (AMP)	• Mapping SDH client signals into OTN • Multiplexing LO ODUk (k = 1, 2, 3) signals into 2.5 Gbit/s tributary slots, and ODU0 into ODU1
Generic Mapping Procedure (GMP)	• Mapping non-SDH CBR clients into OPUk (k = 0, 1, 2, 3, 4) • Multiplexing into 1.25 Gbit/s tributary slots (except ODU0 into ODU1) • Multiplexing into OPUCn • Multiplexing into FlexO • Multiplexing with fgOTN • Mapping FlexE into ODUflex
GFP frames into an OPU payload container	• Mapping packet clients into OTN with GFP (Generic Framing Procedure) encapsulation

[7] Flexible Ethernet (FlexE) is defined in the OIF FlexE implementation agreement as a way to carry multiple Ethernet client signals over groups of Ethernet PHYs. These can be 100, 200 or 400 Gbit/s Ethernet PHYs.

Table 2-5 – OTN mapping method summary

Mapping Method	Application
ODUflex(GFP)	• Mapping packet clients into an ODUflex(GFP) with GFP encapsulation
ODUflex(IMP)	• Mapping streams of 64B/66B clients into an OPU
NOTE – 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 clause 2.3).	

2.4 Flexible OTN (FlexO) frame structure and overhead

The FlexO Recommendation structure is illustrated in Figure 2-17.

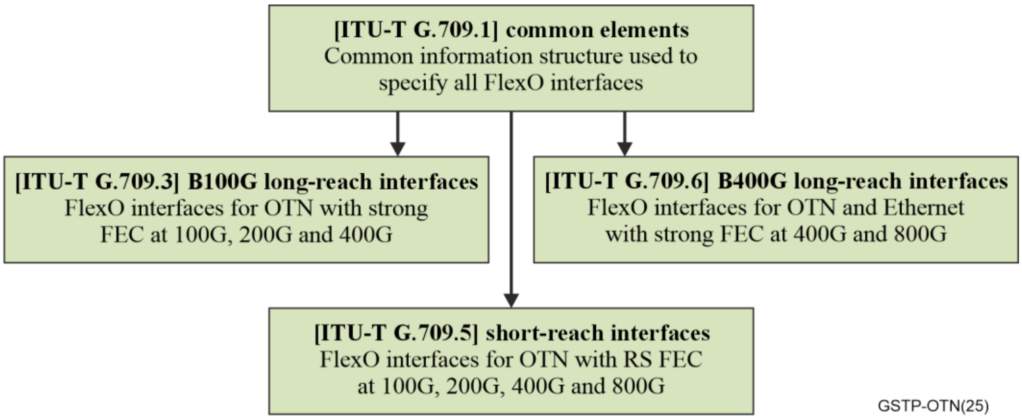


Figure 2-17 – FlexO Recommendation structure

As mentioned above, it was advantageous to map the OTUCn signals into a newly defined modular flexible OTN (FlexO) PHY format rather than directly using the OTUCn signal as the PHY interface signal in the same manner as the OTUk. The FEC was moved from the OTU frame to the FlexO frame, with the OTUCn stream being mapped into the payload area of a stream of FlexO frames. Following an approach that had already become common with single-vendor applications, the alignment between the OTUC frame and the FEC codeword stream is arbitrary[8]. This approach allowed modifying the FlexO frame for different FEC options for different rate and reach applications, as explained below. The flexibility of the FlexO signal format has led to it being adopted by other standards organizations, including OIF and [b-IEEE 802.3] for use with long-reach *x* × 100G interfaces.

FlexO is a modular interface consisting of a set of 100 Gbit/s optical PHY streams that are bonded together to carry an OTUCn. For example, with 100 Gbit/s PHYs, a set of *n* 100 Gbit/s PHYs are bonded together to carry an OTUCn, with each 100 Gbit/s PHY carrying an OTUC slice. This flexibility allows the use of any value of "*n*" for an OTUCn interface rather than defining only certain discrete values of *n* (e.g., just *n* = 4 and 10). Further, if for example a set of *m* 100 Gbit/s PHYs are available, a subset of *n* PHYs can be chosen to carry an OTUCn ($n < m$), which would allow choosing the subset of PHYs with the best optical channel characteristics or carrying multiple OTUCn signals over a set of PHYs.

As explained below, the short-reach version of FlexO (FlexO-x-RS) takes advantage of being able to use existing optical modules that support OTU4 and 100 Gbit/s Ethernet (100GbE) as the individual

[8] This approach was typical for strong soft-decision FEC coding for very long-reach.

FlexO 100 Gbit/s PHYs. As future higher rate Ethernet modules became available (e.g., 200 Gbit/s, 400 Gbit/s and 800 Gbit/s PHYs), FlexO was extended to also make use of them. FlexO makes partial reuse of the lane architecture and FEC structure from 100GbE, 200GbE, 400GbE and 800GbE Ethernet in order to also leverage Ethernet concepts.

The common properties across FlexO interfaces are defined in [b-ITU-T G.709.1]. Figure 2-18 illustrates the flow of mapping relationships for inserting a client signal into FlexO. As noted above, client signals are not mapped directly into an OPUCn but must be first mapped into their own ODUk/ODUflex.

The general text format used to describe a FlexO signal is:

FlexO-(n/ne/x/xe)-<int>-m.

The fields following "FlexO" are defined as follows:

- FlexO instance: A 100G FlexO-1 or FlexO-1e
- FlexO-n: A FlexO information structure with *n* FlexO instances within the FlexO group used to carry one or more OTUCn$_i$ clients ($\sum n_i \leq n$)
- FlexO-ne: A FlexO information structure with *n* FlexO instances within the FlexO group used to carry one or more *y*00Gbit/s Ethernet client signals ($\sum y \leq n$)
  - The FlexO instances have a bit rate optimized for carrying 100 Gbit/s of the $y \times 100$ Gbit/s Ethernet signal
  - Currently $y = 1$, 2, 4 and 8, with 16 in development in [b-IEEE 802.3].
- FlexO-x: A FlexO information structure where x represents the number of interleaved FlexO instances in a PHY. Hence, x indicates the FlexO-x interface bit rate in 100G increments.
  - e.g., x=1 for 100G, x=2 for 200G, x=4 for 400G, etc.
  - For use in carrying client signals (e.g., OTUCn or Ethernet) over a group of *m* instances $(m = \lceil (n/x) \rceil)$
- FlexO-xe: A FlexO information structure where x represents the PHY interface bit rate in 100G increments and where the 100G bit rate is optimized for carrying *y*00 Gbit/s Ethernet client signals.
- FlexO-n(e) and FlexO-x(e) use (e) to indicate that the text describes both the OTN and Ethernet rate optimized versions of the FlexO signal.
- -<int> indicates a specific short or long-reach interface type with respect to FEC or a combination of the FEC and modulation type.
- -m indicates the number of interfaces in the FlexO group
  - Note that m has other meanings within different contexts.

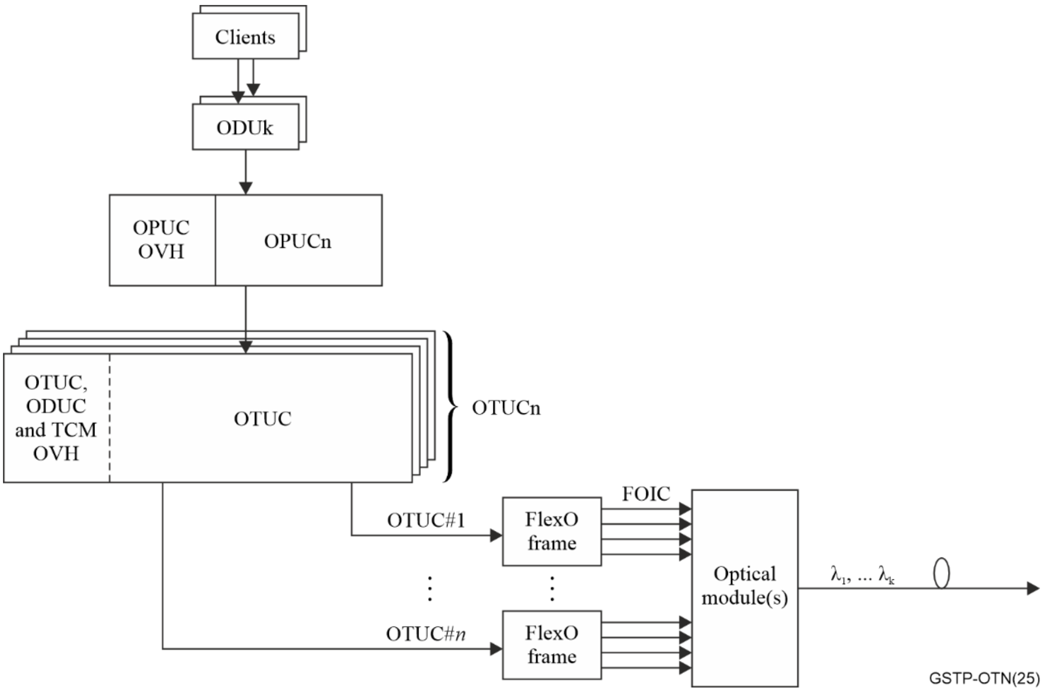


Figure 2-18 – FlexO client mapping relationship example

One or more clients are multiplexed to a FlexO-n(e) group, which can be inverse multiplexed across a FlexO-x(e)-<int>-m group interface where $n \leq (m \times x)$. For example, if $x = 2$ and $m = 2$, the FlexO group consists of two 200G interfaces and is capable of carrying up to 400G of traffic (i.e., $n \leq 4$). For scenarios where the FlexO-n(e) does not fill the entire FlexO-x(e)-<int>-m (i.e., $n < (m \times x)$), some instances are marked as unequipped.

2.4.1 FlexO frame structure and rates

The FlexO frame structure, as shown in Figure 2-19, consists of 5140 bit columns and 128 rows. As described below in clause 2.4.7 for short-reach interfaces, each 5140-bit row is the payload area of an FEC codeword. There are eight frames in the FlexO multiframe. The FlexO overhead fields and payload structures are described below. The overhead is divided into three fields: alignment marker (AM), basic overhead (BOH) and extended overhead (EOH). Each is discussed below.

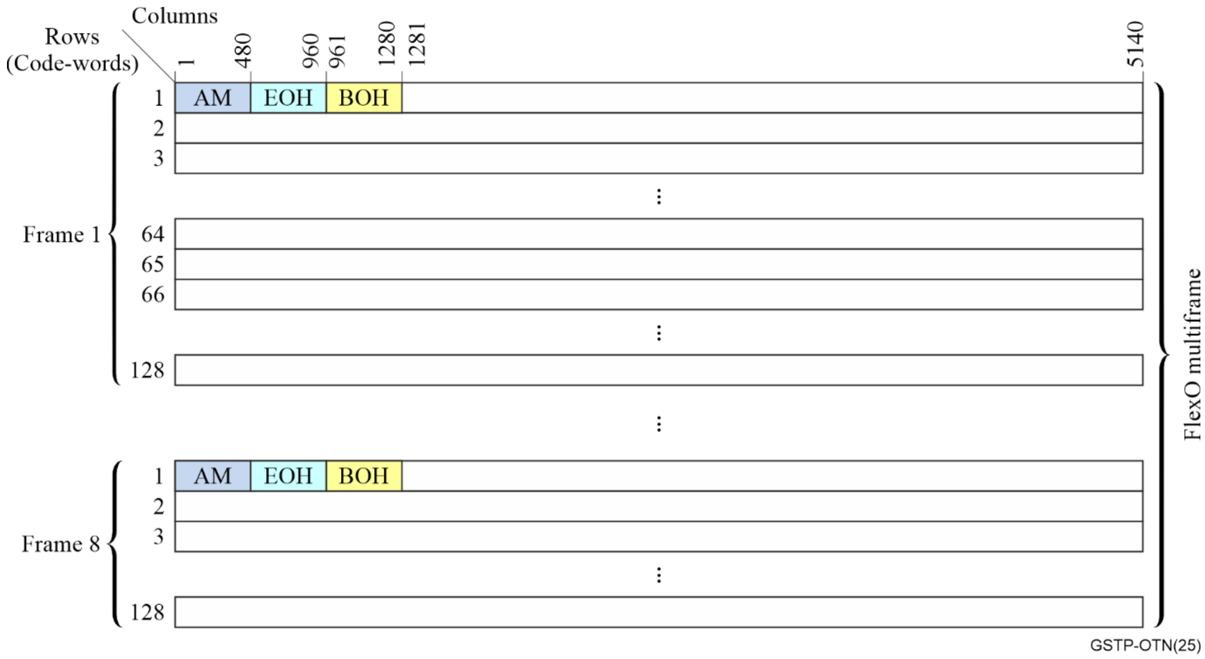


Figure 2-19 – FlexO frame and multiframe illustration

Similar to the OTUk frame, the FlexO signal is scrambled with a frame-synchronous scrambler prior to transmission.

See Table 2-6 for the nominal FlexO-n(e) bit rates.

2.4.2 FlexO-x(e) payload structure

The FlexO-x(e) for x > 1 is constructed from $x \times 64$ rows by 10280 single bit columns[9]. As shown in Figure 2-20, it consists of the interleaving of *x* frame and multiframe-aligned FlexO-1 frames (see Figure 2-19) that are z-bit interleaved in order from lowest to highest identification (IID). See [b-ITU-T G.709.3], [b-ITU-T G.709.5] and [b-ITU-T G.709.6] for the *z* value used with the respective interfaces[10].

Table 2-6 – FlexO-n(e) bit rates

Interface Type	Nominal bit rate	Bit rate tolerance
FlexO-n	≈ n × 105 643 510.782 kbit/s	±20 ppm
FlexO-ne	≈ n × 100 622 438.327 kbit/s.	±20 ppm
NOTE – The actual FlexO-ne rate is derived from a 156.25 MHz clock that is commonly available in Ethernet systems. Specifically, the FlexO-ne rate is n × 511/512 × 514/544 × 1445/1624 × 766 × 156 250 kbit/s.		

[9] $x \times 64 \times 10280 = x \times 128 \times 5140$

[10] $z = 10$ for [b-ITU-T G.709.5] and [b-ITU-T G.709.3] and $z = 128$ for [b-ITU-T G.709.6].

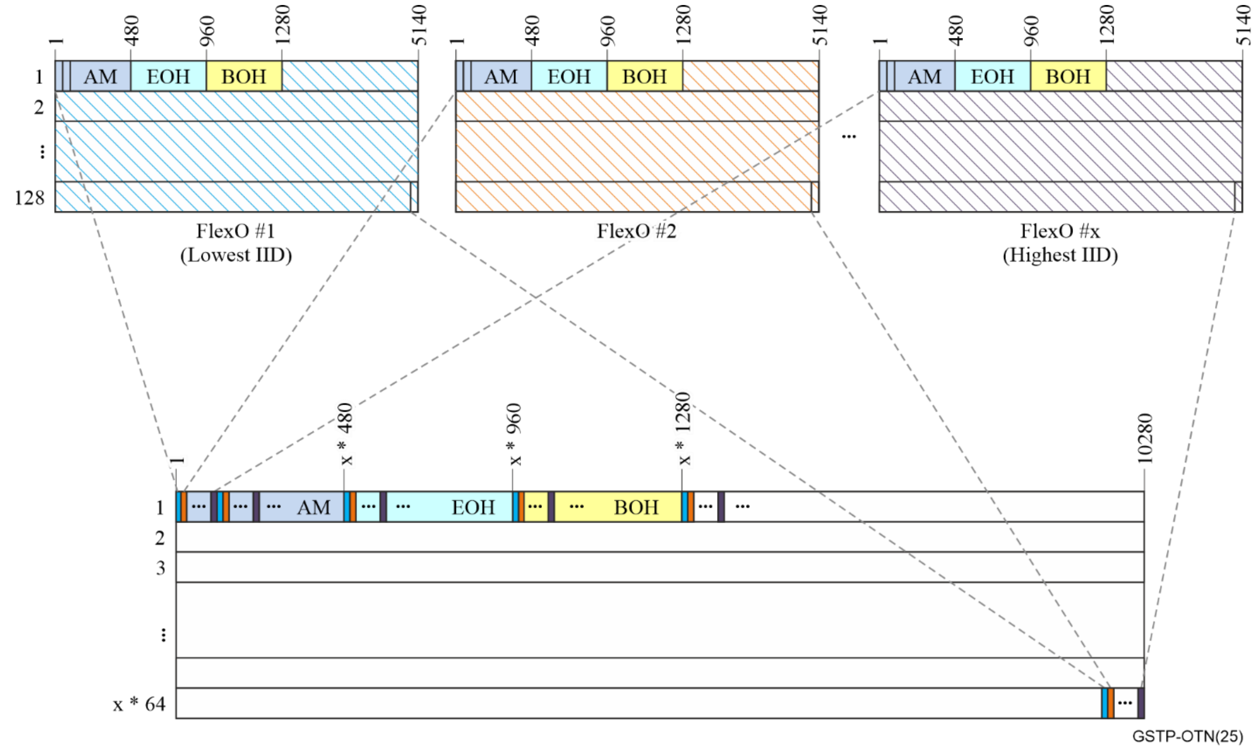


Figure 2-20 – FlexO-x frame and multiframe illustration

2.4.3 FlexO overhead types

As shown in Figure 2-19 and in Figures 2-21 and 2-22 below, FlexO uses many of the same functions defined above in Table 2-2. The overhead specific to FlexO is shown and summarized in Table 2-7.

The basic overhead (BOH) fields contain functions used in multiple FlexO applications. The BOH structure is illustrated in Figure 2-20. The extended overhead (EOH) fields contain functions that are used in specific applications, including security, regen, and FEC block alignment for some long reach FECs.

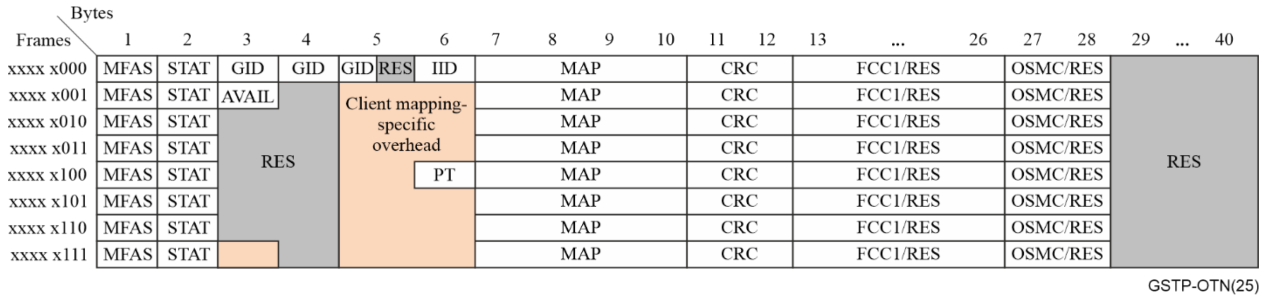


Figure 2-21 – FlexO basic overhead illustration

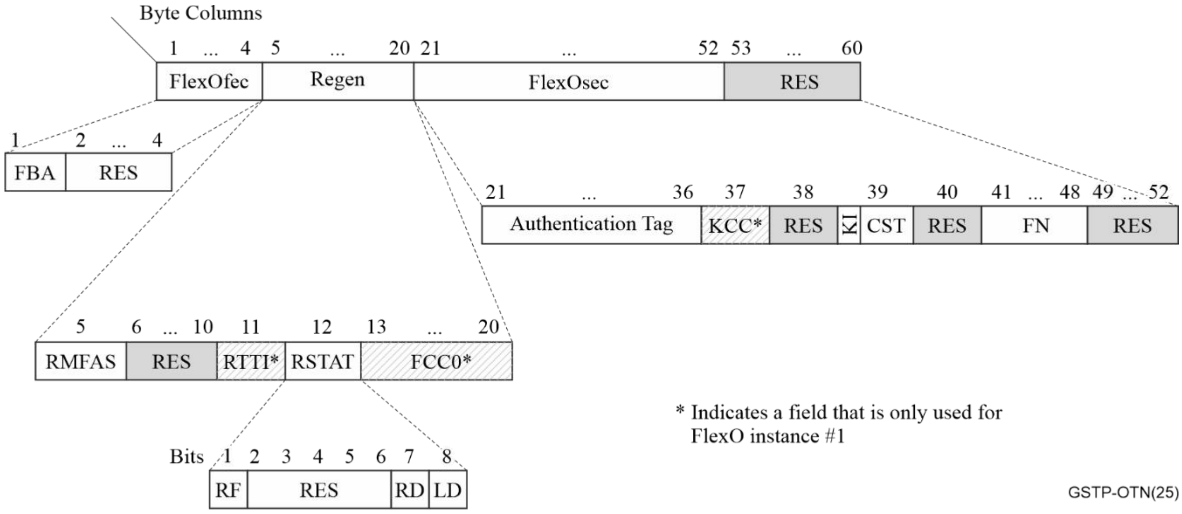


Figure 2-22 – FlexO extended overhead illustration

The first FlexO-x(e)-<int> in a group is indicated by it having the lowest IID value in the group. Since the IID values are not required to be sequential, the MAP conveys to the receiver which IID values it should be receiving for members of that group. In other words, the bit positions with value "1" within the MAP correspond to the IID set for the members of the FlexO-x(e)-<int> interface, and the lowest numbered IID uses the highest numbered MAP location within the group.

Table 2-7 – FlexO-specific overhead

OH Type	Location(s)	Comments
Alignment		
AM	Each FlexO instance	Interface specific alignment mechanism
Basic overhead (BOH)		
GID	FlexO BOH	Group identification: The number indicating the interface group instance (i.e., the group of PHYs) of which a FlexO-x(e)-<int> is a member. The GID allows the receiver to check whether the interface belongs to the intended FlexO group. (GID = 0 indicates that a PHY is not part of any FlexO group, i.e., is unequipped.) The same GID value is used in both directions.
IID	FlexO BOH	FlexO instance identification: The IID field uniquely identifies each of the m interfaces (members) of a FlexO-x(e)-<int>-m interface group, including their order within the group so that they can be correctly reordered at the receiver. The same IID value is used in both transport directions.
MAP	FlexO BOH	Indicates which specific members are used by that FlexO group. The MAP field contains one bit corresponding to each of the 256 possible IID values, and a "1" in a bit position indicates that the corresponding IID is being used by this group.
Extended overhead (EOH)		
FBA	FlexO EOH	FEC block alignment used with the ITU-T G.709.3 long reach interface
Regen	FlexO EOH	Overhead for Flexo regen applications
FlexOsec	FlexO EOH	Security overhead for FlexO

2.4.4 Client mapping into FlexO

2.4.4.1 BMP mapping

The original FlexO application used BMP to map OTUCn clients into the FlexO-n payload area with each OTUC mapped directly into a FlexO instance (i.e., there is a one-to-one relationship between an OTUC and a FlexO instance). The OTUC occupies the entire FlexO frame payload area, minus the FlexO overhead and fixed stuff bytes. Consequently, the FlexO bit rate is determined directly by the OTUC rate, and no dynamic rate justification is required.

As illustrated in Figure 2-23, the payload area of the FlexO multiframe is divided into 128-bit blocks, which are aligned to the start of the FlexO payload area following the 1280-bit set of AM, EOH and BOH fields. The OTUC data is mapped into the payload in 128-bit words that align to the 128-bit structure of the OTUC frame. These 128-bit words align to the 128-bit structure of the OTUC frame. The fixed stuff bytes for the BMP mapping are located in the first 1280 bits of row 65 in frames 1-7 of the 8-frame FlexO multiframe.

While there is an integer number of 128-bit payload blocks within the payload area of a FlexO frame, the payload area of each FlexO frame row is not evenly divisible by 128. Consequently, the last block of most rows will cross over into the next row. Taking into account the eight overhead and seven fixed stuff appearances, the FlexO multiframe contains:

$$((8 \times 128 \times 5140) - ((7+8) \times 1280) \text{ bits/MF}) / (128 \text{ bits/block}) = 40970 \text{ payload blocks}$$

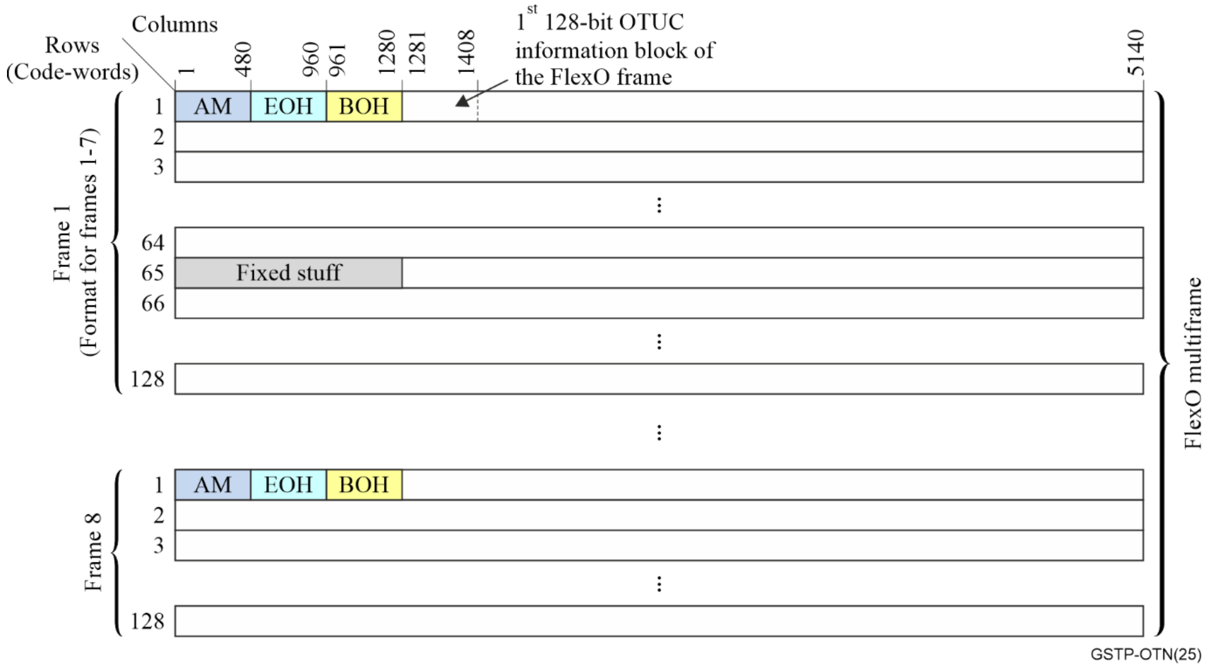


Figure 2-23 – Illustration of the FlexO payload area with BMP

Due to the relative lengths of the OTUC frame and FlexO multiframe payload area[11], the OTUC frame structure floats within the FlexO frame (i.e., there is no fixed positional relationship between the OTUC and FlexO frames).

