International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Series G
Supplement 66
(09/2020)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA,
DIGITAL SYSTEMS AND NETWORKS

5G wireless fronthaul requirements in a passive optical network context

ITU-T G-series Recommendations – Supplement 66

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

For further details, please refer to the list of ITU-T Recommendations.

Supplement 66 to ITU-T G-series Recommendations

5G wireless fronthaul requirements in a passive optical network context

Summary

Supplement 66 to ITU-T G-series Recommendations enumerates the various requirements arising from 5G wireless systems, concentrating on the fronthaul portion of the network and considers how they compare with current and future optical access transport systems. Practical passive optical network (PON) solutions to serve the 5G fronthaul application are hypothesized.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G Suppl. 66	2018-10-19	15	11.1002/1000/13826
2.0	ITU-T G Suppl. 66	2019-07-12	15	11.1002/1000/13991
3.0	ITU-T G Suppl. 66	2020-09-18	15	11.1002/1000/14546

Keywords

5G, fronthaul, PON.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

This is an informative ITU-T publication. Mandatory provision, 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.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this publication may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the publication development process.

As of the date of approval of this publication, ITU had received notice of intellectual property, protected by patents, which may be required to implement this publication. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Supplement 66 to ITU-T G-series Recommendations

5G wireless fronthaul requirements in a passive optical network context

1 Scope

This Supplement considers the requirements for 5G wireless fronthaul specifically in the setting of optical access networks. It synthesizes various specifications from other groups into practically realisable system requirements. This supplement also discusses practical passive optical network (PON) solutions to meet the synthesized requirements.

2 References

[ITU-T G.984.2] Recommendation ITU-T G.984.2 (2019), *Gigabit-capable Passive Optical Networks (G-PON): Physical Media Dependent (PMD) layer specification*.

[ITU-T G.987] Recommendation ITU-T G.987 (2012), *10-Gigabit-capable passive optical network (XG-PON) systems: Definitions, abbreviations and acronyms*.

[ITU-T G.987.2] Recommendation ITU-T G.987.2 (2016), *10-Gigabit-capable passive optical networks (XG-PON): Physical media dependent (PMD) layer specification*.

[ITU-T G.989.2] Recommendation ITU-T G.989.2 (2019), *40-Gigabit-capable passive optical networks 2 (NG-PON2): Physical media dependent (PMD) layer specification*.

[ITU T G.989.3] Recommendation ITU-T G.989.3 (2015), *40-Gigabit-capable passive optical networks (NG-PON2): Transmission convergence layer specification*.

[ITU-T G.8013] Recommendation ITU-T G.8013/Y.1731 (2015), *Operations, administration and maintenance (OAM) functions and mechanisms for Ethernet-based networks*.

[ITU-T G.8271] Recommendation ITU-T G.8271 (2018), *Time and phase synchronization aspects of telecommunication networks*.

[ITU-T G.8271.1] Recommendation ITU-T G.8271.1 (2018), *Network limits for time synchronization in packet networks*.

[ITU-T G.9807.1] Recommendation ITU-T G.9807.1 (2017), *10-Gigabit-capable symmetric passive optical network (XGS-PON).*

[ITU-R M.2083] Recommendation ITU-R M.2083 (2015), *IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond*.

[ITU-R M.2410-0] Report ITU-R M.2410-0 (2017), *Minimum requirements related to technical performance for IMT-2020 radio interface(s).*

[IEEE 802.1Q] IEEE 802.1Q (2014), *Ethernet connectivity fault management*.

[IEEE 1914.1] IEEE 1914.1 (2019), *Standard for Packet-based Fronthaul Transport Networks*.

[3GPP TR 38.801] 3GPP TR 38.801 V2.0.0 (R14) (2017), *Technical Specification Group Radio Access Network; Study on New Radio Access Technology; Radio Access Architecture and Interfaces*.

[3GPP TS 22.261] 3GPP TS 22.261, V16.1.0 (2018), *5G; Service requirements for next generation new services and markets*.

[3GPP TS 36.213] 3GPP TS 36.213 V14.5.0 (2017), *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*.

[3GPP TS 38.101-1] 3GPP TS 38.101-1/2/3 V15.0.0 (2017), *User Equipment (UE) radio transmission and reception (Parts 1, 2 and 3)*.

[3GPP TS 38.211] 3GPP TS 38.211 V15.0.0 (2017), *Physical channels and modulation*.

[3GPP TS 38.214] 3GPP TS 38.214 V15.0.0 (2017), *Physical layer procedures for data*.

[3GPP TS 38.401] 3GPP TS 38.401 V0.2.0 (2017), *5G;NG-RAN; Architecture description*.

[3GPP TR 38.913] 3GPP TR 38.913, V14.3.0 (2017), *Study on scenarios and requirements for next generation access technologies*.

[3GPP R3-161813] 3GPP R3-161813, *Transport requirement for CU&DU functional splits options, CMCC*.

[3GPP R3-162101] 3GPP R3-162101, *CU-DU split: Refinement for Annex A (Transport network and RAN internal functional split), NTT DOCOMO*.

[TS 38.470] 3GPP Release 15, TS 38.470 … 475 series (NG-RAN; F1 interface).

[TR 38.816] 3GPP Release 15, TR 38.816 V15.0.0 (2017-12), *Study on CU-DU lower layer split for NR*.

[Babkin] A. Babkin et al. (2013), *LTE Network Throughput Estimation*, Internet of Things and its Enablers (INTHITEN 2013), pp. 95-104, June.

[Chanclou] P. Chanclou (2017), *How does passive optical network tackle radio access network evolution?* pp. 1030-1040, v9 (11), JOCN, Nov.

[CMRI] China Mobile Research Institute et al. (2016), *Toward 5G C-RAN: Requirements, Architecture and Challenges*, V. 1.0, Nov.

[COMBO] EU Project: COMBO (Convergence of fixed and Mobile Broadband access/aggregation networks), *COMBO_D3.3_WP3_30June2015_EAB_v1.0*, Deliverable D3.3, Sec. 8.3.3, July 2015.

[CPRI] CPRI Specification V7.0 (2015-10-09).

[Doetsch] U. Doetsch, et al., *Quantitative analysis of split base station processing and determination of advantageous architecture for LTE*, Bell Labs Technical Journal 18(1), 105–128 (2013).

[eCPRI] eCPRI Transport Network V1.2 (2018), *Common Public Radio Interface: Requirements for the eCPRI Transport Network*.

[Lee] H. H. Lee et al. (2016), *Real-time demonstration of QoS guaranteed 25-Gb/s PON prototype with Ethernet-PON MAC/PHY and cost-effective APD receivers for 100-Gb/s access networks*, Optics Express, vol. 24, No. 13.

[NGMN] NGMN Alliance (2017), *5G End-to-End Architecture Framework*, v0.6.5, May.

[NGMN2] NGMN Alliance (2011), *Guidelines for LTE Backhaul Traffic Estimation*, July.

[OBSAI]	Open Base Station Architecture Initiative (OBSAI), *BTS system reference document*, Version 2.0, 2006. http://www.obsai.com/specs/OBSAI_System_Spec_V2.0.pdf
[ORAN-CTI-TC]	O-RAN Alliance, *Cooperative Transport Interface Transport Control Plane Specification,* O-RAN.WG4.CTI-TCP.0-v01.00, April 2020.
[ORAN-FH]	O-RAN Alliance, *Control, User and Synchronization Plane Specification*, WG4.CUS.0-v03.00, April 2020.
[SCF]	Small Cell Forum, *Small Cell Virtualization Functional Splits and Use Cases*, document 159.07.02 (2016-01-13).
[Tashiro]	T. Tashiro, et al. (2014), *A novel DBA scheme for TDM-PON based mobile fronthaul*, OFC paper Tu3F.3.
[Tayq]	Z. Tayq et al. (2017), *Real Time Demonstration of the Transport of Ethernet Fronthaul based on vRAN in Optical Access Networks*, Th3A.2, OFC.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Supplement

None.

4 Abbreviations and acronyms

This Supplement uses the following abbreviations and acronyms:

AWG	Array Waveguide Grating
BBU	Baseband Unit
CP	Control Plane
CPRI	Common Public Radio Interface
Co-DBA	Cooperative Dynamic Bandwidth Allocation
CU	Central Unit
DL	Down Link
DU	Distributed Unit
eCPRI	evolved Common Public Radio Interface
eMBB	enhanced Mobile Broadband
EPC	Evolved Packet Core
FAPI	Functional Application Platform Interface
HARQ	Hybrid Automatic Repeat Request
MAC	Media Access Control
MEF	Metro Ethernet Forum
MIMO	Multiple Input Multiple Output
mMTC	massive Machine Type Communication

NGC	Next Generation Core
nFAPI	network Functional Application Platform Interface
NGFI	Next Generation Fronthaul Interface
NR	New Radio
NRT	Non-Real Time
OBSAI	Open Base Station Architecture Initiative
OLT	Optical Line Terminal
OMCI	ONT Management and Control Interface
ONU	Optical Network Unit
OPL	Optical Path Loss
PDCP	Packet Data Convergence Protocol
PON	Passive Optical Network
RAN	Radio Access Network
RF	Radio Frequency
RLC	Radio Link Control
RNL	Radio Network Layer
RRC	Radio Resource Control
RRH	Remote Radio Head
RT	Real Time
RU	Remote Unit
TNL	Transport Network Layer
TTI	Transmission Time Interval
UL	Up Link
UP	User Plane
URLLC	Ultra-Reliable Low Latency Communication

5 Conventions

None.

6 Overview of 5G wireless fronthaul architecture

NOTE – An early draft of this clause was included in the ITU-T Technical Report GSTR-TN5G. This version of the Supplement contains corrections and additional updates.

6.1 Evolution of wireless transport architecture from 4G to 5G

In the evolution from 4G/LTE to 5G new radio (NR) transport architecture, the main change is that the original baseband unit (BBU) function in 4G/LTE can be split into three parts namely a central unit (CU), a distributed unit (DU) and a remote unit (RU). The reasons for this redesign are manifold [NGMN]. The new design will better facilitate radio access network (RAN) virtualization with flexible assignment of computing resources across three different network entities. It also allows for decreased fronthaul line rates, while meeting latency demands.

Specific functions residing in the CU and DU are deployment dependent and still under discussion. Figure 6-1 gives one possible example of the evolution from 4G to a functional split architecture in 5G [CMRI]. The RAN architecture in 4G consists of evolved packet core (EPC), BBU and remote radio head (RRH). When evolving to 5G, in this example, part of the user plane (UP) functions are moved from EPC to CU and DU, Layer 2 (L2) non-real time and Layer 3 (L3) functions from BBU to CU, Layer 1 (L1)/L2 real-time functions from BBU to DU and the rest of L1 functions from BBU to RU. EPC functions are redistributed among the next generation core (NGC), CU and DU. Other distributions of functions between NGC, CU, DU and RU may also be possible, as will be discussed further in this supplement. The two new interfaces between CU and DU, and between DU and RU are called next generation fronthaul interface NGFI 2 and NGFI 1 [IEEE 1914.1], respectively, and the associated transport links are frequently called Fronthaul-II and Fronthaul-I [CMRI].

Note that 3GPP considers a split base station architecture consisting of CU and DU only. In this Supplement, we adopt an architecture consisting of three elements, CU, DU and RU, as this architecture offers more flexibility for distributing functionalities and for meeting mixed requirements on transport bandwidth and latency [IEEE 1914.1].

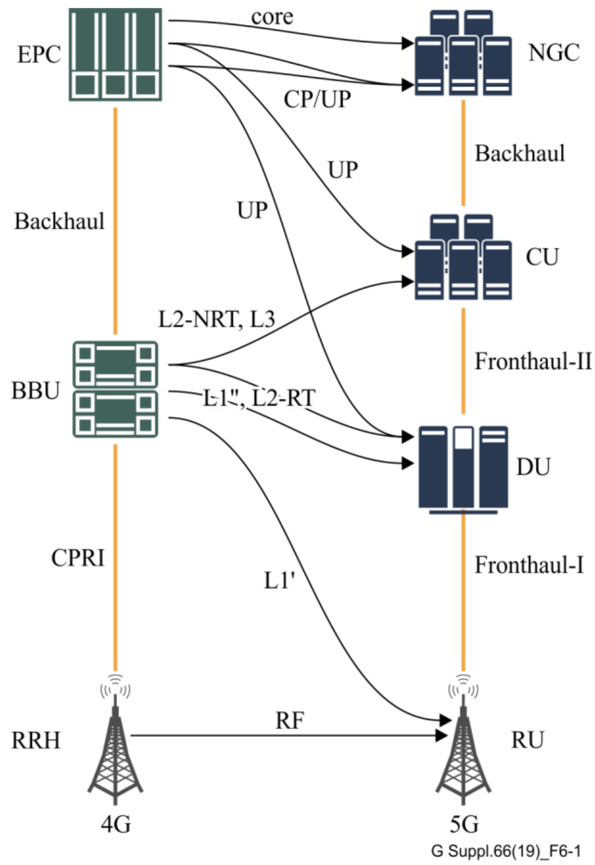


Figure 6-1 – Evolving from single-node BBU in 4G to split function architecture in 5G [CMRI]

6.2 Functional split architecture

In both the upstream and downstream directions, the radio signals go through a series of signal processing blocks. Figure 6-2 shows these functional blocks and potential split points in both 4G and 5G wireless networks [3GPP TR 38.801].

It is important to mention that traditional fronthaul at option 8, using CPRI [CPRI] or OBSAI protocol, requires continuous bitrate transport regardless of user traffic being present or not. However, with the other split points, (options 1-7), the amount of data to be transported scales with the user traffic. More detail of the requirements for different split options will be discussed in clause 7.

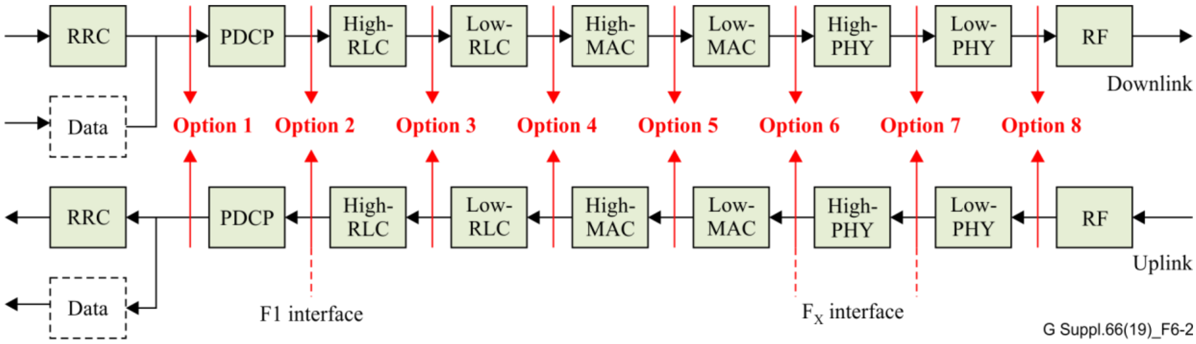


Figure 6-2 – Signal processing chain of 4G and 5G wireless base stations and optional split points [3GPP TR 38.801]

6.2.1 Conventional fronthaul in 4G wireless network

Conventionally in 4G wireless network, the fronthaul link is between RF and the remaining L1/L2/L3 functions using CPRI/OBSAI protocol (option 8 split point). This split point option allows the centralization of all high layer processing functions, at the expense of the most stringent fronthaul latency and bandwidth requirements.

