5G RAN CU - DU Network Architecture, Transport Options and Dimensioning v 1.0

5G RAN CU - DU network architecture, transport options and dimensioning

by NGMN Alliance

Version:	**1.0**
Date:	**12-April-2019**
Document Type:	**Final Deliverable (approved)**
Confidentiality Class:	**P - Public**

Project:	**RAN functional split and x-haul**
Editor / Submitter:	**Richard MacKenzie (BT)**
Contributors:	**Javan Erfanian (Bell Canada); Richard MacKenzie, Neil Parkin, Andy Sutton, Dave Townend, Peter Willis (BT); Jinri Huang (China Mobile); Mark Grayson (Cisco); Josef Roese, Dimitris Siomos, Wolfgang Stoermer (Deutsche Telekom); Kang Wang, Yu Yang (Huawei); Nader Zein (NEC); Philippe Sehier (Nokia); Philippe Chanclou, Cyril Delétré, Charles Hartmann, Steve Jones (Orange); Giwan Choi (SK Telecom); Brett Christian (Sprint); Marc Kneppers (Telus); Jovan Golic (TIM); John Kay (US Cellular); Jian Yang, Yuanbin Zhang (ZTE)**
Approved by / Date:	**NGMN Board, 12th April 2019**

Abstract

The 5G RAN architecture allows for a range of deployment options, supporting a range of 5G services. There are multiple options on functional split (how the RAN can be disaggregated into distributed and centralised components), which offer different trade-offs. As the industry progresses, we are starting to understand more details about these trade-offs. This document aims to provide latest updates on the functional split options that might be considered for 5G, and provide insight into how these splits might be deployed

The 5G RAN has two main functional splits options: high layer split (HLS) and low layer split (LLS). HLS is the more mature solution and we provide recent updates from industry activities, describe examples for transport dimensioning, and also address security considerations.
LLS is developing fast with specifications now published. We provide an overview of recent industry activities on LLS, and examples of transport dimensioning which show significantly improved performance in comparison to compressed CPRI.

Contents

1 Introduction 4
1.1 Motivation 4
1.2 Split options 5
2 Transport options 6
2.1 Dark Fibre 6
2.2 Passive WDM 7
2.3 Active transparent WDM 8
2.4 Ethernet 8
2.4.1 Ethernet Over Dark Fibre 8
2.4.2 Ethernet Services 9
2.4.3 FlexE and G.mtn 9
2.5 Optical Transport Network 10
2.6 Microwave and mm-wave 11
2.7 GPON, XGPON, XGSPON, NGPON2 and beyond 14
2.8 Free Space Optics 15
3 High Layer Split 16
3.1 Industry Updates 16
3.1.1 3GPP 16
3.1.2 O-RAN Alliance 16
3.2 Dimensioning Transport Network 17
3.2.1 Capacity based analysis 17
3.3 Security 18
3.3.1 Threats and needs 18
3.3.2 Solutions and options 20
4 Low Layer Split 22
4.1 Industry updates 22
4.1.1 CPRI Cooperation 22
4.1.2 IEEE 1914 22
4.1.3 IEEE 802.1CM 23
4.1.4 Small Cell Forum 23
4.1.5 Telecommunication Technology Association of Korea (TTA) 23
4.1.6 Telecom Infra Project 24
4.1.7 xRAN Alliance/O-RAN Alliance 24
4.2 Dimensioning - Capacity-based analysis 30
4.2.1 LLS scalability for Massive MIMO 30
4.2.2 User plane 30
4.2.3 Control plane 31
4.2.4 Considerations on peak versus average rate 31
5 Conclusions 32
Abbreviations 33
References 34

1 INTRODUCTION

1.1 Motivation

The 5G RAN has a number of architecture options, such as how to split RAN functions, where to place those functions, and what transport is required to interconnect them. In [1] an overview was provided of the architecture options and the various tradeoffs, along with industry status updates.

This document focusses more on the transport options that may be used within the 5G RAN. We start with an overview of the transport options that could be considered, providing the latest industry status, commenting on the maturity of the solutions, and finally a summary of their pros and cons.

Next, we consider the high layer split (HLS). This solution is more mature than the low layer split (LLS) but there have still been some notable industry updates, which are summarized here. We then provide an example of the capacity-based transport dimensioning for a cell site that uses the HLS. This work then considers security considerations for HLS, first discussing security threats, followed by potential solutions. In particular a comparison of IPSec and SCTP DTLS is provided.

The focus then switches to the LLS. Since the publication of [1] there have been significant industry updates and progress, which are summarized here. We then move onto capacity-based dimensioning for LLS, using the xRAN/O-RAN Alliance latest specification (xRAN is now merged into the O-RAN Alliance). These indicative numbers show the potential decrease in throughput requirements in comparison to compressed CPRI, and are also used in the discussion about traffic dimensioning considerations for the LLS.

This document therefore provides an industry update, but more importantly highlights the transport options that an operator may wish to use in the 5G RAN, along with indicative numbers for transport dimensioning with both HLS and LLS solutions.

1.2 Split options

Figure 1 provides an overview of the different possible split options for a disaggregated RAN. The top part was outlined initially in 3GPP Release 14 study on radio access architecture and interfaces [2] based on an E-UTRA protocol stack. The lower part shows the general characteristics associated with different splits, in particular highlighting the trade-off between cost and complexity versus latency and transport requirements.

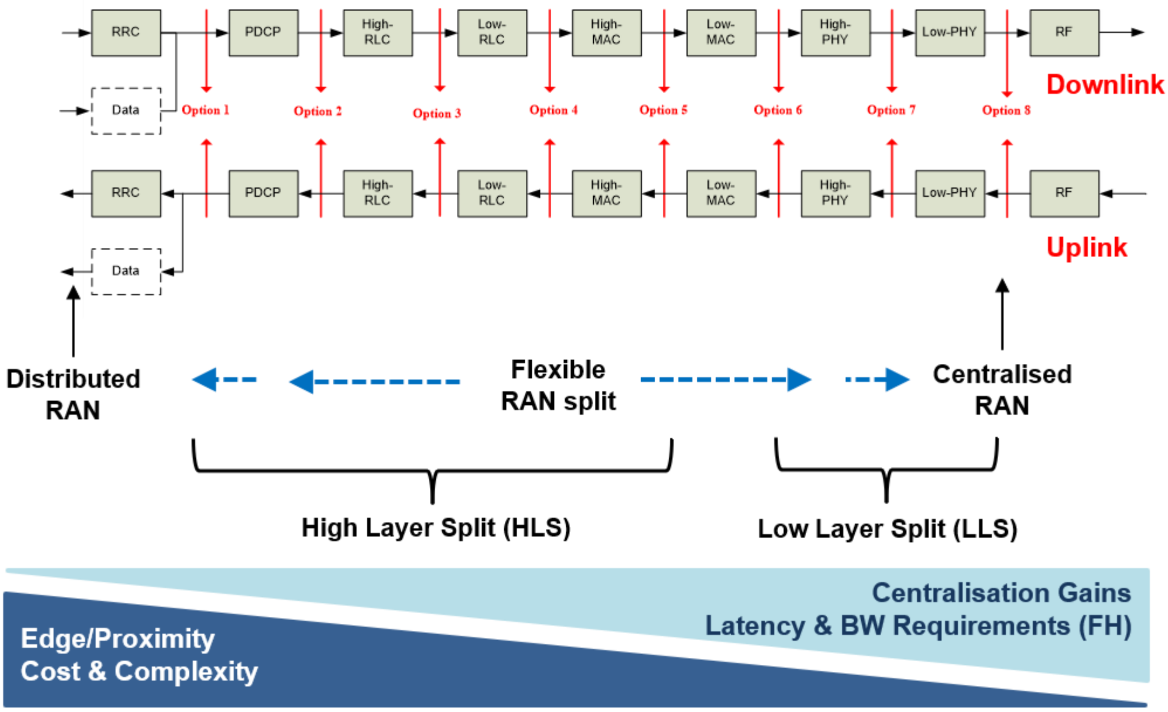

Figure 1 - Functional split options (upper part of figure from [2])

[1] discusses these options in more detail, summarising them as either high layer split (HLS), split options 1-5 using 3GPP terminology, or a low layer split (LLS), split options 6-8 using 3GPP terminology. There are even investigations into an option 9 split [3], where the RF is digitised and centralised, which could have lower transport requirements than CPRI (option 8). Option 7 also has several variants (see Figure 2), commonly referred to as option 7-3, 7-2 and 7-1 in downlink and 7-2 and 7-1 in the uplink. These can offer different benefits and can have significantly different transport requirements.