2.4.4.2 GMP mapping

The GMP mapping was introduced into FlexO to accommodate both clients with clock tolerances outside the FlexO range (e.g., Ethernet clients with a ±100ppm range) and to enable the multiplexing of multiple clients into the FlexO-x(e) payload area where each client may have originated in a

[11] The bit ratio between the FlexO frame and OTUC client is 4112/4097.

different source timing domain. GMP can be used either for mapping or multiplexing y00GBASE Ethernet clients into a FlexO-ne (y = n) or OTUCn$_i$ clients into a FlexO-n.

As explained below, the GMP frame payload area P_{server} is 10260 blocks for the OTUC mapping into FlexO-1 and 10220 blocks for the Ethernet mapping into FlexO-1e, which allowed using the same 14-bit C_m coding as the OTN OPUk. See Figure 2-24 for an illustration of the GMP mapping field structure and specific fields. The field that is unique to FlexO GMP applications is the Client Status (CSTAT) field that communicates fault and maintenance conditions.

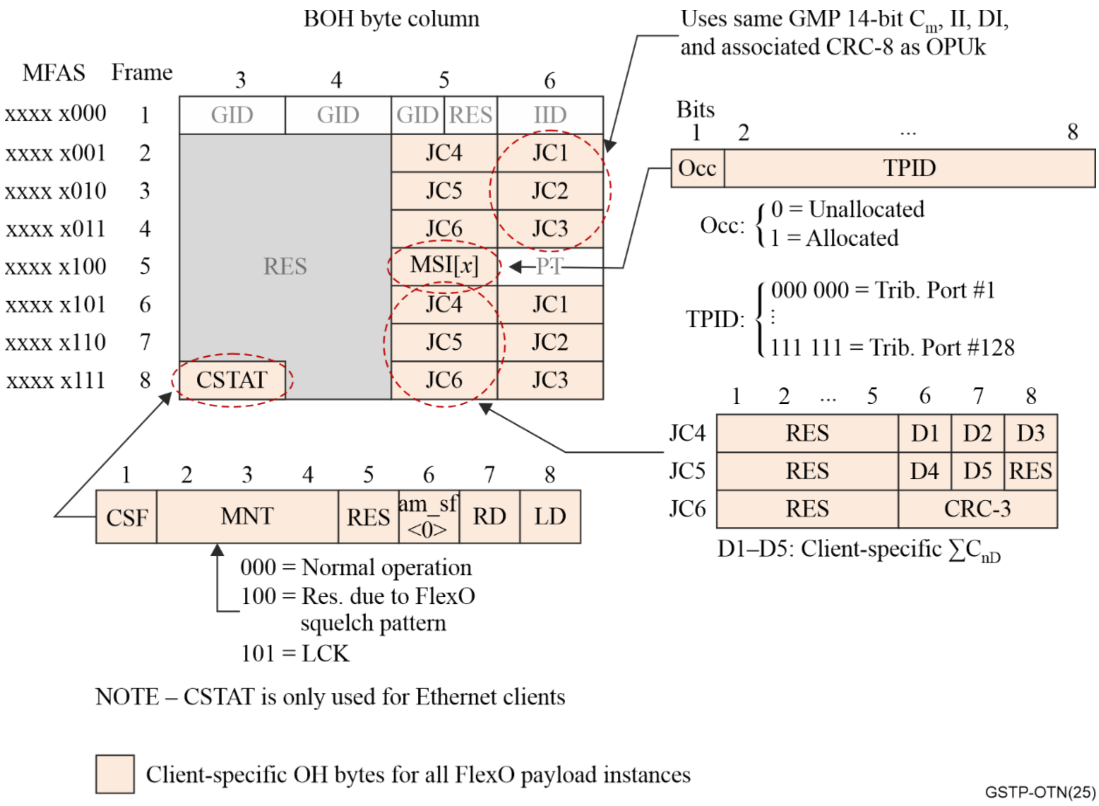


Figure 2-24 – Mapping specific overhead for client mapping with GMP

2.4.4.2.1 GMP mapping and multiplexing of y00GBASE Ethernet clients into FlexO-ne

In the case of y00GBASE Ethernet clients (y = 1, 2, 4, 8), they can be mapped directly into a FlexO-ne of the same nominal rate (i.e., n = y), which avoids the added bit rate and complexity associated with first mapping the Ethernet client into an ODUflex and subsequently into an OTUCn. Using GMP for this mapping also allows multiplexing one or more separate y00GBASE Ethernet clients (y ≤ n) into the n instances of a FlexO-ne. The resulting Ethernet optimized FlexO signal is referred to as a FlexO-ne / FlexO-xe (i.e., a FlexO-n with a rate and format optimized for y00GBASE Ethernet clients).[12]

As illustrated below in Figure 2-25, the y00GBASE Ethernet client GMP mapping or multiplexing into FlexO-ne uses m = y × 257 bits with the 257 bits corresponding to the Ethernet 256B/257B block codes.

The GMP frame portion of a FlexO-ne multiframe is illustrated in Figure 2-25 with the payload area of each FlexO instance divided into 257-bit blocks. A 5-bit fixed pad is added immediately following

[12] The FlexO-ne signal format has been adopted by OIF and OpenROADM as the mechanism for providing the FEC frame for their long reach Ethernet interfaces (e.g., y00ZR / y00ZR+). [b-IEEE 802.3] subsequently adopted the same frame format for some 800G PHYs in 802.3dj. The main difference between the OIF/ [b-IEEE 802.3] signals and the ITU-T FlexO-xe is that the OIF and 802.3 interfaces are only defined for point-to-point applications. In contrast, FlexO includes overhead for telecom networking applications.

the FlexO frame overhead in the first row of the FlexO frame in order to have an integer number of 257b block locations between the pad and the end of the row. Each subsequent row has 5140/257 = 20 257b blocks, which results in the P_{server} being 10220 blocks per FlexO instance as mentioned above.

The FlexO frames are aligned across the y FlexO payload instances of the FlexO-ne (i.e., multiframe aligned and phase-locked). For the 14-bit C_m field, m = y × 257 bits, which allows mapping successive groups of y 257-bit data or stuff blocks into the set of instances. This alignment also allows the GMP overhead of each FlexO instance to use identical C_m and $\sum C_{nD}$ in each frame. In other words, while the C_m is actually the number of y × 257b blocks, the alignment means that each FlexO instance can treat its GMP overhead as pertaining to the 257b blocks in its own payload area.

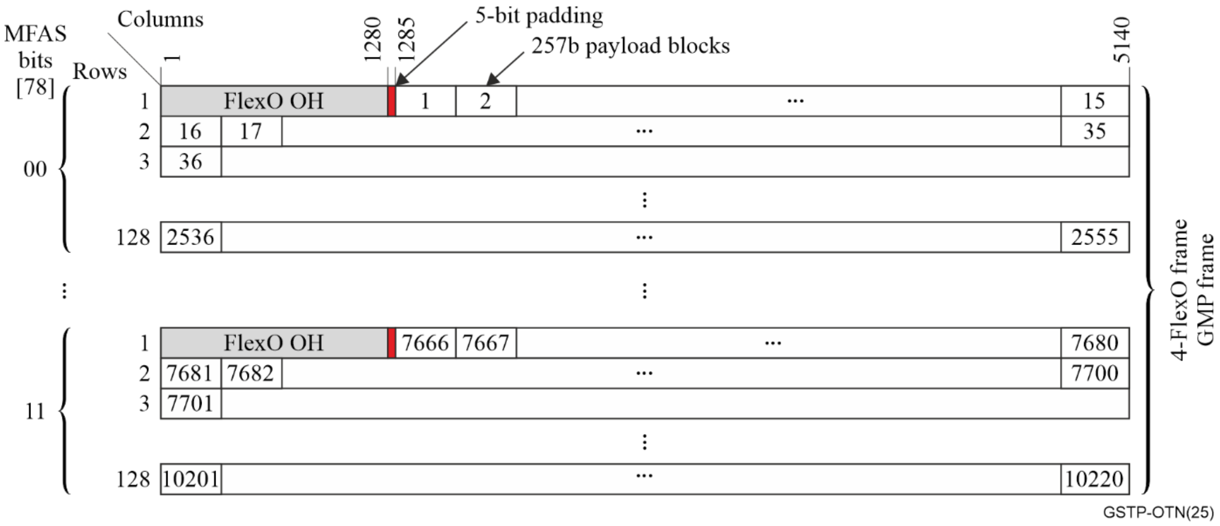


Figure 2-25 – FlexO-ne payload area for a GMP frame

2.4.4.2.2 GMP mapping and multiplexing OTUCni clients into FlexO-n

In the case of $OTUCn_i$ clients, each OTUC is associated with a FlexO instance (i.e., there is a one-to-one relationship between them). The MSI is used to indicate the OTUC content of each FlexO instance, including the associated TPID. An OTUC, which uses a 128-bit-oriented frame structure, is mapped into the FlexO instance payload in successive sets of two 128-bit data blocks. Consequently, each 256b block within the FlexO instance payload area (see Figure 2-26) is filled with either 32 bytes of the OTUC or a 32-byte GMP stuff word.

As illustrated in Figure 2-26, the payload area of the 4-frame GMP frame is divided into 256-bit blocks. Since 5140 is not divisible by 256, the blocks at the end of all but the last row wrap around into the next row. Note also from Figure 2-26 that the GMP mapping does not use the BMP fixed stuff bits shown in Figure 2-22. Using these bits for payload provides sufficient bandwidth to use GMP for the OTUC into FlexO without needing to increase the FlexO rate. As can be seen from Figure 2-26, the resulting P_{server} per instance is 10260 blocks.

As with the Ethernet mapping case, the FlexO frames are aligned across the group of n FlexO payload instances of the FlexO-n, which allows mapping the $OTUCn_i$ as successive groups of n 256-bit data or stuff blocks into the set of instances. For the 14-bit C_m field, m = n × 256 bits, which allows mapping successive groups of y 256-bit data or stuff blocks into the set of instances. Here also, the FlexO instance frame alignment allows the GMP overhead of each FlexO instance to use identical C_m and $\sum C_{nD}$ in each frame. In other words, while the C_m is actually the number of n × 256b blocks, the alignment means that each FlexO instance can treat its GMP overhead as pertaining to the 256b blocks in its own payload area.

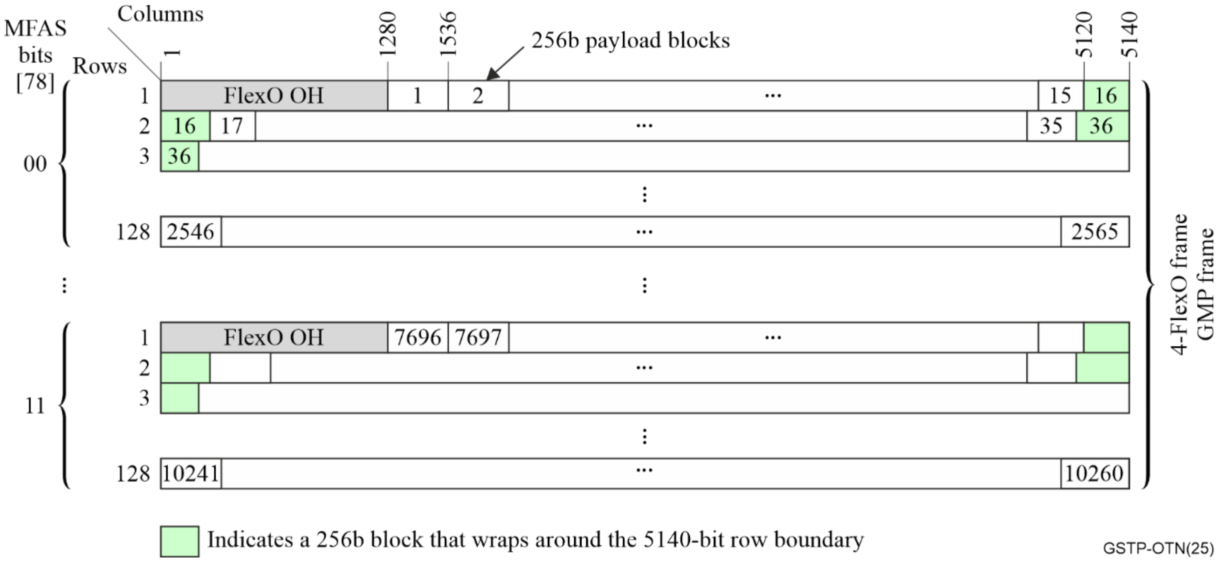


Figure 2-26 – FlexO-n payload area for a GMP frame

2.4.5 Other FlexO applications

The FlexO frame EOH enables functionality for specific applications, including regenerators and link-level security applications.

2.4.5.1 Regenerator applications

As an alternative to, or in addition to the use of FEC discussed below or optical amplification, longer reaches can be achieved by regenerating the signal at one or more points along the span. Regeneration goes beyond simple repeater functionality in that it also involves terminating the optical signal, recovering the digital signal's clock and data, and mitigating jitter that has accumulated on the signal. The "fresh" digital signal is then re-transmitted with the clean clock over the next optical span[13].

Since the regenerator node is an active piece of equipment, it is important to be able to manage it (e.g., detect equipment faults, determine the quality of the signal it receives, and set the parameters for its transmitted signal). This is functionally similar to the OTN OTUk regenerator overhead referred to as "RS" (regenerator section) overhead.

2.4.5.2 FlexO link security (FlexOsec)

When interface confidentiality is required, encryption covers the information in the FlexO frame payload area prior to FEC adaptation and FlexO-x instance interleaving. Encryption is applied per individual FlexO instance. GCM-AES-256 is used for performing the encryption. Authentication covers the payload area, the BOH field and a portion of the EOH field.

Network operators and service providers have various applications that require confidentiality and authenticity of the data transported on their networks, as illustrated conceptually in Figure 2-27. These applications for OTNsec (security) are described in more detail in [b-ITU-T G.sup.76]. FlexOsec focuses on link/span security. Operators looking to secure their infrastructure in the network can use encryption and authentication on a per span (link) basis. The links/spans interconnect the OTN network elements (e.g., OTN switches, OTN regens, etc.) within the same administrative domain. The key management and agreement are owned by the network operator and outside the scope of ITU standardization. All client and ODUk services transported by the links are agnostic to the security application.

[13] In other words, regeneration typically involves O/E, clock and data recovery, jitter filtering, E/O.

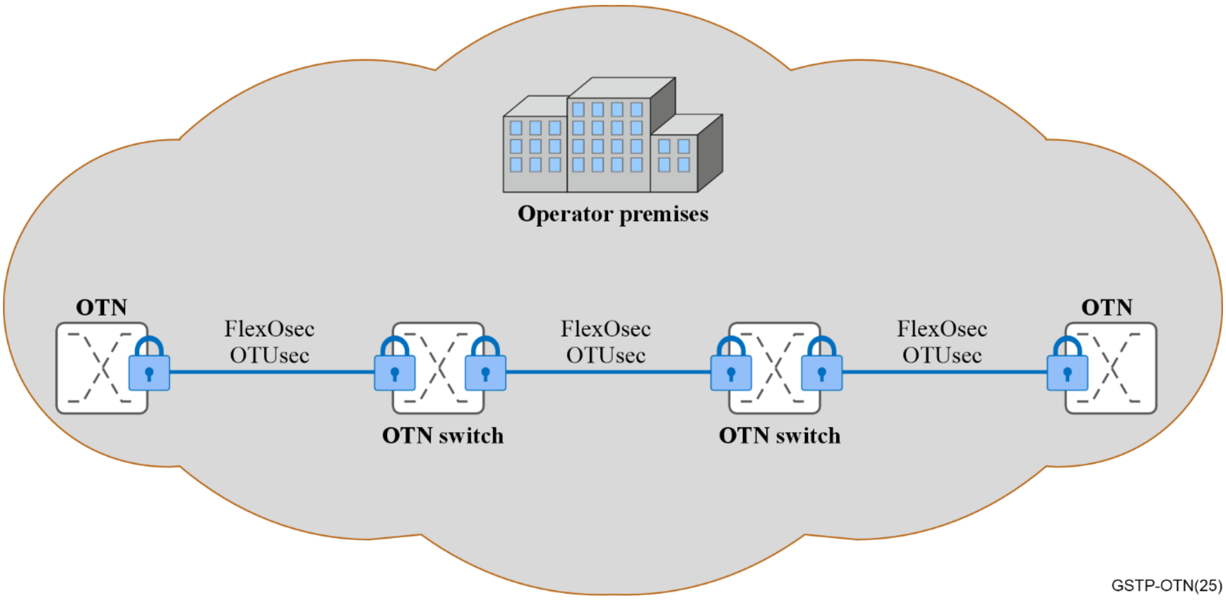


Figure 2-27 – Security conceptual architecture

2.4.6 FlexO interfaces data flow illustration

2.4.6.1 FlexO-n Data flow for OTUCn mappings

The data flow from the OTUCn client mapping to the output FOIC is illustrated in general terms in Figure 2-28 for n = 2, 4, and 8. See clause 2.5 for a discussion and information regarding FOIC. Note that at a high level, the FlexO data flow is conceptually similar to the [b-IEEE 802.3] Ethernet flow for the same rates, especially with respect to how the interleaving of data into the FEC engines is performed. For the case of OTUCn client multiplexing, multiple $OTUCn_i$ clients are GMP multiplexed into the FlexO-n payload area rather than a single OTUCn.

As illustrated in Figures 2-28 transmit direction and discussed above, the OTUCn is divided into its n constituent OTUC signals, each of which is then mapped into its own FlexO-1 frame (see Figure 2-19.) With the exception of the AM fields, the FlexO overhead is added to the frames. The n FlexO frames are then multiplexed together via interleaving 10-bits from each in a round-robin manner to form the input to a FlexO-n scrambler. The scrambler output is distributed in a 10-bit round-robin manner to the set of FEC engines. For n = 2 or 4, the odd numbered FlexO-1 signals are processed in one FEC engine and the even numbered FlexO-1 signal are processed in the other FEC engine. After FEC encoding, multiplexing and 10-bit symbol distribution is performed on the FEC engine outputs to create the 4n logical lanes. The combination of the multiplexing and distribution of the 4n 10-bit symbol streams results in each of the AMs appearing on the appropriate logical lane at this point. Having each AM appear on a separate lane allows the receiver to find and align to each lane, and perform lane deskew and de-interleaving of the 10-bit symbols.

The cases of y = 1 and y = 8 are similar but have some distinctions due to how the equivalent rate Ethernet signals are defined.

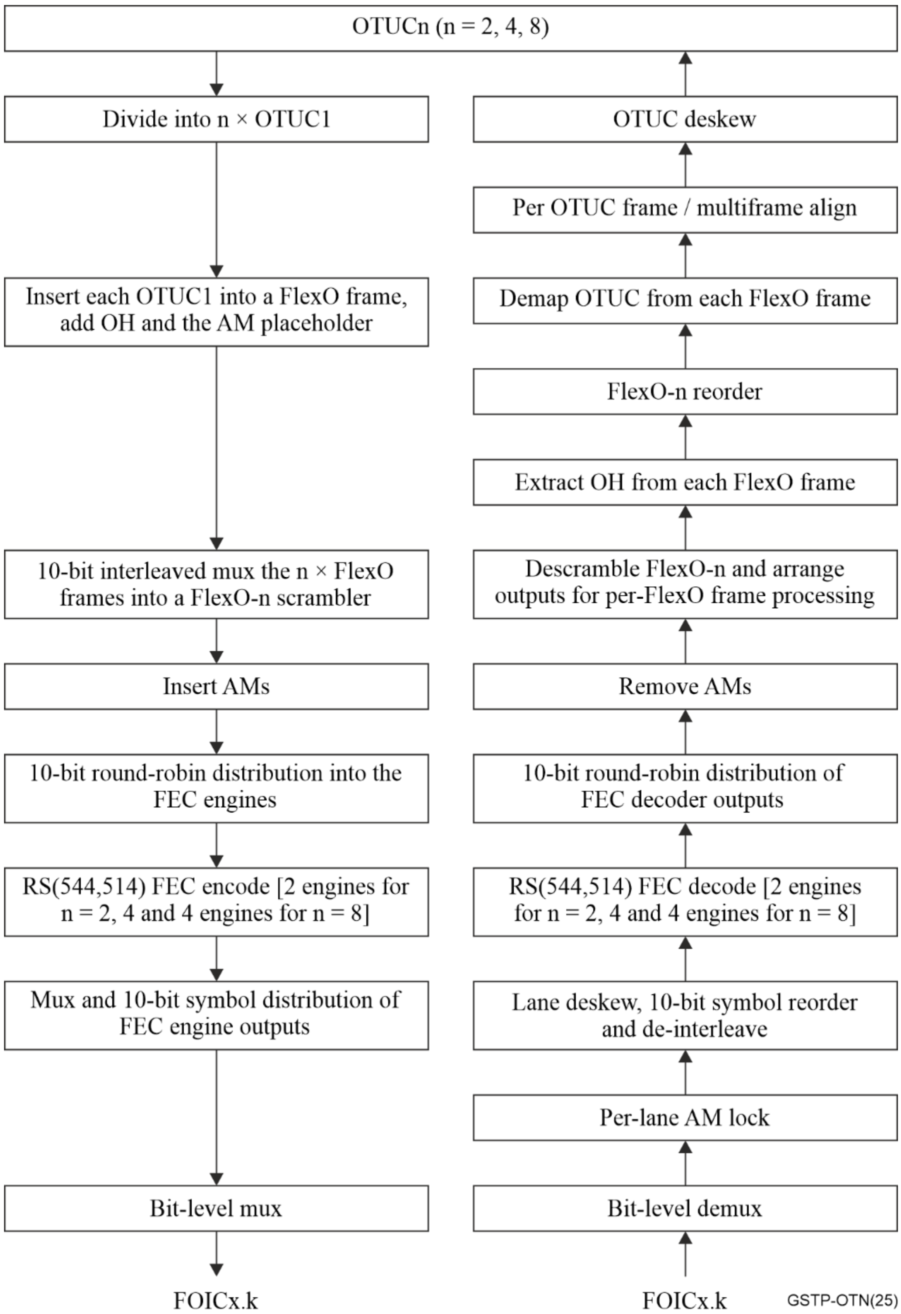


Figure 2-28 – Generic illustration of the FlexO OTN single client mapping data flow

2.4.6.2 FlexO-ne Data flow mapping for y instances of a y00GBASE Ethernet client

As illustrated in Figure 2-29, the Ethernet client signal is terminated up to the point of recovering a stream of 256B/257B blocks (e.g., by performing client FEC decoding, descrambling, AM removal, etc.). This client stream is referred to as the OTN reference signal. GMP encoding occurs at this point with the y00 Gbit/s stream of 257b blocks divided into 100 Gbit/s streams on a round-robin 256B/257B block basis and each of the 100 Gbit/s streams then inserted into the payload of a FlexO-1e instance. When the GMP and other FlexO overhead has been added to the set of FlexO-1e instances, the remainder of the flow (FlexO FEC encoding, AM insertion, scrambling, etc.) is the same as for an OTUCn client. Figure 2-29 illustrates the flow for y = 2 and 4. The cases of y = 1 and y = 8 are similar but have some distinctions due to how the equivalent rate Ethernet signals are defined. For example, 100GBASE Ethernet uses a single FEC engine, does not perform descrambling and rescrambling prior to 256B/257B transcoding, and does not use am_sf<2:0> bits. For 800GBASE Ethernet, the signal flow is effectively constructed by using a pair of the y = 4 flows, with #1-4 using one flow and #5-8 using the other flow. Consequently, there are a total of four FEC engines. Lanes

0-15 result from the first FEC engine pair and lanes 16-31 result from the second FEC engine pair. These 32 lanes are combined to create the FOIC using 32:8 bit multiplexing. The two 400G 257b streams are descrambled without AMs.

As stated above, in general, multiple instances of y00GBASE clients can be multiplexed into a FlexO-xe payload.

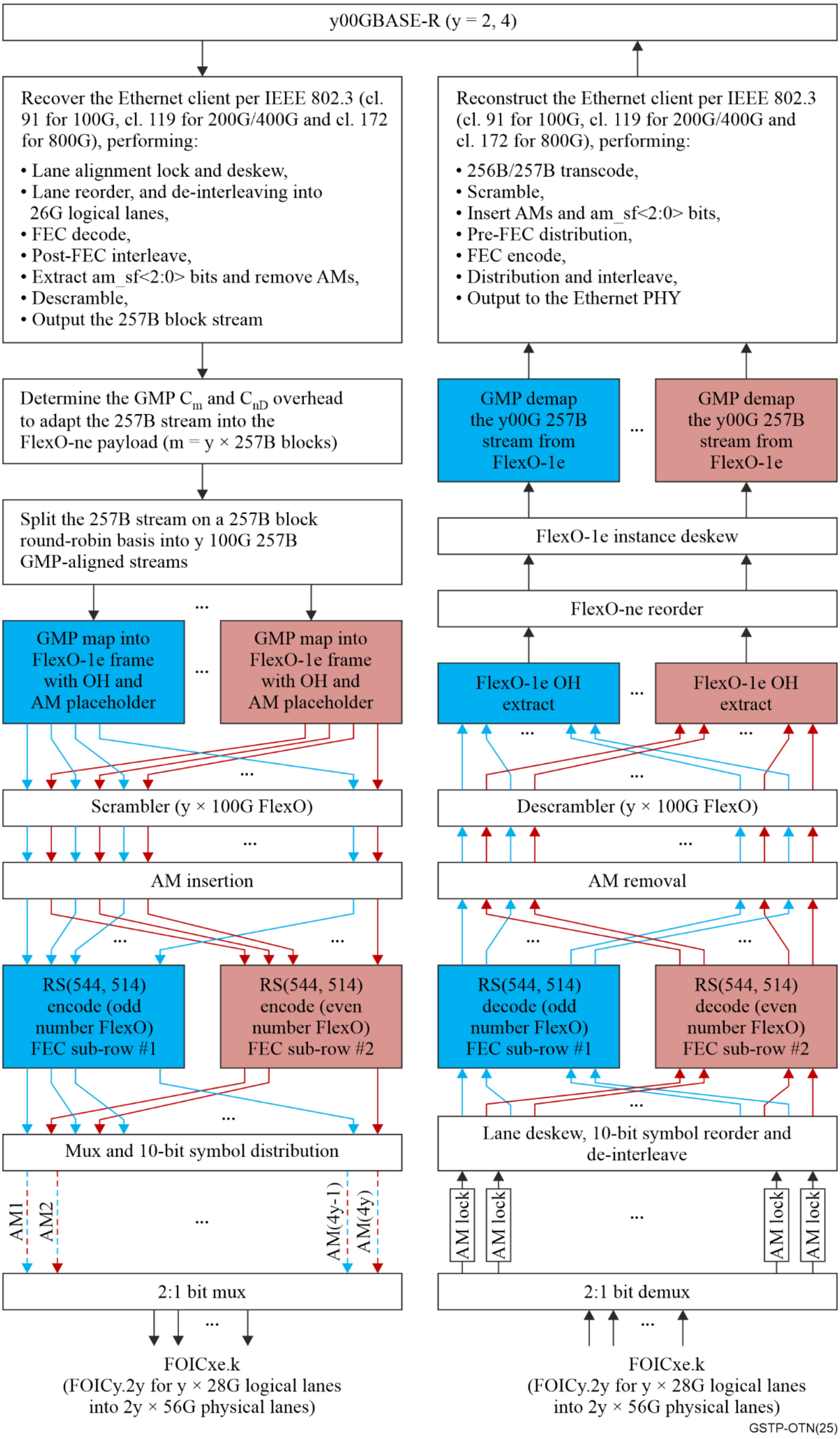


Figure 2-29 – Specific y00GBASE Ethernet client mapping data flow example for y = 2, 4

2.4.7 Short reach FlexO interfaces data flow illustration

The short reach FlexO interface is typically used for the FlexO electrical interface of order C (FOIC) described in clause 2.5. The data flow from the OTUCn client to the FOIC is illustrated in general terms in Figure 2-31, with more detailed examples provided in [b-ITU-T G.709.5]. Note that at a high level, the FlexO data flow is conceptually similar to the [b-IEEE 802.3] Ethernet flow for the same rates, especially with respect to how the interleaving of data into the FEC engines is performed.