This conventional fronthaul is based on transport of digitized time domain IQ data. For very high capacity applications, such as eMBB (enhanced mobile broadband) or for radio sites with many independent antenna elements (massive MIMO or multi-layer MIMO), these fronthaul solutions require unreasonably high transport capacities, while allowing for transport latencies between RU and DU/CU (Option 8 split interface) of only up to a few hundred microseconds.

Table 6-1 shows the approximate data rates for time domain IQ data fronthaul (CPRI rates without line coding) needed to support various radio frequency bandwidths and numbers of antenna ports in wireless networks using parameter ranges given by 3GPP in [3GPP TR 38.801].

Table 6-1 – Required fronthaul data rate in 5G wireless network [3GPP TR 38.801]

Number of antenna ports	Radio channel bandwidth			
	10 MHz	20 MHz	200 MHz	1 GHz
2	1 Gbit/s	2 Gbit/s	20 Gbit/s	100 Gbit/s
8	4 Gbit/s	8 Gbit/s	80 Gbit/s	400 Gbit/s
64	32 Gbit/s	64 Gbit/s	640 Gbit/s	3 200 Gbit/s
256	128 Gbit/s	256 Gbit/s	2 560 Gbit/s	12 800 Gbit/s

Equation 1 shows that transmission over the option 8 interface requires a CPRI net data rate of 491.52 Mb/s per 10MHz radio bandwidth and per antenna port [Doetsch].

$$B_{CPRI} = A \cdot f_s \cdot b_s \cdot 2 \cdot (16/15) \tag{1}$$

Here, A is the number of antennas; f_s represents the sample rate (15.36 MS/s per 10 MHz radio bandwidth) and b_s the number of bits per sample (15 for LTE, expected to remain valid also for 5G eMBB services). The remaining factors take into account the separate processing of I and Q samples (factor 2), and the additional overhead information in the CPRI frames (factor 16/15).

6.2.2 New functional split options in 5G wireless network

The increase in data rates in 5G makes it impractical to continue with the conventional CPRI fronthaul implementation. Moving towards a higher layer split (Figure 6-2) would relax the latency and bandwidth requirements, but then fewer processing functions can be centralized (see clause 7). It is thus critical that the new functional-split architecture takes into account technical and cost-effective trade-offs between throughput, latency and functional centralization.

Several standards bodies have hence moved to identify different split points in the radio processing chain (Figure 6-2) that allow for substantial reduction of the transport capacities in C-RAN architectures compared to the current approach. The choice of optimal 5G NR split points depends on specific deployment scenarios. In April 2017, 3GPP announced the selection of option 2 (PDCP/high RLC) as the high layer split point (F1 Interface) [TS 38.470], while staying open for any of the low layer split points (option 6 for MAC/PHY split or option 7 for intra-PHY split) [TR 38.816]. Here we use Fx as the generic notation for the low layer split points option 6 or option 7 for convenience. Cascaded split architecture is also considered to allow for additional flexibility.

In fact, the option 7 split point has been further diversified by several groups, both in view of different modes of cooperation between multiple radio sites, and in view of fixed network transport requirements.

In this clause, we consider the architecture models and split point definitions provided by four groups: 3GPP, eCPRI, xRAN (became the O-RAN Alliance in June 2018) and the Small Cell Forum.

We take the 3GPP architecture model and the main split options defined therein as the starting point for our discussion. Figure 6-3 maps sub-options defined by 3GPP [3GPP TR 38.801] to the split points defined in the eCPRI specification [eCPRI], in the O-RAN document [ORAN-FH] and in the Small Cells Forum document [SCF].

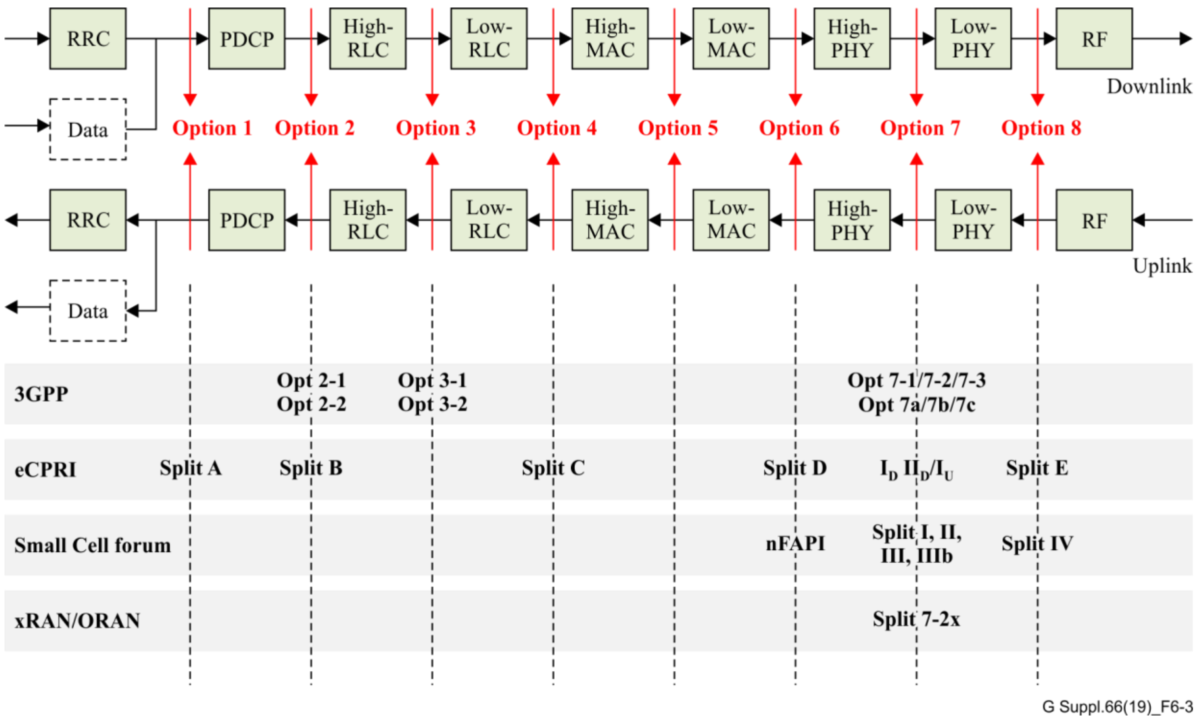


Figure 6-3 – Mapping different split points to the 3GPP model in 3GPP [3GPP TR 38.801], CPRI group [eCPRI], SCF [SCF] and O-RAN [ORAN-FH]

For the split option 7 (intra-PHY), multiple sub-options have been defined, warranting a separate, more detailed mapping, as shown in Figure 6-4.

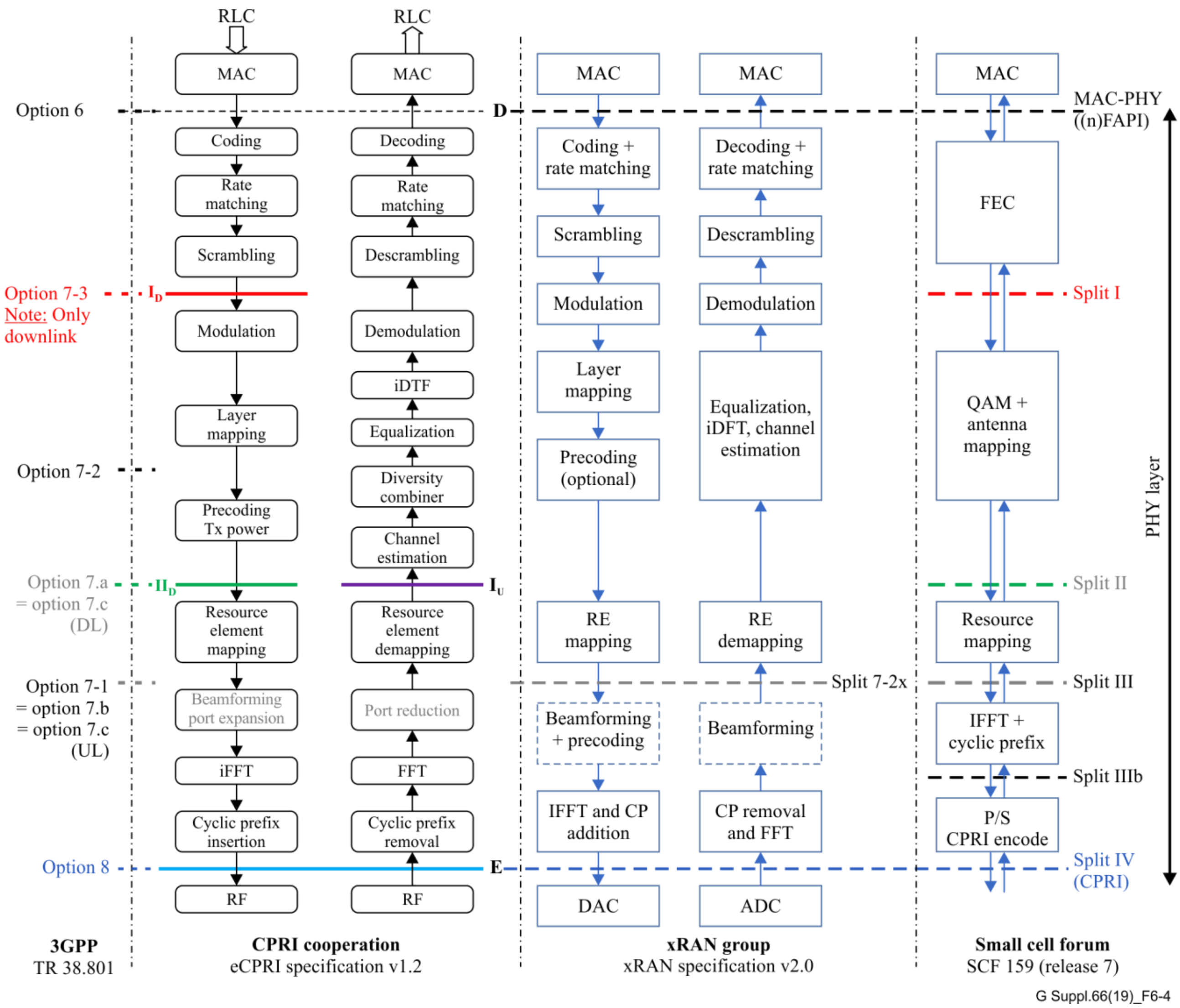
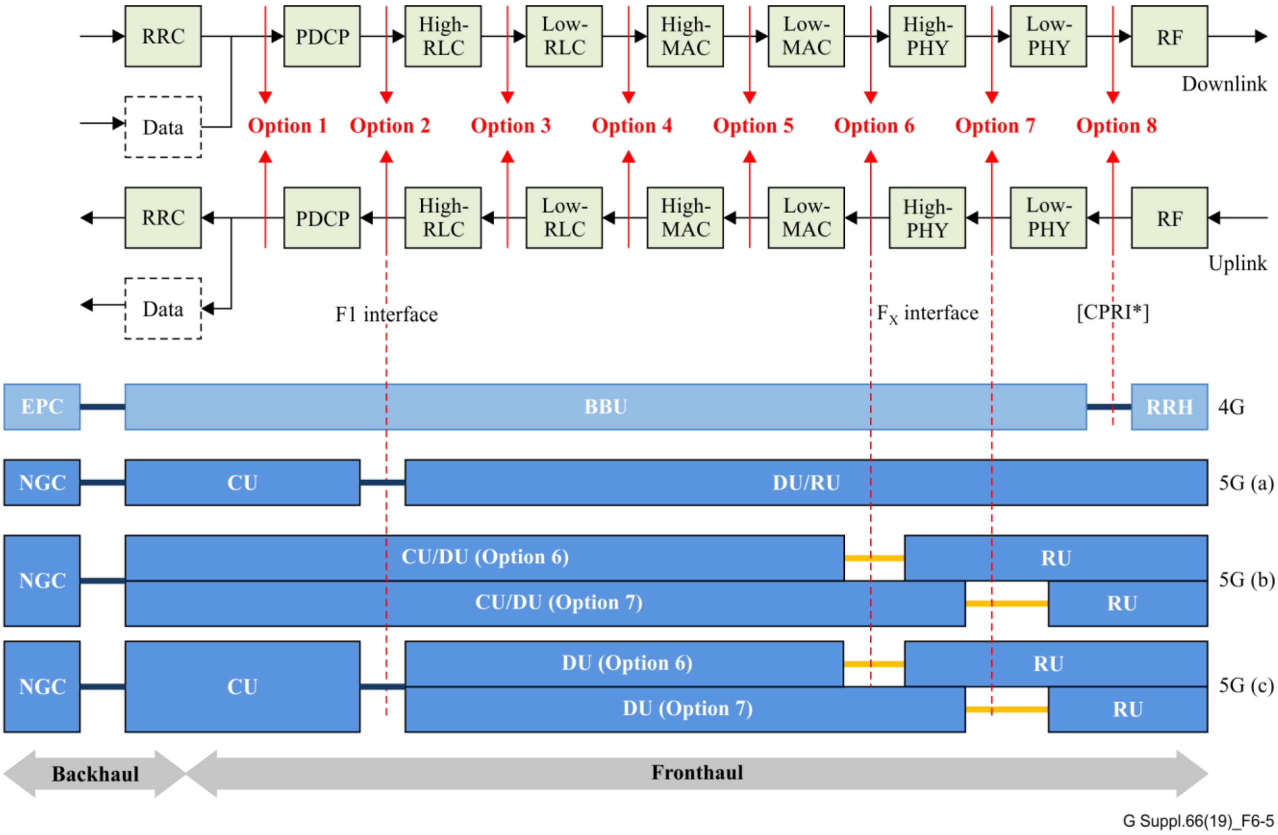


Figure 6-4 – Detailed comparison and mapping of option 7 sub-splits

It must be noted that the options 7a, 7b, 7c used in Table A.1 in Annex A of [3GPP TR 38.801] are not equivalent to options 7-1, 7-2, 7-3 in the main body of [3GPP TR 38.801]. The former had been used only in an early phase of the 3GPP discussion (cf. references in [3GPP TR 38.801]) and hence, going forward, are not used in this document. For the calculation of transport network capacities for options 6 and 7, the eCPRI, O-RAN and/or Small Cell Forum split points shall be used instead.

Possible mappings of the functional split options F1 and Fx to the CU/DU/RU architecture are illustrated in Figure 6-5. Each of the three elements, CU, DU and RU can host any of the signal processing functions.



* CPRI is one possible transport protocol for option 8.

Figure 6-5 – Mapping of CU/DU/RU functions according to the split points. 5G(a) high layer split (F1); 5G(b) low layer split (Fx); 5G(c) cascaded split.

6.3 Deployment scenarios

In general, there are two deployment scenarios for 5G NR fronthaul network depending on the location of the distributed unit (DU):

– **Centralized RAN (C-RAN)**: DU is centralized in an access convergence room or small access room, as shown in Figure 6-6(a). In this scenario, as the distance between DU and RU is typically 10 km or less, using point-to-point fibre, direct connection would need a large number of trunk fibre resources. Hence, the use of an efficient transport system, such as PON, can reduce the demand for trunk fibre resources.

– **Distributed RAN (D-RAN)**: DU is deployed in in the base station room, or RU/DU/CU are integrated and deployed in a base station as shown in Figure 6-6(b). In this scenario, the distance between DU and RU is generally very short such that a direct point-to-point fibre connection is suitable for fronthaul transmission.