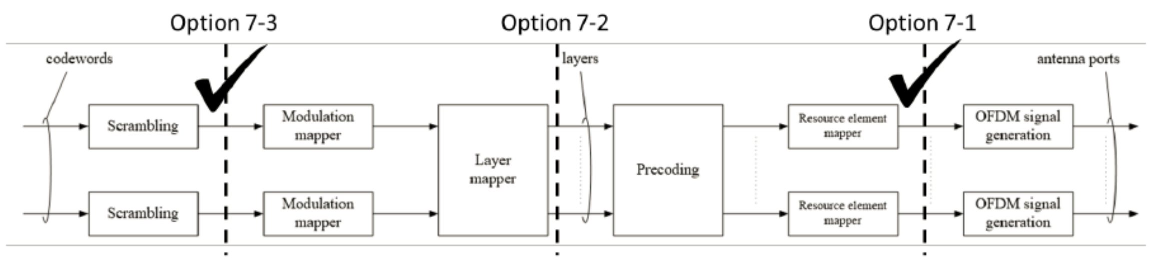

Figure 2 - Possible option 7 functional split possibilities for DL [4]

The HLS is currently focussed on option 2, while LLS is converging, but there are still several variants. In this document we show the latest status on HLS and LLS through the leading industry activities.

2 TRANSPORT OPTIONS

As identified in NGMN's earlier whitepaper [1], an operator has a variety of options regarding the location of RAN functions such as Radio Unit (RU), Distributed Unit (DU) and Centralized Unit (CU). The decision depends on a range of factors including whether to use physical or virtual network functions, compute resource capabilities within the network, the transport options available, and the 5G services to be supported. In [1] some examples of these decisions based on latency considerations and centralization gains were provided, and Figure 3 shows the simplified architecture used for those examples.

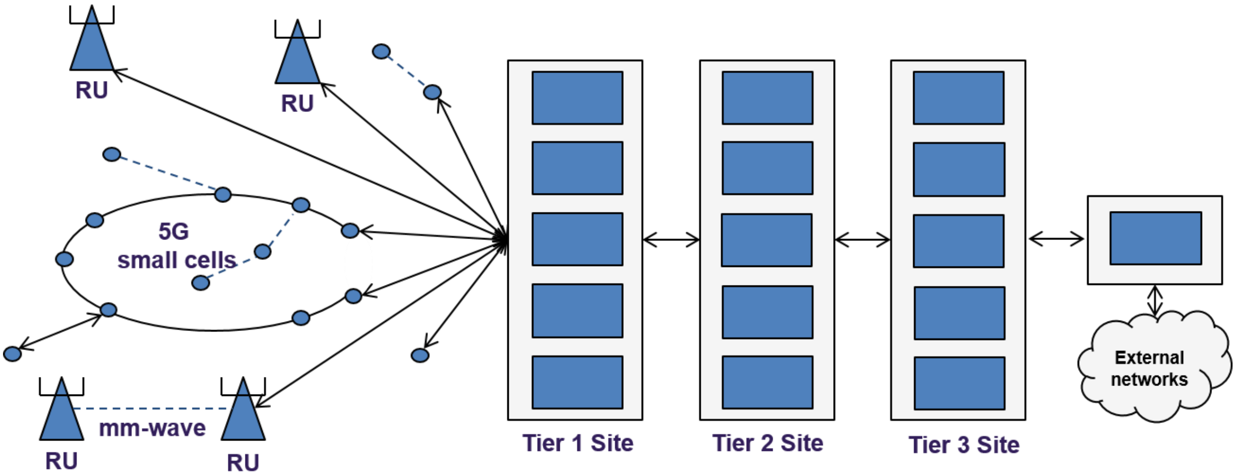

Figure 3 – A generic example 5G RAN with multi-tier aggregation. Blue shapes indicate locations of possible compute capability to support RAN functions (e.g., DU and/or CU) [1]

An overview of some of the key transport options that may be used as part of the 5G access network is provided below. As such, they can be candidate transport options to support interfaces in a disaggregated RAN (e.g., the higher layer split or the lower layer split).

Many of these transport options were reviewed in the 2015 whitepaper as fronthaul options for CRAN [5]. Other standard development organisations dedicated to transport have proposed solutions. For optical access network segment (including dark fibre, WDM and TDM PON), FSAN and ITU-T, SG15 Q2 has published a whitepaper [6] and launched a standardisation specification [7]. Here we consider some of those options again, in the context of supporting 5G services, along with some additional options. Inputs have also been considered from [8] and [9].

2.1 Dark Fibre

For current deployments of 4G CRAN (using CPRI/OBSAI Option 8 functional split) deployments, this is the most common transport option.
Deployment over dark fibre is a straightforward option; however, it can require multiple fibres which could trigger high CAPEX in some deployment scenarios. Standard Optical (small form-factor) pluggables (SFP) allow for transmission up to 80km for 10Gbit/s (for higher data rates, these distances are reduced) or more without amplifiers, which is actually more than sufficient distance for fronthaul due to its delay limitations, plus the fact that the vast majority of sites are well within this distance. Bidirectional transmission (single fibre) will be preferred to simplify fixed network operation (not two fibres, one for downstream, a second for upstream).
This dark fibre could be lit through fixed access equipment like the OLT (Optical Line Terminal) with PtP (Point to Point) interfaces. This solution allows facilitating fixed OAM (operations administration and maintenance) of the dark fibre.

Where phase synchronization (1588 v2) is required single fibre operation has operational advantages compared to dual fibre operation. It is possible to buy single fibre pluggable optics (at least for 1Gbit/s and 10Gbit/s) although they are less common than dual fibre and can limit the reach.

Delay limitations restrict the maximum link length to approximately 20km for 4G. For 5G services, the TTI can be much shorter, and so the maximum link length may be further reduced as a result.

Summary
The most common option for 4G CRAN fronthaul. Operators may need to consider coexistence with existing CRAN deployments when planning transport for 5G. From a 5G point of view this CRAN solution may become impractical when trying to support massive MIMO configurations.

Pros

- Simple to deploy – this however, depends on several considerations including: availability of fibre in the access network; the location/distance to the cell site; and the local regulatory environment
- Not limited to a provider's products. You can upgrade your network at your own pace.

Cons

- Cost as a service
- Availability at every location
- Availability in some regions
- Scaling to massive MIMO

2.2 Passive WDM

To make more efficient use of fibre, while keeping the simplicity of dark fibre, it is possible to deploy CWDM or DWDM pluggable transceivers. These are able to bridge up to 80km (again this may apply to data rates up to around 10Gbit/s, but for higher data rates these distances will be reduced), and extended CWDM pluggables - up to 120km. These distances, however, do not include the additional attenuation of passive WDM filters or other equipment and are based on a typical power budget for pluggables. The latency of the system is caused by the light-propagation delay only—passive WDM filters introduce no additional delay. Passive WDM allows for transmission rates up to 100Gbit/s (all digital data rates including CPRI). Increasing data rates is possible by adjusting interfaces, not infrastructure. The multiplexing factor/splitting ratio is up to 80 wavelengths.