For 100 Gbit/s FlexO-1-RS PHYs, the FEC is applied to each FlexO-1. The higher rate interfaces (FlexO-x-RS, $x > 1$), essentially follow the same 10-bit block (i.e., RS symbol) interleaving format as [b-IEEE 802.3] for the equivalent rate by distributing the stream across pairs of FEC engines. As shown in Figure 2-28 above, when multiple FlexO instances are multiplexed into a FlexO-x frame with x > 1, the constituent instances are combined on a 10-bit round-robin basis, and the frame row consists of 10280 bits (1028 10-bit symbols). Figure 2-30 provides a conceptual illustration of how the FlexO-x row is subdivided on a 10-bit symbol basis into the two sub-rows that provide the input to the respective FEC engines for FlexO-2 and FlexO-4. This subdivision results in the AM overhead being associated with the correct information blocks. Note that this approach is consistent clause 119 in with [b-IEEE 802.3]. Note also that per clause 119.2.4.7, the 10-bit re-interleaving of the FEC encoder outputs alternates between the FEC engines. Following the approach for 800GbE defined in [b-IEEE 802.3], FlexO-8 divides the incoming client stream into two separate 400 Gbit/s streams that each follow the FleO-4 flow until they are merged in the final bit multiplexing to create the FOIC.

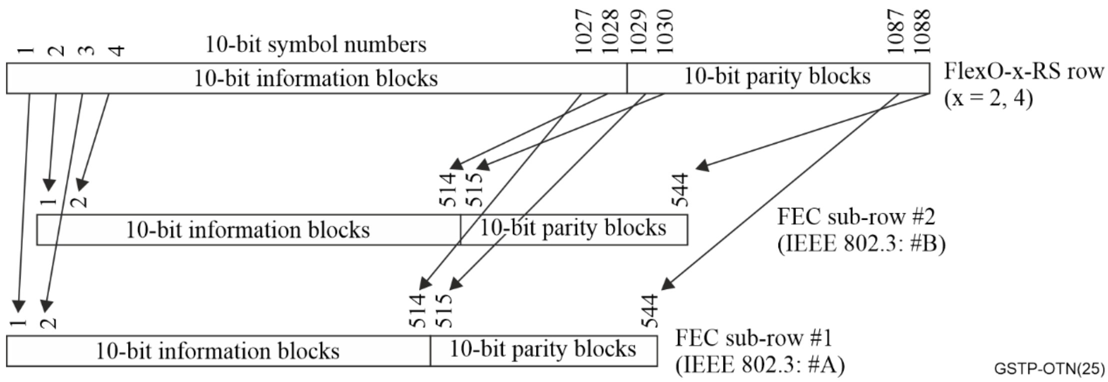


Figure 2-30 – FEC sub-row structure for FlexO-x (x = 2, 4)

2.4.8 Long reach FlexO interfaces

Long-reach FlexO interfaces are defined in [b-ITU-T G.709.3] for rates ≤ 400G and [b-ITU-T G.709.6] for rates ≥ 400G. Each uses a stronger FEC than the RS10 used with the short-reach FlexO interface in order to achieve multi-span metro reaches of up to 450 km. The signal formats for a given FEC have variations based on the specific optical signal modulation technique used to transmit them. Long-reach FEC is typically implemented in the optical modules rather than in framer devices, which only need short-reach FEC for their electrical FOIC interfaces to the optical modules.

The FlexO-x-RS of the FOIC is typically terminated in the optical module and the FlexO signal is encapsulated as the payload of another stronger FEC. Since such optical modules typically include digital signal processing (DSP) to enable coherent interfaces, the frame associated with this strong FEC is called a DSP frame. Consequently, the long reach interfaces are labelled as FlexO-x-D-<fec> interfaces, where D indicates the use of a DSP frame.

2.5 OTN electrical interfaces – MFI FlexO interfaces of order C (FOIC) and OTL (OTN transport lane)

OTL and FOIC are types of an OTN electrical module to framer interface (MFI). Intra-systems interfaces like MFI are generally outside the ITU-T standards scope. However, because such interfaces are crucial to support the OTN ecosystem[14], they were initially described in informational supplement [b-ITU-T G-Sup.58].

The OTN signal on the MFI includes an FEC. In most applications, the optical module terminates the MFI FEC and implements a stronger FEC for the optical span in addition to the module's optical DSP (ODSP) function. While the primary application of the electrical interfaces is MFI, both OTL and FOIC can also support options for the digital portion of the MFI electrical interface to pass through the optical module onto the optical interface.

The OTL naming convention is:

OTLk.n where:

- n is the number of parallel electrical lanes in the interface
- k designates the OTN signal rate (i.e., the rate of the OTUk carried on that interface).

The FOIC naming convention is:

FOICx.k-MFI/FOICxe.k-MFI where:

- x is the FlexO-x interface bit rate (i.e., $x \times 100$ Gbit/s),
- k is the number of parallel electrical lanes in the interface,
- e designates an Ethernet-optimized FlexO.

All FOIC interfaces use the same RS FEC as short-reach ITU-T G.709.5 FlexO optical interfaces.

Note that within this section, the term "logical lanes" is equivalent to the Ethernet BASE-R PCS lanes over which the FlexO-x-RS signal is distributed. An FOIC physical lane can carry one or multiple multiplexed logical lanes.

Note that the OTL and FOIC electrical interfaces defined in [b-ITU-T G-Sup.58] are typically in support of an optical interface application code from either [b-ITU-T G.695] (clause 8) or [b-ITU-T G.959.1] (clause 8).

As shown in Table 2-8, OTL interfaces were defined for OTU3, OTU4 and for use with an OTUC. FOIC interfaces are defined for important network applications, which typically coincide with the equivalent Ethernet interface rates and AUI structures. See Figure 2-31 for FOIC application examples.

Table 2-8 – Defined OTN electrical MFI

Electrical land rate	OTLk.n	FOICx.k-MFI
11 Gbit/s	OTL3.4 OTL4.10	–
28 Gbit/s	OTL4.4 OTLC.4 OTL4.4-SC	FOIC1.4 FOIC2.8 FOIC4.16
56 Gbit/s	OTL4.2 OTL50.1	FOIC1.2 (Note) FOIC2.4 FOIC4.8
112 Gbit/s	–	FOIC1.1 FOIC4.4 FOIC8.8

[14] These MFI interfaces are analogous to the [b-IEEE 802.3] attachment unit interface (AUI) for chip-to-module interfaces.

Table 2-8 – Defined OTN electrical MFI

Electrical land rate	OTLk.n	FOICx.k-MFI
106.5 Gbit/s	–	FOIC1e.1 FOIC4e.4
NOTE – Use of n × FOIC1.2 is also defined for 200, 400 and 800 Gbit/s interfaces.		

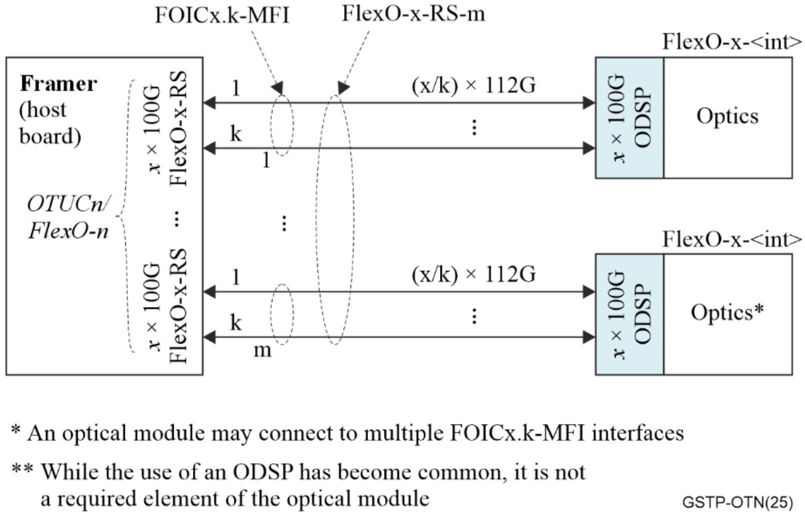


Figure 2-31 – Illustration of a typical FOIC application and structure

2.6 Fine grain OTN

As SDH networks are being retired, some network operators requested adding the ability to carry the primary legacy SDH network clients (e.g., 155 Mbit/s SDH and 2 Mbit/s E1 CBR clients, and packet clients coming from a 10 Mbit/s or 100 Mbit/s Ethernet UNI) directly over OTN with reasonable bandwidth efficiency. This application was accommodated by defining a new ODU variation called the fine grain ODUflex (fgODUflex) and a 10 Mbit/s fine grain TS (fgTS) size for ODU0. ODU1. ODU2 and ODUflex server layers. GMP is used both for multiplexing the fgODUflex into the OPU server and mapping CBR clients into the fgOPUflex. Packet clients are encoded and mapped as a stream of [b-IEEE 802.3] clause 82 64B/66B blocks.

Conveniently, there are 3808 = 32 × 119 byte-columns in a normal OPU payload area and dividing the OPU0 capacity rate by 119 yields a rate of (1.244160 Gbit/s / 119) = 10.455 Mbit/s fgTS capacity, which is adequate for carrying an OTN-mapped 10MBASE Ethernet client signal. An OPU1 supports 2 × 119 fgTS, an OPU2 supports 8 × 119 fgTS, and an OPUflex(fgTS, n) supports n × 119 fgTS.

The fgODUflex rate is chosen to be (p / 119.5) × 1.244160 Gbit/s = p × 10.409203 Mbit/s and the fgOPUflex channel rate is (237 / 239) × (fgODUflex(p) rate) = p × 10.322097 Mbit/s, where p is the number of fgTS that the fgODUflex will occupy in the server.

Using the normal approach of an OPU GMP multiframe length equal to the maximum number of OPU TS would result in a prohibitively long multiframe. However, an important consequence of the fgODUflex rate definition is that since the fgODUflex rate is defined with a fixed nominal relationship to the OPU server rate, the full range of the GMP C_m count is not required. The GMP count MSBs will remain unchanged across all GMP periods, which allows defining an OPU C_m base value (C_mB) and using the GMP overhead to only communicate the small difference between the actual C_m associated with this client fgODUflex and its base value (i.e., $C_mT = C_m – C_mB$). Transmitting C_mT only requires a 6-bit count field that fits into four JC bits/row. Since C_{nD} overhead is not required, it is possible to carry the GMP overhead associated with four separate fgTS in the JC bytes of each frame, shortening the multiframe length by a factor of four. See Figure 2-32.

The fgODUflex frame is based on the normal ODU, except that additional overhead columns have been added to provide the desired overhead bandwidth and latency. See Figure 2-33. The TSOH fields in Figure 2-32 contain the JC bytes for CBR client mapping, an OMFI for packet client mapping, and the PT and CSF for both client types.

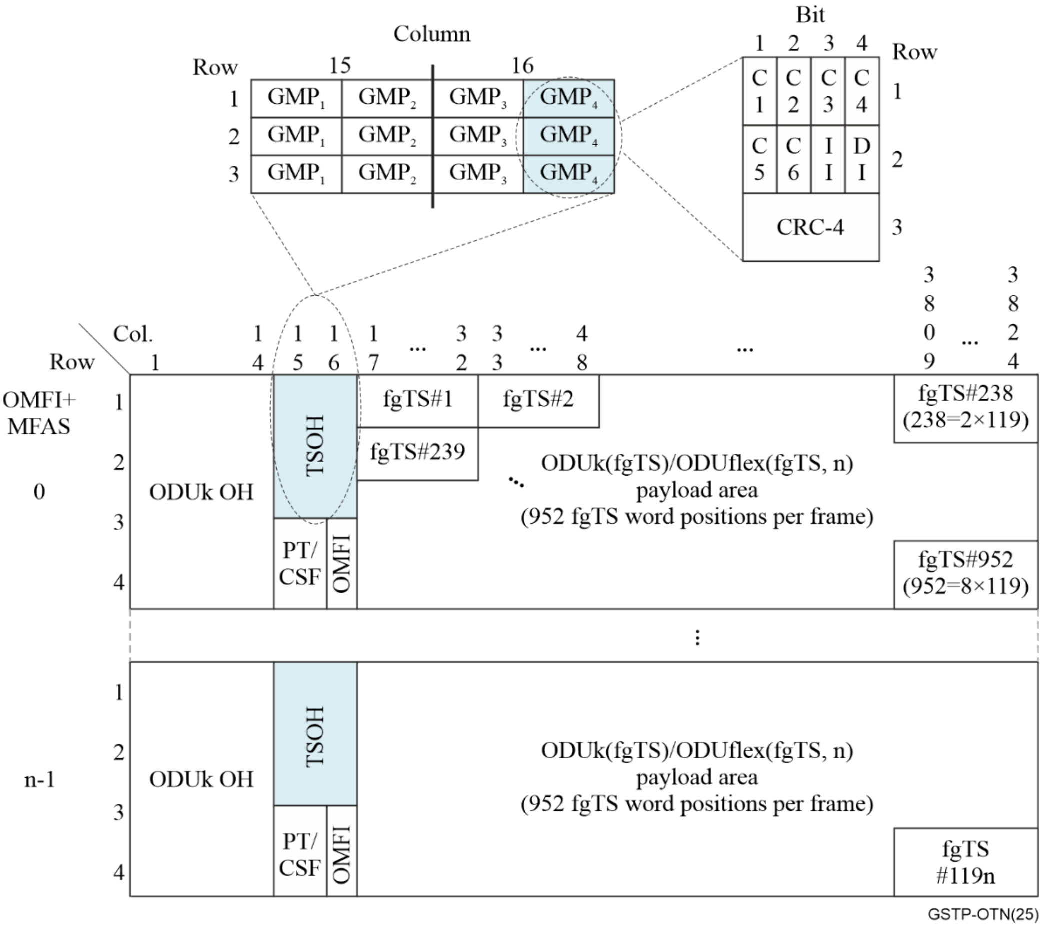


Figure 2-32 – ODUflex(fgTS, n) server frame format for fgOTN

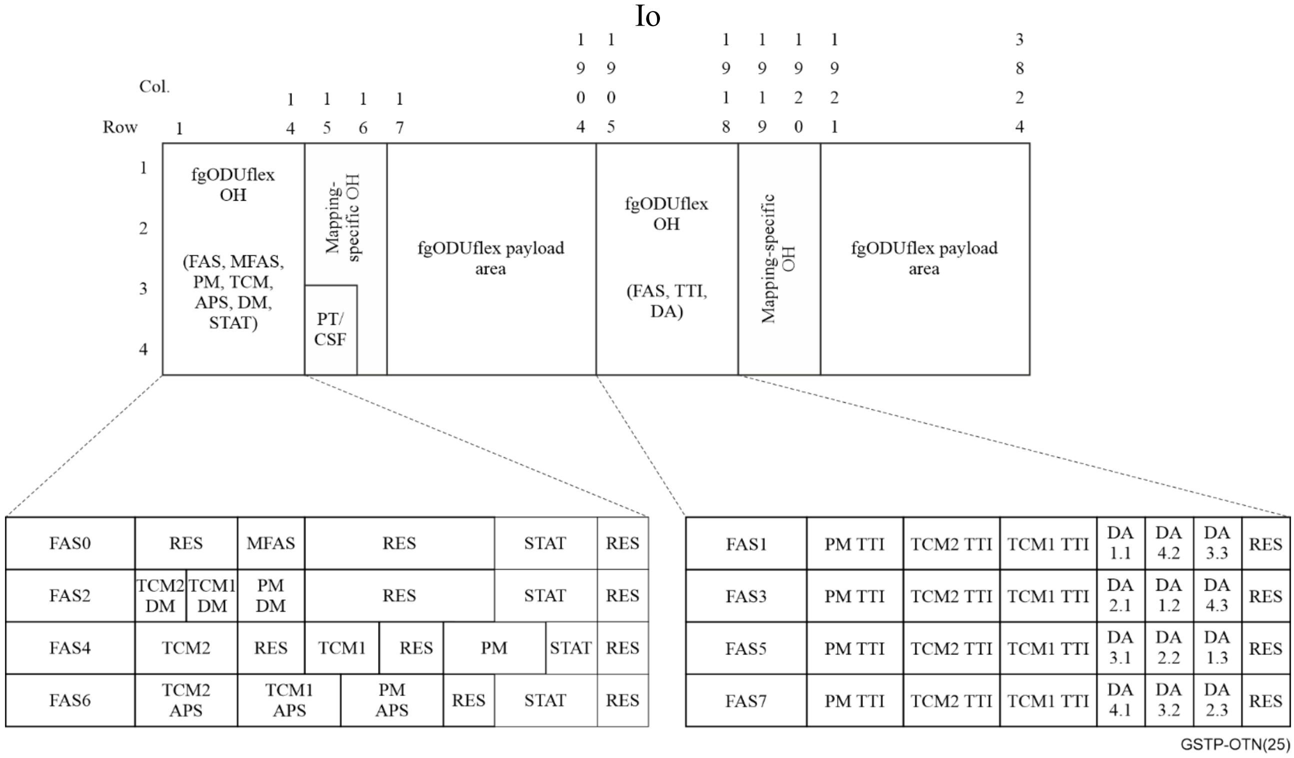


Figure 2-33 – fgODUflex(p) frame format for fgOTN

3 Operations, administration, maintenance (OAM) and provisioning

Transport network architecture breaks the functions of a network into the familiar stack of layered networks. The development of the OAM architecture used in OTN networks started around 1990. The functional modelling methodology used to describe layer networks is described in [b-ITU-T G.806]. The specific details of OTN OAM are described in [b-ITU-T G.709], [b-ITU-T G.709.1], [b-ITU-T G.709.3], [b-ITU-T G.709.5], [b-ITU-T G.709.6], [b-ITU-T G.798] and [b-ITU-T G.807].

Support for robust ("carrier class") OAM at layer 1 and layer 0 enables true networking capabilities at those layers and differentiates OTN from other transport technologies that support only point-to-point connectivity in these layers.

3.1 General OAM concepts

Operations refer to the activities undertaken to keep the transport network up and running. This includes monitoring the performance of the network; detecting, localizing, and repairing failures; and reporting performance and problems to the network management system (NMS) or software-defined networking (SDN) controller. Ideally, problems in the network should be identified before they affect client traffic. Examples of operations tools include proactive measurement of bit error rate or other performance measures and verification of connectivity between endpoints.

Administration concerns activities required to use network resources in an efficient manner. It involves bookkeeping of available resources, collection of performance statistics, planning network configuration, provisioning of network resources, and maintaining network availability. Examples of administration tools are management of the administrative state of resources (i.e., whether resources are allowed to provide service) and configuration of connectivity across the network or subnetwork.

Maintenance activities involve diagnostics and repairs to the network. [b-ITU-T M.20] defines maintenance as all actions required to set up and maintain, within well-defined limits, any element involved in a network connection. The objective of maintenance is to minimize the probability that network failures occur and minimize the impact of failures when they do occur. Examples of maintenance tools include loopbacks and on-demand measurement of performance.

The OAM architecture used in OTN networks has a common behaviour with other transport network technologies such as SDH and Ethernet over transport. Details of the OAM used in those other technologies are outside the scope of this technical paper.

OAM tools provide the capability to support fault management, performance monitoring and configuration management in a layer network. In other words: the OAM toolbox contains the tools and utilities to plan, install, monitor and troubleshoot a layer network.

Fault Management (FM) includes:

- defect detection: anomalies affecting the transport of client information are detected by continuous proactive monitoring. Persistent anomalies are considered to be service affecting defects. Detected defects are correlated with other detected defects to find the most probable fault cause, and consequent actions, e.g., protection switch, are taken;
- fault localisation: if the defect information is insufficient to locate the failure, on-demand OAM functions can be used to determine more accurately the cause of the defect;
- fault reporting: persistent fault causes are reported to the network management system (NMS) or SDN controller to provide the appropriate alarm reports to the maintenance staff for maintaining the quality of service (QoS) level offered to customers;
- protection switching: after a defect has been detected a protection switch can be initiated to restore the interrupted traffic and thus improve the service availability.

Performance monitoring (PM) includes measuring the performance (packet losses, transfer delay, bit errors, etc.) of the transport of client information in order to verify the quality of service (QoS) and to estimate the transport integrity.

Configuration management (CM) includes indicating the operational state of a connection, e.g., whether it can be used to transport client data or whether the set-up of the connection is completed.

3.2 OTN layer networks

At a very high level, the OTN layer network hierarchy consists of one or more path layers, one or more section layers, and the optical media layer, as shown in Figure 3-1.

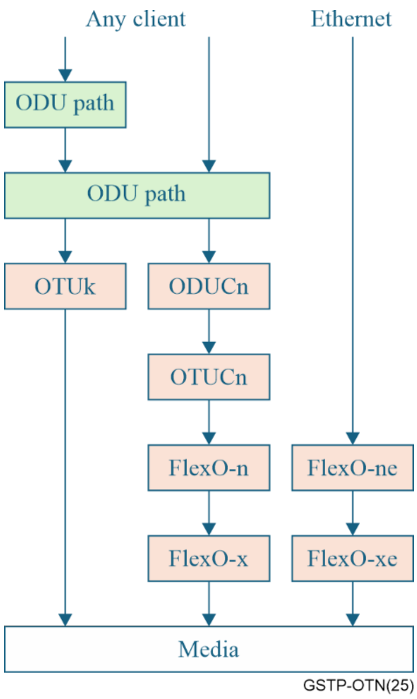


Figure 3-1 – OTN layer networks

3.2.1 OTN path and section layers

The defining characteristic of a path layer is that it supports flexible connectivity (i.e., connection functions). An OTN path layer carries either a client signal (e.g., Ethernet) or a multiplex of lower-speed (also called lower order) OTN paths. The path layer in OTN is called ODUk, with values of k as shown in Table 3-1. When an ODUk carries lower-order OTN paths, they are called ODUj (with j < k).

In contrast, a section layer does not support flexible connectivity. The section layer in OTN was originally OTUk, with values of k as shown in Table 3-1. The ODUk and OTUk layers share a common frame structure (see Figure 2-3), and the structure does not vary with the value of k. Since the frame size does not change with the bit rate, the overhead repetition period shrinks as the value of k increases.

Table 3-1 – OTN digital layer networks

Signal rate	k	Path layer	Section layer	Example client(s)
~1.25 Gb/s	0	ODU0	OTU0	1000BASE-X
~2.5 Gb/s	1	ODU1	OTU1	STM-16
~10 Gb/s	2	ODU2	OTU2	STM-64, ODU0
~10.3 Gb/s	2e	ODU2e	–	10GBASE-R
~40 Gb/s	3	ODU3	OTU3	40GBASE-R
~100 Gb/s	4	ODU4	OTU4	100GBASE-R, low order ODUj
Any rate	flex	ODUflex	–	Packet clients, 200GBASE-R, 400GBASE-R, 800GBASE-R, SAN or video clients
~n00 Gb/s	Cn	–	ODUCn/OTUCn, FlexO	ODUj paths OTUCn
~n00 Gb/s	–	–	FlexO	x00GBASE-R

Scaling the ODUk/OTUk frame structure beyond 100 Gbit/s was technologically challenging due the increasing overhead repetition rate (and there was no benefit from an OAM perspective to processing the overhead at ever faster rates). To support clients and links beyond 100G, some new structures were introduced to OTN. First, two new section layers, ODUCn and OTUCn, were introduced. These two layers share a common frame structure (like ODUk and OTUk do); they differ from ODUk and OTUk in that the ODUCn/OTUCn are based on interleaving 100G ODUC frame instances to form the higher-speed frame (e.g., an ODUC4 is formed by interleaving four ODUC frames). Similar to how SDH overhead was specified, some overhead elements are present in all ODUC instances, while other elements are present only in the first instance. This change to an interleaved structure froze the overhead repetition rate, since higher rates are created by interleaving a larger number of structures rather than by transmitting the frame at a higher rate. The path layer for beyond 100G clients is ODUflex (so for the path layer, the overhead repetition rate does continue to decrease as the client rate increases). At the time the beyond 100G OTN structures were defined, it was expected that most beyond-100G links would be carrying a multiplex of lower-rate OTN clients rather than Ethernet clients with rates faster than 100 Gb/s.

A key difference between the OTUk and the ODUCn/OTUCn section layers concerns the inclusion of Forward Error Correction (FEC) codes. The OTUk includes overhead for the FEC code, whereas the ODUCn/OTUCn frame does not. The OTUk included a FEC overhead field because at that the time it was defined, it was expected that all OTN signals would use a single standard FEC code. This proved to not be the case in actual deployments, as vendors opted to also support stronger FEC codes enabling longer reach links. As a result, when the ODUCn/OTUCn structure was defined, it did not include a FEC area, and it was necessary to specify FEC frames for OTUCn separately.

At the time the OTUCn section was defined, it was also the case that there were no client interface modules with transmission rates faster than 100 Gb/s. To facilitate OTUCn client handoffs between operator networks or domains, FlexOTN (FlexO) was defined as a way to inverse multiplex one OTUCn signal over multiple short-reach optical modules (e.g., the same modules that were being used to support 100GBASE-R Ethernet PHYs). At the time FlexO was defined, the intent was that one OTUCn would be carried by n 100G optical links in short reach applications only. Beyond-100G OTN line interfaces were generally assumed to be proprietary (using vendor-proprietary FEC codes and optical interfaces that were highly optimized to maximize reach).

FlexO has continued to evolve as optical interfaces generally have evolved. Higher rate interfaces (200G, 400G, and 800G) have been introduced. Demand for interoperable line interfaces based on standardized stronger FEC codes led to the definition of long-reach FlexO interfaces. Most recently, FlexO has been evolved to be a pair of section layers rather than an inverse multiplexing technology, and direct mapping of Ethernet to FlexO has been standardized, enabling OTN to deliver point-to-point Ethernet services without the complexity of the path layer overhead.