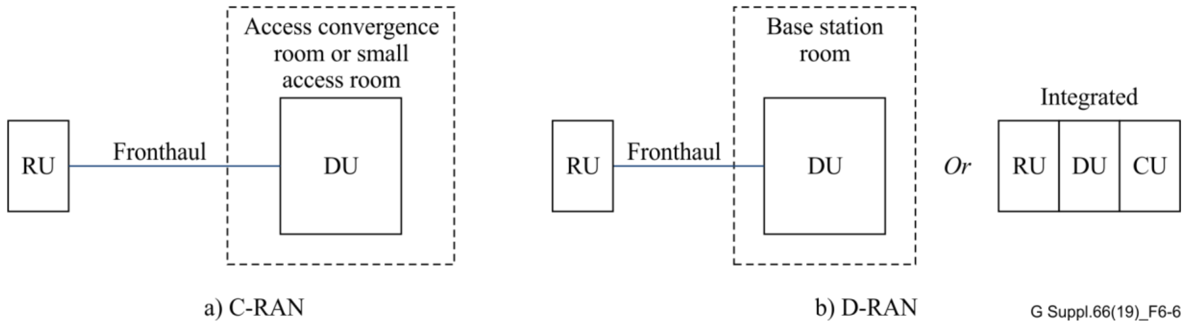


Figure 6-6 – RAN deployment scenario schematic diagram

Furthermore, the C-RAN deployment can be divided into two categories: large concentration and small concentration, as shown in Figure 6-7. For the large concentration mode, DU is generally deployed in the access convergence room, whereas for the small concentration mode, DU is in small access rooms.

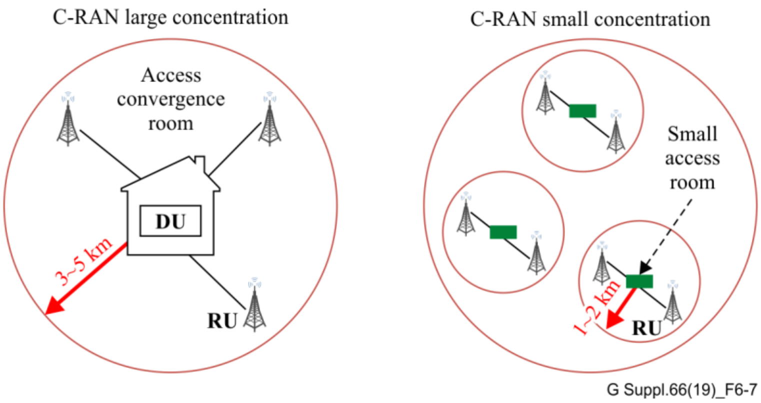


Figure 6-7 – Two categories of C-RAN deployment

7 Requirements for 5G fronthaul transport

In this clause those aspects of 5G NR that affect the requirements for the fronthaul transport layer are discussed. After a short description of the RAN and service level background, bandwidth, latency, synchronization/jitter and OAM requirements are addressed in more detail.

7.1 RAN and service level background

3GPP considers RAN architectures that include both 4G and 5G radio access technologies (RAT) coexisting and cooperating in one common network [3GPP TR 38.801]. Aside from this mixed network architecture, 5G networks alone will comprise a variety of services with substantially different traffic characteristics (Table 7-1) [ITU-R M.2083], as well as a variety of radio technologies with substantially different RF configurations (below 6 GHz, above 6 GHz, massive MIMO, multi-layer MIMO). However, not all of these services and technologies will necessarily be provided and used at the same time in the same network. For example, massive machine type communication (mMTC) applications will in many cases be used in closed networks, such as in manufacturing sites, and will not need to coexist with eMBB applications. In other scenarios, multiple services may use the same radio hardware (antenna and RF equipment), but the fronthaul transport may be different for different services, depending on traffic and latency requirements.

Table 7-1 – High-level overview of expected traffic characteristics for various 5G services, see for example Figure 3 in [ITU-R M.2083]

Radio technology	Peak rate	Average rate	e2e delay (service level)
Enhanced mobile broadband (eMBB)	5-10 / 20 Gbit/s (UL/DL)	100 Mb/s per user in urban/suburban areas 1-4 Gbit/s (hot spot areas)	10 ms
Ultra-reliable low latency communication (URLLC) / Critical machine type communication (incl. D2D)	much lower than in eMBB: N × Mbit/s	much lower than in eMBB: n × Mbit/s	1-2.5 ms
Massive machine type communication (mMTC)	much lower than in eMBB: N × Mbit/s	much lower than in eMBB: n × kbit/s - n × Mbit/s	1-50 ms

7.2 Transport bandwidth requirements

From a transport bandwidth perspective, the most important characteristic of the higher layer split options 1-7 is the fact that the amount of data to be transported scales with the user traffic on the air interface. Hence, the transport at these split points can benefit from statistical multiplexing gains in aggregating network architectures. By contrast, option 8 requires continuous bitrate transport at very high rates, whether user traffic is present or not. So, there is no opportunity to take advantage of statistical multiplexing.

To provide a rough insight into possible scaling of bit rates when going through the different split options, Table 7-2 shows the transport data rates calculated for a particular cell scenario [3GPP TR 38.801]. The parameters used for the evaluation are: 100 MHz radio bandwidth, 256-QAM modulation, 8 MIMO layers, 32 antenna ports (same for downlink and uplink), varying bit widths for IQ data representation for options 7a/b/c (7 to 16 bits per I and per Q sample).

It is important to note that the bit rates shown in Table 7-2 only apply for the very particular cell site configuration described above. The bit rates at one or more split points sensitively change with any modification of the cell site configuration and shall hence not be taken literally for other cell sites. They only serve for providing a raw impression of the orders of magnitude. More differentiated assessments of required transport capacities at the F1 and Fx split points are presented in clause 10.

Table 7-2 – Transport bit rates and latency ranges at different functional split interfaces, adapted from Annex A in [3GPP TR 38.801] (Note caveat in the text above)

Protocol split option	Required downlink bandwidth	Required uplink bandwidth	One way latency (order of magnitude)
Option 1	4 Gbit/s	3 Gbit/s	1-10 ms
Option 2	**4016 Mbit/s**	**3024 Mbit/s**	
Option 3	[lower than Option 2 for UL/DL]		
Option 4	4000 Mbit/s	3000 Mbit/s	100 to few 100 µsec
Option 5	4000 Mbit/s	3000 Mbit/s	
Option 6	**4133 Mbit/s**	**5640 Mbit/s**	
Option 7a	**10.1-22.2 Gbit/s**	**16.6-21.6 Gbit/s**	
Option 7b	**37.8-86.1 Gbit/s**	**53.8-86.1 Gbit/s**	
Option 7c	**10.1-22.2 Gbit/s**	**53.8-86.1 Gbit/s**	
Option 8	157.3 Gbit/s	157.3 Gbit/s	

The numbers for the example in Table 7-2 represent peak values for the required transport bit rates under optimal conditions on the radio channel. They are calculated using 4G models for one particular cell configuration as described above. With focus on split options 2, 6 and 7 (bold numbers), the transport bit rate for this particular example:

– increases by less than one percent for split option 2 (F1 interface) as compared to option 1 (backhaul);
– decreases by 2–15 times for split option 7.a/b/c as compared to option 8 (conventional fronthaul) and decreases by 30-40 times for option 6.

It must be noted, however, that in general there is not a fixed ratio of transport bandwidth between different split options. For instance, if in the above example each antenna port served one MIMO layer (i.e., one individual user data stream per antenna) instead of 32 antennas serving only 8 MIMO layers altogether (i.e., 8 individual user data streams per 32 antennas), then the backhaul capacity would be 16 Gbit/s in the downlink, instead of 4 Gbit/s as shown in the table. But still, the aggregate rate at split Option 8 will remain at 157.3 Gbit/s (CPRI rate without line coding).

In realistic deployments, the throughput on the air interface changes with the actual conditions of the radio channel (interferences, reflections, environmental conditions, etc.). This throughput variation will in turn require varying fronthaul transport capacities at the different split options (except for option 8). Regardless of these details, the analysis shown in Table 7-2 provides a useful qualitative indication of how the chosen functional split will affect the required bit rates on the transport layer.

7.3 Latency requirements

In the following next generation fronthaul latency requirements as given by the lower layer split options in radio architectures, and separately from that, latency requirements from the services provided by the radio network are discussed.

RAN related latency requirements

With regard to latency the transport at the interfaces of split options 1-8 falls into two categories:

– "non-real-time" transport with latencies in the range of several milliseconds (options 1-3);
– "real-time" transport with latencies in the range of a few hundred µsec (options 4-8).

In LTE networks, the differentiator between categories is whether the hybrid automatic repeat request (HARQ) loop crosses the split interface or not. If it does, then the µsec requirement applies; if it does not, then transport latency is specified solely by the requirements of the application layer which is typically in the millisecond range. In 4G networks, the HARQ loop is a synchronous process, its duration being strictly linked to 8 times the transmission time interval (TTI) length of 1 ms, i.e., 8 ms round trip time (RTT). Taking into account typical hardware and software implementations of today, there are usually only a few hundred microseconds RTT left for transport at the option 8 interface.

NOTE – Different values for this transport latency figure are found in the literature; also in [3GPP TR 38.801] one-way latency of 250 µs is mentioned as an example. There is, however, no clear specification given in standards. The value only depends on the specific vendor implementation of the HARQ loop.

In 5G networks, the subframe length of 1 ms can be divided into 1, 2, 4, … 32 slots (instead of the TTI in 4G), each thus lasting 1 ms or as short as 31.25 µs [3GPP TS 38.211]. Besides this, the HARQ process will be changed from synchronous to asynchronous HARQ.

The low layer split latency values have been addressed in two recently released documents:

– In [eCPRI] fronthaul transport networks are categorized according to supported data flow latency classes, with maximum one-way frame delays of 25, 100, 200 and 500 µs, respectively.
– The O-RAN group has taken an approach in which the latency is derived from the processing capabilities of the radio equipment at either end of the fronthaul 1 link [ORAN-FH]. The equipment is categorized into different classes, depending on the combination of the equipment, the residual latencies can be as large as 350 µs or even higher (one way).

Service related latency requirements

Latencies of 5G services are addressed in several industry reports, white papers and standards:

– In Table 1 of NGMN 5G White Paper [NGMN]: user experience requirements are collected including latency expectation;

– In Table 7.2.2-1 of TS 22.261, V16.1.0 [3GPP TS 22.261]: performance requirements for low-latency and high-reliability scenarios;

– In clauses 7.4, 7.5, 7.6 of TR 38.913, V14.3.0 [3GPP TR 38.913]: latency for control plane, user plane, and infrequent small packets;

– In clause 4.7 of ITU-R M.2410-0 (11/2017) [ITU-R M.2410-0]: user plane latency and control plane latency.

Although these documents discuss many similar aspects, they are not completely aligned in the latency requirements. The only alignment is for eMBB services, where the latency requirement of user plane is agreed to be 4 ms (one way). There is no common latency requirement for mMTC and URLLC services.

7.4 Synchronization requirements

The time synchronization requirements relevant to mobile networks are provided in [ITU-T G.8271] Tables II.1 and II.2.

Analysis on the deployment of these requirements to the fronthaul scenarios is under study for [ITU-T G.8271.1].

The analysis below is based on the work in CPRI group [eCPRI]. Some of the aspects indicated below may need to be revised to take into consideration the results of the [ITU-T G.8271.1] analysis. The readers should check these references for any updates.

The eCPRI group, adopting several 3GPP and ITU-T standards, specifies four Classes of timing accuracy requirements [eCPRI], as shown in Table 7-3. Definitions for the parameters in Table 7-3 are illustrated in Figure 7-1. Among these four classes, Class A+ and Class B are applicable for the inter-cell site scenario, which could be transported over a PON system, thus are relevant to our discussion. Note that if there are multiple hops (E-O-E conversion) in the inter-cell link, then the timing accuracy requirement could be even more stringent than what is described in Table 7-3.

As indicated in Table 7-3, timing error requirements for the transport network range from 20 ns (Class A+, Case 2) and 100/190/200 ns (Class B). Class A+ Case 2 is for supporting emerging new use cases such as high accuracy positioning services.

Table 7-3 – Timing accuracy requirement as specified in eCPRI [eCPRI]

Category (note 1)	Time error requirements at UNI, |TE|			3GPP Time alignment error (TAE) requirements at antenna ports
	Case 1 (note 2)		Case 2 (note 3)	
	Case 1.1 (note 4)	Case 1.2 (note 5)		
A+	N.A.	N.A.	20 ns (relative)	65 ns (note 6)
A	N.A.	60 ns (relative) (note 7)	70 ns (relative)	130 ns (note 6)
B	100ns (relative) (note 7)	190 ns (relative) (note 7)	200 ns (relative)	260 ns (note 6)
C (note 8)	1100 ns (absolute) (note 9)		1100 ns (absolute) (note 9)	3 us (note 6) (note 10)

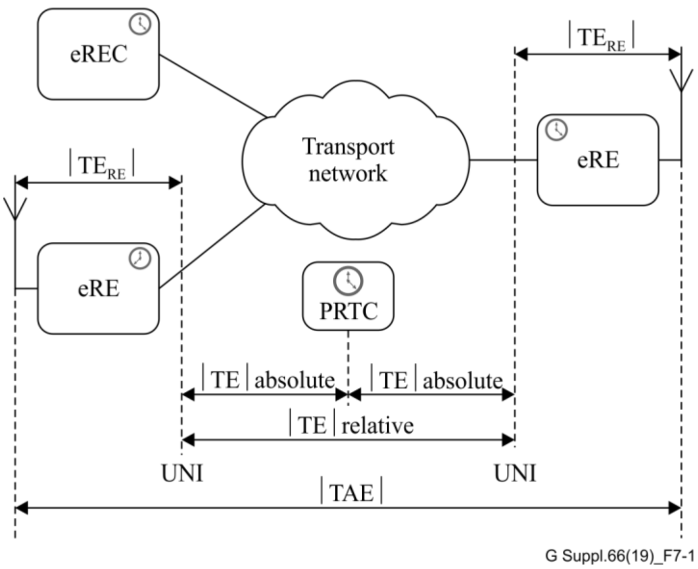


Figure 7-1 – Timing accuracy definitions [eCPRI]. eRE: eCPRI radio equipment. eREC: eCPRI radio equipment control

7.5 OAM functions

Traditionally in 4G LTE, OAM functions are provided natively by Ethernet for the backhaul link, but for the fronthaul link, there is no monitoring function in CPRI protocol. This is not a concern as BBU and RRH are collocated at the cell site and managed by the same mobile operator.

In the 5G NR architecture, network centralization will require coordination between the radio network layer and the transport network layer, which are likely under different jurisdictions – mobile operator and fixed network operator. OAM functions, such as detection of faults and performance degradation, will need to be implemented and coordinated so that each operator will be able to monitor its own network segment and hand over the information to the other.

7.5.1 OAM requirements from the wireless perspective

The wireless OAM functions are mainly of interest to the wireless operator. As long as the transport network can carry them transparently, there is no further interaction needed. However, the following text gives a short summary of the kinds of OAM implemented in OBSAI and eCPRI.