If single fibre operation is required for phase synchronization then this can be achieved using the CWDM/DWDM pluggables but different passive WDM filters.

Summary
Similar to the dark fibre option but with better reuse of facilities, i.e. less fibres but more expensive interfaces.

Pros

- Relatively simple to deploy
- Reduces the number of fibre pairs required for macro base stations deployment
- Easy to upgrade without requiring new infrastructure

Cons

- Uses more expensive pluggable optics than dark fibre
- Not presently manageable (but potentially) as a network (cf. G.989 – PtP WDM of NG-PON2) and still costly as a service

2.3 Active transparent WDM

An active or classical WDM system may be used in the access network for bridging longer distances, saving fibres as well as providing reliable optical layer capability by using Optical Supervisory Channel (OSC), etc. For fronthaul network, saving fibres and optical layer reliability are valuable, but extending distance is not necessary since the fronthaul distance is limited by HARQ. Active WDM electro-layer process may introduce latency and jitter which is more sensitive to fronthaul transport solutions.

Summary

Similar to the passive WDM option but with better optical OAM capability.

Pros

- Mature optical layer operation and maintenance capability
- Saves the fibre demand of the 5G macro base station

Cons

- Can be more expensive compared to passive WDM

2.4 Ethernet

Ethernet is now becoming a likely candidate for transport, for both HLS and LLS. Recent specifications such as IEEE 802.1CM, IEEE 1914.3, eCPRI, and xRAN (now part of O-RAN Alliance), allow for Ethernet to be the transport for LLS.

2.4.1 Ethernet Over Dark Fibre

If dark fibre or a managed point-to-point Ethernet service is available then Ethernet using the standards shown in Table 1 below is an option (Ethernet service providers may only provide a subset of these options). The advantages of using these Ethernet standards is the availability of low cost optical transceivers that are compatible with a wide range of low cost Ethernet CPE and switches. Care should be taken when mixing transceivers from different vendors as they are not always optically compatible. Some Ethernet CPE and switches are also not compatible with all transceiver vendors.

Table 1 - Physical Ethernet Standards

Standard	Speed Gbit/s	Distance km	Wavelength nm	Notes
1000Base-LX/LH	1	10	1310	Typical distance
1000Base-ZX	1	70	1550	Not standardized
10GBase-LR	10	10	1310	Sometimes 15Km
10GBase-LX4	10	10	1300 CWDM	
10GBase-ER	10	40	1550	
10GBase-ZR	10	80 – 120	1550	Not standardized
40Gbase-LR4	40	10	4x ~1310	4 10G waves
40Gbase-FR	40	10	1550	1 40G wave
100Gbase-LR4	100	10	4x ~1310	4 25G waves @ 800 GHz

100Gbase-ER4	100	40	4x ~1310	4 25G waves @ 800 GHz
100Gbase-LR10	100	2	10x ?	10 non-standard WDM waves
100GBase-ZR	100	80	1550	Coherent. Standard still developing
200Gbase-FR4	200	2	4x CWDM	Ratified Dec 2017
200Gbase-LR4	200	10	4x CWDM	Ratified Dec 2017
200Gbase-ER4	200	40	4x CWDM	Standard still developing
400Gbase-FR8	400	2	8x CWDM	Ratified Dec 2017
400Gbase-LR8	400	10	8x CWDM	Ratified Dec 2017
400Gbase-ER8	400	40	8x CWDM	Standard still developing
400G ZR	400	80	1550	OIF

New Ethernet standards have been developed that should address the timing & sync, low latency, high reliability requirements for 5G transport. The relevant IEEE standards are: For Timing & sync: P802.1AS-Rev. For Low latency 802: .1Qav, .1Qbu, .1Qbv, .1Qch, .1Qcr. For High Reliability 802: .1CB, 1Qca, 1Qci, 1Qcz. For slicing i.e. Dedicated Resources & APIs 802: .1Qat, .1Qcc, .1Qcp, P802.1CS, P802.1DC

Summary
Widely available solution, with many enhancements available, or being developed, to help support 5G requirements.

Pros

- Low cost & ubiquitous transceivers and equipment
- Technology widely understood (little to no training required)
- Wide range of carrier capabilities e.g. Ethernet OAM, Ethernet protection, now available
- Ethernet network service providers have mature and competitive offerings
- low latency support

Cons

- Limited distances and speeds
- Newer low latency Ethernet technologies immature

2.4.2 Ethernet Services

Many network operators offer a range of network services that operate at various distances and speeds using a variety of network technologies. For example, it is common to deliver Ethernet services over long distances by the network service provider encapsulating Ethernet in MPLS Pseudowires and transporting the traffic over a MPLS network. The performance of the Ethernet service will be dependent on the underlying technology and design of the MPLS network which may vary widely from network operator to network operator. The MEF (Metro Ethernet Forum) has defined a set of common specifications and certification regimes for carrier Ethernet services that assures the capability and performance of Ethernet services across competing providers.

2.4.3 FlexE and G.mtn

Notable recent enhancements which use Ethernet 64B/66B blocks are defined outside the IEEE bring new enhanced capabilities to support slicing, ultra low latency and high availability. These are

FlexE defined by the OIF (Optical Internetworking Forum) and supported by vendors targeting data centre interconnects and large network operators. FlexE has three general capabilities, such as bonding Ethernet PHYs, sub-rating of Ethernet PHYs and channelization. So FlexE allows a single physical network interface to be hard bandwidth partitioned or channelized to support the slicing of point to point links. Flex-E currently supports up to 400 GE PHYs in 5 Gbit/s granularity.

The ITU-T Study Group 15 has a G.mtn (Interfaces for a metro transport network) standard in definition based on standard Ethernet and FlexE that essentially switches Ethernet 64B/66B line codes almost as if they were TDM time slots. G.mtn provides a cross connected Ethernet network that can deliver TDM like low latency and jitter. 64B/66B cross connection and multiplexing introduces additional latency like OTN does. G.mtn will define two new layer networks (path and section) that will be used in the metro networks, including the transport of D-RAN and C-RAN traffic and run over 50GBASE-R, 100GBASE-R, 200GBASE-R, 400GBASE-R Pluggable Ethernet Module. The path layer provides flexible connections that carry client data and path OAM in 64B/66B blocks. The section layer frame format will be defined in a way that maximizes reuse of FlexE implementation logic including support for bonding homogenous groups of 50GBASE R, 100GBASE R, 200GBASE R, 400GBASE R PMDs (Physical Medium Dependent). So G.mtn will provide an end to end transport network with the ability of slicing, ultra low latency and high availability. Some early Chinese implementations of G.mtn-like are already available.