3.2.2 OTN optical media layer

At the bottom of the stack is the optical media layer, consisting not only of optical fibres, but also optical amplifiers, optical mux/demux devices (e.g., simple passive filters or active flex-grid filters), and optical switching systems (e.g., reconfigurable optical add-drop multiplexers (ROADMs) or optical cross-connect systems). The configuration of the active components and the concatenation of components and fibres create network media channels between optical transmitters and receivers. The signals produced by optical transmitters are called optical tributary signals (OTSi).

Because of the flexible connectivity that exists in ROADMs and optical cross-connects, or in the cabling between different optical layer devices, the media layer has many properties of a path layer network; the OTSi suffers impairments between the transmitter and receiver, and it is possible for OTSi to be misconnected. At the same time, an individual segment of fibre between two optical devices resembles a section layer (i.e., a portion of the end-to-end connectivity between the optical transmitter and receiver). This makes it tempting to think of the optical layer as if it were a stack of layers like the digital network layers.

However, there is one extremely important difference between the media layer and the digital section and path layers. In the digital layers of the OTN, overhead to support OAM functions is added by encapsulating the payload in a digital frame structure that includes the overhead. In the optical media layer, there is no such encapsulation possible. The OTSi emitted from a transmitter is a stream of photons with a particular wavelength; there is no frame structure in the optical layer and there is no ability for intervening components like optical filters, couplers, or amplifiers to 'add overhead' to the signal they receive.

Nonetheless, it is convenient to think about the media layer as if it were a stack of digital layers: a path layer between transmitter and receiver, with section layers between points where optical multiplexing and demultiplexing occurs, and between points where optical amplification occurs. To facilitate this management view of the media layer, the concept of non-associated overhead is used to superimpose the familiar hierarchical management structure of digital layers onto the media layer (which as noted above does not have any hierarchy). A set of digital overhead layers is defined:

- the optical transmission section (OTS), extending between amplifiers or between an amplifier and an optical mux/demux device,
- the optical multiplex section (OMS), extending between optical mux/demux points,
- the optical channel (OCh) or optical transport signal group (OTSiG), extending between a transmitter and a receiver. The OCh is the original specification and remains the most widely implemented. The overhead defined for the OCh assumes an OTUk is carried over the OCh and that the OTUk and OCh always terminate in the same location; as such, the OCh overhead consists of only maintenance signals, with all the monitoring overhead in the digital structure of the OTUk. As OTN evolved, the OTSiG added additional overhead, allowing additional monitoring of the OTSiG without any dependency on the digital client layer.

These digital overhead layers consist only of the overhead that would be used if the optical media truly was a digital hierarchy, and the resulting frame structure (with only the digital overhead related to the media layer) is carried over a separate OTSi called the optical supervisory channel (OSC). The OSC is terminated at every active OTN node. Processing the appropriate overhead from the OSC at the appropriate nodes enables management of the optical media layer as if it were a stack of three digital layers. For example, an optical amplifier (optically) demultiplexes the OSC from the group of optical signals it receives and extracts the digital frame structure from the OSC. The OTS overhead is extracted, processed, and replaced by new OTS overhead for the next span, while the OMS and OTSi overhead is passed through. The digital frame structure is then modulated onto a new OSC and (optically) multiplexed into the group of optical signals at the output of the amplifier.

Because the optical media overhead is carried over a different OTSi than the payloads for which it is providing the overhead, the payload and overhead can fail independently. As a result, there are separate payload and overhead versions of most maintenance signals for the optical media layer.

The optical media overhead and OSC are functionally standardized – meaning the information content is specified in OTN Recommendations, but the specific frame format and optical wavelength of the OTSi are not specified.

The optical media overhead is not used on all OTN links. In particular, single-client ("grey") links will generally not have this overhead. Small OTN networks based on passive optical components also will generally not have this overhead.

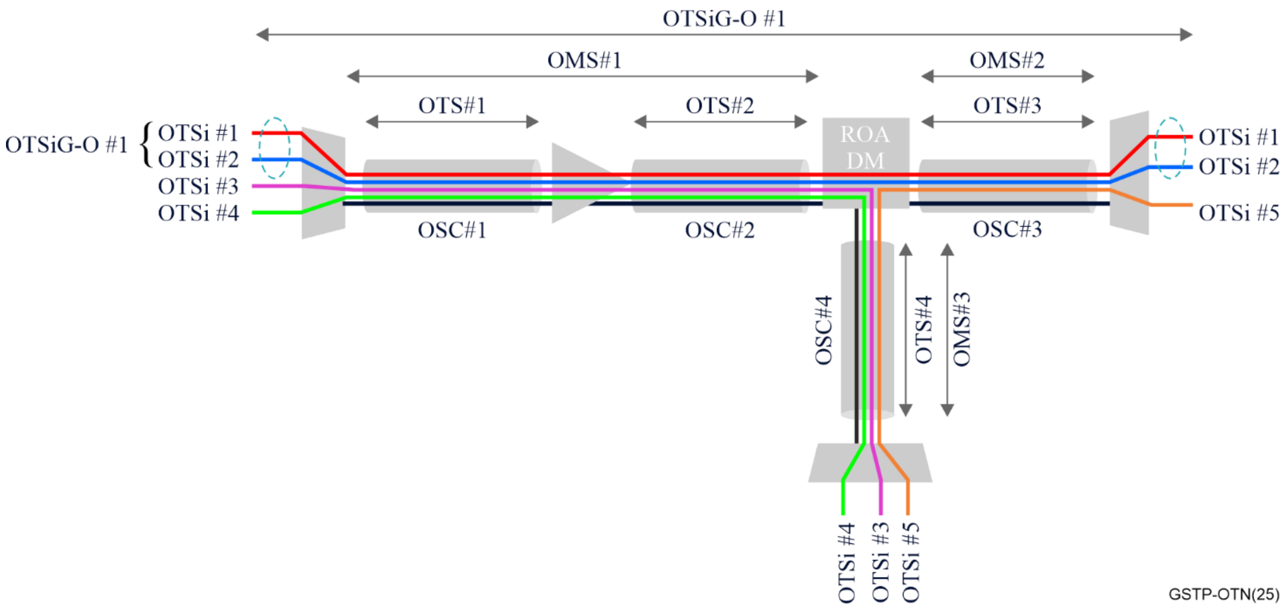


Figure 3-2 – Illustration of optical signal maintenance entities (OSME)

3.3 OTN OAM tools

OTN provides the following set of OAM capabilities:

Continuity supervision is the proactive monitoring of the integrity of a trail. In OTN, lack of continuity is detected via loss of signal (LOS), open connection indication (OCI), and loss of tandem connection (LTC) defects.

LOS indicates a defect in the fibre connectivity between two OTN ports or a failure of an optical transmitter. Due to the use of non-associated overhead to monitor the media layer, LOS is detected separately for the payload of an OTSi (dLOS-P) and for the non-associated overhead in the OSC (dLOS-O). dLOS-P is detected in the adaptation function between an OTSi or OTSiG and the lowest digital layer in the stack (OTUk or FlexO). dLOS-O is detected at the adaptation function between the OTSi that carries the OSC and the digital structure of the OSC.

OCI indicates that a connection function (switch fabric) upstream from a path endpoint is missing a connection between an output connection point and an input connection point. This defect is monitored for all path layers in the OTN. In the context of the media path layer, OCI may also be accompanied by LOS, and serves to further qualify the fault cause. In the context of the digital path layers, OCI functions like a 'digital loss of signal'. A path endpoint that is in service expects to receive a signal from some other path endpoint; OCI indicates that this is not happening.

LTC indicates a loss of signal for a tandem connection. This is caused by a misconfiguration in an upstream node (e.g., the other endpoint of the tandem connection may not be configured correctly, or there may be a misconnection between the endpoints of the tandem connection).

Connectivity supervision monitors that two endpoints that are connected (i.e., have continuity) are intended to be connected to each other; in other words, it detects misconnections in a layer network. In OTN, trail trace identifier (TTI) overhead is used to provide connectivity verification. Each endpoint is configured to transmit a particular TTI value and expects to receive a particular TTI value. If the actual received value does not match the expected value, a defect is detected, and consequent actions may include insertion of maintenance signals (to prevent delivery of misconnected information to a client layer) or protection switching.

Signal quality supervision monitors the quality of transmission and detects the signal degradation defect. Mechanisms for detecting signal degradation in OTN include use of uncorrected FEC errors or additional overhead to directly measure error performance of the OTN path. Signal degradation defects are typically used to trigger protection switching or control plane restoration in order to ensure that the required quality of service is delivered. Signal quality supervision also includes measurement of the latency of a path or tandem connection.

Payload mismatch supervision monitors that the payload inside the OTN container is what is expected. The OTN overhead includes a payload label information element that allows a sink node to determine if the client at the source node matches what the sink node expects to receive. If the mapper at one end of a path and the de-mapper at the other end of the path are configured differently, in most cases the de-mapped signal would be unrecognizable to the client layer, but in some cases, it could still be recognizable, so maintenance signals are typically inserted as a consequent action when a payload label mismatch is detected.

In the case where the payload of an OTN container is a multiplex of lower speed OTN containers, the OTN overhead also includes information elements to describe the multiplex structure, i.e., tributary slots and which signals occupy them; this overhead is monitored, to ensure that the expected tributary signals are present in the expected tributary slots (and only the expected tributary slots) and that no signals are present in tributary slots that are expected to be empty.

Alignment supervision monitors the signal for valid frame and multiframe alignment overhead to ensure that overhead and payload information are correctly processed. If frame or multiframe alignment is lost, maintenance signals are inserted.

Maintenance signal supervision monitors the presence of maintenance signals that may have been inserted by a lower layer or an upstream node. OTN maintenance signals in each layer of the network vary depending on whether the layer is a path or section layer. All layers generally include a forward defect indication (i.e., a maintenance signal inserted by a node toward downstream nodes due to a failure that was detected) and backward defect indication (i.e., a maintenance signal inserted by a path endpoint in response to detecting a defect or maintenance signal, to enable single-ended maintenance of the path), as well as a "lock" signal that indicates the entity is administratively out of service.

Digital path layers and the media layer include the previously discussed OCI maintenance signal.

The media layer also includes a payload missing indication (PMI) maintenance signal that is essentially equivalent to OCI and is used to suppress dLOS-P defects when there is no payload expected to be in an OTSi.

Protocol supervision refers to monitoring protocols that may be carried in overhead channels, such as an automatic protection switching protocol or a security protocol.

3.4 OTN overhead to support OAM

3.4.1 Client-specific OAM

Client-specific OAM provides information about how the client signal is mapped into the path and identifies the type of client signal. The information elements are:

- A Payload structure identifier (PSI) that consists of up to 256 bytes of information. This is carried over a single byte with a 256-frame multiframe. The first frame of the multiframe always includes a payload type indication (PT) that identifies the type of client and how it is mapped. This enables payload mismatch supervision. The value of the PT also provides the definition for the other 255 frames in the multi-frame as well as for the other 15 bytes of mapping-specific information.
  – In the case that the client is a multiplex of lower-speed path layers, some of those additional 255 frames will contain a multiplex structure identifier (MSI). The payload area in this case is divided into a set of tributary slots. The MSI indicates which tributary slots carry which client signals and which are unoccupied.
- 15 bytes of mapping-specific information.

– In the case of a multiplex of lower-speed path layers, these bytes are shared among all the clients via the multiframe.
– The information in these bytes depends on the type of client mapping, and serves to ensure that the client can be correctly de-mapped from the ODU path at the sink node.

3.4.2 Common aspects of digital layer OAM

There are several OAM information elements that are common to most of the digital path and section layers of the OTN:

- A trail trace identifier (TTI) to support connectivity supervision. The format of the TTI generally includes source and destination access point identifiers and operator-specific information.
- A BIP-8 error detection code to support quality of service monitoring. At the time OTN was originally defined, this code enabled estimation of the bit error ratio for the entity being monitored. With complex FEC codes used in modern OTN links, this code provides error detection, but not an accurate estimate of the BER.
- Status bits to support communication and detection of maintenance signals.
- Bits to support remote defect indication maintenance signals and remote error counts to facilitate single-ended management of a section.

3.4.3 Path layer OAM

ODUk provides the path layer for OTN. The path layer frame structure is the same as the section layer frame. The location of the path and client-specific OAM are shown in Figure 2-7.

The information elements of the ODU overhead are:

- End-to-end path monitoring (PM) using the common OAM elements.
- Six sublayers of tandem connection monitoring (TCM) using the common OAM elements. A tandem connection is a subset of an end-to-end path. Tandem connections can be nested or cascaded to support a variety of applications where it is beneficial to monitor a portion of an end-to-end path separately from the end-to-end path.
- Delay measurement overhead for the end-to-end path and all six TCM sublayers.
- Two general communications channels (GCC1 and GCC2) to support management/control connectivity. These channels may be used individually or combined to create a single higher-speed channel. These channels also support access at intermediate nodes along the path in addition to at the path endpoints.
- A protection channel is provided for automatic protection switching (APS) protocols and general communications related to protection (PCC). Protocol supervision is provided for the APS channel. The protection channel is multi-framed to provide support for protection of the end-to-end path, all six TCM sublayers, and a single section link.

Figure 3-3 illustrates the use of nested and cascaded tandem connections to support a service where the OTN path layer is carried by two operators.

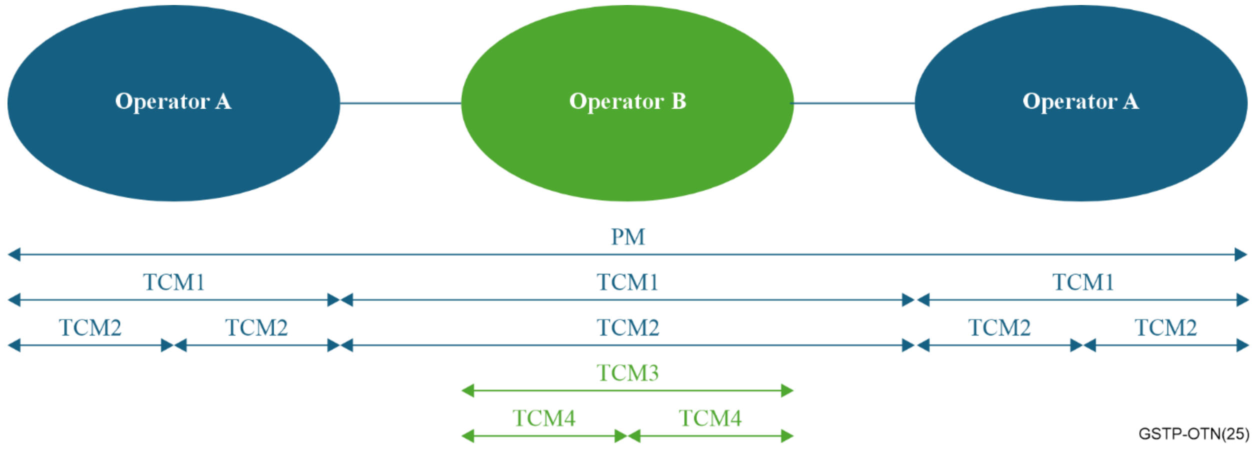


Figure 3-3 – Tandem connection monitoring example with four TCM levels

3.4.4 OTUk and OTUCn section layer OAM

OTUk provides the section layer for signals up to ~100 Gb/s. OTUCn provides equivalent capabilities for higher rate signals. The frame structure is shown in Figure 2-7.

The information elements are:

- Frame and multi-frame alignment signals provides alignment supervision for the section and path overheads (which use a common frame structure). The one-byte MFAS provides for up to a 256-frame multi-frame, although many fields use shorter multi-frame lengths based on using a subset of the MFAS bits
- Section monitoring (SM) using the common OAM elements
- A general communications channel (GCC0) to support management/control connectivity
- A synchronization management channel (OSMC) to support communication of synchronization information, including both synchronization status messages and PTP, to support synchronous operation of an OTN.

The OTUk definition included a FEC area because at the time OTN was originally defined, it was expected that all OTN links would use the same standard FEC. This never proved to be the case, with vendors offering "enhanced FEC" codes that provided substantially more coding gain that then standard, enabling much longer reach. In recognition of the value of supporting a variety of FEC codes, the OTUCn section does not include a FEC area. Rather, the FEC frame is part of the adaptation bottom of stack digital layer to the optical media. In most cases the OTUCn is carried over FlexO, and the FEC code is added to the FlexO frame structure. Proprietary implementations may also map OTUCn directly into a FEC frame for transmission over the media layer.

3.4.5 ODUCn section OAM

ODUCn is a second section layer for beyond-100G OTN signal. In regeneration applications, the ODUCn provides an end-to-end section monitoring capability, while the OTUCn supports monitoring each individual link. While the ODUCn is a section, it has some information elements that are similar to the path layer due to possibility that an ODUCn section spans multiple OTUCn sections.

The information elements are the same as for the path layer (even though the ODUCn is a section layer):

- End-to-end monitoring (somewhat misleadingly called PM) using the common OAM elements
- Six sublayers of tandem connection monitoring (TCM) using the common OAM elements. Tandem connections can be used in cases where there are multiple regenerators and it is desirable to monitor multiple OTUCn sections as a single entity.

- Two general communications channels (GCC1 and GCC2) to support management/control connectivity. These channels may be used individually or combined to create a single higher-speed channel. These channels also support access at intermediate nodes along the path as well as at the path endpoints.

3.4.6 FlexO section OAM

As described previously, FlexO was originally defined as an inverse multiplexing adaptation of OTUCn to the media layer. A frame structure and overhead for supporting the inverse multiplexing were defined. Since the adaptation to the media also must include the FEC for OTUCn, multiple FlexO FEC frames are also specified.

In the time since the original definition, direct mapping of Ethernet clients to FlexO in support of point-to-point connections has been added, and along with this, FlexO has evolved to be a full section layer, including overhead to support regenerator applications (so effectively FlexO can function like two section layers, similar to how ODUCn and OTUCn provide two section layers).

The location of the FlexO OAM is shown in Figure 2-18 and Figure 2-20.

The information elements in the FlexO OAM are:

- Group management overhead provides protocol supervision related to the inverse multiplexing that FlexO can provide. This includes an identifier for the group as a whole, an identifier for each instance in the group, and an indication of which instances are expected to be part of the group. Collectively these elements enable detection of misconfigurations and incorrect cabling or connectivity in the media layer.
- End-to-end section monitoring overhead provides maintenance signal and remote fault indications for the FlexO section. Unlike the other section layers, there is no error detection code for the payload (however, the FlexO section always uses FEC, which may have error marking capabilities). Also, unlike the other section layers, there is no trail trace identifier. This traces back to FlexO's origin as an inverse multiplexing adaptation process rather than a section layer; the OTUCn has the TTI to assist with connectivity supervision within the FlexO section.
- Regenerator section monitoring overhead provides maintenance signal and remote fault indications for the portion of the FlexO section between two regenerators or between an endpoint and a regenerator. This overhead does include a TTI. Use of the regenerator section is optional.
- Mapping-specific overhead provides information about the client mapping into the FlexO group. It contains one byte that provides a payload structure identifier (PSI) and 13 bytes of mapping-specific information.
- Two general communications channels (FCC0 and FCC`) to support management/control connectivity. FCC0 is for communication to regenerator nodes, while FCC1 is for end-to-end communication across the FlexO section layer. Use of the FCC channels is optional.
- A synchronization management channel (OSMC) to support communication of synchronization information, including both synchronization status messages and PTP, to support synchronous operation of an OTN. Use of the OSMC is optional.

3.4.7 Optical media layer OAM

As described in clause 3.2.2, overhead for the optical media layer is provided via a non-associated channel and consists of 3 digital layers that create a management hierarchy on top of the otherwise flat optical media layer. The information elements for the optical media layer are shown in Figure 3-4.

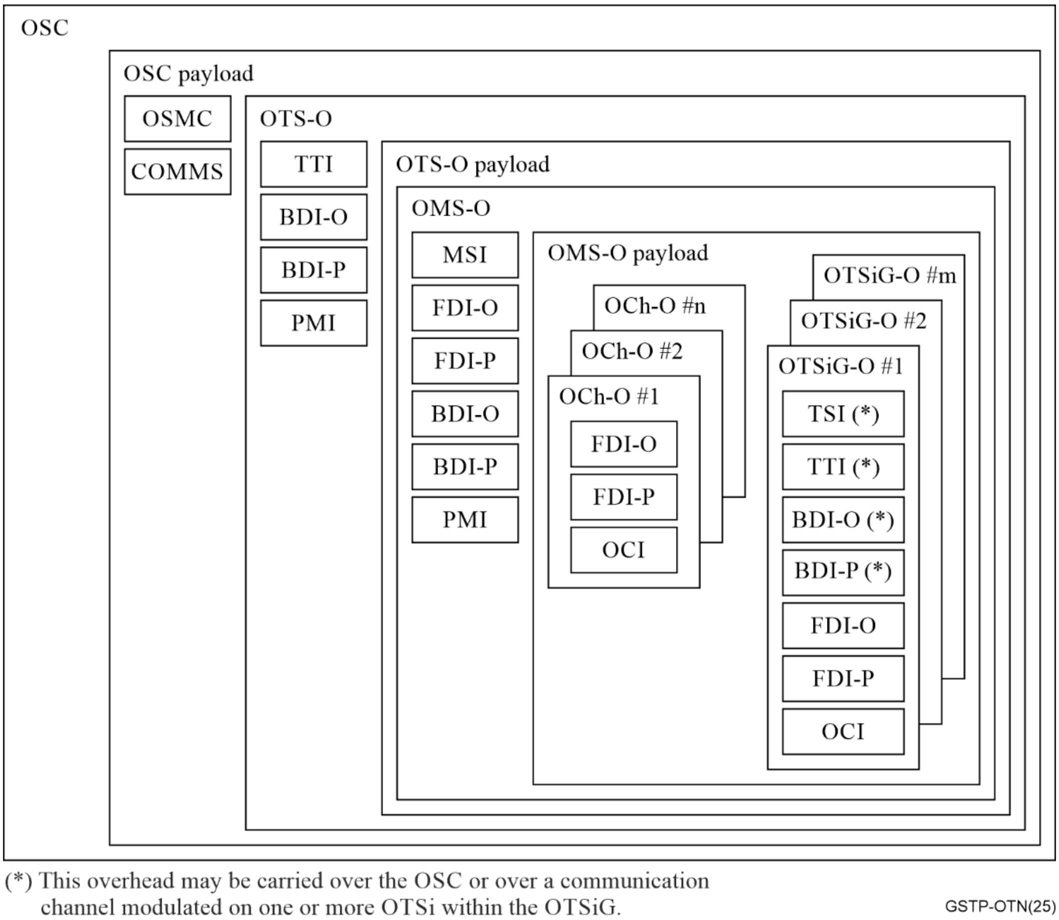


Figure 3-4 – Media layer overhead in the OSC

The OTS overhead (OTS-O) includes:

- A trail trace identifier (TTI) to support connectivity supervision for each individual link in the media layer. The format of this TTI is not specified.
- A payload missing indication (PMI) maintenance signal to support continuity supervision. A source inserts the PMI signal if the associated payload signal has not been transmitted by the source (e.g., due to a transmitter failure or a failure at the input of an optical amplifier).
- Backward defect indication maintenance signals for both the payload (BDI-P) and overhead (BDI-O) that are used to indicate that the OTS payload or overhead, respectively, was not received by the sink function that is co-located with the source that inserts these maintenance signals (i.e., the same function as RDI in a digital layer network). The "payload" in this case refers to the aggregate optical signal carried by the other wavelengths in the fibre (i.e., all the wavelengths other than the OSC). The "overhead" refers to the OTS-O signal within the OSC.

The OMS overhead includes:

- A multiplex structure identifier (MSI) that is used in flexible grid applications to encode the OTSiG multiplex structure, occupied frequency slots, and media channel structure. This overhead enables detection of configuration mismatches between the source and sink nodes.
- Forward defect indication maintenance signals for both the payload (FDI-P) and overhead (FDI-O) that are inserted by intermediate nodes in an OMS to indicate that a failure was detected (i.e., the same function as AIS in a digital network). The OMS termination uses this information to qualify other defects it may detect in the signal it receives. The "payload" in this context is the aggregated set of signals at the output of an optical multiplexer or input of an optical demultiplexer. The "overhead" refers to the OMS-O signal within the OSC.

- BDI-P and BDI-O signals that are used to indicate that the OMS payload or overhead, respectively, were not received.
- A payload missing indication (PMI) maintenance signal that is used to indicate if the associated payload signal has not been transmitted (e.g., due to a transmitter failure or a failure at the input of an optical multiplexer).

The OCh overhead consists of:

- Forward defect indication maintenance signals for both the payload (FDI-P) and overhead (FDI-O) that are inserted by intermediate nodes in an OCh to indicate that a failure was detected (i.e., the same function as AIS in a digital network). The OCh termination uses this information to qualify other defects it may detect in the signal it receives. The "payload" in this context is the OTSi from an optical transmitter. The "overhead" refers to the OCh-O signal within the OSC.
- BDI-P and BDI-O signals that are used to indicate that the OCh payload or overhead, respectively, was not received.
- An open connection indication (OCI) is transmitted when there is no (optical) matrix connection for a particular media channel at an output of the optical matrix. In this configuration, the OCh sink node would detect LOS for the OCh in question. The OCI provides additional information that the media channel is intentionally not occupied.

The OTSiG overhead adds two additional overhead elements to what the OCh overhead provides:

- A trail trace identifier (TTI) to support end-to-end connectivity supervision across the optical media layer between a transmitter and a receiver (without relying on the digital layer being carried to provide this).
- A transmitter structure identifier (TSI) that indicates the type of signal being carried by the OTSi. This information enables monitoring of the consistency of the configuration at the source and sink ends of the OTSiG.

In addition, the OSC may contain a COMMS channel supporting management/control communication and/or an optical synchronization management channel (OSMC) for carrying synchronization status messages for frequency synchronization and/or PTP messages for phase synchronization applications.

4 Protection for survivability

4.1 Introduction

Protection for survivability refers to the ability of the network to maintain an acceptable level of service during a network or equipment failure or traffic signal degradation.

ITU-T has developed a number of Recommendations for generic and technology-specific protection schemes. This clause discusses these Recommendations and their applications.

Generally, there are two types of network survivability techniques:

Network Protection is the replacement of a failed or degraded working resource with a pre-assigned standby resource. Protection mechanisms tend to be deterministic in nature. Generally, the protection action is completed in tens of milliseconds. Protection mechanisms are autonomous and, once configured, operate without intervention from the control or management planes. Protection schemes used in OTN are defined in the following ITU-T Recommendations:

– G.808.1 "Generic protection switching – Linear trail and sub-network protection"
– G.808.2 "Generic protection switching – Ring protection"
– G.873.1 "Optical Transport Networks (OTN) – Linear protection"

– G.873.2 "ODUk shared ring protection"

Network Restoration is the replacement of a failed or degraded working resource by re-routing using dynamically allocated spare resources. Restoration is non-deterministic and the network behaviour in failure conditions is less predictable. The restoration action may take seconds to complete depending on the amount of traffic being restored. Restoration requires control plane actions. There are a number of ITU Recommendations defining control plane architectures, signalling and routing. Network restoration techniques and the related Recommendations are described in more detail in clause 8.