OBSAI

The OBSAI system specification document [OBSAI], published in 2006, described high-level concepts of OAM and provisioning (OAM&P) functionalities of the OBSAI base transceiver stations (BTS). The specifications support only BTS internal functions and interfaces and provide a set of expected OAM&P capabilities:

– Configuration management functions for enabling deployment of OBSAI BTS modules;

– Fault Management capabilities for monitoring the health of OBSAI BTS modules;

– Performance management functions for collecting performance metrics of OBSAI BTS modules;

– Software management functions related to installing and updating software.

eCPRI

The latest eCPRI document does not specify network connection maintenance and control, but provides a number of methods and standards that can be used [eCPRI]. There are basically two methods:

1) Ethernet OAM: use monitoring elements natively provided by Ethernet in the backhaul network according to IEEE 802.1Q (Ethernet connectivity fault management) [IEEE 802.1Q] and ITU-T G.8013/Y.1731 (OAM functions and mechanisms for Ethernet based networks) [ITU-T G.8013].

2) Through messages defined by eCPRI: Internet Control Message Protocol (ICMP) for IPv4 and IPv6, and in particular, message types 2,3,6,7 and/or 64-255 in Table 7-4.

Table 7-4 – eCPRI message types (reproduced from Table 4 in [eCPRI])

Message Type #	Name	Section
0	IQ Data	3.2.4.1
1	Bit Sequence	3.2.4.2
2	Real-Time Control Data	3.2.4.3
3	Generic Data Transfer	3.2.4.4
4	Remote Memory Access	3.2.4.5
5	One-way Delay Measurement	3.2.4.6
6	Remote Reset	3.2.4.7
7	Event Indication	3.2.4.8
8 - 63	Reserved	3.2.4.9
64 - 255	Vendor Specific	3.2.4.10

7.5.2 OAM requirements from the transport perspective

In general for the transport link, from the perspective of service level agreement (SLA), OAM should include the following parameters:

– Optical link monitoring: optical transceiver parameters, channel discovery and registration, wavelength allocation, loop-back test, etc.

– Performance monitoring: throughput, frame loss rate, fault identification and management, availability, latency, jitter, etc.

Two concepts for implementing OAM functions in the fronthaul link are discussed in [COMBO]. Although in [COMBO] they mainly focus on the CPRI/eCPRI protocol, the concepts are applicable for other functional split interfaces as well.

1) Out-of-band monitoring using a dedicated management channel that does not share bandwidth with the payload (different from the antenna site management channel);

2) In-band monitoring within the overhead of protocols such as Ethernet, OTN, light-weight CPRI framing, or PON. In-band schemes do share bandwidth with the payload.

In a hierarchical network, the upper layer's overhead becomes the lower layer's payload, so the usage of out-of-band and in-band needs to take that into consideration.

Another concern is that the OAM signalling should not introduce too much latency into the fronthaul network. It is recommended to insert OAM messages in low layers near the physical line, e.g., PMD, PCS, or MAC layers, instead of the IP or upper layers.

8 PON architecture for 5G fronthaul transport

In this clause how a passive optical network (PON) can support 5G NR fronthaul and how the radio network elements (CU, DU and RU) can be mapped to the transport elements in PON (OLT, ONU) are discussed. Both 3GPP and IEEE describe the concept of layered network architecture in [3GPP TS 38.401] [IEEE 1914.1]. According to their definition, CU/DU/RU belong to the radio network layer (RNL), while OLT/ONU belong to the transport network layer (TNL). This concept is illustrated in Figure 8-1.

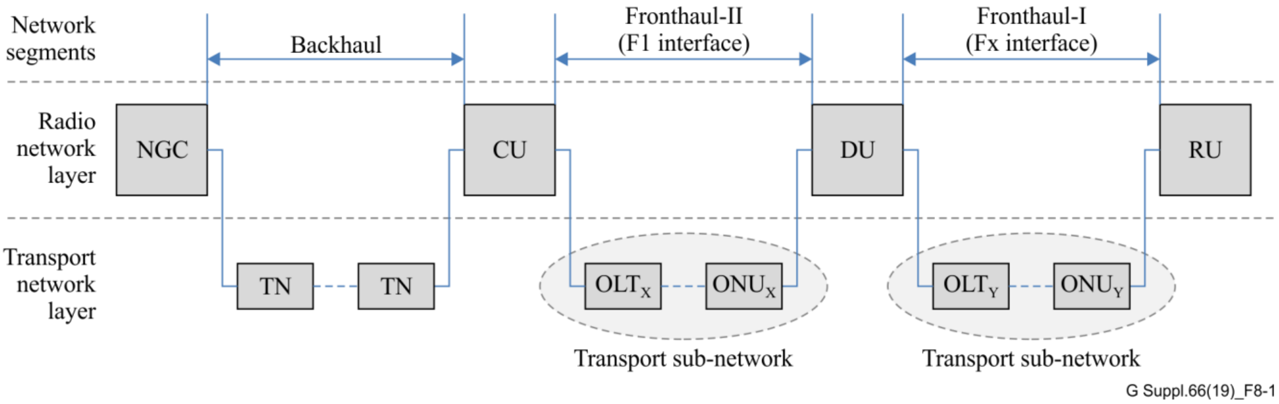


Figure 8-1 – Concept of layered structure showing RNL (CU, DU, RU) and TNL (OLT, ONU). TN: Transport network [IEEE 1914.1]

Figure 8-2 shows four example scenarios of mapping the CU/DU/RU to OLT/ONU. Here the F1 and Fx interfaces are used as examples only.

High layer split over the F1 interface is represented in Figure 8-2(a), which is applicable for very low latency between DU and RU. Low layer split over the Fx interface is represented in Figure 8-2(b). This scenario is applicable for general purpose cloud networks. Figure 8-2(c) represents the cascaded split scenario and is useful for small cell deployment in dense urban areas. Parallel split over both F1 and Fx interfaces is shown in Figure 8-2(d). This scenario is applicable for a mixed deployment to meet high and low latency requirements. The specific choice of the most suitable architecture would depend on specific deployment scenarios, as well as service-based latency and performance requirements. In addition, the ODNs in the basic structures in Figure 8-2 can be further expanded to include point-to-point (PtP), star, or tree topologies.

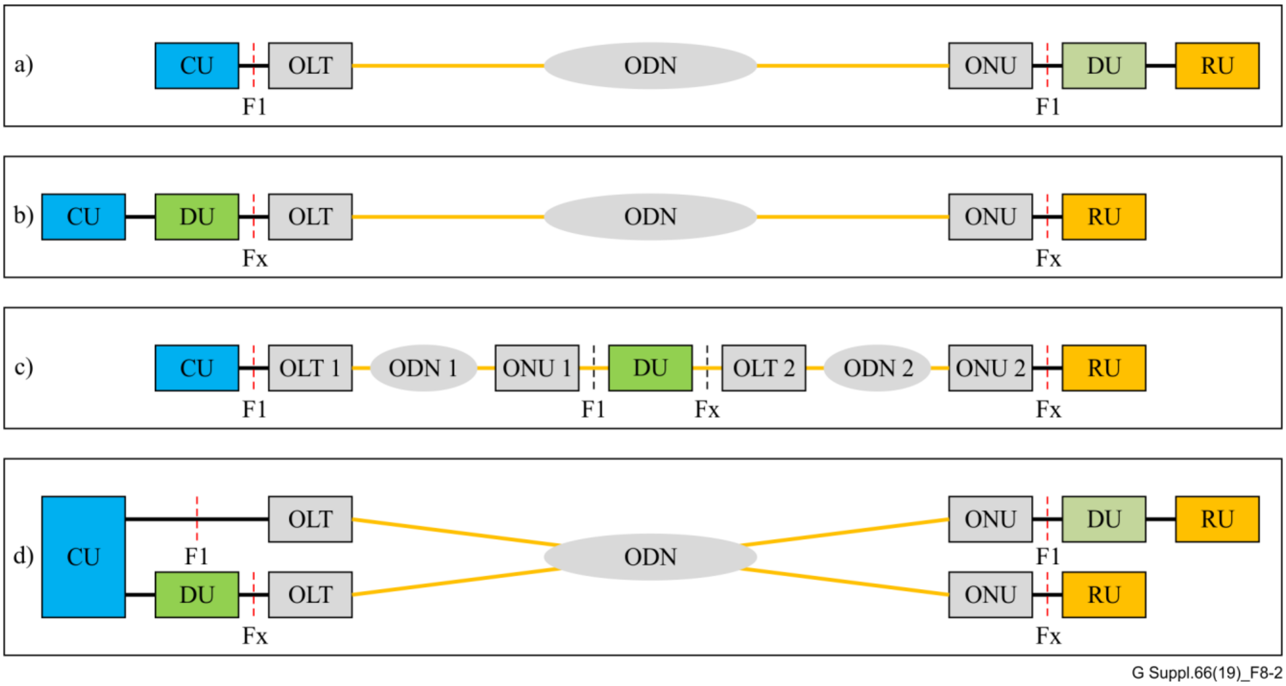


Figure 8-2 – Mapping of CU/DU/RU to PON in optical fronthaul architecture. (a) high layer split; (b) low layer split; (c) cascaded split; (d) parallel split

9 Practical PON system solutions

Depending on each operator's deployment requirements, in some situations, the coming 5G antenna site is already supporting 2G/3G/4G, so the same access fibre system has to support multiple F1 and Fx interfaces as below:

1) Several F1 (one per 5G carrier, which will be deployed sequentially) plus several Ethernet backhaul for 2G/3G/4G;

2) Several Fx (one per 5G RUs) plus several Ethernet backhaul for 2G/3G/4G (or 4G CPRI fronthaul).

In other situations, when 5G networks use separate fibre systems and fronthaul networks from legacy RANs, the above cases do not apply.

In this clause, six potential PON use cases and their challenges to support 5G fronthaul are discussed. Table 9-1 provides a summary of these use cases.

Table 9-1 – Summary of potential use cases

Clause #	Associated scenario	Functional split supported	Services
9.1.1	Legacy TDM PON (for fixed access services) with PtP WDM overlay for both fixed and wireless services	Fx	Low latency services, cloud/ virtualization
9.1.2		F1`	More latency tolerant
9.1.3		F1, Fx	Mixed services
9.2.1	Dedicated PON for wireless services only	Fx	Low latency services, cloud/ virtualization
9.2.2		F1	More latency tolerant
9.2.3		F1, Fx	Mixed services

9.1 Legacy PON with WDM overlay

One immediate solution to support wireless fronthaul is to overlay new wavelengths in a legacy PON, without sharing bandwidth with legacy fixed access services. Both NG-PON2 TWDM and PtP WDM could be used for this scenario.

9.1.1 Low layer split

A schematic of the use case to support low layer split is shown in Figure 9-1. Since low layer split has very strict latency requirements, more wavelength resources are needed to reduce the bandwidth sharing between RUs when several ONUs/RUs share one TDM wavelength.

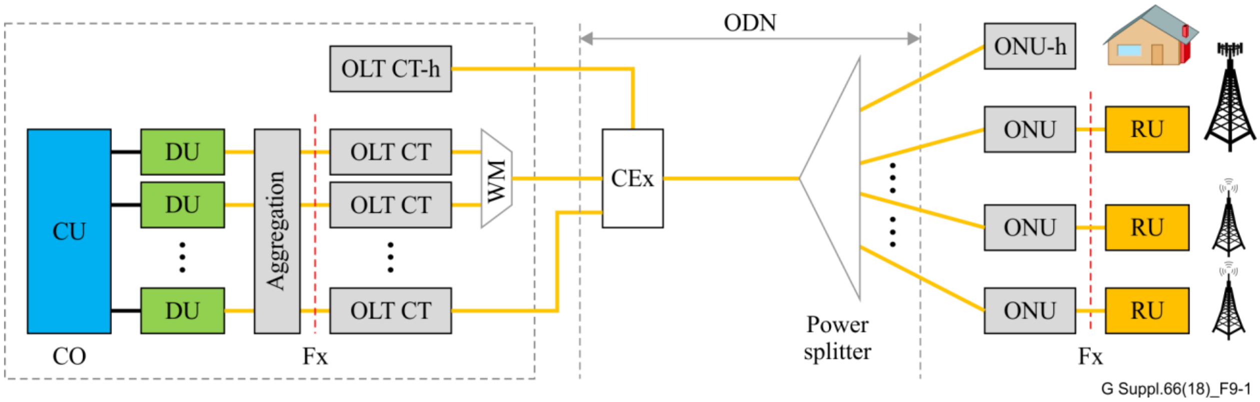


Figure 9-1 – Low layer split fronthaul for legacy PON with WDM overlay. Note that each OLT CT can support multiple ONUs, which is not shown in the figure for simplicity

9.1.2 High layer split

High layer function split is much more tolerant to latency and bandwidth sharing. A schematic of the use case to support low layer split is shown in Figure 9-2. The cascaded split can also be supported when a second ODN is appended as shown in the bottom part of the figure. Compared with low layer split, fewer wavelength resources are expected. However, fewer processing functions can be centralized in the central office.

Note that each OLT CT can support multiple ONUs, which is not shown in Figure 9-2 for the sake of simplicity.

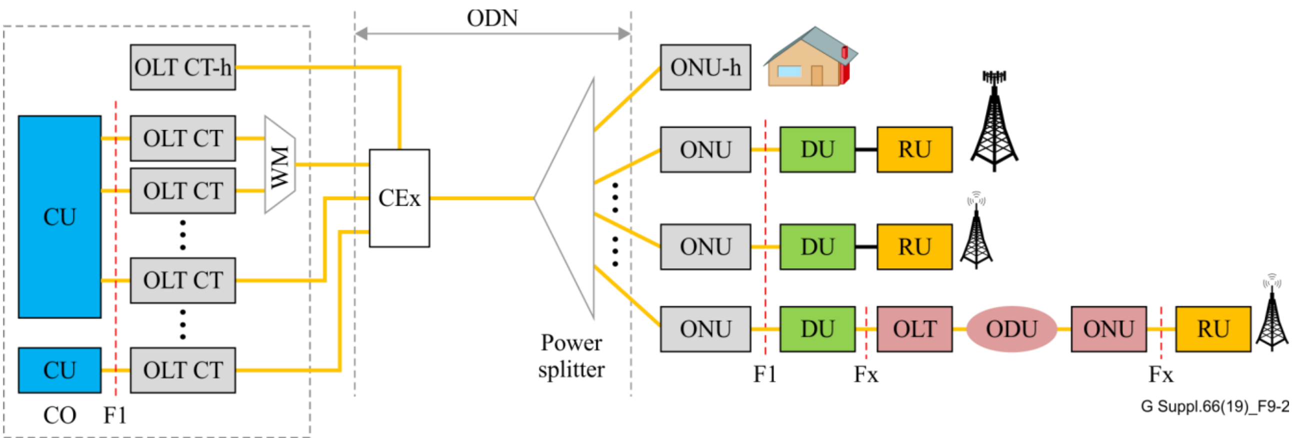


Figure 9-2 – High layer split fronthaul based for legacy PON with WDM overlay

9.1.3 Mixed high layer and low layer split

A mixed setup can be realized by WDM overlay, such as PtP WDM in NG-PON2, with both low layer and high layer splits. This heterogeneous setup allows the distributed RAN with sub-tended RU from a central site as shown in Figure 9-3. Note that the CUs could be separate units as shown in Figure 9-3 or share the same unit. The challenge of the mixed setup is to support the wide variation of latency and bandwidth requirements on the same PON. Note also that each OLT CT can support multiple ONUs, which is not shown in Figure 9-3 for the sake of simplicity.