Pros

- Meets the strict jitter requirements, even for legacy CPRI
- Aggregation of links
- multiplexing
- A complete networking solution
- Scalable and manageable
- Low latency

Cons

- no statistical multiplex gain/no bandwidth reuse between channels
- Requires large physical links (50Gbps and above) with large b/w increments (5Gbps)
- Requires an underlying electrical transport switching layer to support channelization

2.5 Optical Transport Network

Optical Transport Network (OTN) standards and applications are very mature. One benefit of OTN is that LLS RUs, e.g. option 8 CPRI/OBSAI signals of LTE RUs and option 7 eCPRI signals of 5G LTE and NR RUs, can be transported by multiplexing onto a single wavelength using OTU1~OTU4(2.5G~100G, 25G and 50G OTN interface is under development in G.709.25-50).Another form of short reach OTN is FlexO-x-RS (x=1, 100G; x=2, 200G; x=4, 400G). However mapping of the baseband signal onto ODU introduces some latency which, in turn, reduces the maximum fibre length. Generally, the latency introduced by mapping can be less than 10us per OTN node (Note, OTN standards don't quote any latency values and there are many modes, device forms, and ways to implement OTN, so this is a very approximate number).
Another benefit of OTN is Forward Error Correction (FEC) which makes links less sensitive to bit errors and improves reach. However, as already mentioned, the fronthaul connection length is not only limited by transport technology but also by the HARQ. In the case of fronthaul, FEC will even

reduce the achieved distance through introduced latency for which the introduction of 5us latency may reduce the distance by 1 km.
Last of all, with the help of OTN's rich operation and maintenance ability, the fault can be demarcated very well and recovered through protection path efficiency.

Summary
Provides a common and complete optical transport solution for native fronthaul transport

Pros
- Meets the strict jitter requirements, even for CPRI
- Aggregation of links
- Reduces the number of fibre pairs required for macro base station deployment
- A complete networking solution
- Scalable and manageable

Cons
- Transponder needed in addition to standard pluggables
- no statistical multiplex gain

2.6 Microwave and mm-wave

Microwave (MW) and millimeter wave (mmW) radio transmission solutions aim to offer fibre-like capabilities where fibre is not considered economically viable or is too time-consuming to deploy. The exact frequencies that are defined as MW and mmW vary, but a typical definition is depicted below:

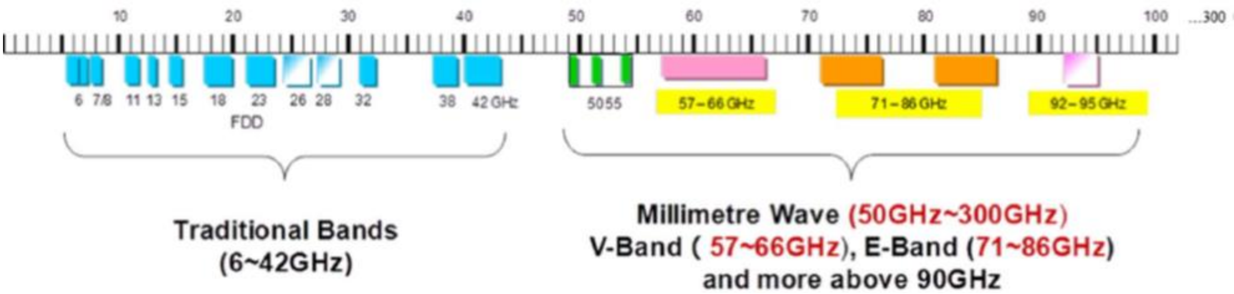

Figure 4 - MW and mmW spectrum [8]

As per Figure 4, MW spectrum ranges from 6GHz to 42GHz, whilst mmW spectrum covers frequencies ranging from 50GHz upto 300GHz and specifically includes V-band (a.k.a. 60GHz), E-band (a.k.a. 70/80GHz), W-band (a.k.a. 100GHz) and, going beyond Figure 4, D-band (a.k.a 150GHz) [8].

Evolutions in the access network and fronthaul architectures drive more demanding transmission requirements on both MW and mmW spectrum technologies to cover use cases in dense urban/urban to rural environments.

In general practical distances, due to atmospheric attenuation and link availability, range from a few hundred meters (dense urban/urban) up to tens of kilometers (rural) depending on frequency band, required link availability, product capabilities, and regulations/licensing conditions. For instance,
- Few hundred meters up to 1km: V-band, E-band, W-band, D-band
- 1-3km: Traditional MW (≥23GHz), E-band, W-band, D-band

- 3-10km: Traditional MW (e.g. 15-23GHz), E-band, Band & Carrier Aggregation
- >10km: Traditional MW (e.g. 6-15GHz), Band & Carrier Aggregation

Further detail on Band & Carrier Aggregation options are provided later in this section.

In terms of capacity, wireless backhaul solutions typically accommodate symmetrical data rates up to 10Gbit/s, where the upper limit comes from port specifications. Further capacity increases are anticipated due to enhancements [8], in terms of:

- **Spectrum:**
  - New bands above 90 GHz, namely W-band and D-band (offer > 1.7x and > 3x times more spectrum compared to E-band)
  - Wider channels of 112/224 MHz at traditional MW bands
  - Band and Carrier Aggregation
- **Efficiency:**
  - LOS (Line of Sight) MIMO/OAM (Orbital Angular Momentum) combined with XPIC (Cross Polarisation for Interference Cancellation)
  - Smart antennas and/or antennas of higher directivity
  - Advanced interference cancellation techniques
  - Multi-band/-carrier/-RF

These advanced techniques are anticipated to increase capacity to nx10Gbit/s even reaching 100Gbit/s performance. This could be made possible, for example, by utilizing the potential of high multiplexing order of 8 (16 when combined with XPIC) offered by OAM transmission technology [10]. By applying the concept of the Band and Carrier Aggregation, higher capacities can also be reached by layer-1 link aggregations and exploitation of higher bands, such as W-band (92 to 115GHz) and D-band (130 to 175GHz) where standardization activities are under way, could offer the highest capacities. It is worthwhile mentioning that CEPT has defined technical conditions for use of both D-band and W-band during 2018.

The main Band and Carrier Aggregation solution of interest currently, is the traditional band + E Band combined link where the E-Band link runs at a lower availability than usually associated with MW link deployments. This allows for the E-Band link to work over longer hop distances, matching that of the traditional link, and providing a much higher aggregated capacity (a target of 10Gb/s over 10km is often proposed). In this way high priority traffic can be carried over the traditional MW capacity during outages of the E Band (due to rain induced fades for instance).

Latency is dependent on the service processing time (service interface unit), switching and packet processing (packet switching unit), modem configuration (modem unit) and service's frame size. The industry is improving the designs in terms of device architecture, forwarding technologies and better modem operation to meet the future requirements of 5G networks. MW and mmW radio links have lower propagation delay (3.34us/km) than optical transmission over fibre (5us/km). Typical one way latency of a mmW radio link is less than 50us, and below 20us is achievable on some very high capacity links.

Based on current capabilities, MW/mmW should be able to support early deployments of 5G base stations in a D-RAN or HLS configuration [8].

Generally speaking, the main fronthaul application for MW / mmW currently is the distributed RAN architecture based on new functional split options (HLS) and a key enabler to urban small cell densification.

It is however feasible, that MW and mmW could be used for LLS. In general, this is likely to rely on future MW/mmW enhancements rather than current solutions that shall provide data rates even up to 100Gbit/s. The feasibility also depends both on the future capacity requirements of the 5G base station, as well as the details of the LLS. Today there are E-band solutions that could provide CPRI transport (using CPRI option 3), but there is limited deployment up to this point.

Today, V Band (60GHz) and sub-6GHz (for near LOS or Non LOS) may provide effective solutions for small cells with HLS or backhaul. V Band licensing is still largely a mix of unlicensed, light licensed, licensed and block allocated regulation, while sub-6GHz bands are limited in terms of spectrum availability and channel bandwidths and are now under pressure from additional spectrum needs for Radio Access Networks.

For 5G small cells, the D band is anticipated to be promising as it could offer a very small form factor and high capacity capability. The small form factor for antennas in D band is very adapted to this scenario. It allows physically separate transmit and receive antennas which provide sufficient isolation for use of full duplex transmission on the same radio channel

Where MW and mmW is used for transport, an operator should ensure that RF emissions safety is carefully assessed, including where members of the public or occupational personnel come into proximity, to remain within the limits set out in the ICNIRP guidelines [11] and conditions defined by local regulations.