4.2 Network objectives

The network objectives that are considered important and should be met for any protection switching schemes are listed below.

Switching time is the time taken to initiate and complete the protection switching event. It excludes fault detection time and any hold-off time. A value of 50 ms has been standardized for cases where there are no pre-existing failures and the transmission delay meets certain constraints.

Transmission delay is determined by the physical length of the fibre (propagation delay) and the processing time of the protection protocol along the path. Generally, the propagation delay on 1000 km of fibre is approximately 5 ms and processing delay is in the range of hundreds of microseconds.

Hold-off time. In nested protection switching schemes, it is beneficial to impose a hold-off time on the outer protection mechanism to allow the inner protection mechanism time to restore the traffic before the outer protection is activated. OTN protection schemes support a hold-off timer that is configurable from 0 to 10 seconds in steps of 100 ms.

Switch initiation. Transmission and connectivity defects are the criteria for initiating protection switching. Transmission defects refer to the error performance of the path. A threshold is configured (this could be expressed as a bit error ratio or some other measure of performance based on the type of errors that are expected to occur). When the threshold is exceeded, a protection switch is triggered. Connectivity defects refer to whether the correct endpoints are connected across the network; this is monitored via overhead such as the trail trace identifier. A persistent connectivity defect triggers a protection switch.

4.3 Protection switching classes

Generically. protection switching can be performed on subnetwork connections or on trails. Within the OTN, only subnetwork connection protection is used.

4.3.1 Sub-network connection protection (SNC)

SNC protection is used to protect a segment of a trail within a network or multiple networks. The sub-network can be between two connection points (CP), between CP and termination connection Point (TCP), or the entire end-to-end network connection between two TCPs. SNC is the most common type of protection.

SNC protection is further characterized by the type of monitoring used to control the protection switch.

SNC/I – inherent monitoring. In this variant, signal fail (SF) and signal degrade (SD) conditions are determined by the server layer termination and adaptation functions (Figure 4-1). Defects within the layer being protected do not cause protection switches; applying SNC/I to a path layer means connectivity defects within that path layer will not cause protection switches. Because the switching is based on server layer defects, an SNC/I scheme can only protect a single link connection within the end-to-end trail.

SNC/Ne – non-intrusive monitoring of end-to-end overhead. This requires the use of non-intrusive monitoring functions observing the trail overhead to determine the SF/SD conditions for invoking protection switching (Figure 4-2). SNC/Ne can be operated across an entire network connection or over smaller subnetwork connections within the network connection. Because the protection switch is triggered based on non-intrusive monitoring of the trail overhead, the use of concatenated SNC/Ne protection domains may result in cascading switches in each domain.

SNC/Ns – non-intrusive monitoring of sublayer overhead. This requires the use of non-intrusive monitoring functions observing sublayer (i.e., tandem connection) overhead to determine the SF/SD conditions for invoking protection switching (Figure 4-2). SNC/Ns can be operated across an entire sublayer connection or a smaller portion of a sublayer connection. Because the protection switch is triggered based on non-intrusive monitoring of the sublayer overhead, the use of concatenated SNC/Ns protection domains may result in cascading switches in each domain.

SNC/S – sublayer monitoring. Tandem connection/segment sublayer functions are deployed to determine the SF/SD condition. The sublayer overhead/OAM is used. SNC/S protection supports network architectures which make use of concatenated protected subnetworks without interaction between the concatenated protection domains. Such network architectures are able to restore traffic for the case of multiple faults without cascaded switching (Figure 4-3).

In all SNC schemes, concatenation of multiple SNC domains results in protection against multiple faults (one per protected subnetwork), but each interconnect node remains a single point of failure. However, most network nodes have equipment redundancy, most offices have backup power available, and in any case node failures are less common than link failures, so the "single point of failure" represented by an interconnect node is generally quite reliable. In the example in Figure 4-3, if the interconnect nodes (in the middle of the figure) are physically separated nodes, the link between them would also be a protection domain.

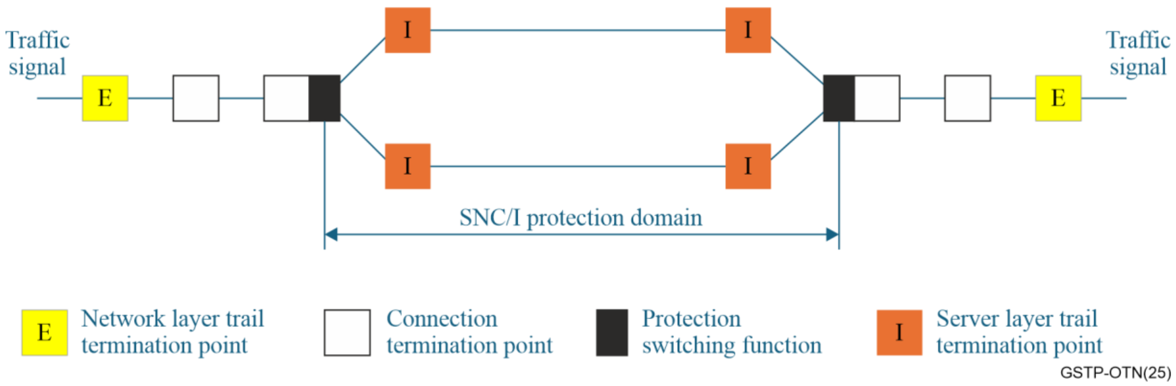


Figure 4-1 – SNC/I protection

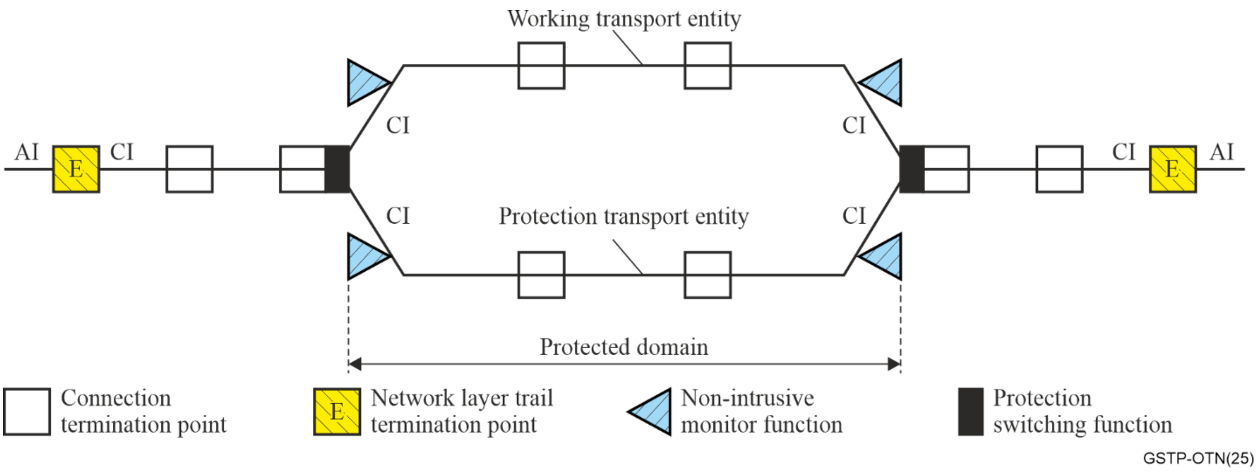


Figure 4-2 – 1 + 1 SNC/N protection

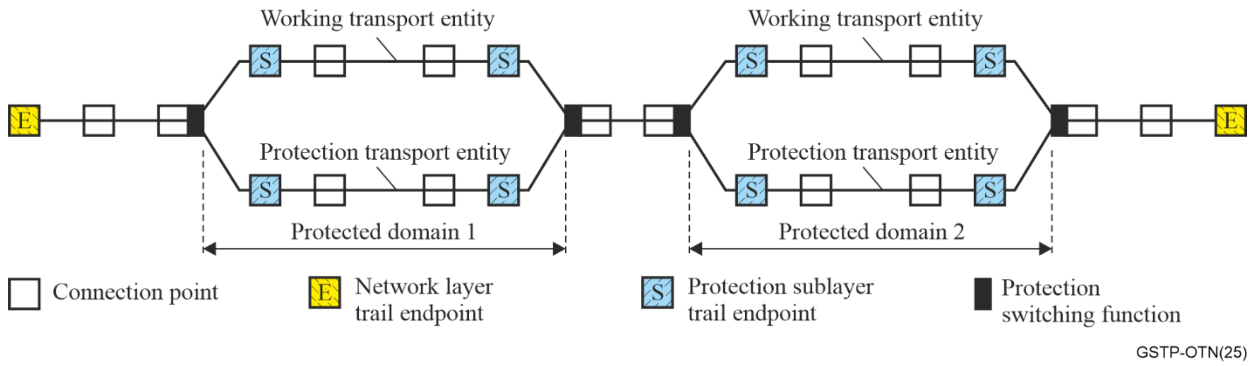


Figure 4-3 – Cascaded SNC/S protection

4.3.2 Trail protection

Trail protection replaces a working trail with a protection trail in the event of a failure, as illustrated in Figure 4-4. It is essentially equivalent to SNC/Ne protection when the endpoints of the SNC/Ne domain are the trail termination points. Trail protection differs from SNC/Ne in that the working and protect paths have separate trail termination functions (which increases cost relative to SNC/Ne protection). As such, there is no monitoring of the performance of the 'combined' working and protect paths other than at the client layer. Because of those drawbacks, trail protection is not standardized for OTN.

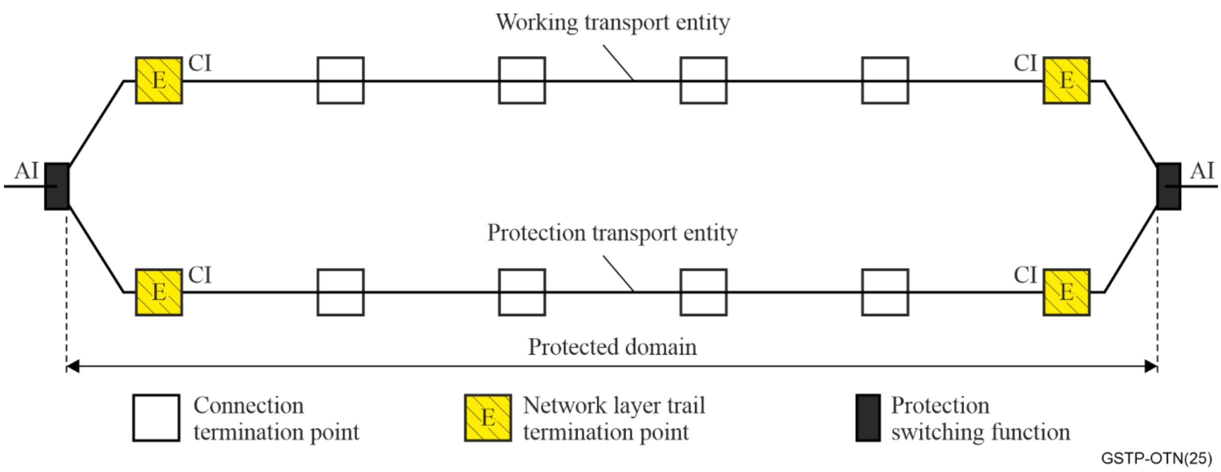


Figure 4-4 – Generic concept of trail protection

4.4 Protection switching architectures

OTN protection includes both linear and ring architectures.

Each type has its own advantages and disadvantages as detailed below.

4.4.1 1+1 linear protection architecture

In this protection scheme, a protection channel is dedicated as a backup facility to the working channel. The normal traffic signal is permanently bridged onto the protection channel at the source endpoint of the protection domain, as shown in Figure 4-5. At the sink endpoint, a selector chooses between the working and protection signals.

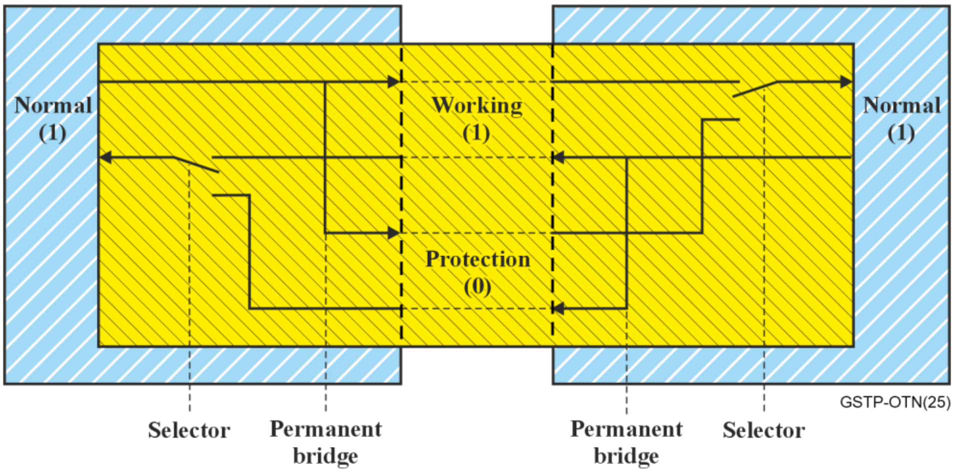


Figure 4-5 – 1+1 Protection architecture

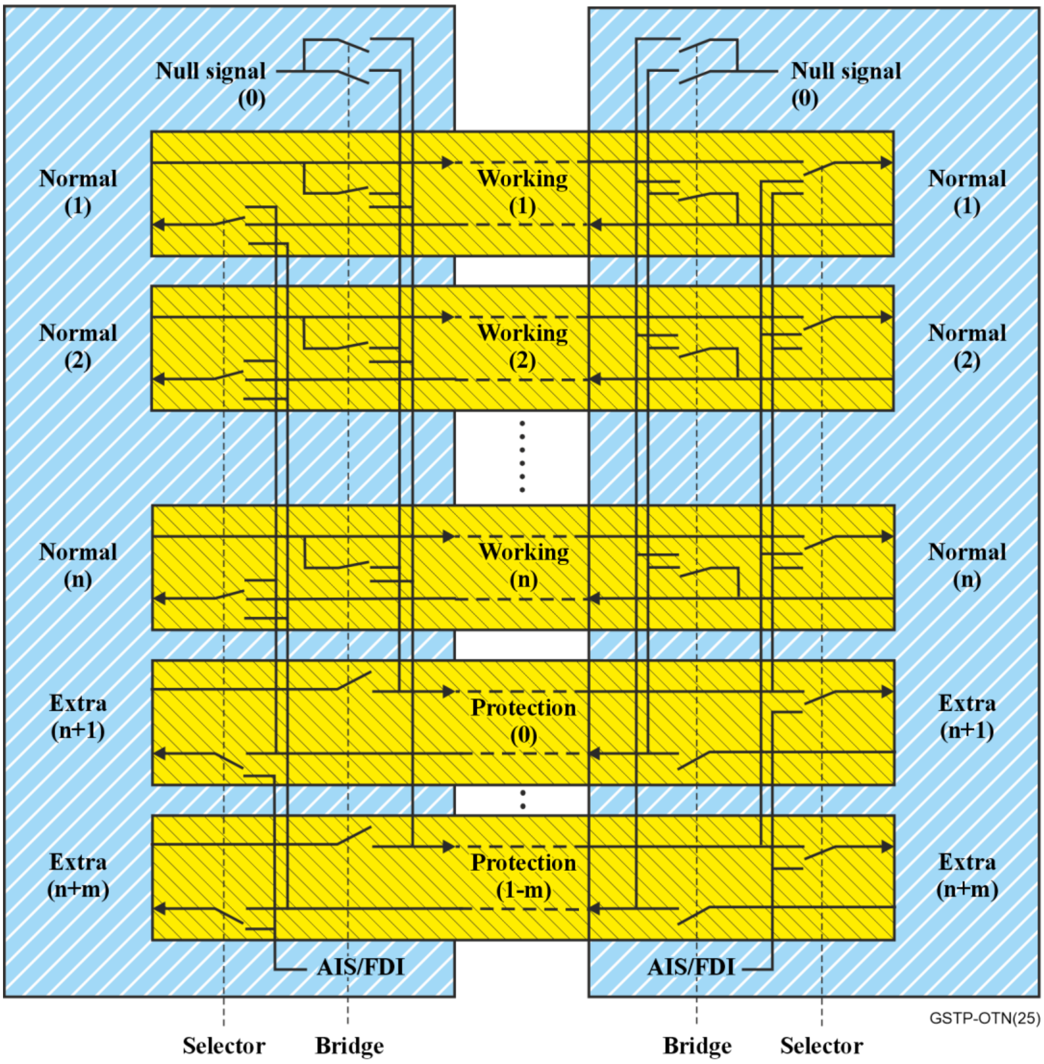
The advantage of this scheme is its simplicity. The main disadvantage of this type of protection is the 100% extra capacity needed in the network.

4.4.2 *m*:*n* linear protection architecture

This is an alternative linear protection architecture where m protection channels are shared by n working channels, with $m \leq n$, as shown in Figure 4-6. The bandwidth of each of the protection channels must be adequate to protect any of the n working channels. When a defect is detected on a working channel, a protocol is used to signal to the source end, requesting that the impaired channel be bridged to an available protection channel. Once the bridge is in place, the sink end selector is operated to choose traffic from the protection channel.

The most common implementation of this type of protection has $m = 1$ (i.e., one protection channel for n working channels).

The main advantages of this architecture are that the amount of protection bandwidth in the network is less than 100% of the working bandwidth, and that the protection channels can be used to transport extra (unprotected) traffic when they are not used to protect working traffic. When there is neither extra traffic nor working traffic using a protection channel, a null signal (either the ODUk Open Connection Indication maintenance signal or a test signal) is transmitted. The disadvantage is that the mechanism is more complicated than 1+1.



Broadcast bridge option: Normally permanently connected to working and occasionally to protection.

Figure 4-6 – m:n Protection architecture

4.4.3 Shared ring protection architecture

Shared ring protection schemes divide the bandwidth of a physical ring (which may consist of one or two fibre pairs) into three groups of channels: protected traffic, non-preemptible unprotected traffic, and a protection channel. The protection channel is shared by all protected traffic on the ring, and must be large enough to protect all of the protected traffic. Similar to *m*:*n* linear protection, a protocol is required to coordinate the use of the protection channel. Switching to the protection bandwidth takes two forms. All rings can support *wrapping*, where traffic that would traverse a failed span is "wrapped" around the ring in the opposite direction over the protection channel at the nodes that are adjacent to a fault. Rings that have two pairs of fibres can also support *steering*, where traffic is redirected around a failure at the point where it enters and exits the ring.

Wrapping protection is illustrated in Figure 4-7, while steering is illustrated in Figure 4-8.

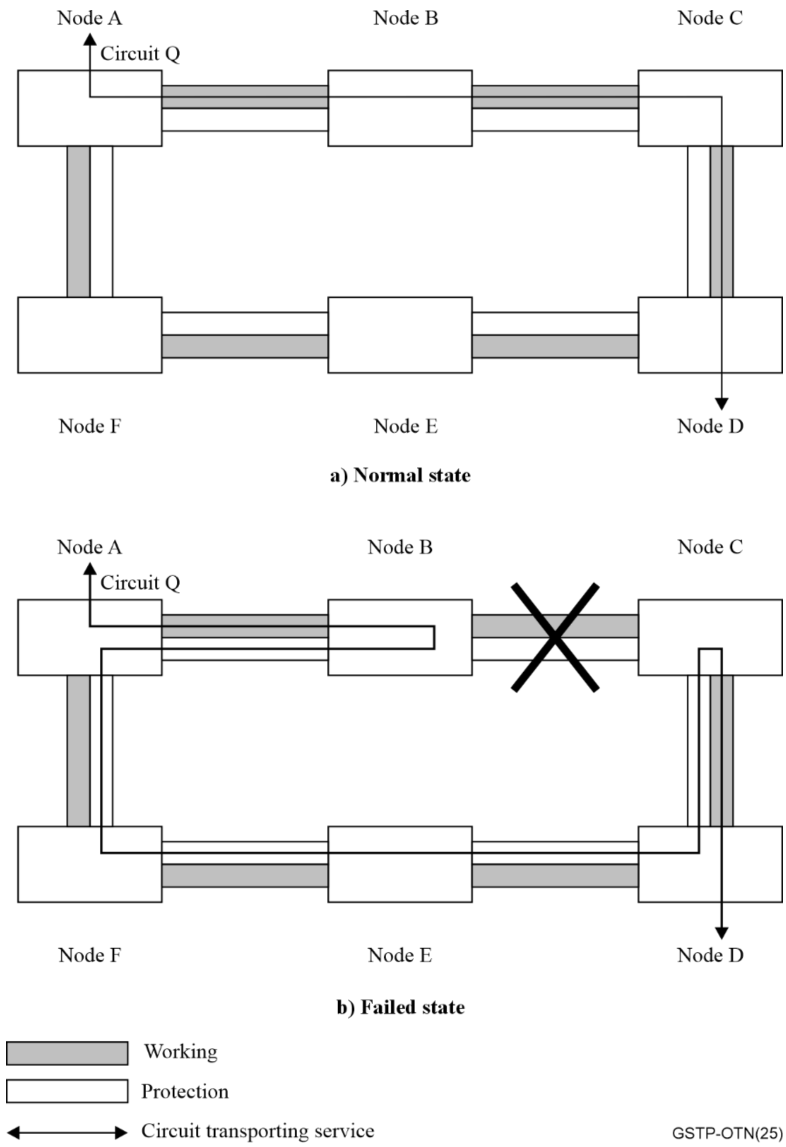


Figure 4-7 – Wrapping ring protection

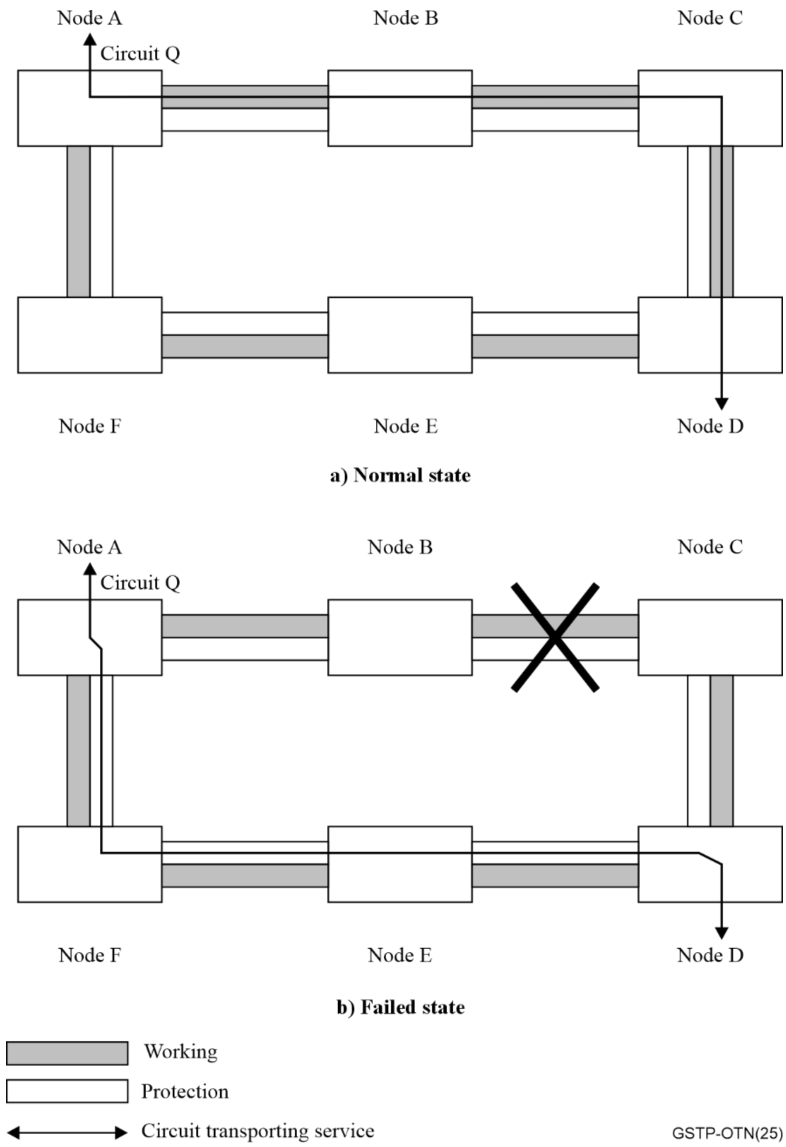


Figure 4-8 – Steering ring protection

Note that the linear protection schemes discussed previously can also be used with network topologies that are physical rings. 1+1 SNC protection is commonly used in such networks, where the working channel takes one direction around the ring and the protection channel takes the other direction around the ring.

4.5 Protection switching parameters

4.5.1 Directionality

Protection switching may be unidirectional or bidirectional.

Unidirectional switching. In unidirectional switching, the normal traffic signal is switched to the protection channel only at the endpoint that detects a fault. For 1+1 linear architecture, the sink end selector is activated only and no communication with source end is required. For other protection switching architectures, the operation of the sink end selector and the source end bridge need to be coordinated via a communication channel between the two ends of the protected domain, which is referred to as the automatic protection switching (APS) channel. Unidirectional protection switching is a simple scheme to implement and under multiple failure conditions there is a greater chance of restoring traffic faster than bidirectional switching.

Bidirectional switching. In this type of switching, the normal traffic signal is switched from the working channel to the protection channel at both ends of the protection span. In all architectures, an APS channel is needed. For 1+1 linear architecture, the selectors at sink and source ends may be operated independently (if both ends detect a defect), or the APS protocol may be needed to coordinate the operation at one end (if only one end detects a defect). For other protection switching architectures, the operation of the selectors and bridges at both sink and source ends must be coordinated via the APS channel.

Bidirectional protection switching ensures that both directions of traffic take the same path through the network (and thus experience similar latency) and may result in a minimum number of severely errored seconds (SES) during repair and maintenance of the working path due to a smaller number of switching events.

4.5.2 Revertive operation

Protection switching may operate in a revertive or non-revertive manner.

Revertive operation. When the defect or the external command are cleared on the working channel the normal traffic signal returns to the working channel that was in use prior to these events. Revertive operation is likely to be favoured because of the following reasons:

– in any architecture other than 1+1 linear, the protection channel is shared, and thus may be needed to protect against other faults, or may be used to carry extra traffic in the absence of faults;
– protection channel may have lower performance (e.g., it may have larger latency);
– to simplify network management.