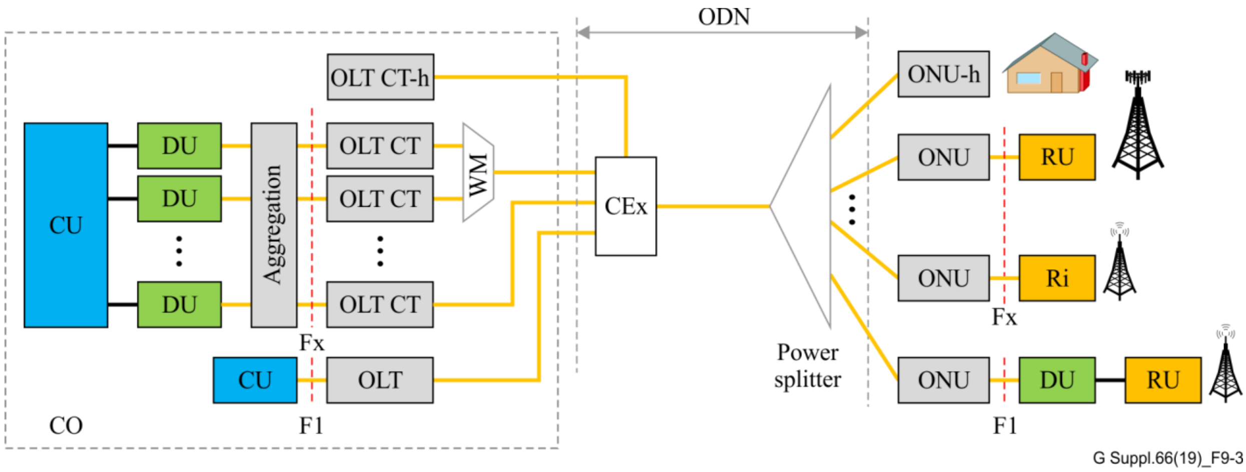


Figure 9-3 – High layer and low layer split fronthaul based for legacy PON with WDM overlay

9.2 Dedicated PON for wireless services

In order to avoid any degradation to fixed user services, a more practical scenario is to build dedicated PONs specifically for mobile fronthaul.

9.2.1 Low layer split

Figure 9-4 shows a use case for dedicated PON to support low layer split. Since low layer split has strict latency requirement, WDM-PON is a good candidate for this use case. A dedicated TWDM-PON would be more resource efficient due to its ability of statistical multiplexing. It would however need improved bandwidth allocation and ranging schemes. Note that each OLT CT can support multiple ONUs, which is not shown in Figure 9-4 for the sake of simplicity.

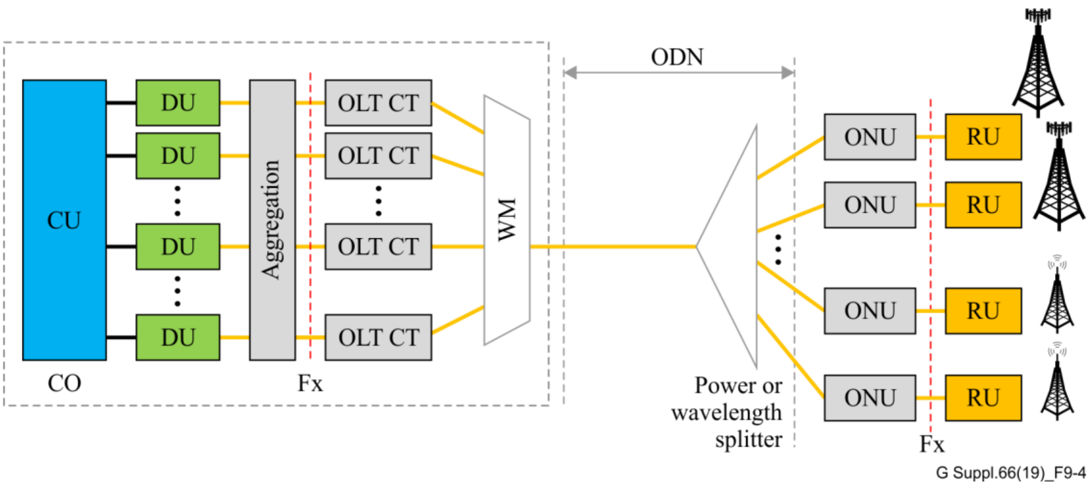


Figure 9-4 – Low layer split fronthaul based on dedicated PON

9.2.2 High layer split

High layer split can be supported by dedicated PON for wireless services as shown in Figure 9-5. Compared with low layer split, the requirement of both bandwidth and latency are much relaxed. Similar to the use case in clause 9.1.2, cascaded split can also be supported when a second ODN is appended as shown in the bottom part of the figure. Note that each OLT CT can support multiple ONUs, which is not shown in Figure 9-5 for the sake of simplicity.

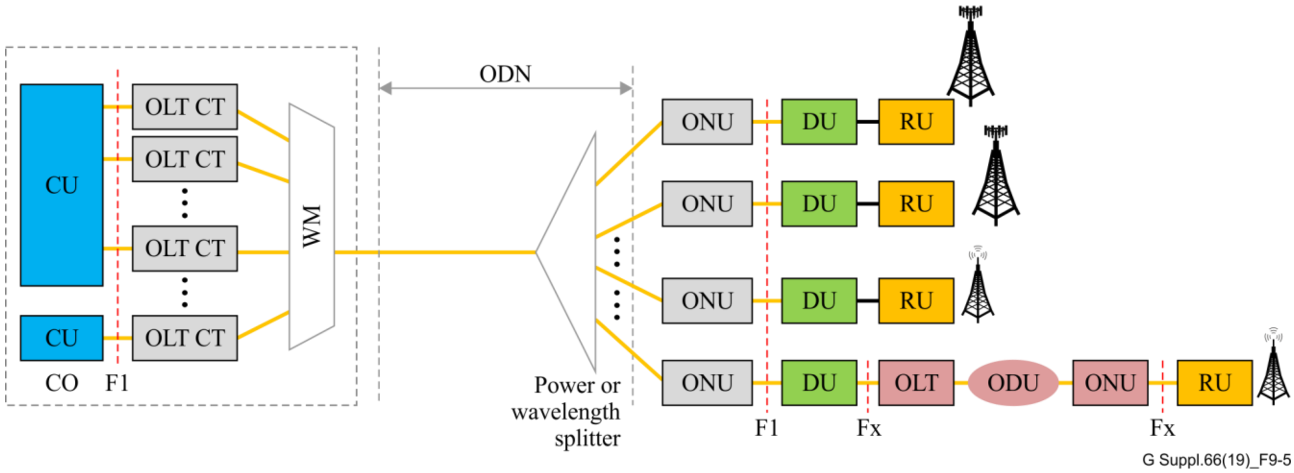


Figure 9-5 – High layer split fronthaul based on dedicated PON

9.2.3 Mixed high layer and low layer split

A PON with mixed low layer and high layer splits is possible when considering sub-tended RU from a central site as shown in Figure 9-6. However, the different requirements on bandwidth and latency could pose some challenges. Note that the CUs could be separate units as shown in Figure 9-8 or share the same unit. Note also that each OLT CT can support multiple ONUs, which is not shown in Figure 9-6 for simplicity.

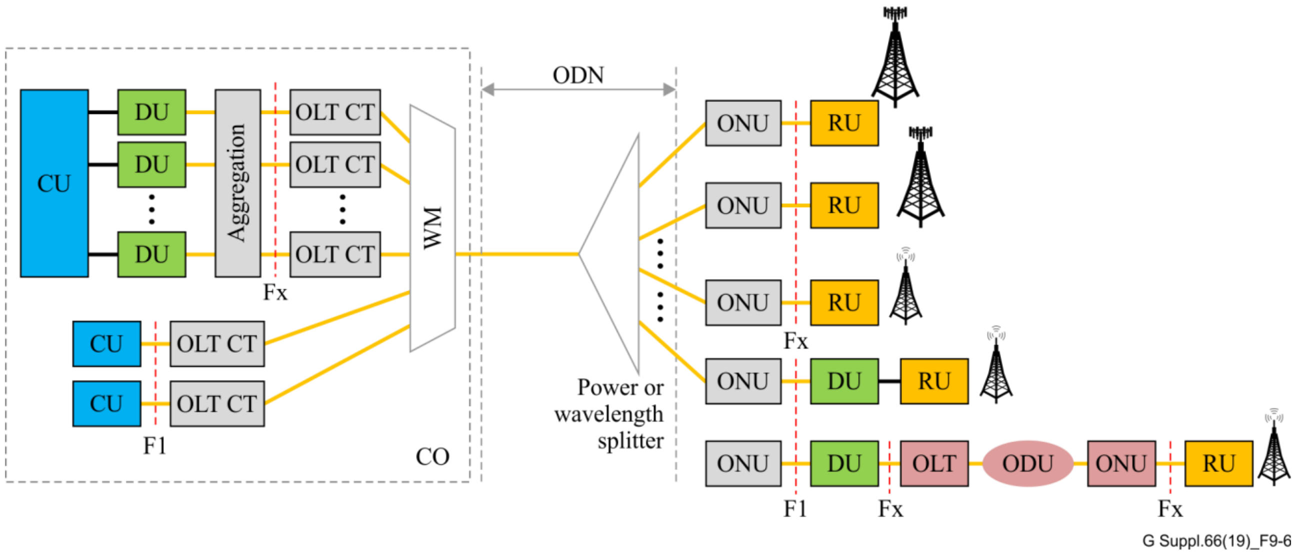


Figure 9-6 – Mixed Low layer and high layer splits fronthaul based on dedicated PON

9.3 PON system implementation examples

For the F1 and backhaul interfaces, TDM-PON with data rates over 10 Gbit/s should be sufficient to meet both the bandwidth and latency requirements. In a TDM-PON, the conventional dynamic bandwidth allocation (DBA) mechanism can cause delay in the order of ms, which is not compatible with delay sensitive 5G services, nor with latency sensitive fronthaul transport at the Fx interface.

Many latency improving mechanisms have been proposed to mitigate the issue, such as the cooperative DBA (CO-DBA) [Tashiro] or using a dedicated activation wavelength to eliminate the quiet window on the latency sensitive channel during ONU activation. On the other hand, WDM-PON does not require DBA or ranging, thus can be used for low latency services and fronthaul transport without modifications. This clause provides implementation examples of low latency TDM-PON, TWDM-PON and WDM-PON for 5G services. It should be kept in mind there is no one-size-fits-all implementation as the specific solution will depend on each operator's deployment needs.

9.3.1 TDM-PON with low latency bandwidth assignment

When TDM-PON is used for 5G fronthaul transport (see Figure 9-7), the downstream latency is low, whereas upstream latency could be in the order of several milliseconds, caused by the DBA cycle time. Each ONU must send a request to an OLT first, and then the OLT grants an upstream bandwidth to each ONU to avoid any upstream data collisions. For letting new ONUs join the system, the activation phase imposes a 250 µs delay (for 20 km ODN). The subsequent ranging phase also imposes a delay in the order of up to 200 µs. Note that these two delays do not usually add up, as the respective phases are separated in time. In order to use TDM-PON for low-latency demanding fronthaul transport, it is therefore necessary to reduce upstream latency.

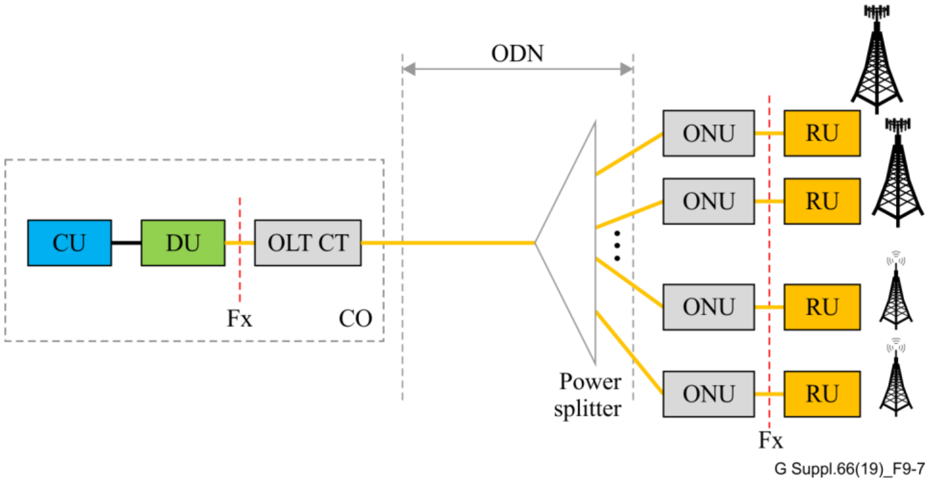


Figure 9-7 – TDM-PON for 5G fronthaul transport (example of low layer split fronthaul)

This clause introduces three potential latency reduction methods for TDM-PON: differentiated service class (clause 9.3.1.1), CO-DBA (clause 9.3.1.2), and dedicated activation wavelength (DAW) for quiet window elimination (clause 9.3.1.3). This clause only provides a high level overview (more details of the latter two methods are discussed in other ITU-T ongoing projects). Note that the DAW method is applied during ONU activation, while the other two methods are applicable after an ONU is successfully registered. The three methods can be used in combination with each other.

9.3.1.1 Differentiated service classes

One potential latency reduction method consists in differentiating service classes in the upstream direction, and to assign mobile traffic in highest priority [Lee]. For example, as shown in Figure 9-8, the service classes typically could be composed of fixed bandwidth, assured bandwidth, non-assured bandwidth and best-effort service. The fixed bandwidth class with highest priority reserves and cyclically allocates upstream bandwidth regardless of demand. On the other hand, the assured bandwidth class is similar to fixed bandwidth, but the bandwidth could not be given without demand. Thus, the mobile traffic generated from DU and/or RU can be connected the fixed bandwidth class and guaranteed bandwidth and low-latency upstream transmission are possible.

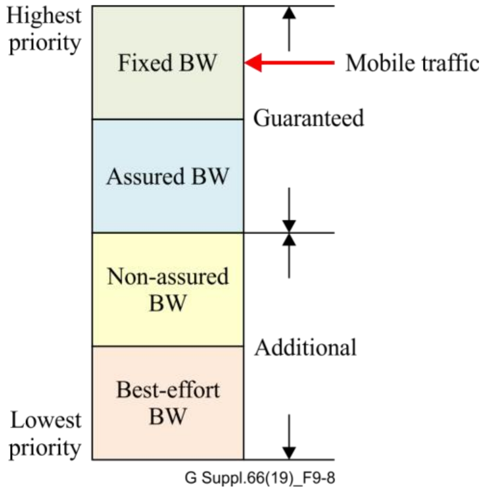


Figure 9-8 – Low latency scheduling for 5G fronthaul transport according to differentiated service classes

9.3.1.2 Optical-mobile cooperative DBA (CO DBA)

When applying the highest priority to mobile traffic (fixed bandwidth), the downside is that such bandwidth must be dimensioned to the peak rates, and any unused portion of this bandwidth cannot be reallocated to other nodes or to other services.

With DBA in TDM-PON, the OLT sends bandwidth allocation information to each ONU, and each ONU can send upstream data only in the allowed time slot. Traditional DBA methods take into consideration the dynamic upstream traffic and configured traffic contracts. They are done in a reactive way based on monitoring of the upstream traffic and buffer status report of each ONU. As a result, upstream data from DU and/or RU wait in the ONU until the completion of bandwidth allocation since optical and mobile equipment do not exchange information with each other.