Summary
Accommodates macro-cell and small-cell fronthaul scenarios from dense/urban to rural environments. For the higher bands, solutions are maturing and under technical evaluation, so the capabilities are still evolving. Typically MW could support 5G backhaul and HLS, but mmW is likely required before LLS could be considered.

Pros

- Current performance meets early 5G deployments and anticipated technological advancements will satisfy full 5G deployments (macro-cells, small-cells)
- Bridges fibre deployment gaps cost-effectively
- Instant and familiar in similar ways to backhaul applications
- Evolution paths can exist from existing solutions
- Flexible and quick to deploy
- Ultra-low and deterministic transmission latency (a few tens of us) and jitter
- Key enabler to urban small cell densification.

Cons

- Due to 5G deployments, regulatory bodies should allow new spectrum and more bandwidth to be accessible for wireless backhaul. Requires wide channels (new mmW spectrum above 90 GHz can solve this issue) which would come at the cost reduced link length
- Need coordination and interference management when operating in unlicensed/license-exempt bands
- No interoperability between different vendors' radio kit, which could be an issue for multi-vendor disaggregated RANs that integrate MW/mmW transport.
- Requires LOS or near LOS operation.
- Dependence on meteorological conditions (heavy rains, storms)

2.7 GPON, XGPON, XGSPON, NGPON2 and beyond

GPON, and its enhanced variants, is used to provide FTTP (Fibre to the Premise) and is widely available in many urban areas. The optical budget of a typical GPON system limits distances between the central office (e.g. a tier-1 aggregation site) and customer premises to approximately 20km.

The potential GPON variants that could be used for 5G transport are XG-PON (10 Gbit/s down / 2.5 Gbit/s up) XGS-PON (10 Gbit/s down / 10 Gbit/s up) and NG-PON2 (Aggregate capacity is 40Gb/s from four channels at 10Gbit/s. This can be configured in the following combinations: 10 Gbit/s down / 10 Gbit/s up, 10 Gbit/s down / 2.5 Gbit/s up). 25 Gbit/s solutions are now being considered. NGPON-2 also allows for the provision of 10Gb/s point-to-point wavelengths (PtP WDM) on the same network as the TDMA link. Bitrate in the access network above 25Gbit/s is still a research topic [12][13]. 50G TDM-PON (ITU-T G.hsp.50Gpmd being specified now with DL of 50Gbps and UL to be specified later)

A typical latency for GPON and its variants is on the order of 1.5ms as detailed in [14]. There are activities in FSAN/ITU looking to reduce PON latencies. This is explained in detail in [15]. In a TDM PON downstream latency is low and is not a problem, but as each ONU (Optical Network Units) has to send a request to the OLT for a grant to obtain bandwidth to transmit data to prevent any collisions from other ONU's on the PON. To reduce this latency one option is what is termed cooperative DBA. In Co-DBA the mobile scheduler (CU/DU) can share information with the PON Scheduler, when the mobile equipment requires upstream bandwidth to the mobile scheduler, the scheduler not only sends allocations back to the mobile devices but also to the PON Scheduler. This means the OLT can work out upstream bandwidth allocations in advance. This allows low latency upstream traffic for the mobile device.

As shown in section 3.2.1, for the HLS or backhaul, a 1 Gbit/s interface is usually sufficient for LTE, but will not be sufficient for a cell site that supports NR (or NR plus LTE). 10 Gbit/s could be sufficient transport for a single site, however 25 Gbit/s would be a more future proof interface to deploy. Reasons for this include (a) the likelihood that more carriers will be added in the future and (b) as small cells are deployed in higher densities, several small cells may use wireless links to aggregate to the nearest fibre point-of-presence, thus a single ONT may actually be aggregating the traffic for several 5G cells.

If an operator is interested in using PON to provide transport for a LLS, the increased throughput requirements mean that a 10 Gbit/s interface will already be limiting as soon as NR is available, thus 25 Gbit/s is probably the lowest that should be considered. Another significant challenge for transporting a LLS over PON is the latency of PON, which may be too high for most of the common LLS interface options. Some implementations of nFAPI option 6, and TIP implementations of 7-3, should be possible. If lower latency solutions do become available for PON, this will make it more open to support a LLS more widely.

At the time of writing, the cost of 25 Gbit/s interfaces might seem like a significant barrier to wide deployment in a PON system. However, given the volumes anticipated for 25 Gbit/s in data centres, the costs are likely to drop. This could allow for 25 Gbit/s interfaces to be considered for FTTP environments in the future. It is expected that for a 25G PON the upstream direction would be 10G in the first deployments. The burst mode technology will come in this instance from XGSPON developments. 25Gb/s for symmetrical 25G PON using ,for example PAM4, has been demonstrated and would be used for businesses [16].

Summary
Provides Fibre-to-the-premise (FTTP) transport. Several variants from below 1 Gbit/s to over 25 Gbit/s

Pros

- PON is typically the chosen FTTP technology and is widely available in many urban areas
- Lowest FTTP deployment cost (compared to fibre point-to-point)
- Perfectly suited to bursty service traffic
- PON enables smooth capacity upgrade (evolution to higher capacity PONs) by using Standardised system elements (e.g. co-existence elements in ITU-T G.984.5)

Cons

- Current PON systems show latencies in the upstream in the order of milliseconds. These are now being addressed by the ITU using Co-DBA.
- Single wavelength TDM-TDMA PON offer relatively small capacity (2.5G and 10G), and multi-wavelength PON (NG-PON2) development is in its infancy
- (Common to any optical fibre technology is) the high cost of using line rates higher than 25G per wavelength (currently a core network technology)
- PON does not suit fixed bandwidth connections such as split option 8 interfaces.

2.8 Free Space Optics

Free Space Optics (FSO) is the use of light waves rather than radio waves to form a link in free space. This gives FSO its unique features of large bandwidth, licence-free spectrum and high data rates. Conversely the use of light rather than radio provides two main drawbacks; relatively high dependence on favorable atmospheric conditions to maximise the data rate transmission and the need for a continuous clear line of sight which limits deployments and complicates the setup and real time tracking. As we move to higher frequency radio spectrum, e.g. mmW, such limitations also become significant.

Typical link lengths are from 50m – 5km. data rates are typically in the range of 100 Mbps – 10 Gbit/s.

FSO systems can require additional maintenance in comparison to RF-based systems since the lenses at each end may accumulate dirt, dust, or other pollutants that would require occasional cleaning in order to maintain link thoughput performance.

The reliability of the link is an obvious concern for many 5G services. Often FSO comes as part of a hybrid system, which can alleviate the reliability concern, and allow FSO to act simply as capacity enhancement. Given the data rate requirements for a 5G base station, FSO is likely to restrict the capacity of a base station. As a result the HLS is more likely to be deployed over this form of transport, rather than the LLS.

Summary
Easy to deploy and, in terms of infrastructure, low cost. Short links up to around 10 Gbit/s.

Pros

- Narrow beam profile
- No spectrum licenses
- Easy to install and low cost

Cons

- Reliability. Significantly affected by atmospheric conditions
- Data rates still a challenge for LLS

3 HIGH LAYER SPLIT

The high layer split interface (HLS) (3GPP Option 2) between the RLC and PDCP scales with user data and can have a latency tolerance in the order of several milliseconds, although this depends on the latency requirements of the 5G services being delivered and the CU function requirements. This interface can, therefore, either be node internal or a networked interface between nodes and even between sites that are not typically more than 3-5ms apart.