However, revertive operation does cause a second traffic disruption due to the switch back to the working channel. To prevent frequent operation of the protection switch due to intermittent defects, such as bit error ratio fluctuation, a fixed period of time is allowed to elapse before a normal traffic signal can use a recovered working channel, where the signal fail (SF) and signal degrade (SD) have been cleared. This period is called a wait-to-restore period and it is generally configurable from 5-12 minutes.

Non-revertive. The normal traffic signal stays on the protection channel and does not return to the working channel after the defect or external commands are cleared. The advantage of non-revertive operation is that, in general, it will have less impact on traffic performance. In the case of 1+1 architecture, the second traffic disruption that occurs in revertive operation can be completely avoided. In other architectures, the switch back to the working channel can be delayed to a maintenance window (as long as no other faults occur that require use of the protection channel prior to that maintenance window).

4.5.3 Hold-off timer

In some applications, it can be beneficial to delay protection switching when a defect is detected. The operation of a hold-off timer is such that the detection of a defect that would trigger a protection switch instead triggers the hold-off timer. When the timer expires, the current defect status is passed to the protection mechanism. A hold-off timer is typically configurable over a range of several seconds, in 100 ms increments.

One example application for hold-off timers is a network that has protection mechanisms at multiple layers. In such a network, a single failure may trigger multiple recovery mechanisms, which may interact with each other, resulting in undesirable and non-deterministic network states. When configured properly, a hold-off timer allows the innermost (lowest layer) protection switching to operate first and restore traffic before the outermost protection switching is activated.

A second example application is 1+1 configurations, where hold-off timers can be used to prevent early and ineffective switching due to the differential delay between the long and short path in cases where both paths have failed. Applying a hold-off timer to the shorter path that is at least as long as the differential delay between the two paths ensures that the correct status for both short and long path is known at the time the protection logic is evaluated.

4.6 Protection switching trigger criteria

[b-ITU-T G.808.1] defines generic signal fail (SF) and signal degrade (SD) conditions that are associated with a layered network. The signal degrade condition is typically determined based on the performance of the path layer (e.g., it may be triggered by an increased bit error ratio or other performance measurement). The signal fail condition is typically the result of defects detected via the layer network overhead. For OTN, such defects include alarm indication signal (AIS), trace identifier mismatch (TIM), and lower layer failures like loss of signal or loss of frame. The details regarding SD and SF can be found in [b-ITU-T G.798].

4.7 Protection switching temporal model

Figure 4-9 illustrates the temporal model for protection switching.

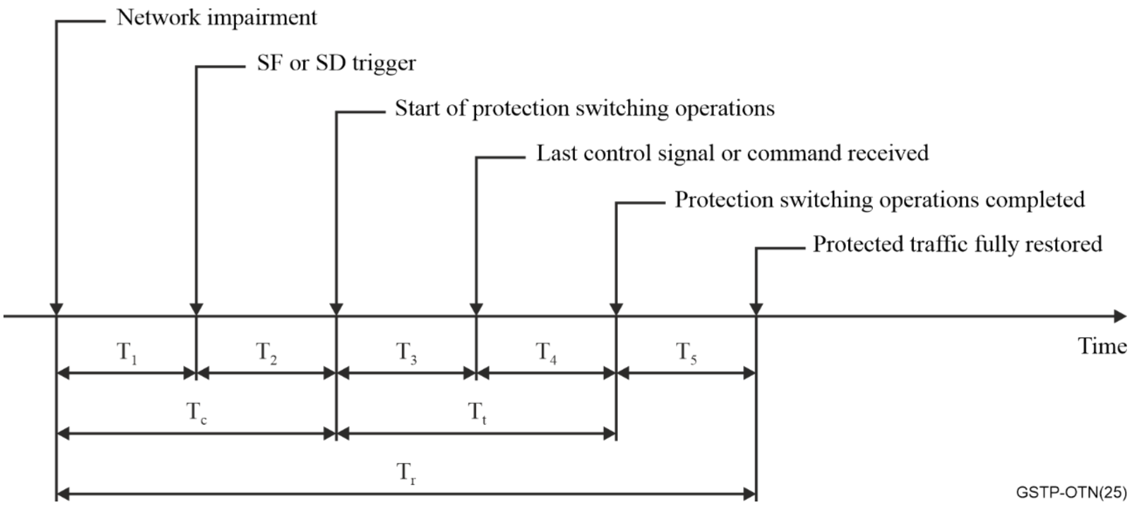


Figure 4-9 – Protection switching temporal model

Detection Time T_1. The time interval from the onset of the network impairment to its detection.

Hold-off time T_2. The time taken to confirm that the defect condition(s) require protection switching.

Switching operation time T_3. The time taken to complete the processing and transmission of the APS messages to invoke protection switching.

Switching transfer time T_4. The time needed to operate the selector(s) to complete the protection switching operation

Recovery time T_5. The time needed after completion of the protection switching operation to fully restore the protected traffic.

The protection switching time (T_3+T_4) is standardized to be less than 50 ms. Furthermore, restoration is the total time taken to restore protected traffic and is the summation of all the time intervals above (i.e., $T_1+T_2+T_3+T_4+T_5$).

4.8 Automatic protection switching (APS) protocol

4.8.1 APS protocol types

All protection switching schemes other than 1+1 unidirectional linear protection require an APS protocol to coordinate the bridge and selector actions at both ends of the protection domain. Different APS protocols are required depending on the type of protection.

For linear protection, the protocol may use between 1 to 3 phases (meaning protocol exchanges from one end to the other end). Additional phases ensure that the two endpoints agree on the highest-priority request is before operating the bridges and selectors, which in the case of shared protection helps prevent misconnected traffic. 1+1 architectures have a permanent bridge, so fewer phases can be used. *m*:*n* or shared ring protection requires at least 2 phases.

4.8.2 APS protocol information elements

The APS protocol contains the following information elements:

– Request/state type: identifies the highest priority fault conditions, external commands or the state of the protection process; states can include automatic switch (due to a fault), manual or forced switch requested by management/control, lockout of protection (indicating that no traffic is allowed to use the protection channel), or other values based on the configuration of the protection group (e.g., do not revert, in the case of a non-revertive configuration, or wait to restore in the case of a revertive configuration).

– Requested signal: identifies the signal that the endpoint wants to be bridged to the protection channel (which can be a null signal, a normal/protected signal, or the extra traffic signal, depending on the protection architecture and configuration).

– Bridged signal: identifies the signal that is currently bridged to the protection channel.

– Protection configuration: indicates the protection architecture, switching and operation type.

4.8.3 External commands

In addition to the automatic switching requests that are communicated by the APS protocol, there are external requests that can be invoked via management/control. Generally, only one external command is issued per protection group. External commands can be pre-empted or denied by higher priority failure conditions, requests or states. These commands are useful to do the following:

– to change equipment configuration for maintenance purposes;

– to disable access to protection;

– to test the protection protocol;

– to freeze the current state of the protection process to prevent further action;

– to clear previous external switch commands.

4.9 Multi-layer survivability

Current networks support a variety of technologies such as OTN, Ethernet, and IP, resulting in multi-layer transport networks with a variety of layer nesting depending on technology deployment and network evolution. Each technology layer may support a variety of protection and restoration schemes. Therefore, multi-layer survivability uses two or more nested protection mechanisms while single layer survivability employs a single end-to-end or cascaded protection scheme.

A good survivability strategy in a multi-layer network is to select the survivability schemes associated with these technologies that offer optimal performance to deliver the desired quality of service more cost-effectively than can be achieved in a single layer network.

Multi-layer survivability can effectively combine the merits of protection mechanisms of its constituent's layers such as the transport and service layers protection schemes. However, it is necessary to ensure that a single failure cannot trigger multiple protection mechanisms, which may interact with each other resulting in undesirable and unpredictable network states. A good approach to the design of a multi-layer network survivability strategy is to identify the usefulness of nesting protection mechanisms and to identify cases where interworking should be avoided.

5 Synchronization

5.1 Introduction

One of the main initial needs for synchronization in telecommunication networks was to allow clocks in transmission and switching equipment to operate at equal or almost equal frequencies in order to transport signals between them that carry digital information and to do so without introducing single bit errors or bursts of errors. Synchronous operation of equipment that is spread out over a large geographic area required a distribution network for synchronization information.

The effects which cause degradation of this synchronization information can be divided into two categories. First, there is a continuous degradation of these signals due to the accumulation of phase noise, caused by imperfect components and design. This causes jitter and wander on the digital signals that are transported over the network. Excessive level of jitter and wander can cause bit errors, loss of frame, or controlled slips. Second, there may occasionally be a complete failure of a synchronization link, leaving network elements or entire network parts without synchronization information.

The design of the synchronization network tries to minimize the effects both of the continuous phase noise accumulation and of incidental loss of synchronization.

With the deployment of mobile networks, and other applications that themselves require both frequency and time synchronization, distribution of accurate and resilient synchronization became increasingly an important requirement for the transport network. The End application can in this way get synchronization from the transport network, possibly combined with a local timing reference (typically based on GNSS technology).

This clause describes the synchronization aspects in transport networks based on various technologies. In clause 5.2, the aspects that are common to all types of networks are outlined. In clause 5.3, the specific aspects related to OTN networks are presented. Some of the principles adopted in OTN networks are based on solutions developed for the distribution of timing in Ethernet networks over the physical layer (synchronous ethernet) and the packet layer (in this case the timing is delivered via specific timing packet protocols such as PTP, precision time protocol). Therefore, some of the general principles related to the transport of timing over Ethernet networks are also presented in clause 5.2.

The ITU-T has dedicated many Recommendations to the synchronization aspects of the transport networks, as shown in Figure 5-1. This clause is mainly based on the content of these Recommendations.

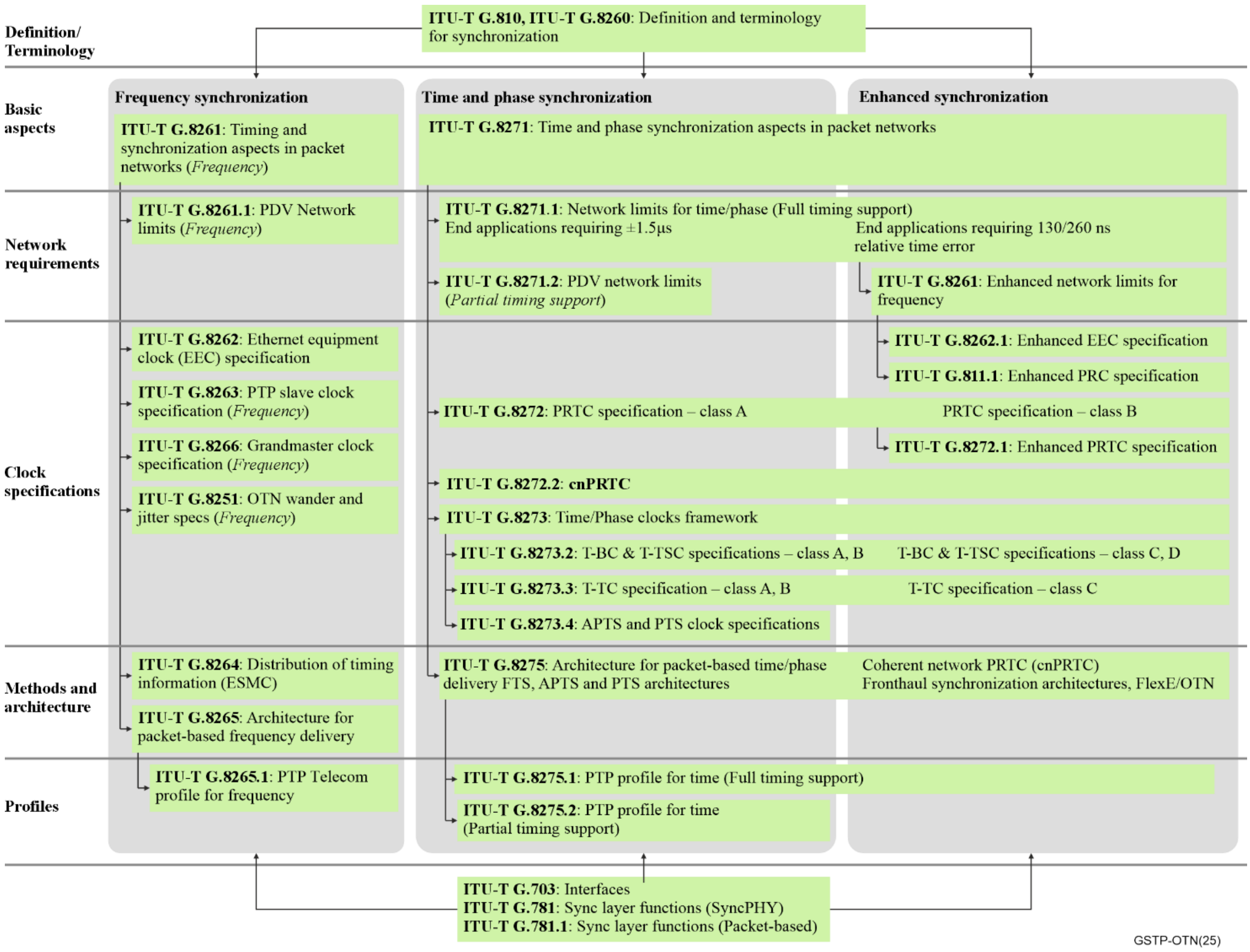


Figure 5-1 – ITU-T Recommendations on synchronization

5.2 Common Aspects

5.2.1 History

The early history of the specification of network synchronization, is closely related to the history of the digital transmission. It is only when transmission became digital that the timing of the symbols on the line became important. The "modern" synchronization studies have started with the definition of PDH (plesiochronous digital hierarchy) and PSTN (public switch telephone networks); before that, transport networks were mainly concerned with the frequency accuracy of analogue carriers.

The first digital hierarchy was PDH, which defined signal at different rates, 2, 34 and 140 Mbit/s for the European hierarchy and 1.5, 6.3 and 44.7 Mbit/s for the North American hierarchy. All these rates were defined with a certain frequency accuracy in the ITU-T G.7xx series of Recommendations. The reference frequency was generated by a clock defined in [b-ITU-T G.811]. The slave clocks widespread all over the PDH networks are specified in [b-ITU-T G.812]. The PDH hierarchy has introduced the concepts of digital multiplexing and frequency justification; these techniques raised problems such as frequency accuracy, jitter, wander and slips.

The next step in the development of the digital transport networks was the SDH network with a multiplexing system which needed a complete network synchronization system in order to avoid accumulation of large amounts of jitter and wander in its payload signals.

A further step was the introduction of OTN. This network can operate asynchronously and still transport synchronous SDH payloads, so the synchronization aspects have an impact also on the OTN specification.

Finally, the introduction of packet networks such as Ethernet, raised new concerns in the distribution of timing over the transport network due to the traditionally asynchronous nature of Ethernet networks.

Methods have been defined to distribute timing (both frequency synchronization and time synchronization) over Ethernet. Finally, the OTN technology was also further enhanced to optionally allow the distribution of accurate timing over a synchronous OTN layer.

5.2.2 Terminology

The definition of the terms used in this clause can be found in [b-ITU-T G.810]. In particular [b-ITU-T G.810] defines terms related to the transport of timing through telecommunication networks, synchronization networks, clock equipment, clock modes of operation and clock characterization. All these terms are used in the ITU-T Recommendations dealing with synchronization.

The meaning of some parameters related to the synchronization networks and based on [b-ITU-T G.810] are given in the following.

Bit errors may occur at points of signal regeneration, also as a result of timing signals being displaced from their optimum positions in time. Synchronization is robust against bit errors. A very high error rate is required to degrade the quality of timing; in general timing is correctly recovered until a loss of signal is declared following an excessive error ratio (e.g., worse than 1×10^{-3}).

A slip is the repetition or deletion of a block of bits in a synchronous or plesiochronous bit stream due to a discrepancy in the read and write rates at a buffer.

Slips arise as a result of the inability of an equipment buffer (and/or other mechanisms) to accommodate differences between the phases and/or frequencies of the incoming and outgoing signals in cases where the timing of the outgoing signal is not derived from that of the incoming signal. Slips may be controlled or uncontrolled depending on the slip control strategy. [b-ITU-T G.822] specifies the number of slips tolerated by the different portions of transport networks (international, national and regional).

It should be noted that it is impossible to eliminate slips when there is a frequency difference between the incoming and outgoing timing signals. Slips accumulate along networks without any "compensation". Slips may occur in various synchronous equipment. Given the specified levels of phase variation, slip occurrences may be minimised in synchronous equipment by appropriate choice of buffer capacity as well as rigorous specification of clock performance. The introduction of pointer processors in SDH avoids such impairment in transport networks.

Propagation delay does not affect the transport of synchronization timing; this might be a problem only if the data and the timing are transported on different media, or in case of transport of time and phase information when the delay differs in the transmitting and receiving direction. The reason is that transport of time synchronization relies on a two-way exchange of timing messages that are used to estimate the path delay based on half the round-trip time (i.e., assuming a symmetric link). Otherwise, an asymmetric link would result in time errors.

Jitter is the short-term variations of the significant instants of a timing signal from their ideal positions in time (where short-term implies that these variations are of frequency greater than or equal to 10 Hz).

Wander is the long-term variations of the significant instants of a digital signal from their ideal position in time (where long-term implies that these variations are of frequency less than 10 Hz).

Jitter and wander are the main sources of problems in transmission. Excessive levels of jitter and wander are a source of degradation and may result in bit errors or even loss of signal. This is very important for all transport networks: PDH, SDH and OTN. [b-ITU-T G.823] and [b-ITU-T G.824] specify the jitter and wander requirements of PDH signals, [b-ITU-T G.825] specifies the jitter and wander requirements of STM-N signals in SDH networks and [b-ITU-T G.8251] specifies jitter and wander requirements in OTN.

5.2.3 Metrics specifications

[b-ITU-T G.810] also contains the definition of metrics used to characterize the quality of timing signals: peak-to-peak jitter, root-mean-square (RMS) jitter, MTIE, MRTIE, TDEV and TVAR.

Peak-to-peak jitter over a specified time interval is the difference between the maximum jitter value over that time value and the minimum jitter value over that time interval. The RMS jitter is the RMS of the jitter statistical distribution.

Maximum time interval error (MTIE) and maximum relative time interval error (MRTIE) characterize the delay caused by buffers and help to define the right size of buffers. These metrics were defined since PDH, but they are very important also for the characterization of SDH networks where the pointer processor buffers have a huge effect on the quality of timing.

Time deviation (TDEV) and time variance (TVAR) provide a statistical analysis of the quality of clocks, which is very useful in SDH to characterize the noise accumulated along long chains of clocks.

Appendix II of [b-ITU-T G.810] provides important material related to metrics.

The transport of timing through packet networks and the transport of time synchronization requires specific definitions and metrics. They are defined in [b-ITU-T G.8260].

5.2.4 Transport of frequency synchronization over the physical layer

Following the principles defined for SDH in [b-ITU-T G.803], [b-ITU-T G.781], [b-ITU-T G.813] and [b-ITU-T G.825], a solution has been specified for the transport of frequency synchronization over the physical layer of Ethernet. This is usually referred to as synchronous Ethernet (SyncE).

SyncE clocks are specified in [b-ITU-T G.8262] and [b-ITU-T G.8262.1], and SyncE network limit is specified in ITU-T G.8261. The network synchronization architecture is still based on [b-ITU-T G.803]. The SSM (synchronous status message) and functional models are specified in [b-ITU-T G.8264] and [b-ITU-T G.781].

As described in clause 5.3.5, similar principles have been applied to the transport of frequency synchronization over the OTN physical layer.

5.2.5 Transport of time synchronization

The basic principles and solutions for the distribution of time synchronization are specified in [b-ITU-T G.8271]. An example of the architecture to deliver time synchronization is shown in Figure 5-2.

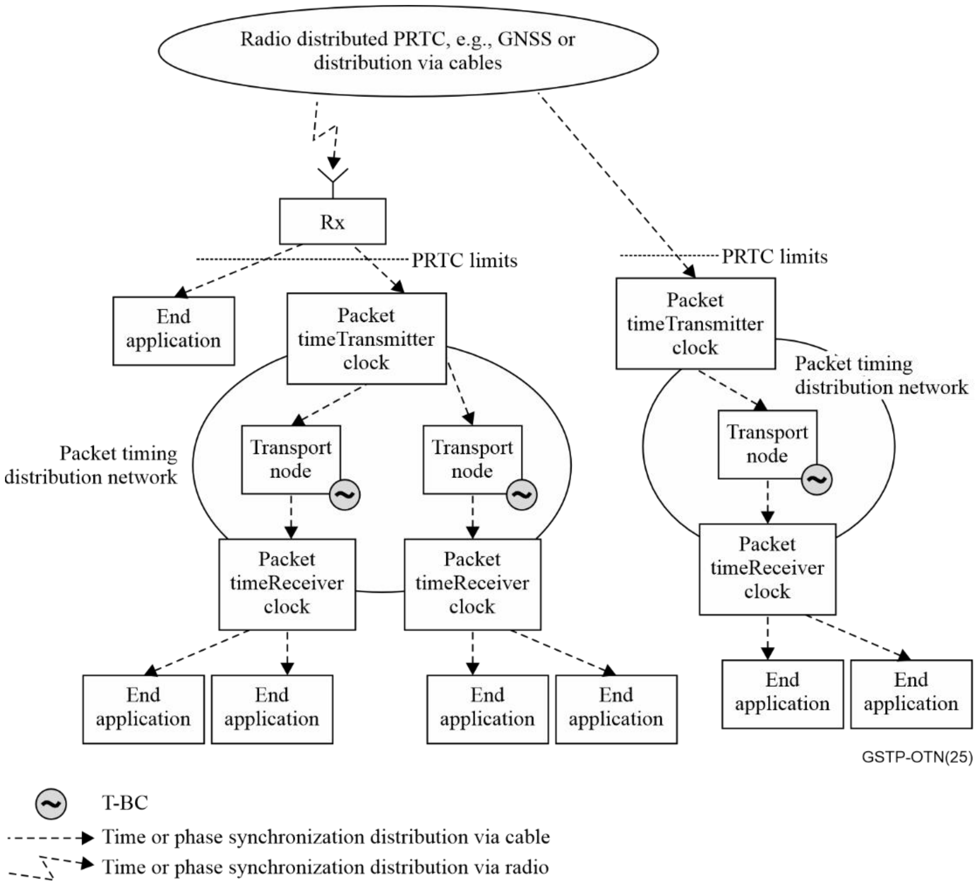


Figure 5-2 –Example of time synchronization distributed via packet-based methods (see Figure 3 in [b-ITU-T G.8271])

The transport of time synchronization requires the exchange of messages carrying timestamps. The protocol used for the distribution of accurate time synchronization is PTP as specified by [b-IEEE 1588].

The application of this protocol to a specific industry of this protocol requires the definition of so called "profiles", i.e., specification on which of the [b-IEEE 1588] options are relevant, and specifications on the details that meet the needs of the specific application, such as the specific transport layer (i.e., OTN, Ethernet, etc.), and PTP message rates. Telecom profiles for time synchronization have been specified in [b-ITU-T G. 8275.1] (for networks where every node is required to process the PTP messages – "network with full timing support") and [b-ITU-T G.8275.2] (for networks where not all nodes are required to process the PTP messages – "network with partial timing support").

Clocks to be deployed at the top of the synchronization network have been specified in [b-ITU-T G.8272] (primary reference time clock, PRTC), [b-ITU-T G.8272.1] (enhanced PRTC, ePRTC) and [b-ITU-T G.8272.2] (coherent network PRTC, cnPRTC).

Clocks forming part of the distribution chain have been specified in [b-ITU-T G.8273.2] (telecom boundary clock, T-BC) and [b-ITU-T G.8273.3] (telecom transparent clock, T-TC). The clock at the end of the PTP timing distribution is also specified by [b-ITU-T G.8273.2] (telecom time synchronous clock, T-TSC).

The transport of time synchronization over OTN has been specified using these technologies as described in clause 5.3.6.

5.2.6 Synchronization network design

The proper design of the synchronization network is a fundamental step to prevent that the above quoted impairments from occurring or to ensure that they are limited. One of the critical aspects to be considered in the network design is preventing the creation of timing loops in particular after reconfiguring of the network. In addition, it is important for the correct operation of the synchronization network that clocks of a lower hierarchical level only accept timing from clocks of the same or higher hierarchical level.

5.2.7 Testing

Testing the quality of timing in transport networks using the metrics defined in [b-ITU-T G.810] is possible with testing equipment specified by ITU-T. These testing specifications are specified in:

– [b-ITU-T O.171] for PDH

– [b-ITU-T O.172] for SDH

– [b-ITU-T O.174] for SyncE

– [b-ITU-T O.173] for OTN

The description of this equipment is outside the scope of this Technical Paper.

5.3 Specific OTN aspects

NOTE – Understanding this clause requires knowledge of the OTN frame, which is described in clause 2 of this Technical Paper.

ITU-T initial decision was that there is no need for a new synchronization layer in addition to the SDH existing one, and the OTN physical layer was not required to transport network synchronization. More precisely, neither the ODUk nor any layers below it were required to transport synchronization. The network synchronization distribution could remain a function of the client layer, e.g., SDH. Any SDH signal (which must meet [b-ITU-T G.825] that provides the jitter and wander requirements for SDH clients) is suitable for providing synchronization (see [b-ITU-T G.803]). SDH clients must meet the requirements defined in [b-ITU-T G.825] for both asynchronous and bit-synchronous OTN mappings.

In order to meet these requirements, the OTN frame has been designed so that the client mapping/demapping and transfer through OTN network equipment and 3R regenerators do not prevent SDH tributaries from meeting the jitter defined in [b-ITU-T G.825] and wander requirements.

An OTN network element (NE) was not required to implement synchronization interfaces, complex clocks with holdover mode, or SSM processing.

OTN transports client signals into an OTN frame (see the frame structure of clause 3) that is transported by an optical tributary signal (OTSi) on one lambda of the optical transport module (OTM). Each lambda can carry its frame based on [b-ITU-T G.709] with its own frequency (i.e., there is no requirement for a common clock for the different OTN frames of the OTM).

As a typical case, a trail through OTN is generated in an OTN NE that maps the client into an ODUk and is terminated in another OTN NE that demaps the client signal from the ODUk. Between the two OTN trail terminations, there might be 3R regenerators, which are pieces of equipment that perform complete regeneration of the pulse shape, clock recovery and retiming within required jitter limits.

The option to distribute timing (both frequency and time synchronization) over the OTN layer was added in a further revision of the OTN. This is described in clauses 5.3.5 and 5.3.6.

5.3.1 Network requirements

The jitter and wander control philosophy are based on the following network requirements needs:

– to recommend a maximum network limit that should not be exceeded at any relevant node network interface (NNI);
– to recommend a consistent framework for the specification of individual digital equipment (i.e., jitter and wander transfer, tolerance and generation requirements);
– to provide sufficient information and guidelines for organizations to measure and study jitter and wander accumulation in any network configuration.