9.3.1.2.1 CO DBA mechanism example for mobile fronthaul

With CO DBA, an information exchange is introduced between the mobile scheduler (CU/DU) and the PON scheduler (DBA) in the OLT, as shown in Figure 9-9. Mobile equipment (UEs) request required upstream bandwidth to the CU/DU, the CU/DU then sends allocation decisions back to the UEs and also signals corresponding information to the OLT. This allows the OLT to determine upstream bandwidth allocations in advance [Tashiro]. The OLT then applies these bandwidth allocations around the arrival time of upstream mobile traffic. In this way, it avoids the DBA to spend time to detect the presence of the mobile traffic by means of ONU buffer feedback. This enables low-latency upstream transmission.

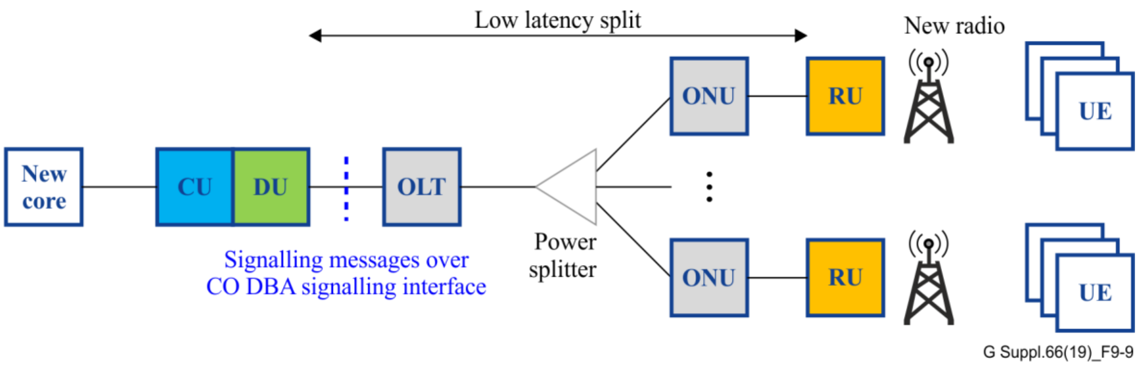


Figure 9-9 – CO DBA signalling interface for mobile fronthaul

Figure 9-10 shows an example of the mobile scheduling process time diagram, which consists of the following phases:

– UEs send data in a given 5G NR slot(N) and request for air interface capacity for a future slot (N+A, e.g., for LTE this is N+8).

– CU/DU takes scheduling decisions and notifies each UE about the allocated air interface resources to that UE for that future slot (N+A).
  - In parallel, CU/DU deduces the corresponding fronthaul traffic load per RU, based on the scheduling allocations to the corresponding UEs;
  - The CU/DU notifies the traffic load per RU for the given slot (N+A) to the OLT, by means of dedicated signalling messages that include an identifier for this traffic;
  - The OLT adapts its DBA for the given slot (N+A) for the T-CONT corresponding to the traffic identifier.

– UEs send uplink traffic over the air in the given slot (N+A), which is processed in RU and carried as fronthaul packets over the PON. The UEs also send requests for another future slot (N+2*A).

The repetition rate of the UE – CU/DU interactions is once every slot. The repetition rate of the CO DBA interaction between CU/DU and OLT depends on the variability of the traffic, with a maximum of once every slot.

Note that here RU refers to the logical interface connecting to the UNI in an ONU. In an actual implementation, it is possible to have multiple RU interfaces integrated in a physical RU equipment.

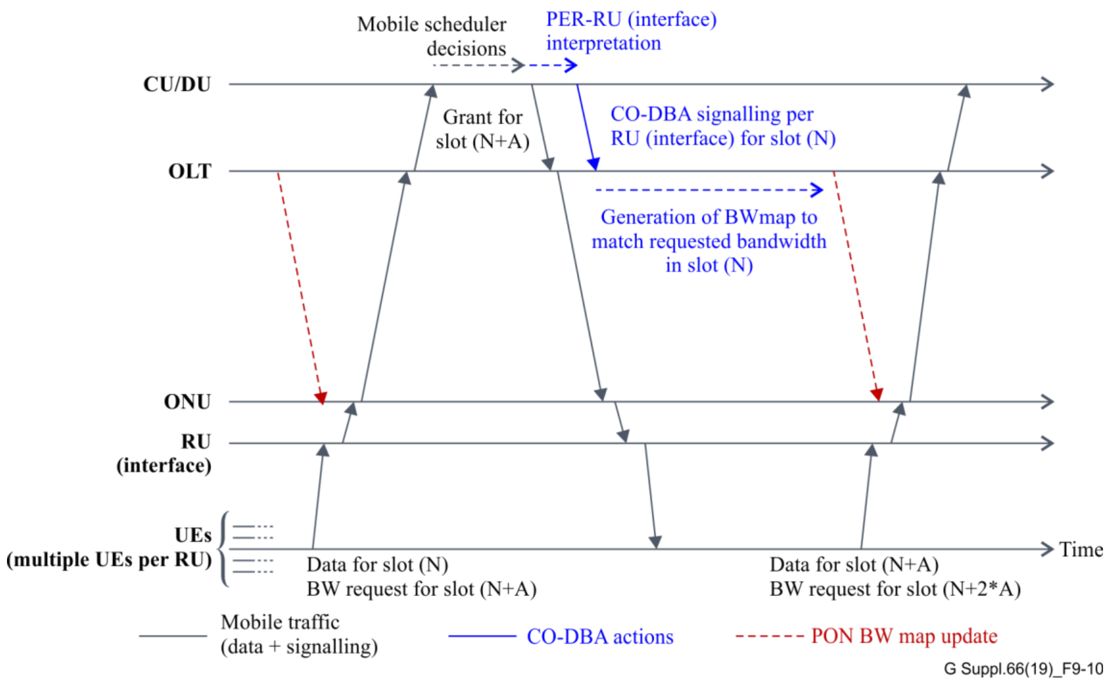


Figure 9-10 – An example of the optical-mobile cooperative DBA time diagram

9.3.1.2.2 CO DBA architecture examples for mobile fronthaul

When considering the possible variants, there are several basic cases of connectivity between ONU UNI(s) and RU interface(s) that can be encountered, as shown in Figure 9-11, considering that:

– each OLT can interact with multiple CU/DUs, and hence can connect to multiple CU/DUs;

– each CU/DU can have RUs over multiple PONs on multiple OLTs, and hence can connect to multiple OLTs;

– each PON can serve a mix of RUs pertaining to different CU/DUs;

– each RU pertains to only one CU/DU;

– each RU can have multiple interfaces, each interface connects to a ONU UNI.

There are four possible ways to connect RU interface(s) to ONU UNI(s), as indicated in Figure 9-11. In practice, some ways make more sense than others.

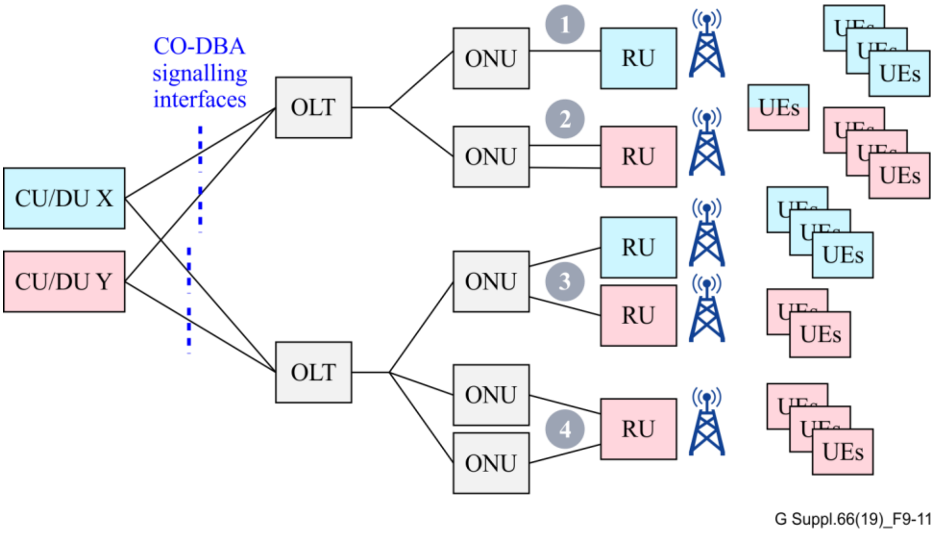


Figure 9-11 – Architecture variants for CO DBA

Additionally, CO DBA can support a mix of different low-latency services with different latency requirements to be supported on the same PON, and a mix of low-latency and non-low latency services on the same PON.

9.3.1.2.3 CO DBA functional roles

Major functional roles in CO DBA are as follows:

– CU/DU and OLT are connected by a logical CO DBA signalling interface, which can share the same physical interface as the data traffic.

– The CU/DU determines how much traffic will be required for given time intervals and given services, and the required maximal upstream latency of this traffic. The CU/DU communicates such reports in signalling messages to corresponding OLT(s) on which the corresponding RUs are connected to a PON via ONUs.

– The CU/DU equipment adds an ID per report to identify the service and its corresponding RU interface.

– The CU/DU updates this information to follow variations in expected bandwidth of the RU.

– The OLT accepts and parses these signalling messages, using the ID to link a report to its corresponding T-CONT. The OLT adapts PON bandwidth allocations according to the reports in the signalling messages.

– The OLT calculates the unused period of upstream signals for each mobile scheduling cycle by using wireless bandwidth requirements of each ONU received over CO DBA signalling interface. When this unused time is enough for the quiet window, the OLT opens it.

9.3.1.2.4 CO DBA with O-RAN cooperative transport interface (CTI)

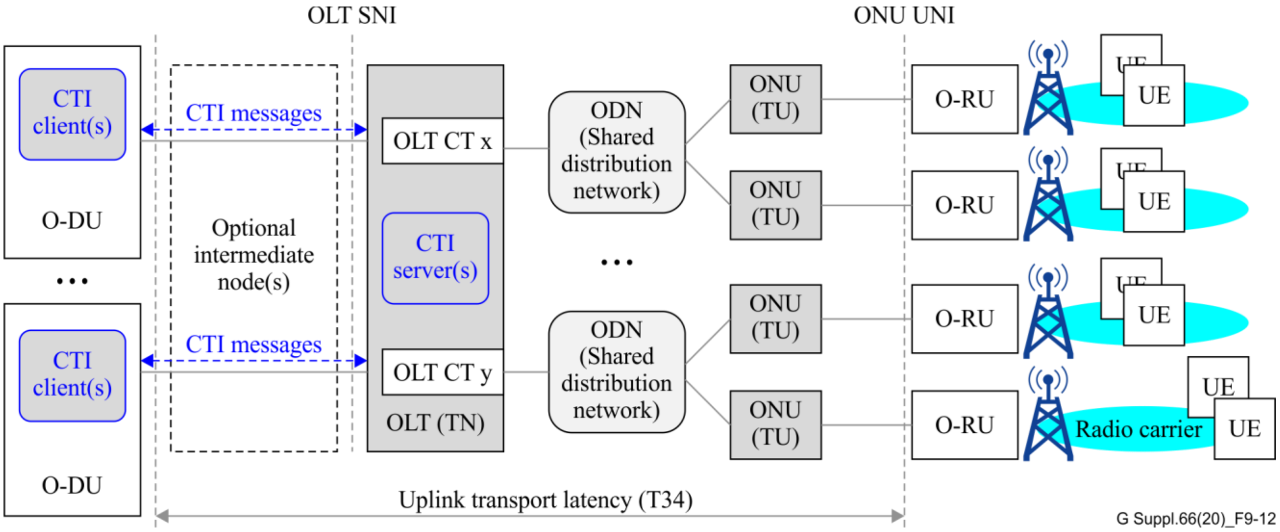
For the transport of mobile fronthaul traffic, the signalling interface for CO DBA has been specified by the O-RAN alliance in the cooperative transport interface (CTI) [ORAN-CTI-TC].

This clause describes the framework of the CTI and gives an example of how the OLT can execute the CO DBA to reduce the latency of upstream signals due to opening the quiet window. In the example the terminology of O-RAN is adopted by using O-DU and O-RU in place of DU and RU, respectively.

The [ORAN-CTI-TC] specifies the protocol and exchange of messages between a CTI server at the transport node (TN) side and a CTI client at the O-DU side. One or multiple CTI clients send bandwidth reports to one or multiple CTI servers. The TN connects transport units (TUs) over shared distribution networks. The TUs are connected to the O-RUs.

In the case of PON based access networks, the TN represents the OLT node containing one or multiple TDM or TWDM PON channel terminations (CTs), connecting one or multiple ONUs (TUs) on one or multiple ODNs (shared distribution networks). Figure 9-12 shows the framework as used in O-RAN and the related PON elements (for simplicity the figure does not show all the possibilities of the WDM overlays).

For each corresponding OLT CT, CO DBA then uses the information received through the CTI interface to update its bandwidth allocations to the corresponding T-CONTs of the corresponding ONUs.



NOTE – Uplink transport latency is called T34 in [ORAN-CTI-TC

Figure 9-12 – O-RAN CTI Framework

As shown in Figure 9-12, there can be multiple O-DUs exchanging CTI messages with the OLT, and there can be multiple corresponding CTs, ODNs and ONUs.

The uplink latency for fronthaul transport spans from the O-RU network interface to the O-DU interface. This covers delays in the ONU, ODN, OLT and any intermediate aggregation nodes between the OLT and O-DU.

The OLT can execute the CO DBA to reduce the latency of upstream signals due to opening the quiet window.

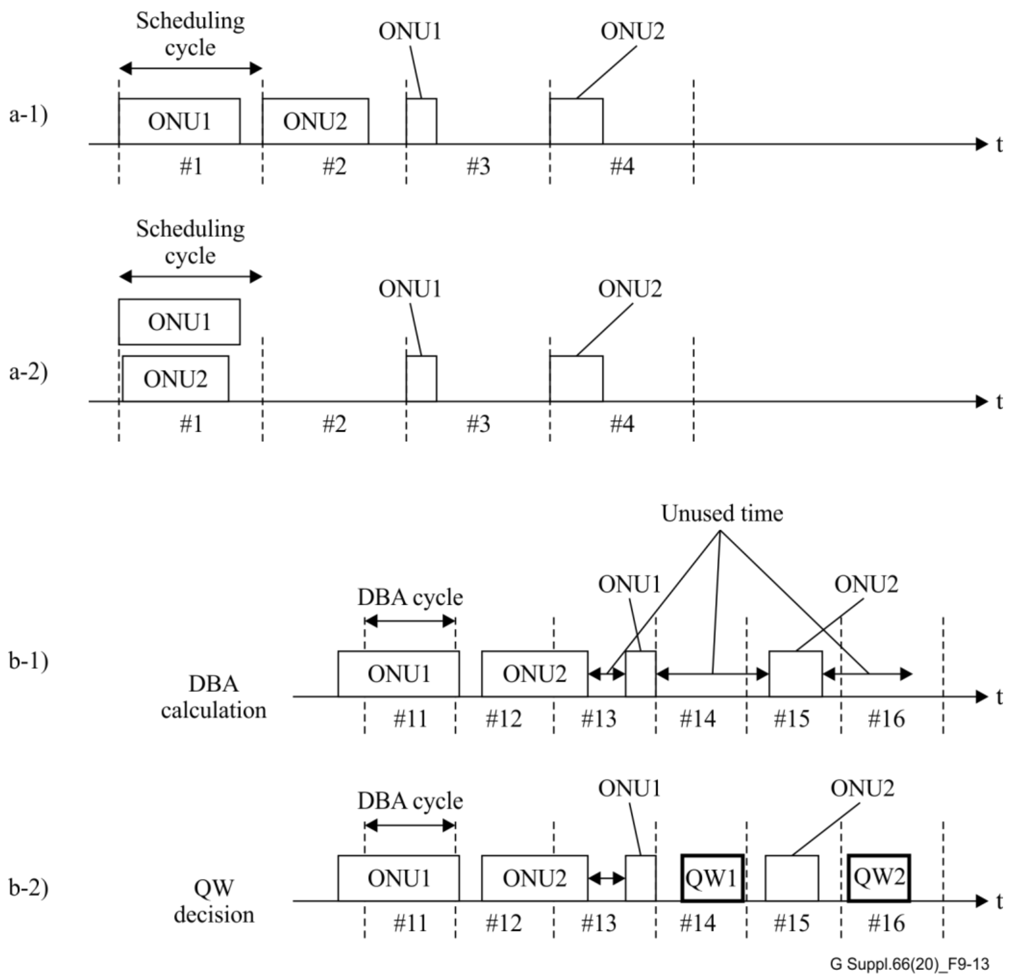
Below is an implementation example of how the OLT makes use of the CTI and CO DBA to reduce latency due to quiet window opening.