3.1 Industry Updates

3.1.1 3GPP

3GPP Rel15 specified a first release of 5G specifications (EN-DC (E-UTRA - NR Dual Connectivity) and SA (Standalone) operation). Included in these technical specifications was interfaces to support a disaggregated gNB. The F1 interface for the HLS where a gNB-CU (hosting RRC, SDAP and PDCP protocols) and a gNB-DU (hosting the RLC, MAC and PHY layers) is connected via F1 interface. The E1 interface where a gNB-CU-UP (hosting the RRC and the CP part of the PDCP) and gNB-CU-CP (hosting the SDAP and the UP part of PDCP) is connected via E1 interface. The E1 & F1 interface objectives are to facilitate inter-connections of logical nodes supplied by different manufacturers via an open interface. The R15 late drop will bring further dual connectivity options for the NG-RAN and is expected to be fully specified in June 2019.

3GPP Rel 16 is currently specifying the higher layer split for eNB architecture evolution E-UTRAN and NG-RAN, where the W1 interface is being specified for HLS between ng-eNB-CU and ng-eNB-DU. The E-UTRAN aspects of the HLS are still being debated. Rel 16 will also specify support for Integrated Access & Backhaul (IAB) for NR, this may need to enhance to the functionality of E1 and F1 and will add a new architecture for IAB nodes.
In R16 study items include solutions for NR to support non-terrestrial networks, investigating adaptations to allow the operation of NR protocol in non-terrestrial networks. NR Industrial Internet of Things investigating enhancements to URLLC (Ultra Reliable Low Latency Communications). Future R16 Study items will investigate enhancements for disaggregated gNB architecture focusing on enhancements to PDCP flow control with a gNB CU/DU.

3.1.2 O-RAN Alliance

The O-RAN Alliance was formed in 2018 as a merger between xRAN Alliance and C-RAN Alliance, and aims to evolve a more open and intelligent RAN. One of the new workgroups, WG5, will focus on open interfaces that cover the HLS, as well as interfaces between CU entities [17]. WG5, "Open f1/W1/E1/X2/Xn Interface Workgroup", will be compliant with 3GPP specifications, but will go further to try to ensure openness. In particular this could increase the possibility of multi-vendor interoperability across the F1/W1/E1/X2/Xn interfaces. So far, this workgroup has not published any specifications, but a first release of open X2 is anticipated by the end of April 2019.

3.2 Dimensioning Transport Network

The bandwidth requirements on the HLS-interface are similar as for backhaul. For dimensioning the transport network assuming a high layer split interface between the radio site and, for example, an aggregation site of the transport network, similar methods as for backhaul dimensioning can be applied.

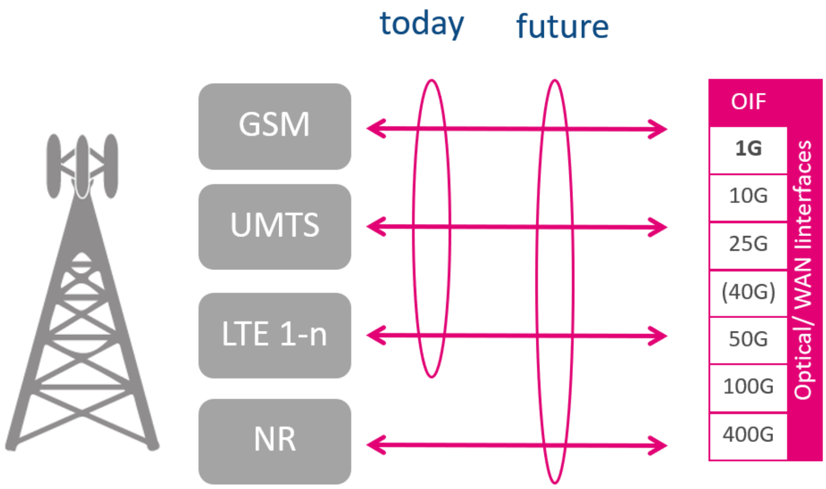

Figure 5 - Example of transport options to support future radio requirements

The major challenge will come from the data rates on the air interface which will be delivered with the new NR carrier, as shown in the Figure 5 example, at larger carrier bandwidth which offers peak data rates up to 4 Gbit/s (assuming 100 MHz carrier bandwidth, 8 MIMO layer).
This will mainly drive the transport network dimensioning e.g. by upgrading the optical interface of the fibre link from 1 Gbit/s to 10/25 Gbit/s.

3.2.1 Capacity based analysis

Dimensioning the transport network can be done in different ways. Each Operator may have its own philosophy how to do it depending on the capabilities of their transport network.
NGMN has described one way of doing it in the white paper “Guidelines for LTE Backhaul Traffic Estimation” [18]. Another methodology is to apply a rule of thumb saying, the transport network should be dimensioned in a way that at least 1 sector with peak rates plus the other two sectors with average data rate is supported.

An example is given in Figure 6, where an existing site running multiple LTE carriers will be upgraded with a NR carrier. The existing LTE carriers could rely on a fibre with 1G interfaces, however as soon as the NR carrier kicks in, 1G wouldn’t be sufficient anymore and new interfaces would be needed.

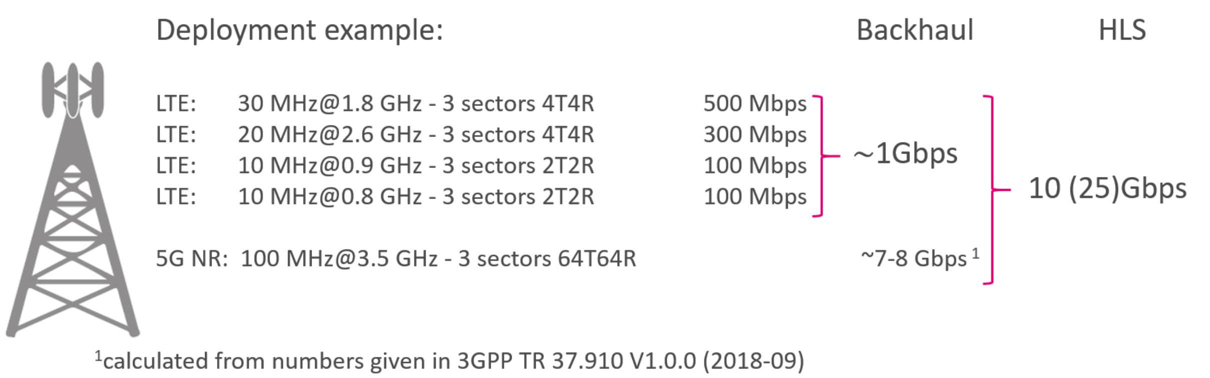

Figure 6 - Deployment example, site with LTE carriers upgraded with a NR carrier

The next step would be a 10G optical interface, which would be also sufficient following this rule of thumb. However, providing some readiness for further extension with additional carriers, a 25G interface might be the better choice.

3.3 Security

3.3.1 Threats and needs

The 5G RAN architecture offers different functional split configurations of the gNB as presented in [1]. Figure 7, taken from [1], shows different placement options for RAN functions at the cell site, aggregation site and edge site.