In an OTN, jitter and wander accumulate on the transmission path according to the generation and transfer characteristics of interconnected equipment, 3R regenerators, client mappers, demappers, multiplexers and demultiplexers.

[b-ITU-T G.8251] specifies the maximum network limits of jitter and wander that shall not be exceeded and the minimum equipment tolerance to jitter and wander that shall be provided at any relevant interfaces based on the optical transport network.

A hypothetical reference model for 3R regenerators has been defined in [b-ITU-T G.8251] and simulations have shown that the network limits can be met at the OTUk interfaces. They have also shown that it is possible to keep jitter within network limits for an OTN network composed of 50 3R regenerators.

[b-ITU-T G.8251] also addresses the clock for ODUk (and other OTN frames) multiplexing defined in [b-ITU-T G.709]. Two HRMs (hypothetical reference model), HRM1 and HRM2, representing two extreme network cases that include ODUk multiplexing were defined in order to perform network simulations and verify that the jitter and short-term wander requirements are met for various types of clients.

For example, when SDH clients are carried by the OTN network, the specifications defined in [b-ITU-T G.825] must be met at the SDH interfaces so that the SDH clients might still be synchronization interfaces. Appendix II of [b-ITU-T G.8251] provides a provisional adaptation of the ITU-T G.803 SDH synchronization reference chain to include OTN islands. Considering that SDH may be transported by OTN islands, the SEC (SDH Equipment Clock) will no longer be present and will be replaced by OTN NEs. This leads to the definition of a reference chain where one or some SECs located between two SSUs are replaced by multiple OTN islands. Each of these islands may be composed of OTN NEs performing mapping/demapping or multiplexing/demultiplexing operations. This adaptation of the reference chain raises a buffer size constraint for the OTN NEs to maintain the overall network wander and jitter performance within specified limits. Predominantly the mapping and the demapping functions of the OTN contribute to wander accumulation due to the buffers being involved in these functions. The size of the buffers in the mapping and the demapping functions as specified in [b-ITU-T G.798] allows up to 10 mapping/ multiplexing nodes per OTN island. A total of 100 mapping/demapping functions can be performed on the multiple OTN islands.

These specifications have been established together with the definition of the clocks required by all functions defined for OTN, i.e., client mapping/ demapping, ODUk multiplexing/demultiplexing and 3R regenerators.

5.3.2 Mapping and multiplexing

These topics are presented in clauses 2 and 3 of this Technical Paper related to the frame specified in [b-ITU-T G.709] where the justification process is specified. Two types of mapping have been specified for the transport of CBR payload (e.g., SDH/SONET).

– The first one is the asynchronous mapping (AMP, GMP and fgGMP), which is the most widely used, where the payload floats within the OTN frame. In this case, there is no relationship between the payload and the OTN frame frequencies, thus simple free-running oscillators can be used to generate the OTN frame.
– The second is the synchronous mapping (BMP) where the timing used to generate the OTN frame is extracted from CBR client tributary (e.g., SDH/SONET); in the case of loss of signal

(LOS) from the input client, the OTN frequency that does not transport payload is generated by a free running oscillator, without the need for a holdover mode.

These mappings have been specified so that an asynchronous demapper may extract the CBR client signal mapped according to both techniques.

The multiplexing process is based on an asynchronous multiplexing scheme.

It is important to know that the justification process has been defined so that the OTN signals defined in [b-ITU-T G.709] may have a frequency range of up to ±20 ppm.

Detailed specifications related to atomic functions contributing to timing and jitter are defined in [b-ITU-T G.798]. An important requirement is that the demapping function must have a maximum clock filtering bandwidth of 300 Hz for most clients.

The bandwidth requirements for the following clients are more stringent in order to meet the applicable requirements: 1GE (100 Hz), DVB_ASI (200Hz), fgODUflex and E1 (both 10Hz).

5.3.3 Clock specification

Four ODC (ODUk clock, where ODUk stands for optical channel data unit-k) types are defined in [b-ITU-T G.8251], for different applications:

1) ODCa for asynchronous mapping (AMP, GMP and fgGMP) of clients into ODUk, and asynchronous multiplexing of ODUj into ODUk;
2) ODCb for bit-synchronous mapping (BMP) of clients into ODUk;
3) ODCr for 3R regeneration;
4) ODCp for demapping of constant bit rate (CBR) clients and demultiplexing of ODUk clients.

NOTE 1 – The ODUk and ODUj in the above bullets are just examples, and other OTN frames described by clause 2 are also applicable.

The ODCa and ODCb generate the timing signal for the ODUk and OTUk signals produced by an OTN network element. The ODCr generates the timing signal for the ODUk and OTUk produced by a 3R regenerator. The ODCp generates the timing signal for a demapped CBR client signal or a demultiplexing ODUk client signal. The free-run accuracy of these clocks should be within ±20ppm or ±100ppm, depending on the specific client signals.

The ODCa, used for asynchronous mapping, is free-running and the bit rate offset is accommodated by appropriately controlled stuffing.

The ODCb, used for bit-synchronous mapping, is locked to the bit rate of the incoming payload signal and the bit rate offset is accommodated by a fixed stuff pattern. The synchronous operation is continued even if the received payload content is AIS. If the incoming signal fails, the ODC enters the free-run condition.

The ODCr, used for 3R regeneration, is locked to the bit rate of the incoming OTSi_AP (optical tributary signal access point) signal (including AIS, alarm indication signal). If the incoming signal fails, the ODCr enters the free-run condition.

The ODCp, used for the CBR demapper (or the ODUk demultiplexing), is locked to the bit rate of the gapped OPUk (Optical Channel Payload Unit-k) clock (i.e., the timing of the signal that results from taking the OPUk payload and applying the justification control). If the incoming signal fails, the ODCp enters a free-run condition.

Annex A in [b-ITU-T G.8251] specifies the jitter generation of these clocks and, when applicable, noise tolerance, jitter transfer and transient response.

NOTE 2 – All these clock functions are used for clock recovery and clock filtering of a particular signal.

They never serve as an equipment synchronization source. Therefore, there is no holdover mode specified for these clocks since there is no need for an accurate clock when the input signal disappears. This is a major difference compared to SDH.

5.3.4 Transport of synchronous Ethernet through OTN

In the year 2000, when the first version of OTN was specified, the only synchronization signals were the STM-N signals. In that period Ethernet signals were generated by free running oscillators and did not carry reference synchronization timing on their physical layer. Therefore, the mapping of Ethernet signals had not been specified to be timing-transparent due to the lack of application for this feature.

The introduction of synchronous Ethernet (SyncE) added a new requirement to OTN mappings of Ethernet: the timing transparency of SyncE is similar to what was required to OTN for the transport of synchronization carried by SDH.

[b-ITU-T G.8251] was revised to allow the transport of synchronous Ethernet with an analysis of the effect of OTN on the distribution of synchronization via synchronous Ethernet clients. The additional analysis and simulations have been performed with 1 GE Synchronous Ethernet and higher synchronous Ethernet signals, and they have demonstrated that the requirements applicable to the timing of synchronous Ethernet carried by OTN are met. Especially for 1GE, it has nominal rates that are significantly less than the nominal rate of the CBR2G5 client considered in the previous release of OTN and causes a slightly higher timing error.

5.3.5 Transport of frequency synchronization through OTN layer (SyncO)

Because of the need to support the distribution of accurate time synchronization across the OTN network, it was required to define a solution to also distribute accurate frequency synchronization over the physical layer at every OTN node. In fact, the solution for time synchronization is based on the full timing support of the network (see clause 5.3.6), that combines the packet layer (PTP) with the frequency synchronization distributed via the physical layer.

The solution for frequency synchronization distribution through the OTN layer (also named SyncO) is based on the architecture defined in [b-ITU-T G.8261] and follows the same principles defined for SyncE.

The SyncO specifications are specified in [b-ITU-T G.8262] and [b-ITU-T G.8262.1]. The SSM and eSSM are specified in [b-ITU-T G.8264] and [b-ITU-T G.781].

The SSM and eSSM messages are encapsulated into GFP-F frames as specified by [b-ITU-T G.7041] and are carried in the OSMC channel (OTN synchronization message channel) as specified by [b-ITU-T G.709].

5.3.6 Transport of time synchronization through OTN layer

The transport of time synchronization over OTN is based on the general principles defined in [b-ITU-T G.8271] (high-level synchronization solutions and end application requirements), [b-ITU-T G.8274] (network synchronization architecture) and G.8275.1 (full timing support PTP profile). The clock specification is provided in [b-ITU-T G.8272] (PRTC) and [b-ITU-T G.8272.1] (ePRTC) respectively, and in [b-ITU-T G.8273.2] for telecom boundary clock (T-BC). No transparent clock solution is defined yet for this application.

PTP messages are carried in the OTN overhead and are terminated and regenerated at every OTN node that integrates a T-BC function. The PTP mapping is specified in [b-ITU-T G.8275.1] as follows: GFP-F encapsulation as per [b-ITU-T G.7041], and insertion of the GFP-F frames into the OTUk OSMC as per clause 15.7.2.4 of [b-ITU T G.709] or into the FlexO OSMC as per clause 9.2.10.1 of [b-ITU-T G.709.1], or the OSC OSMC as per clause 14.1 of [b-ITU T G.709], in which case the encapsulation is vendor-specific.

The PTP event messages are timestamped as specified in clause 15.7.2.4.1 of [b-ITU-T G.709] and in clause 9.2.10.1 of [b-ITU T G.709.1]. Figure 5-3 from [b-ITU-T G.709] illustrates how to generate timestamps for PTP event messages.

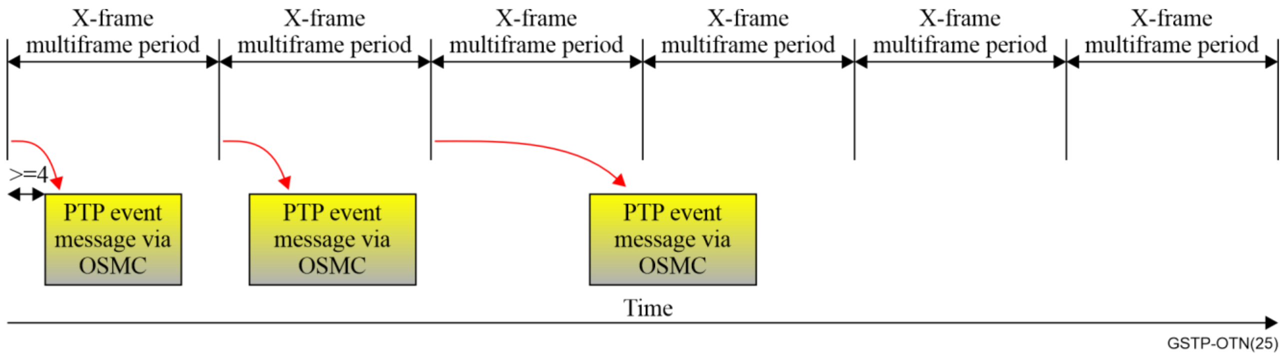


Figure 5-3 –Event timestamp generation, timing diagram example for OTUk based on Figure 15-16 of [b-ITU-T G.709]

6 Management and control

6.1 Introduction

Up through the late 1990s, transport networks and equipment were wholly configured and provisioned via centralized management systems. Transport networks expanded greatly in capacity and scope in the 1990s. This stimulated interest in a more dynamically configurable transport infrastructure and the automation of connection management-related operations was recognized as being a critical requirement. From this, the concept of the automatically switched optical network (ASON), software defined networking (SDN) and the management control continuum were developed one after another.

6.1.1 Architectural approaches – ASON and SDN

Development of the ASON Recommendations was driven by the vision of enabling multi-vendor and multi-service provider interoperable networking that supports end-to-end switched connection services on a global scale. ASON related ITU-T Recommendations started with the architecture defined in ITU-T G.8080 (renumbered as [b-ITU-T G.7703] after 2018) and expanded to signalling and routing requirements, followed by signalling protocols.

Interest in software defined networking (SDN) for transport networks resulted in work starting on [b-ITU-T G.7702] in 2014. A considerable amount of ITU-T G.8080 was found to be applicable to SDN architecture. The common aspects to both ASON and SDN architectures were captured in [b-ITU-T G.7701] which has grown as more functions have been realized to be in common. [b-ITU-T G.7702] contains SDN specific aspects beyond those contained in [b-ITU-T G.7701].

With these three Recommendations, SDN architecture for transport is described using [b-ITU-T G.7701] and [b-ITU-T G.7702]. ASON architecture is described using [b-ITU-T G.7701] and [b-ITU-T G.7703]. Both ASON and SDN architectures are applicable to connection-oriented circuit and packet transport networks, as defined in [b-ITU-T G.800].

6.1.2 Management-control continuum

Management Recommendations predating ASON (ITU-T G.8080) including those comprising FCAPS, are still needed for equipment, alarms and notification functions. It is recognized that network management (e.g., performed by an OSS), ASON control, and SDN controllers all perform the same operations on the transport resources. This leads to the concept whereby management and control (MC) functions are considered to be in a continuum, known as the management-control continuum.

Figure 6-1 illustrates the management-control continuum concept and its relationship with transport resources, whereby MC functions operate on transport resources and receive state information about resources. The MC function (in MC components) may be contained in MC systems, of which three types are shown: SDN, ASON, and others.

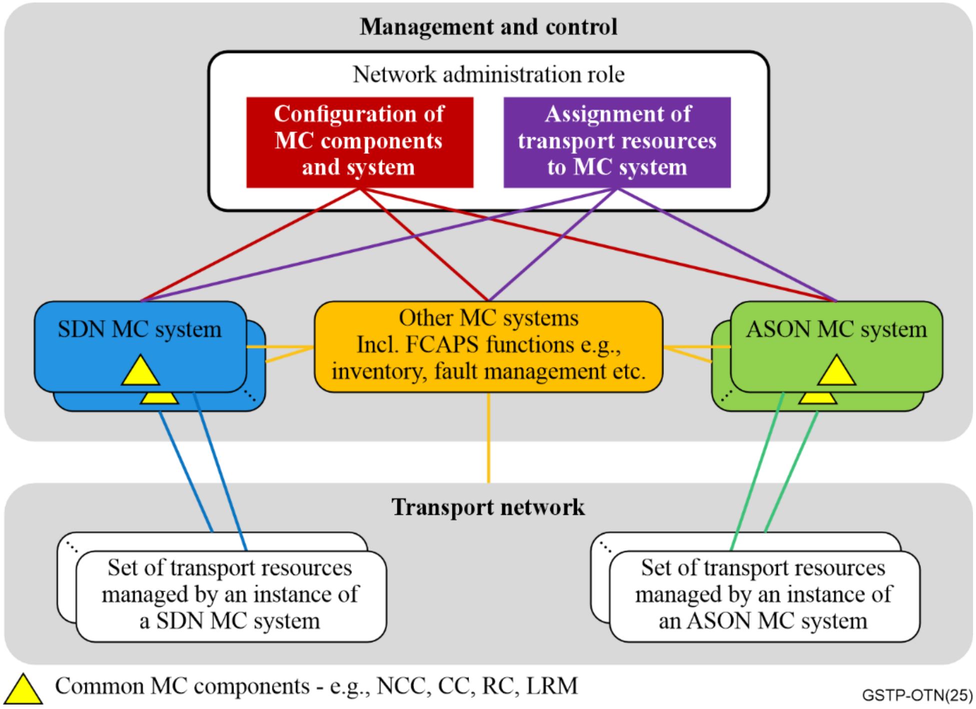


Figure 6-1 – Management-control continuum

6.2 Basic principles

MC architecture principles follow transport architecture defined in [b-ITU-T G.800] and [b-ITU-T G.807] in that networks can be modelled using recursive building blocks, and support the provisioning of various services. Reference points in the architecture enable business boundaries to be reflected and result from technology or administrative/managerial boundaries.

A well-designed architecture should give service providers control of their network, while providing fast and reliable call set-up. Management-control systems should be reliable, scalable, and efficient. They should be sufficiently generic to support different technologies, differing business needs and different distributions of functions by vendors (i.e., different packaging of the MC components).

Principle 1: Use the "call" construct, which reflects a service association that is distinct from connections that transfer information. A call is associated with zero or more connections.

Principle 2: Establish an architecture with reference points at domain boundaries.

Principle 3: Provide naming/identifier types to distinguish transport resources from control plane components such as signalling and routing control, and SCN (signalling communication network) addresses.

The architecture supports the establishment of services through the automatic provisioning of end-to-end transport connections across one or more domains.

It is a requirement to be able to support integrated call and connection management in a hybrid environment of MC enabled network elements. Thus, consistent and cohesive MC interactions and behaviour are critical with respect to the underlying transport resources.

6.3 Architecture and control components

In terms of architectural models, the transport plane refers to the transport resources described by [b-ITU-T G.805] and [b-ITU-T G.800], and MC refers to the managed components and objects and the systems that operate on them.

The MC Architecture Recommendation is composed of [b-ITU-T G.7701], [b-ITU-T G.7702] and [b-ITU-T G.7703] [b-ITU-T G.7702] describes the MC in a centralized manner while [b-ITU-T G.7703] describes the MC in a distributed manner, and [b-ITU-T G.7701] illustrates the commonality of the components usage in the SDN and ASON.

The following MC components are described in [b-ITU-T G.7701], as illustrated in Figure 6-2.

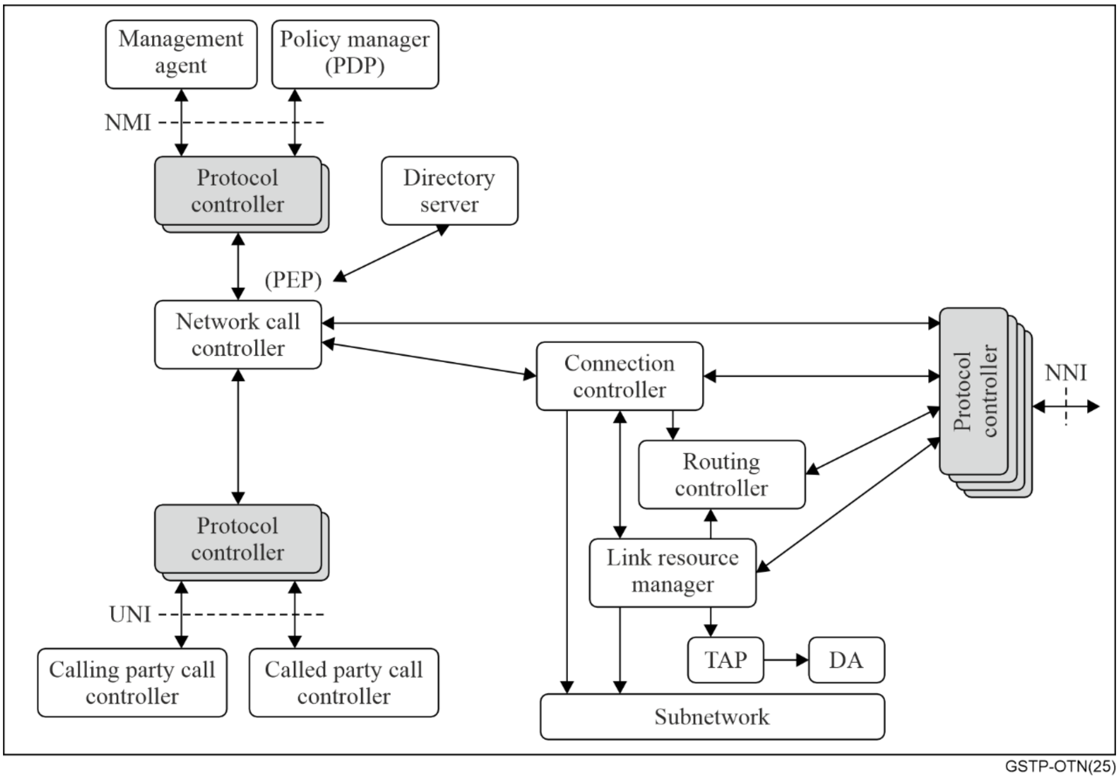


Figure 6-2 – Relationship among MC components

– The calling/called party call controller (CCC) and network call controller (NCC) components are responsible for controlling the setup, release, and modification of calls at domain boundaries, and may be invoked by management request or by signalling messages.

– The connection controllers (CC) component is responsible for establishing connections across a domain, requesting the route to use from the routing controller (RC) component, and requesting specific local link resources from the link resource manager (LRM) component. A call may have zero or more connections.

– The RC component is responsible for maintaining a view of the topology (nodes and links) within a routing area (RA) for the purpose of supporting path computation needed to establish connections.

– The LRM component is responsible for the control plane local connection inventory, managing the resources available to the link and allocating a specific link connection when requested.

– Protocol controller (PC) components communicate over abstract interfaces using primitives, so-called primitives, to distinguish these communications from real communications over physical interfaces.

– The termination and adaptation performer (TAP) component is the only ASON component that understands functions that can forward information (usually hardware, but could be virtual) and must therefore be collocated with forwarding functions.
– The discovery agent (DA) component is responsible for running auto discovery processes.
– The directory service (DS) component is responsible for identifier resolution and coordination among peer components, to provide mappings between identifier spaces for other components.

6.4 Key control functions

This clause provides description of the key control functions for ASON and SDN.

6.4.1 Data communication network/signalling communications network (DCN/SCN) architecture

DCN/SCN architecture enables the interactions of protocols running across the MCs, which may include the ASON UNI, NNI and SDN CPI. Resilience for DCN/SCN is required to support call/connection control performance.

[b-ITU-T G.7712] describes the data communications network, which is the IP communications infrastructure for optical control plane signalling communications network (SCN).

Several telecommunication technologies can support the DCN functions, such as circuit switching, optical transport network (OTN), and packet transport.

[b-ITU-T G.7712] considers networks that are IP-only, OSI-only, and mixed (i.e., support both IP and OSI), and includes interworking between parts of the DCN supporting IP-only, parts supporting OSI-only, and parts supporting both IP and OSI. The DCN provides layer 1 (physical), layer 2 (data-link) and layer 3 (network) functionality and consists of routing/switching functionality interconnected via links.

6.4.2 ASON signalling

The primary function of ASON signalling is call and connection management for establishing connections across the network. [b-ITU-T G.7713] provides protocol neutral specifications encompassing UNI, I-NNI and E-NNI reference points, supporting both soft-permanent and switched connections.

[b-ITU-T G.7713] provides distributed call and connection management requirements encompassing operations procedures, signalling network resilience to user and network defects, signal flow exception handling, and restoration for single and multiple rerouting domains. It also includes attribute specifications, message specifications, state diagrams, Call and connection controller management, and provides the basis for mapping to specific protocol solutions (e.g., ITU-T G.7713.x series).

The ASON components most relevant to signalling are the CCC/NCC, CC, and PC. The CCC is relevant to the client facing side of the UNI, while the NCCs are relevant to the network facing side of the UNI and the E-NNI, and allow each domain to have independent functions. Connection Controllers (CCs) establish the connections associated with each call segment.

[b-ITU-T G.7713] also supports inter-layer signalling control, assuming that client layer routing can compute a route that involves representations of server layer connectivity to which the client can be mapped. It also specifies the resilience of the signalling controller (SC), which contains the functions of connection control and/or call control. It also provides some information on the management of the call and connection controller function, including setting up and releasing a call and associated connections, as well as modifying a call.

6.4.3 ASON routing

The primary function for ASON transport routing is to provide path computation for connection management. ASON routing refers to the dissemination of reachability, topology, and resource/capability information throughout the control plane routing topology. ASON routing, defined in [b-ITU-T G.7715], provides architecture, requirements, high-level attributes, messages, and state diagrams from a protocol-neutral perspective. It is applicable after the network has been subdivided into routing areas, and necessary network resources are accordingly assigned.

Protocol-neutral routing requirements include support for hierarchically contained routing areas, non-congruent routing adjacency topology and transport network topology, independence from intra-domain protocol and control distribution choices, policy constraints on information exchange (e.g., imposed at E-NNI), and architectural evolution (levels, aggregation, segmentation).

It supports multiple links between nodes (allowing for link and node diversity), encompasses different classes of protocols (e.g., link-state, path vector), and facilitates comparison of specific inter-domain routing protocol proposals against quantifiable requirements. The primary ASON components relevant to routing are the RC and PC that instantiates a routing protocol.

[b-ITU-T G.7715.1], based upon [b-ITU-T G.7703] and [b-ITU-T G.7715], provides further architectural analysis for link state routing, encompassing the exchange of routing information between hierarchical routing levels, including visibility reachability and topology. ASON link state routing relies upon the basic link state functions of adjacency, database synchronization, and periodic or event-driven advertisements.

Path computation and routing are impacted by layer specific, layer independent, and client/server adaptation information elements. The routing protocol must be applicable to any transport layer network, and representation of routing attributes should not preclude their applicability to other transport network layers. It must also allow for dissemination of layer specific characteristics, and per-link attribute. Layer-specific link characteristics are provided in [b-ITU-T G.7715.1] to support multi-layer routing applications.

[b-ITU-T G.7715.2] specifies the requirements and architecture for the functions performed by routing controllers (RCs) during the operation of remote route query (i.e., a path computation request). The purpose of the remote route query, which is executed before signalling is used, is to establish one or more routes for a switched connection or a soft permanent connection.

In an SDN controller, the RC component computes routes based on the resources in the scope of the client's virtual network. It maintains topology information that is logically in the resource database and the relationship between a virtual network and its underlying topologies.

6.4.4 ASON auto discovery

Commissioning and installing new equipment can be a tedious error-prone process, if performed utilizing manual operations. Automated approaches are a key element in facilitating this process. The connectivity information derived from discovery is crucial for accurately building the network topology database used for computing the path for a connection.

[b-ITU-T G.7714] provides an architecture and protocol neutral descriptions of the discovery process for transport entities, their sub-processes and basic interactions. The architecture is that of a federation of discovery agents (DA) collaborating over all the network elements by means of protocol messages. The process is generic and is applicable to any layer of multilayer networks as described in [b-ITU-T G.800], and allows the discovery process to be used by the MC.

[b-ITU-T G.7714.1] describes the methods, procedures and transport plane mechanisms for discovering layer adjacency for ASON according to the requirements of [b-ITU-T G.7714] and the architecture of [b-ITU-T G.7703].

6.4.5 SDN virtualization

SDN architecture is described in terms of a hierarchy of SDN controllers that are inter-connected by a control plane interface (CPI), as shown in Figure 6-3. An SDN controller may have multiple clients and multiple servers; n+1 level controllers are clients of the level n controller and level n-1 controllers are servers (for the level n controller).