The OLT can also use the scheduling information received via CTI and the resulting CO DBA scheduling decisions to find time intervals that are unused by upstream signals. Such intervals are suitable for opening quiet windows so as not to cause extra latency. The process is described below.

1) The OLT gets the scheduling information via the CTI from the O-DU. The starting time and the duration of the upstream signal from each ONU, where they are extracted from the scheduling information, are illustrated in Figure 9-13 (a-1). As shown in scheduling cycle #1 in Figure 9-13 (a-2), upstream signals may arrive at different ONUs simultaneously.

2) The OLT executes the CO DBA based on the scheduling information. Figure 9-13 (b-1) shows the calculated allocation with the case of Figure 9-13 (a-1). Even with the case of Figure 9-13 (a-2), the CO DBA assigns the bandwidth to ONUs in a TDM manner like Figure 9-13 (b-1).

3) After these allocations, CO DBA compares the unused time with the required quiet window size in each DBA cycle. For example, in Figure 9-13 (b-1), when the required quiet window size is 250 μs for 20 km and the DBA cycle is 1 ms, the unused time in DBA cycle #14 and #16 are long enough to open a quiet window. In this way, the OLT detects the candidate timings for quiet windows as shown in Figure 9-13 (b-2).



NOTE – (a-1) (a-2) Upstream bandwidth estimated from scheduling information. (b-1) DBA calculation results. (b-2) Quiet window detection

Figure 9-13 – CTI scheduling information

Figure 9-14 illustrates that the ONU3 performs serial number acquisition and ranging to fit within quiet windows opened by the CO DBA as above.

Scheduling cycle means the cycle in which the O-DU allocates the bandwidth to the O-RU. In the case of the mobile system, it is the mobile slot duration.

DBA cycle means the cycle in which the OLT allocates the bandwidth to the ONU, which typically covers a few 125 μs TC layer frames.

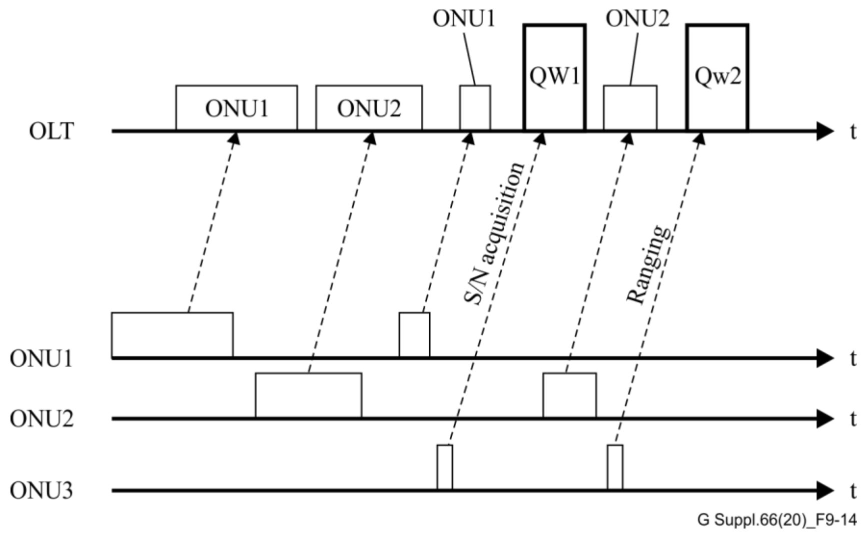


Figure 9-14 – Actual allocation results

9.3.1.3 Quiet window elimination using a dedicated activation wavelength (DAW)

The concept of this method is for an OLT to open the quiet window on a dedicated wavelength channel (i.e., DAW), in order to obtain the required ONU information (ONU-ID, serial number) and perform ranging for ONU activation. By doing so, low latency services on the operating wavelength channel would not be disrupted.

Three scenarios are envisioned with some more practical than others.

– **Scenario A**: A brand new wavelength, not currently in legacy PONs, defined as DAW in the upstream direction. As can be imagined, a brand new wavelength is difficult to come by in the over-crowded PON spectrum.

– **Scenario B**: Re-using a legacy upstream wavelength as DAW. A low latency ONU would have the DAW integrated from the start. Coordination between the operating wavelength and DAW can be done via an integrated MAC to control both wavelengths, or via communication between the legacy OLT and the low latency OLT.

– **Scenario C**: Re-using a pair of legacy downstream and upstream wavelengths for low latency activation. A legacy ONU is designated as a DAW ONU for low latency ONU activation. Coordination between the ONUs is done via communication between the legacy OLT and the low latency OLT.

Note that scenarios A and B use a single DAW in the upstream, whereas scenario C uses a pair of downstream and upstream wavelengths as DAW. This can be a separate PON system such as GPON, or another wavelength pair of a TWDM-PON system (see clause 9.3.2).

This latency reduction method is currently under study in ITU-T.

9.3.2 TWDM-PON for low latency services

While TWDM-PON systems, such as NG-PON2, share all the low latency capabilities of TDM-PON, it exhibits a unique capability of activation overhead elimination, which distinguishes it from all standard-based single channel PON systems.

The single-channel TDM PON systems inherently require a quiet window to allow activation of new or returning ONUs. In [ITU-T G.987] XG-PON and [ITU-T G.9807.1] XGS-PON with 20 km differential fibre distance, the standards calls for a 250 μs general quiet window for ONU discovery and for a 200 μs targeted quiet window for each discovered ONU. During the quiet window, the OLT CT temporarily suppresses upstream transmission by the in-service ONUs, thus contributing to the instantaneous latency and jitter experienced by all traffic flows on the PON.

The multi-channel TWDM PON systems allow to allocate a subset of wavelength channel pairs to perform new and returning ONU activation, while reserving one or more wavelength channel pairs for low latency operation unimpeded by regularly occurring quiet windows. In ITU-T G.989 NG-PON2, the wavelength reservation is achieved with the Serial number grant type indication parameter of downstream wavelength channel profile, see clause 11.3.3.14 of [ITU-T G.989.3]. Once an ONU is activated in an allocated activation wavelength channel pair, it is handed over to the operation low latency wavelength channel pair. The active ONU handover does not impede services of other ONUs in the low latency operation channel, as long as the system implements consistent ranging or other method of equalization delay coordination, see Appendix VII of [ITU-T G.989.3].

9.3.3 WDM-PON for low latency services

A WDM-PON design example is shown in Figure 9-15. Signals from the OLTs, each on a different wavelength channel, are combined in a wavelength multiplexer before transmitting to the cell sites. In the ODN, a wavelength splitter, typically an array waveguide grating (AWG) device, routes the individual wavelengths to different ONUs, each of which is connected to an RU supporting one of the three sectors of an antenna.

The initial design parameters for this example are in Table 9-2. This design assumes a very conservative optical path loss (OPL) of 14 dB to support 18 ONUs in 6 cell sites. Most links are within 10 km distance, with some up to 20 km. Note that this OPL is a rough estimate without taking into account any transmitter dispersion penalty. Some of the parameters in this table could be further adjusted based on specific technology choices.

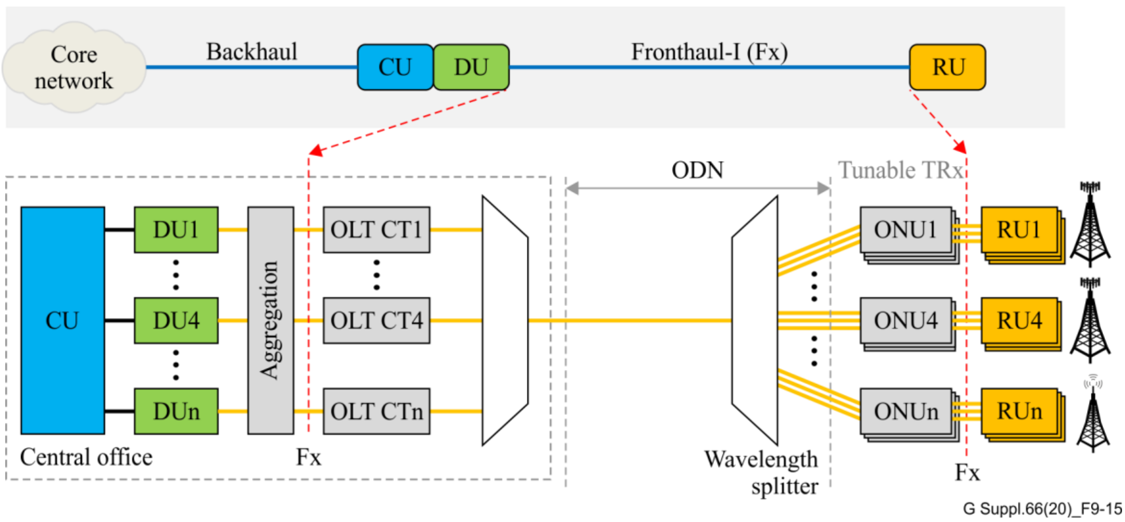


Figure 9-15 – An example of WDM-PON implementation for Fx fronthaul interface

Table 9-2 – Design parameters for the WDM-PON example in Figure 9-9

OPL: 14dB	Data Rate and Reach	20 ONUs per PON port
• Fibre loss: 3.5dB (10 km) • Connector loss: 2dB (4*0.5) • AWG loss: 5.5 dB • Operational margin: 3 dB	• 25 Gbit/s per channel • Most distance < 10 km • A few links at 20 km	• 15-18 wavelength pairs for 5-6 cell sites; • 2-5 wavelength pairs as spare

In addition to the design parameters in Table 9-2, the following key requirements are to be considered.

– Transparent transmission of 25 Gbit/s fronthaul signal;

– Trade-off between colourless ONU wavelength tuning ability and cost;

– Management capability for the WDM-PON system including traffic monitoring, and transmission channel test such as traffic channel loop test;

– Transparent transmission of frequency and time synchronization signal (SyncE and IEEE 1588v2) from SNI to UNI with quality guaranteed to satisfy DU and RU requirements;

– Delay (including both fibre transmission delay and processing delay of the WDM-PON system) to meet fronthaul requirements in clause 7;

– Single fibre bidirectional transmission to conserve fibre consumption and reduce construction costs.

10 PON physical layer requirements

10.1 Capacity

10.1.1 Throughput requirements for the F1 interface

The peak user data rates for the F1 interface can be calculated using the formula published by the Small Cell Forum in the Appendix C of [SCF] for the PDCP-RLC split point and using radio channel parameters taken from the 3GPP documents TS 36.213 for LTE [3GPP TS 36.213] or TS 38.214 for 5G [3GPP TS 38.214]. This model yields the maximum data rate that needs to be transported at the F1 interface in case there is only one UE in the cell, communicating with the cell under perfect channel conditions at maximum possible rate (peak rate at quiet time). This rate scales approximately linearly with the RF bandwidth of the radio link, with the number of independent data streams (MIMO layers) and with the QAM order (i.e., binary logarithm of x-QAM). It is higher than the backhaul bandwidth by up to 3%.

The aggregate data rate for multiple UEs communicating simultaneously in the cell will be less than this peak rate due to interferences, non-optimal channel conditions, dynamic traffic variations and more. For dimensioning the transport network capacity, these effects can be accommodated for by assuming an "average rate at busy time" which we set to 20% of the peak rate at quiet time, following a proposal from NGMN. In [NGMN2] NGMN suggests that in a typical operating condition, the ratio of average-rate and peak-rate per cell site is between 4 and 6.

The bandwidth required for transporting the data at the F1 interface over a fixed line network must take into account the overhead introduced by the transmission protocols on the fixed line, as well as additional capacity for control, scheduling and synchronization mechanisms. The increase of capacity relative to the above bare user data rates at the F1 interface varies with the traffic characteristics and packet size statistics. For simplicity an average increase by 20% over the user data rates is assumed [Chanclo].

Two different network scenarios are analysed in the following two sections, applying the approach described above to an aggregate of cell sites operating with various radio link configurations.

10.1.1.1 Dimensioning of aggregate F1 interface data rates – Example 1

Figure 10-1 displays an x-haul network in which a single PON connects a CU (connected to the OLT) to multiple DUs (each connected to an ONU), each of which serves multiple RUs via dedicated Ethernet point-to-point links.

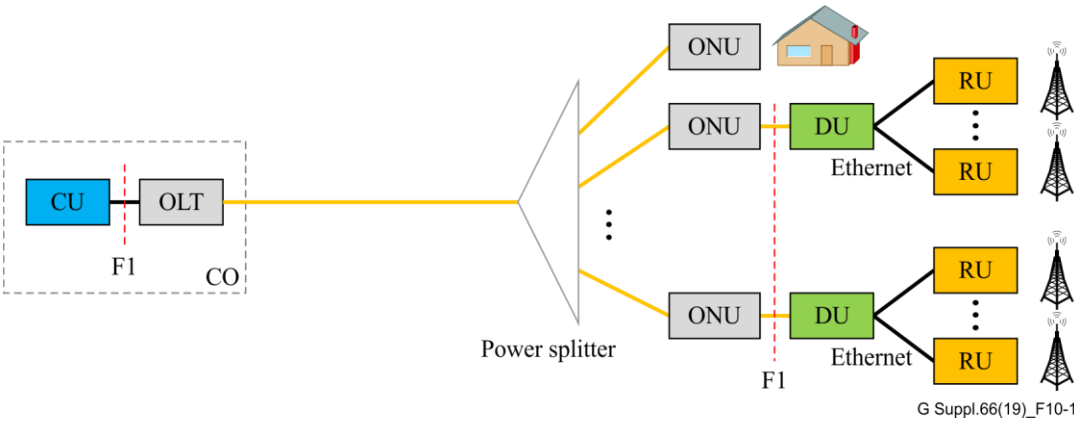


Figure 10-1 – Using a single PON for transport at the F1 interface between CU and multiple RUs

First, the peak transport capacity for backhaul is evaluated for a single RU operating on various MIMO layers, RF bandwidths and different QAM orders (Table 10-1). The figures in the table include 20% overhead for fixed line transport added to the peak air interface data rates, as explained further above. For sub-6 GHz operation, the carrier components are assumed to have RF bandwidths up to 100 MHz and to be modulated at 64-QAM (columns to the left of the vertical separation line). For mm-wave operation above 6 GHz the bandwidths are assumed to be 200 MHz or higher, and the modulation is set to 256-QAM (columns to the right of the separation line).