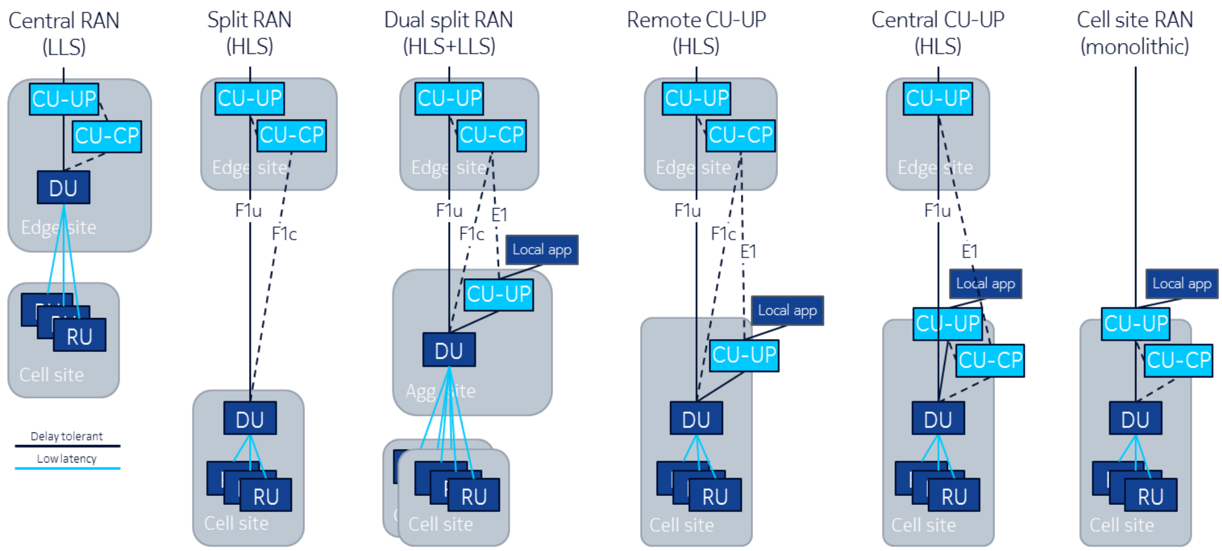

Figure 7 - Example functional placement scenarios [1]

Here we focus on the HLS split option 2, pushed forward by 3GPP. Its main configuration, denoted as “Split RAN (HLS)” in Figure 7, is represented in more detail in Figure 8 taken from [1].

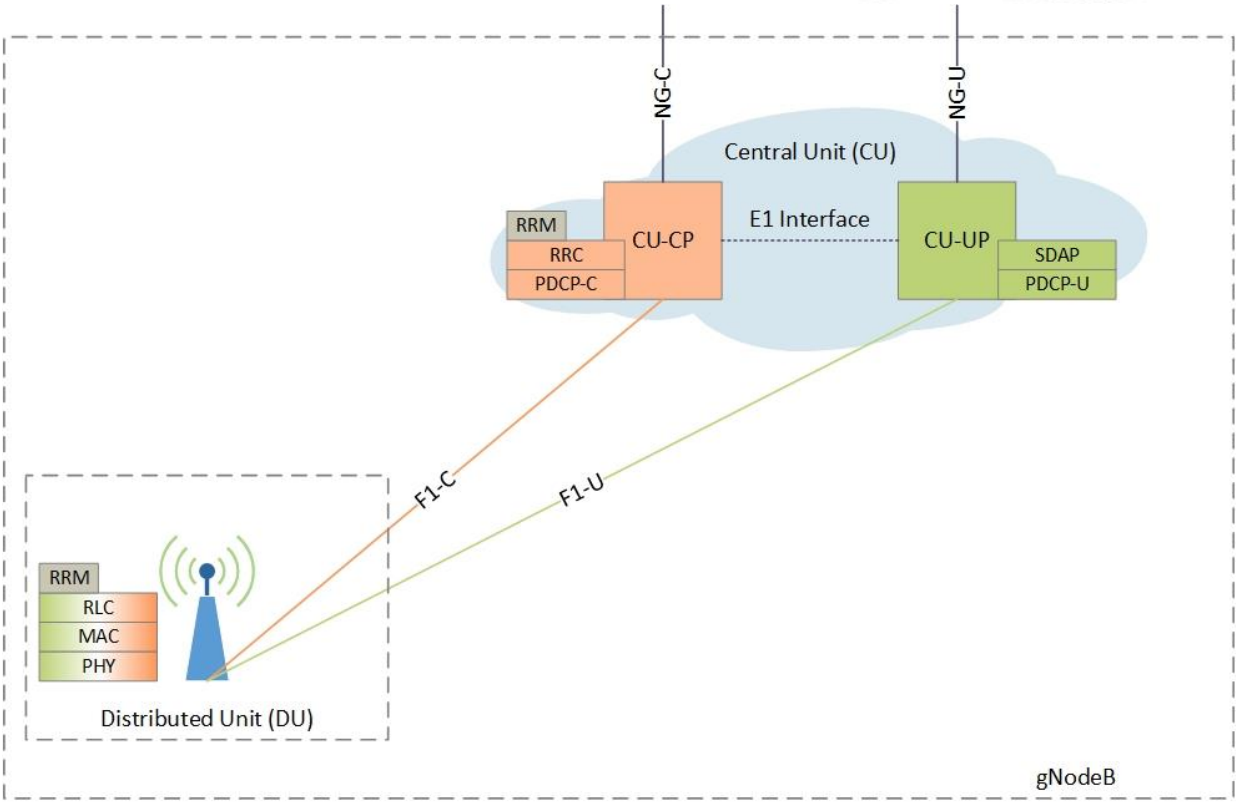

Figure 8 - Possible architecture option for separated Control and User plane logical entities [1]

In the case of the HLS split, the E1 and F1 interfaces may get exposed either physically on the cell sites or on the network side, especially if the network uses resources from different security domains. The threats depend on particular split configurations and implementations of the interfaces.

In split option 2, the PDCP protocol, specified in 3GPP TS 38.323 [19], is executed in CU and enables the protection of the User Plane and Control Plane confidentiality and integrity (including anti-replay) on the F1 interface. More precisely, there is a separation between the Control Plane (CP) traffic on F1-C and User Plane (UP) traffic on F1-U, with the protection by PDCP-C and PDCP-U, respectively, as depicted in Figure 8.

According to 3GPP TS 33.501 [20], the 5G network mandatorily supports confidentiality and integrity protection of both CP and UP over the air interface, by using the corresponding encryption and integrity keys derived from the UE subscription credential. Here, the data integrity is assumed to also include anti-replay protection. In split option 2, this protection relates not only to the air interface, but also extends to F1, since PDCP-protected traffic passes over F1. Note that possibly different security domains and respective regulations may impose the security requirements on F1 that are different from those on the air interface.

For the air interface, according to [20], the use of integrity protection of RRC signalling by PDCP-C is mandatory, and the use of encryption of RRC signalling by PDCP-C is optional. In split option 2, this protection then also extends to F1-C. For the air interface, the use of encryption and integrity protection of UP by PDCP-U is optional, which also extends to F1-U. It should be noted that integrity protection increases the packet size, which may have a significant impact on UP, especially on the air interface. Accordingly, PDCP-U already enables the integrity and confidentiality on F1-U, both of which are mandatory to implement and optional to use. However, this is not the case with PDCP-C and F1-C, as explained below.

F1AP (F1 Application Protocol), specified in [21], is the radio network layer protocol on F1-C providing the signalling messages on F1 which support the procedures for F1 management, UE context management and RRC message transfer. The signalling messages/services can be Non UE-associated and UE-associated. F1AP runs on top of the SCTP (Stream Control Transmission Protocol) transport layer protocol, which itself offers no data protection. Since the PDCP-C PDUs for UL or DL RRC message transfer are contained in F1AP, the RRC signalling messages in F1AP are thus protected by PDCP-C. However, other signalling messages in F1AP, related to F1 management and UE context management, are not protected by PDCP-C.

3GPP TS 33.501 [20] specifies mandatory integrity protection on F1-C, whereas confidentiality protection remains optional for the operators. For such protection, there is thus a need for an additional security channel between DU and CU on F1-C. The operator decides which solution to use for this purpose (see section 3.3.2). In any case, whatever the solution chosen, it shall allow for independent configuration and settings for the transport security on F1-C and F1-U. More precisely, it shall be possible to independently activate/deactivate confidentiality/integrity protection on these two interfaces. The objective is to allow operators to maintain strong security levels on F1-C, while supporting different security options on F1-U. In particular, additional security on F1-U may not be needed if UP traffic is adequately protected on the radio interface (PDCP-U) and the application (e.g., https) levels.