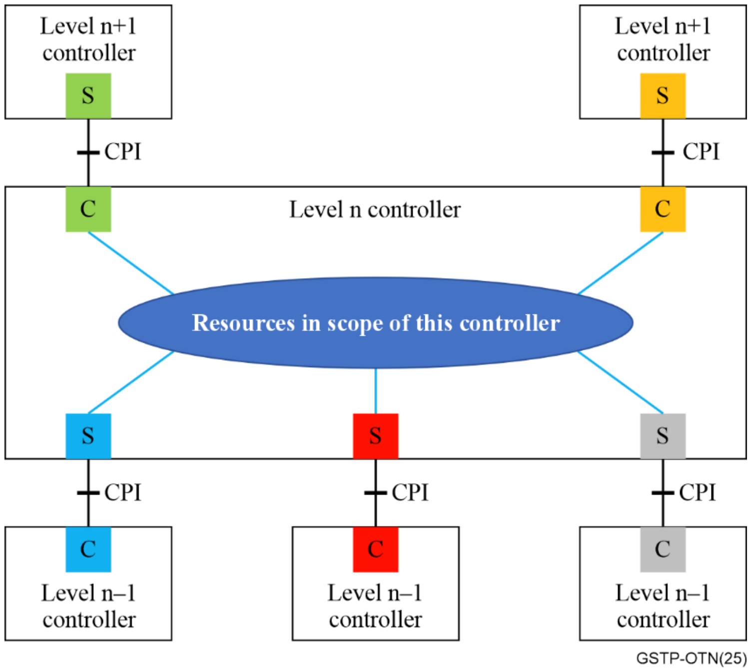


Figure 6-3 – Client/server relationships in the SDN architecture

Within a SDN controller, a particular client is supported by a set of information including the transport network resources (presented in a virtual network, or VN), as well as the MC components required to support that client. Together, these are known as the client context (labelled "C" in Figure 6-3).

The transport network resources managed by (in the scope of) a SDN controller are provided in one or more server contexts (labelled "S" in Figure 6-3). Each server context is supported by a set of information relating to transport resources provided by a server as well as the associated MC components.

The VN represents a part of the network resources information contained in a client context. Transport network resources are assigned to a VN in a client context by administrative or other means. The network resources information in the server context of a client controller is the same as the network resources information in the VN of the server controller.

6.5 Key management functions

6.5.1 Optical network management

The management aspects of optical transport network (OTN) elements are addressed in [b-ITU-T G.874], which contains the transport functions of one or more layer networks of the OTN as described in [b-ITU-T G.709]. The management of optical layer networks is separable from that of its client layer networks. [b-ITU-T G.874] specifies the management functions for fault management, configuration management (CM), account management, performance management (PM) and security management. [b-ITU-T G.874] describes the management network organizational model for communication between an element management layer (EML) operations system (OS) and the optical equipment management function (EMF) within an OTN network element (O.NE).

The management requirements of optical media layer are described in [b-ITU-T G.876], which contains the optical media layer functions defined based on the architecture from [b-ITU-T G.807] and [b-ITU-T G.798]. The management of the optical media layer is independent of the digital clients carried across the media network.

[b-ITU-T G.875] provides a management/control-protocol-neutral information model for managing/controlling network elements in the optical transport network (OTN) specified in [b-ITU-T G.872], [b-ITU-T G.709] and [b-ITU-T G.798]. It identifies the managed entities required for the management/control of OTN network elements. These entities are relevant to information exchanged across standardized interfaces defined in the TMN architecture ITU-T M.3010. The management/control-protocol-neutral information model could be used as the base for defining management-protocol-specific data models, for examples, YANG data models, enabling common management of OTN networks from different vendors regardless of the management protocol and data models being used.

The information model defined in [b-ITU-T G.875] is an augmentation to the generic core information model specified in [b-ITU-T G.7711] for managing OTN transport resources. The core information model defined in [b-ITU-T G.7711] can be used as the base for the extension of OTN-specific information models.

[b-ITU-T G.875] applies to OTN network elements and those systems that manage/control OTN network elements. The management/control system could be an NMS, EMS, SDN controller or a hybrid of these. Functional capabilities of OTN equipment are defined in [b-ITU-T G.798], and requirements of the management of OTN equipment are provided in [b-ITU-T G.7710] and [b-ITU-T G.874]. The information model specified in this Recommendation applies to the management/control interface, as shown in Figure 6-4, specifically for managing/controlling the OTN functional capabilities of the NE.

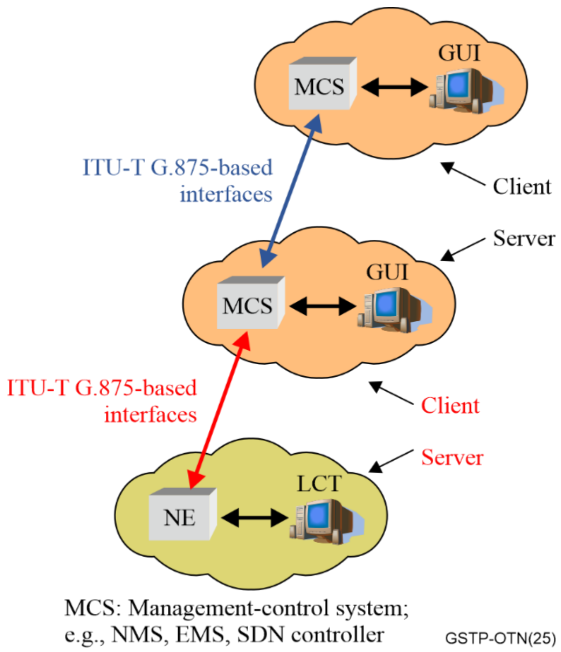


Figure 6-4 – Scope of interface

The UML diagram shown in Figure 6-5 provides a high-level overview of most of the OTN-specific managed object classes without showing the details, such as the attributes and operations of the object classes.

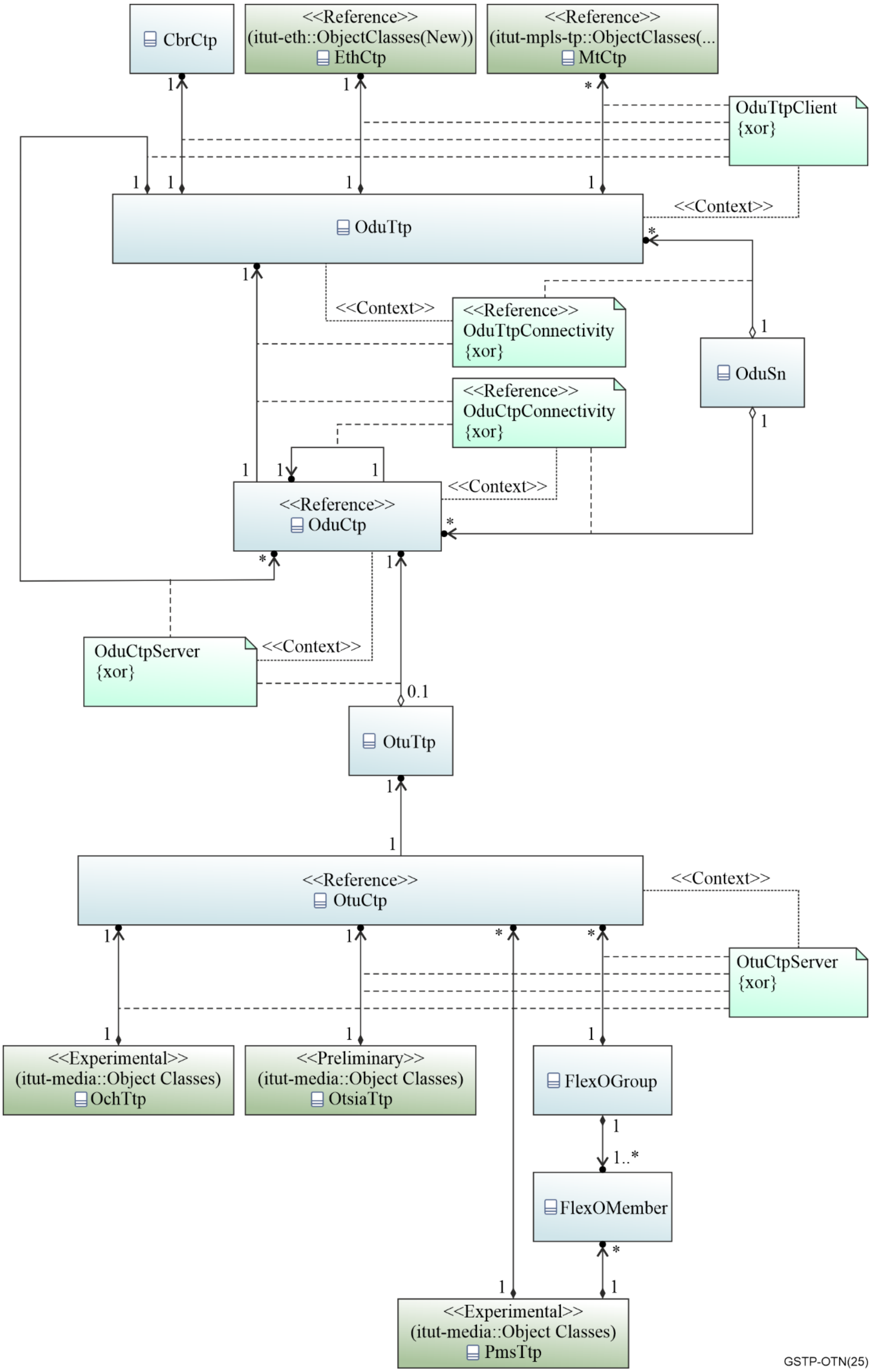


Figure 6-5 – OTN model high level overview

[b-ITU-T G.876] provides a management/control-protocol-neutral information model for optical media layer elements defined in the optical transport network (OTN) from [b-ITU-T G.807] and [b-ITU-T G.798]. Figure 6-6 provides the mapping between the object classes and the functions model.

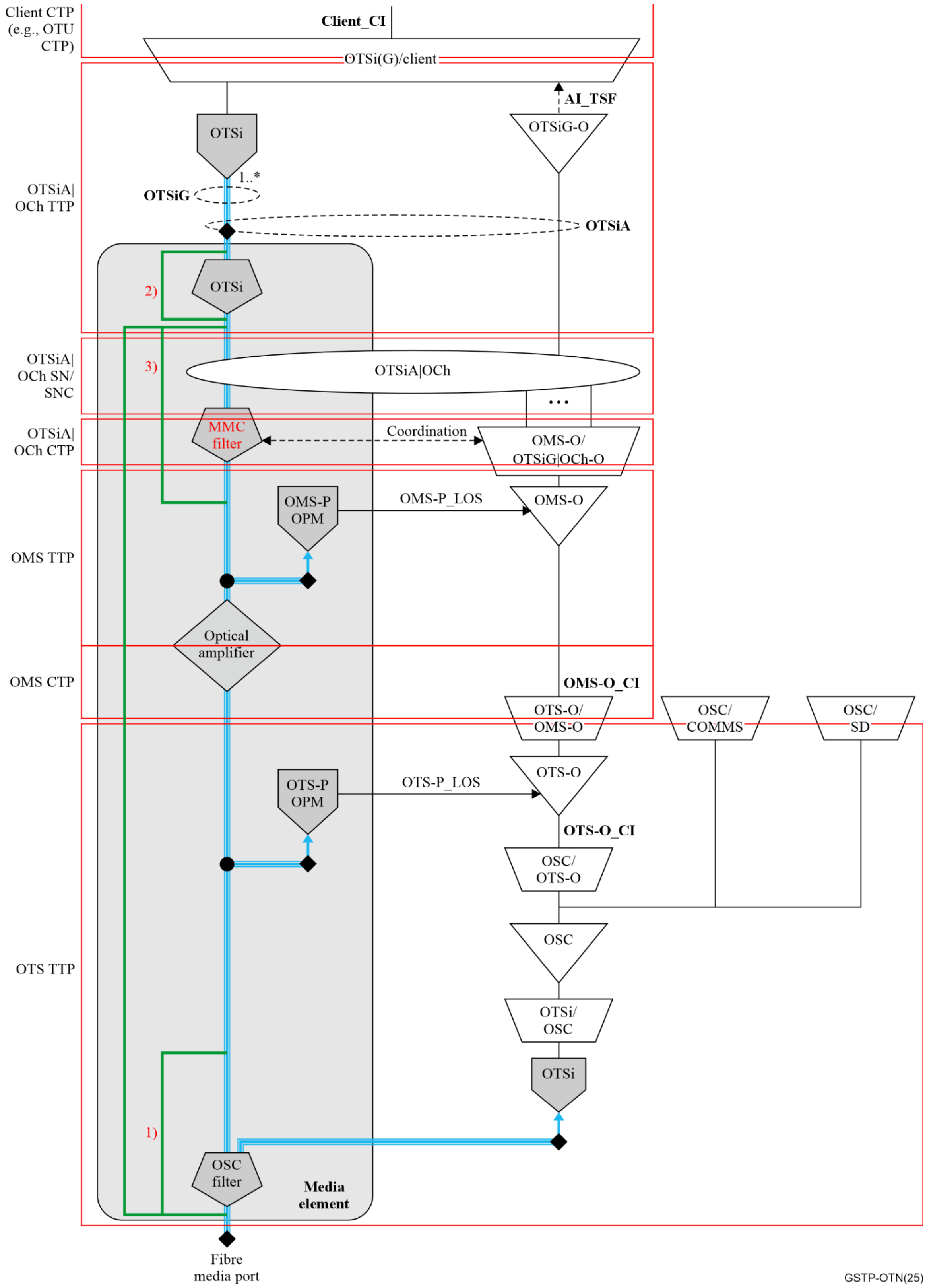


Figure 6-6 – Overview of the object class mapping to a media element and OTN atomic functions

6.5.2 MC component management

Management and control are a continuum and MC systems may contain collections of MC components as shown in Figure 6-1. These MC components need to be managed together with the transport resources, in a coordinated manner.

[b-ITU-T G.7718] contains the framework for MC management. It places MC within the telecommunications management network (TMN) context and specifies how the TMN principles may be applied. A management view is developed, which provides the basis for the MC requirements. Extensive configuration management requirements have been specified in [b-ITU-T G.7718], which include:

– Identifier management, which refers to the administration of identifiers.
– Resource management, which includes allocation/de-allocation of transport resources within the scope of an MC system.
– Routing area (RA) configuration, which assigns MC components to RA and hierarchical levels.
– Protocol controller configuration, which supports the assignment of a point of attachment to the DCN.
– Inventory, topology and transport entity capability exchange, which supports automatic discovery and notification of the addition/removal/upgrade of MC objects.
– Call/connection management, which supports the creation, removal and modification of call/connection.
– Soft permanent connection (SPC) and switched connection (SC) management.
– Fault management using alarm and notifications.
– Performance management based on current and historic usage data.
– Accounting management related to the representation, storage and communication of data.
– Protection and restoration management, which includes notification of MC restoration failure.

[b-ITU-T G.7719] defines a protocol-neutral information model for managing the components that provide the functions of managing and/or controlling the transport network resources. The management requirements for such management and control (MC) components are based on [b-ITU-T G.7718]. The protocol-neutral management information model is specified using the unified modelling language (UML).

6.5.3 MC operation

[b-ITU-T G.7716] addresses the architecture of MC operations, providing guidance for service providers on the transport network planning, initialization, performing typical operations and maintenance in the network. The following processes are illustrated:

– The transport network planning, which includes resource planning and infrastructure considerations.
– The transport network initialization, which includes commissioning and reconfiguration.
– The transport network operation and maintenance, which includes provisioning, resource assignment, discovery and authentication process, link configuration, routing adjacency configuration, reconfiguration, auditing and recovery.
– The operations on virtual networks (VN), which includes the resource assignment, name mapping and reconfiguration.

6.6 Considerations on OTN deployment with MC

6.6.1 Trends on the OTN networks

The implementation of an ASON/SDN MC systems offers to the operators several advantages including:

Closer to the user: the deployment of fibre is making the transport network closer to the user. Shorter distance transmission technologies, and new connectivity types (such as PtMP) are also under discussion. With these, dedicated connectivity can be provided directly to the users.

Service-oriented: Multiple dimensions of parameters are investigated to specify the user's experience for network services. It is possible both to supply high quality services guaranteeing the user experience such as high-reliability and low-latency, with differentiated pricing.

Higher quality: The transport technology development is providing networks with higher bandwidth, lower-latency and higher-reliability, which provides higher-value network connectivity to users.

6.6.2 Future challenges on OTN MC

The introduction of a transport network with MC provides advantages illustrated in the previous section, but faces upfront challenges as well. Some of these difficulties are listed in the following.

Larger-scale networking MC: The scalability of transport network keeps increasing with emerging technologies like fine-grain OTN. MC systems need to scale accordingly.

More dynamic changes in the network: once the transport network becomes closer to the user, the demand for the connectivity is more dynamic. Timely adjustment is required to satisfy the requirement from user/services, which requires dynamic MC techniques.

Coordination and interaction between the networking and other ICT capabilities. According to new network applications, the transport network is not only carrying data, but also processing data. Coordination between the networking, computation and storage is under investigation, and the coordination between the network MC and other ICT management will be needed.

Full life-cycle intelligence: The emergence of AI/ML is helping all aspects of network operation and maintenance. Once each single step in the procedure is covered by the intelligence, there could be end-to-end intelligence in the full life cycle of network and its services. Specifications would be needed to describe these procedures.

Bibliography

[b-ITU-T G.695] Recommendation ITU-T G.695 (2025), *Optical interfaces for coarse wavelength division multiplexing applications.*

[b-ITU-T G.959.1] Recommendation ITU-T G.959.1 (2024), *Optical transport network physical layer interfaces.*

[b-ITU-T G.698.1] Recommendation ITU-T G.698.1 (2023), *Multichannel DWDM applications with single-channel optical interfaces.*

[b-ITU-T G.709] Recommendation ITU-T G.709/Y.1331 (2020), *Interfaces for the optical transport network.*

[b-ITU-T G.709.1] Recommendation ITU-T G.709.1 (2024), *Flexible OTN common elements.*

[b-ITU-T G.709.3] Recommendation ITU-T G.709.3 (2024), *Flexible OTN B100G long-reach interfaces.*

[b-ITU-T G.709.5] Recommendation ITU-T G.709.5 (2024), *Flexible OTN short-reach interfaces.*

[b-ITU-T G.709.6] Recommendation ITU-T G.709.6 (2024), *Flexible OTN B400G long-reach interfaces.*

[b-ITU-T G.781] Recommendation ITU-T G.781 (2026), *Synchronization layer functions for frequency synchronization based on the physical layer.*

[b-ITU-T G.798] Recommendation ITU-T G.798 (2026), *Characteristics of optical transport network hierarchy equipment functional blocks - Overview.*

[b-ITU-T G.7041] Recommendation ITU-T G.7041/Y.1303 (2016), *Generic framing procedure.*

[b-ITU-T G.7701] Recommendation ITU-T G.7701 (2022), *Common control aspects.*

[b-ITU-T G.7702] Recommendation ITU-T G.7702 (2022), *Architecture for SDN control of transport networks.*

[b-ITU-T G.7703] Recommendation ITU-T G.7703 (2021), *Architecture for the automatically switched optical network.*

[b-ITU-T G.7710] Recommendation ITU-T G.7710 (2025), *Common equipment management function requirements.*

[b-ITU-T G.7711] Recommendation ITU-T G.7711 (2025), *Generic protocol-neutral information model for transport resources.*

[b-ITU-T G.7712] Recommendation ITU-T G.7712/Y.1703 (2019), *Architecture and specification of data communication network.*

[b-ITU-T G.7713] Recommendation ITU-T G.7713/Y.1704 (2009), *Distributed call and connection management (DCM).*

[b-ITU-T G.7714] Recommendation ITU-T G.7714/Y.1705 (2005), *Generalized automatic discovery for transport entities.*

[b-ITU-T G.7714.1] Recommendation ITU-T G.7714.1/Y.1705.1 (2017), *Protocol for automatic discovery in transport networks.*

[b-ITU-T G.7715] Recommendation ITU-T G.7715/Y.1706 (2002), *Architecture and requirements for routing in the automatically switched optical networks.*

[b-ITU-T G.7715.1] Recommendation ITU-T G.7715.1/Y.1706.1 (2004), *ASON routing architecture and requirements for link state protocols.*

[b-ITU-T G.7715.2] Recommendation ITU-T G.7715.2/Y.1706.2 (2007), *ASON routing architecture and requirements for remote route query.*

[b-ITU-T G.7716] Recommendation ITU-T G.7716 (2025), *Architecture of management and control operations.*

[b-ITU-T G.7718] Recommendation ITU-T G.7718/Y.1709 (2020), *Framework for the management of management-control components and functions.*

[b-ITU-T G.7719] Recommendation ITU-T G.7719 (2021), *Management information model for management-control components and functions.*

[b-ITU-T G.800] Recommendation ITU-T G.800 (2016), *Unified functional architecture of transport networks.*

[b-ITU-T G.803] Recommendation ITU-T G.803 (2000), *Architecture of transport networks based on the synchronous digital hierarchy (SDH).*

[b-ITU-T G.805] Recommendation ITU-T G.805 (2000), *Generic functional architecture of transport networks.*

[b-ITU-T G.806] Recommendation ITU-T G.806 (2012), *Characteristics of transport equipment – Description methodology and generic functionality.*

[b-ITU-T G.807] Recommendation ITU-T G.807 (2024), *Generic functional architecture of the optical media layer.*

[b-ITU-T G.808.1] Recommendation ITU-T G.808.1 (2014), *Generic protection switching – Linear trail and subnetwork protection.*

[b-ITU-T G.810] Recommendation ITU-T G.810 (1996), *Definitions and terminology for synchronization networks.*

[b-ITU-T G.811] Recommendation ITU-T G.811 (1997), *Timing characteristics of primary reference clocks.*

[b-ITU-T G.812] Recommendation ITU-T G.812 (2004), *Timing requirements of slave clocks suitable for use as node clocks in synchronization networks.*

[b-ITU-T G.813] Recommendation ITU-T G.813 (2003), *Timing characteristics of SDH equipment slave clocks (SEC).*

[b-ITU-T G.822] Recommendation ITU-T G.822 (1988), *Controlled slip rate objectives on an international digital connection.*

[b-ITU-T G.823] Recommendation ITU-T G.823 (2000), *The control of jitter and wander within digital networks which are based on the 2048 kbit/s hierarchy.*

[b-ITU-T G.824] Recommendation ITU-T G.824 (2000), *The control of jitter and wander within digital networks which are based on the 1544 kbit/s hierarchy.*

[b-ITU-T G.825] Recommendation ITU-T G.825 (2000), *The control of jitter and wander within digital networks which are based on the synchronous digital hierarchy (SDH).*

[b-ITU-T G.872] Recommendation ITU-T G.872 (2024), *Architecture of the optical transport network.*

[b-ITU-T G.874] Recommendation ITU-T G.874 (2020), *Management aspects of optical transport network elements.*

[b-ITU-T G.875] Recommendation ITU-T G.875 (2024), *Optical transport network: Protocol-neutral management information model for the network element view.*

[b-ITU-T G.876] Recommendation ITU-T G.876 (2021), *Management requirements and information model for the optical media network.*

[b-ITU-T G.959.1] Recommendation ITU-T G.959.1 (2024), *Optical transport network physical layer interfaces.*

[b-ITU-T G.8010] Recommendation ITU-T G.8010/Y.1306 (2004), *Architecture of Ethernet layer networks.*

[b-ITU-T G.8251] Recommendation ITU-T G.8251 (2022), *The control of jitter and wander within the optical transport network (OTN).*

[b-ITU-T G.8260] Recommendation ITU-T G.8260 (2022), *Definitions and terminology for synchronization in packet networks.*

[b-ITU-T G.8261] Recommendation ITU-T G.8261/Y.1361 (2019), *Timing and synchronization aspects in packet networks.*

[b-ITU-T G.8262] Recommendation ITU-T G.8262 (2024), *Timing characteristics of synchronous equipment clocks.*

[b-ITU-T G.8262.1] Recommendation ITU-T G.8262.1 (2025), *Timing characteristics of enhanced synchronous equipment clocks.*

[b-ITU-T G.8264] Recommendation ITU-T G.8264 (2025), *Distribution of timing information through packet networks.*

[b-ITU-T G.8271] Recommendation ITU-T G.8271/Y.1366 (2020), *Time and phase synchronization aspects of telecommunication networks.*

[b-ITU-T G.8272] Recommendation ITU-T G.8272 (2025), *Timing characteristics of primary reference time clocks.*

[b-ITU-T G.8272.1] Recommendation ITU-T G.8272.1 (2024), *Timing characteristics of enhanced primary reference time clocks.*

[b-ITU-T G.8272.2] Recommendation ITU-T G.8272.2 (2024), *Timing characteristics of coherent network primary reference time clocks.*

[b-ITU-T G.8273.2] Recommendation ITU-T G.8273.2/Y.1368.2 (2023), *Timing characteristics of telecom boundary clocks and telecom time synchronous clocks for use with full timing support from the network.*

[b-ITU-T G.8273.3] Recommendation ITU-T G.8273.3/Y.1368.3 (2020), *Timing characteristics of telecom transparent clocks for use with full timing support from the network.*

[b-ITU-T G.8274] Recommendation ITU-T G.8274 (2025), *Architecture and requirements for packet-based time and phase distribution.*

[b-ITU-T G.8275.1] Recommendation ITU-T G. 8275.1 (2026), *Precision time protocol telecom profile for phase/time synchronization with full timing support from the network.*

[b-ITU-T G.8275.2] Recommendation ITU-T G.8275.2 (2026), *Precision time protocol telecom profile for phase/time synchronization with partial timing support from the network.*

[b-ITU-T G.8310] Recommendation ITU-T G.8310 (2020), *Architecture of the metro transport network.*

[b-ITU-T G-Sup.58] Supplement 58 to ITU-T G-series Recommendations (2024), *Optical transport network module framer interfaces*.

[b-ITU-T G-Sup.76] Supplement 76 to ITU-T G-series Recommendations (2021), *Optical transport network security*.

[b-ITU-T M.20] Recommendation ITU-T M.20 (1992), *Maintenance philosophy for telecommunication networks.*

[b-ITU-T O.171] Recommendation ITU-T O.171 (1997), *Timing jitter and wander measuring equipment for digital systems which are based on the plesiochronous digital hierarchy (PDH).*

[b-ITU-T O.172] Recommendation ITU-T O.172 (2005), *Jitter and wander measuring equipment for digital systems which are based on the synchronous digital hierarchy (SDH).*

[b-ITU-T O.173] Recommendation ITU-T O.173 (2012), *Jitter measuring equipment for digital systems which are based on the optical transport network.*

[b-ITU-T O.174] Recommendation ITU-T O.174 (2009), *Jitter and wander measuring equipment for digital systems which are based on synchronous Ethernet technology.*

[b-ITU Handbook] ITU-T Manual (2009), *Optical Fibres, Cables and Systems*.
https://www.itu.int/dms_pub/itu-t/opb/hdb/t-hdb-out.10-2009-1-pdf-e.pdf

[b-IEEE 802.3] IEEE 802.3-2022, *IEEE Standard for Ethernet*.

[b-IEEE 1588] IEEE 1588-2008, *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems.*

____