Table 10-1 – Peak backhaul data rate from a single DU serving a single RU

MIMO	Peak backhaul data rate from a single DU serving a single RU (Mbps						
16	718	1436	2872	7180	19008	38016	76032
8	359	718	1436	3590	9504	19008	38016
4	180	359	718	1795	4752	9504	19008
2	90	180	359	898	2376	4752	9504
1	45	90	180	449	1188	2376	4752
	10	**20**	**40**	**100**	**200**	**400**	**800**
							RF Bandwidth (MHz)

For the case of traffic from multiple RUs being aggregated towards a common DU (cf. Figure10-1), NGMN has proposed two different ways for combining the individual RU data rates into an aggregate data rate for N cells [NGMN2]. Here, equation formula 2 is applied:

$$B_{aggr}(N) = \max(B_{peak}, N \times B_{avg}) \quad (2)$$

where $B_{aggr}(N)$ is the aggregate rate for N cells, B_{peak} is the peak rate for single cell at quiet time, and B_{avg} is the average rate for single cell at busy time.

For 10 cells (or antenna sectors), each connected to a dedicated RU, the aggregate transport capacity at the common DU is thus calculated to be 2 times the peak backhaul rate for a single RU

(10*average rate = 10*0.2*peak rate*1.02). The factor 1.02 accounts for the average overhead at the F1 interface as compared to backhaul. Table 10-2 shows the resulting transport rates corresponding to the cases shown in Table 10-1. Depending on the radio configuration considered, one or multiple DUs can be served by a single optical channel operating at 2.5 Gbit/s (no colour), 10 Gbit/s (green), 25 Gbit/s (yellow), or 50 Gbit/s and beyond (red).

Table 10-2 – Aggregated F1 interface rate from a single DU serving 10 RUs

MIMO	aggregated F1 interface data rate from a single DU, serving 10 RUs (Mbps)						
16	1465	2930	5860	14649	38780	77560	155120
8	732	1465	2930	7324	19390	38780	77560
4	366	732	1465	3662	9695	19390	38780
2	183	366	732	1831	4848	9695	19390
1	92	183	366	916	2424	4848	9695
	10	**20**	**40**	**100**	**200**	**400**	**800**
							RF Bandwidth (in MHz)

10.1.1.2 Dimensioning of aggregate F1 interface data rates – Example 2

This clause provides another example calculation of the aggregate F1 interface data rate requirements based on an alternative formula (3) proposed by the NGMN Alliance [NGMN2]:

$$B_{aggr}(N) = \max(B_{peak} + (N-1) \times B_{avg}, N \times B_{avg}) \qquad (3)$$

According to NGMN, a cell site is considered operating at its peak capacity when one of its antenna sectors (RUs) is running at peak rate and the other two at average rate. In some 4G deployments, the radio unit and antennas are separate, while in 5G they could be integrated in a single RU. The investigated scenario is illustrated in Figure 10-2.

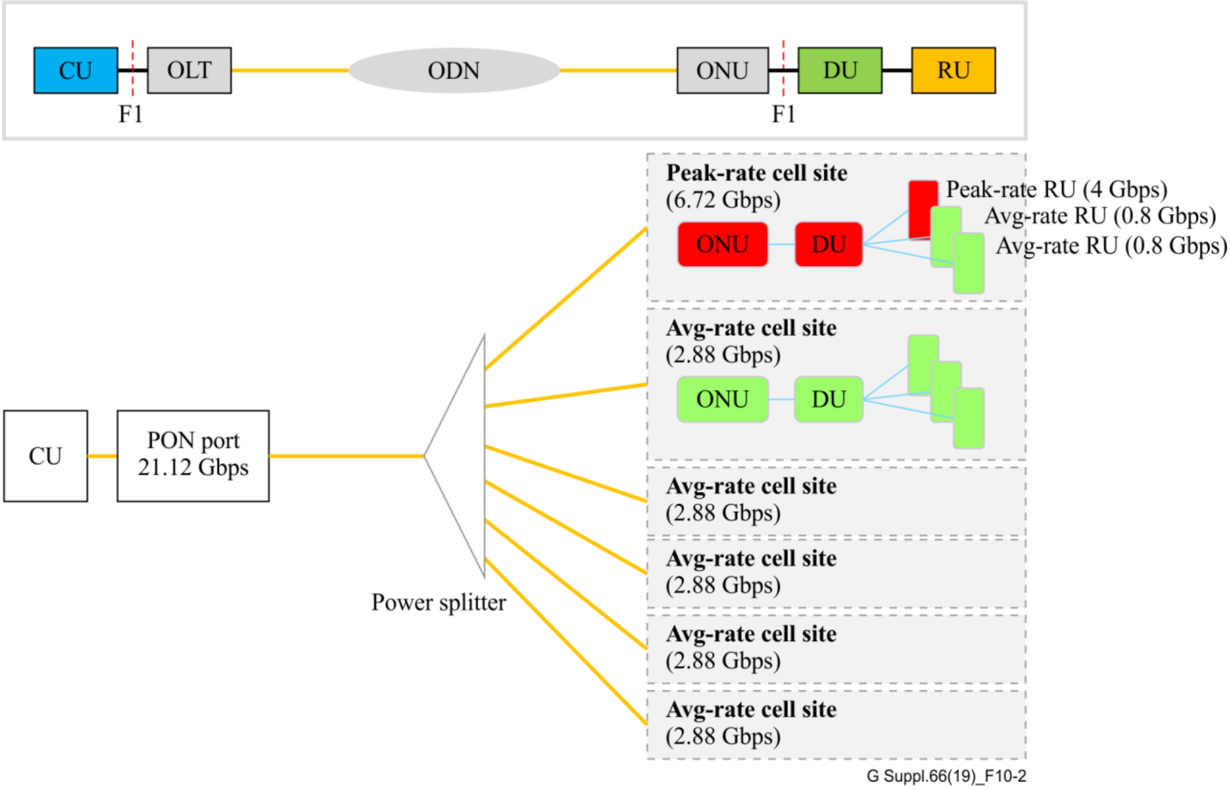


Figure 10-2 – Network architecture used for calculation of capacity requirements per PON port

To improve coverage and increase cell site density in 5G New Radio, it is envisioned that both high and low radio frequency bands will be used. The low frequency band (e.g., 3.5/3.7 GHz) will be for macro cells to provide general coverage, while high frequency band (26/28 GHz) will be mainly for microcells in hot spot areas. Table 10-3 provides a few examples of estimated data rate requirements applying the rules described in the previous clauses.

Table 10-3 – Estimation of F1 signal bandwidth requirements for 256-QAM per carrier

Radio frequency band	Number of Tx/Rx antennas	MIMO layers	Radio channel bandwidth (MHz)	Peak data rate per RU (Gbit/s)	Average data rate per RU (Gbit/s)	F1 data rate per cell site: 1*peak+2*avg (Gbit/s)	Transport data rate for F1 (Gbit/s)
5G, low freq (3.5/3.7 GHz)	16T16R	4	100	2	0.4	2.8	3.36
	64T64R	8	100	4	0.8	5.6	6.72
5G, high freq (26/28 GHz)	4T4R	2*2	2*400	8	1.6	11.2	13.44

Using the example of 64T64R in Table 10-3, the total transport data rate for this (peak-rate) cell site is 6.72 Gbit/s (= (1*4 Gbit/s + 2*0.8 Gbit/s)*1.2). On the other hand, a cell site is considered operating at the average value when all its RUs are running at average rate, which would be 2.88 Gbit/s (= 0.8 Gbit/s*3*1.2) in our example. The capacity requirement for a CU port in Figure 10-2 would thus be 21.12 Gbit/s – supported by a single PON port. Such capacity requirement can be supported by a 25 Gbit/s PON.

10.1.2 Throughput requirements for low layer split

After the TS38.101 [3GPP TS 38.101-1] study, 3GPP has decided to also specify a low layer split point for the gNodeB architecture based on an Option 6 (MAC-PHY) or on Option 7 (Mid PHY) split. The decision has not yet been taken though.

The peak throughput for an option 6 split is comparable to that for an option 2 split, (cf. considerations for F1 as developed above). The formulae in [3GPP TS 36.213] can be extrapolated using the 5G numerology [3GPP TS 38.214]. The same NGMN Alliance aggregation dimensioning as used above [NGMN2] can also be used for option 6 split architectures.

The peak throughput dimensioning of possible option 7 split, as included in the Annex A of [3GPP TR 38.801], is based on contributions by CMCC on a definition of possible option 7 splits (option 7a , 7b, 7c) [3GPP R3-161813], while peak data rates are estimated via formulae given in a contribution by NTT DoCoMo [3GPP R3-162101]. These splits do not correspond to the intra-PHY split points defined in the eCPRI specification [eCPRI]. It is therefore proposed for calculations of the peak transport capacities in G.sup.5GP to apply the formulae provided by the Small Cells Forum [eCPRI] for Split II and Split I (for DL only) corresponding to the II_D and I_U or I_D of eCPRI, respectively. This will give an estimate of the quiet time peak throughputs for eCPRI-based splits.

We note that the ratio between peak and mean of the option 7 splits (SCF Split I and Split II) do not follow the same analysis as described in [3GPP TS 36.213] and as was used for the F1-interface above (cf. 20% as a rule of thumb). Therefore, the aggregation algorithm and the average throughput calculation for intra-PHY splits needs further study.

As the work is still in progress in various SDOs for the low layer split, more details will be provided in future revisions of this Supplement.

10.2 Fibre reach and split ratio

When using PON for 5G wireless transport, the fibre reach is limited by the latency requirements of the service (clause 7.3) and the split ratio is limited by the bandwidth usage (clause 7.2). The typical PON reach and split numbers in residential implementations may not apply.

For the Fx interface, the tight latency requirement between DU and RU could limit the fibre reach to be shorter than the typical reach of residential implementations. The bandwidth requirement could limit the TDM-PON split ratio to be much lower. For WDM-PON, the limiting factor is likely the number of wavelengths that can be afforded by the operators.

For the F1 interface, the latency requirement is in the order of ms such that the maximum fibre reach can be long. Its bandwidth requirement is not as stringent as the Fx interface.

In summary, when designing a new PON system for 5G wireless transport, requirements on latency and bandwidth need to be carefully considered to decide the PON fibre reach and split ratio.

10.3 Optical spectrum

Wavelength plans of G-PON, XG-PON, XGS-PON and NG-PON2 are specified in [ITU-T G.984.2], [ITU-T G.987.2], [ITU-T G.9807.1] and [ITU-T G.989.2], respectively. The EPON and 10G-EPON wavelength plans reuse that of G-PON and XG-PON, respectively.

When designing a PON system for 5G wireless fronthaul, fibre characteristics, such as attenuation and chromatic dispersion, should be taken into account. Single mode fibre (SMF) performance is wavelength dependent. The attenuation of an optical signal transmitted in SMF is the lowest in the C-Band and slightly higher in the L-Band. The zero chromatic dispersion (CD) region is at ~1310 nm for SMF. The optical spectrum selection should also be driven by the availability of opto-electronic components. When coexistence with legacy systems are required, legacy filter characteristics and guard bands to the deployed systems are the other two deciding factors.

11 PON system requirements

System requirements for PON to support 5G transport generally fall into three categories as described in clauses 11.1 to 11.3.

11.1 Requirements set by the wireless networks

For requirements set by the wireless networks, such as latency and synchronization, PON will need to comply with the 5G NR specifications. For example, for 100 μs one-way latency (Table 7-2), it is not possible to support 20 km reach because the propagation time in fibre alone (5 μs/km) would exceed 100 μs.

For synchronization requirements, PON also needs to follow the specifications described in clause 7.4. In order for PON to be a viable solution for 5G NR transport, it is critical that PON meets the synchronization timing error requirements. Several factors affecting the synchronization timing precision are discussed below [ITU-T G.9807.1]:

1) **Fibre propagation delay of different wavelengths upstream and downstream wavelengths**: Using XG-PON as an example, the difference in the index of refraction of the downstream (1577 nm) and upstream (1270 nm) wavelengths result in a systematic error of 61.2 ns when transmitting over 20 km;
2) **Equalization Delay (EqD) accuracy**: as limited by drift of window (DOW) threshold, the EqD accuracy should stay within ±3 ns for XGS-PON;
3) **Internal timing correction**: these are delays due to logical computation and/or other events inside OLT and ONU. One large contributing factor is the downstream SerDes delay, which is about ±6.4 ns for XGS-PON;
4) **System hardware internal error**: different signals may have different transmission paths due to the printed circuit board design. These errors can generally be calibrated in the system level.

As discussed earlier in this Supplement, supporting the Fx interface will require PONs with data rates higher than 25 Gbit/s. For these higher speed PONs, constraints in the synchronization timing requirements would impose additional challenges that need to be solved in order to use PON for 5G transport.

11.2 Requirements to coordinate the PON-wireless interface

For requirements at the PON-wireless interface, the cooperative DBA interface no doubt needs to be supported by TDM-PON. As WDM-PON provides point-to-point connections in the physical layer, its interaction with wireless network is easier than that of TDM-PON. For both TDM-PON and WDM-PON, multiplexing schemes to interconnect OLT and CU/DU need to be selected so that one CU/DU can flexibly support more than one OLT wavelength channel. Liaison communications with wireless SDOs are currently in progress to determine the next steps.

11.3 PON internal requirements not visible to the wireless network

For PON *internal* requirements not visible to the wireless networks, such as the service data encapsulation, PON channel management and monitoring, OLT and ONU timing relationship, message and data security, and protection mechanisms, will follow the existing PON specifications as much as possible. New specifications will need to be developed as needed.

11.3.1 Implementations in PON to support OAM functions

Traditionally in TDM based PON systems, ONT management and control interface (OMCI) is used to provide native OAM functions. For PtP WDM-PON in NG-PON2, AMCC supports per-wavelength monitoring. Here how PON can support both out-of-band and in-band monitoring methods is discussed.

Out-of-band monitoring

The most straightforward implementation of out-of-band monitoring is to use separate wavelength channels for monitoring and data transmission. This can be implemented in WDM-PON and TWDM-PON by allocating one OAM wavelength to control all the other wavelengths on the PON. This management channel could also be used for non-fronthaul uses (as they are tolerant of temporary congestion). However, this method requires two transceivers instead of one. One potentially attractive scheme would be to use a low-cost coexistent system (e.g., G-PON) as the OAM for the main wireless system (e.g., XGS-PON or NG-PON2).

An alternative method is to establish an out-of-band channel that uses the same wavelength as the payload. For P2P WDM-PON in NG-PON2 or generic WDM-PON, AMCC can provide OAM signalling channel for each wavelength. Note that there are two alternatives for AMCC described: one uses low speed envelope modulation on the ordinary high-speed payload signal, while the other uses reserved bits in the transport FEC framing.

For specific WDM-PON payload types, it is possible to use other approaches for OAM signalling, such as inserting OAM messages into the idle fields of the Ethernet protocol rather than into their own Ethernet payloads.

For all the above implementations, the OAM data remains out-of-band, as it does not reduce payload bandwidth.

In-band monitoring

For in-band monitoring over PON, the OMCI provides native OAM functions and could be used to carry the transport OAM information. The only caveat is the necessary encapsulation of wireless signals (such as CPRI frames) into GEM frames and upstream bursts would affect the latency and synchronization requirements, such that additional care is needed. For P2P Ethernet links, the typical OAM functionality could likely be used, because the bandwidth usage of those protocols is small (no more than 10 small packets a second).

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2020