Not using an adequate security mechanism for protecting F1AP on F1-C may put at risk the CU in the transport network or the DUs. In particular, the risks include opportunity to:
- attack the CU and use it as a malicious relay to attack the core network;
- disable NR service for an area by DoS/DDoS on the CU which makes the associated DU(s) unable to operate;
- disable NR service for targeted DoS/DDoS on selected DU(s) which makes them unable to operate.

3.3.2 Solutions and options

The required security on F1 can be provided at the network layer or higher layers. In either case, it is required to provide separate CP and UP data protection. As pointed out above, PDCP-U already provides UP protection on both the air interface and F1-U. However, on F1-C, although PDCP-C provides RRC signalling protection, it does not protect F1 management and UE context management signalling messages within F1AP. In order to avoid duplicate protection, this would then also require the traffic separation within the CP on F1-C.

IPSec is used in LTE networks to add security on S1 and X2 interfaces, by working on the network layer. IPSec could be as well used in the same manner to add security on the F1 and E1 interfaces. In particular, for RAN split applications, it is required that the vendors shall mandatorily implement IPSec on the DU side, as specified in 3GPP TS 33.501 [20] (§9.8.2).

IPSec can provide full data protection at the network (IP) layer, by using either the AH (Authentication Header) protocol for data integrity and entity authentication or the ESP (Encapsulating Security Payload) protocol for data confidentiality, data integrity and entity authentication. Both make use of the authenticated key-exchange protocol IKE (preferably, IKEv2), which uses either pre-shared secret keys or public-key certificates and public-key cryptography. Consequently, the possibility to choose between ESP and AH is an advantage. Protocol numbers, 51 or 50, in the IP header indicate if AH or ESP has been used, respectively.

IPSec can be implemented in the tunnel mode, typically between two security gateways, or in the transport mode, typically between an endpoint and a server or between two endpoints. Transport mode provides end-to-end protection of the payload and is more routing-friendly as it keeps and exposes the original IP header. In the tunnel mode, in addition to the payload protection, the original IP header is protected by being encapsulated within a new IP header, which adds a total overhead of 85 bytes. In 5G, this can affect the dimensioning of the backhauling transport by about 10%.

The operator decides if IPSec will terminate on CU or on the security gateway (SEG). The latter is preferable if there are many DUs for a single CU (for heavy load on CU) or if CU can change (e.g., for failover or load balancing). A single SEG can implement multiple IPSec channels coming from multiple CUs.

IPSec on CU or SEG can be implemented in software (SW) or hardware (HW), where HW is preferable if accelerated computation and/or secure data storage and processing are desired. HW can be a specific element, such as a card, or a dedicated appliance. Both SW and HW implementations should be resistant to implementation-based attacks such as side-channel attacks.

To achieve the separation between CP and UP at the IP layer, it may be required to use multiple virtual interfaces and multiple IP addresses on the CU and/or DU sides, which may be impractical, especially for the tunnel mode, in which the IP addresses correspond to IPSec gateways. Usage of other, higher level identifiers for separation (e.g., port numbers) may be an option for the transport mode. The transport mode seems to be more suitable for multiple DUs.

The separation of different types of signaling traffic on CP (i.e., RRC, F1 management, UE context management) may possibly be achieved at higher layers, on top of the IP layer. As mentioned above, for CP, SCTP is used as the transport protocol on top of IP. SCTP is a message-oriented protocol, which enables multi-streaming and multi-homing, where the messages can optionally be ordered or not. SCTP DTLS (Datagram Transport Layer Security) running on top of SCTP is a TLS (Transport Layer Security) protocol adapted to SCTP which provides essentially the same level of confidentiality, integrity and authentication protection as IPSec, if implemented securely. As such, it can be implemented directly on each CU and it offers the possibility of establishing multiple associations with different security options without the need for multiple IP addresses. However, unlike IPSec, protecting data integrity only is not an option, and that can be considered as a disadvantage.

SCTP DTLS shall comply with 3GPP TS 33.310 [22] (Annex E), should be securely implemented to resist cyber attacks including DDoS, as well as side-channel attacks, and must fulfil the operational requirements. Such SCTP DTLS is added to the supported options on the DU side, apart from IPSec, as specified in 3GPP TS 33.501 [20] (§9.8.2).

The following table summarizes the estimated throughput and latency performance of the best and most cost-effective implementations of IPSec tunnel mode (in HW), IPSec transport mode (in HW) and DTLS (in HW and SW).

Table 2 - Comparison of IPSec and DTLS (HW and SW)

	IPSec Tunnel mode	IPSec Transport mode	DTLS SW	DTLS HW
Throughput performance	Optimal	Good	Good	Optimal
Latency performance	Good	Optimal	Optimal	Optimal
Overhead	~**85** bytes	~**60** bytes	~**48** bytes	~**48** bytes
Overhead impact	**Whole F1-C signalling (F1AP + PDCP-C)**	**F1AP only**	**F1AP only**	**F1AP only**
Interoperability	Complex	Complex	Optimal	Good

Source for IPSec and DTLS overhead: [23]

The throughput and latency performance must be compatible with the gNB requirements, especially the throughput on UP and the latency on CP.

The SCTP DTLS implementation will offer the possibility of monitoring the status of each CU-DU association in terms of connectivity as well as security performance indicators like authentication failure, ciphering errors, unexepected and abnormal events, as well as availability of security logs.

In addition, SCTP DTLS should give the flexibility to configure different kinds of CU-DU associations, especially multi-homing on CU (with the objective to support CU in pools or for use in disaster recovery plans). Auto-configuration of SCTP DTLS association by using CMP or other standardized protocols should be available to ease the deployment of gNB in virtualized infrastructures.

4 LOW LAYER SPLIT

4.1 Industry updates

Since the review of industry activities in early 2018 [1], there have been some significant developments regarding the lower layer split (LLS). Below we summarise those developments.

4.1.1 CPRI Cooperation

The eCPRI specification has evolved to version 1.2. This contains some clarifications and other small editorial changes to improve clarity and readability of the document. The eCPRI Transport Network V1.2 introduces new relaxed latency classes specific to certain use cases. Focus of the upcoming eCPRI V 2.0 Specification will be interworking between CPRI V 7.0 and eCPRI to enable inclusion of installed RRHs with the CPRI interface to an eCPRI fronthaul network.

4.1.2 IEEE 1914

IEEE 1914 NGFI (Next Generation Fronthaul Interface) WG continues their efforts in the NGFI standards development. In particular, the P1914.1 TF has developed a stable draft and begun its 3rd round of WG-level review process. In the stable draft, the NGFI network requirements are developed from two dimensions. From the network perspective, various requirements in different aspects such as the network slicing, protection, convergence networks etc. are developed. In addition, the requirements on NGFI transport nodes, including the temperature, the line rate etc. are also specified. The P1914.1 specification is expected to get its publication this year.

As for another project P1914.3^{TM}, the standard, “Radio over Ethernet Encapsulations and Mappings”, has been approved and published as IEEE 1914.3-2018 TF 2018TM, in October 2018. The following enhancements were added to the standard as part of its final sponsor recirculation ballot:
• an OUI/CID-based mechanism that allows other entities to define payloads for the 1914.3 RoE (Radio over ethernet) header
• TLV control message to request the reading of an RoE node’s parameters
• TLV control message to respond to the aforementioned read request control message
The OUI/CID mechanism has already been used by xRAN (now part of O-RAN Alliance) in their latest fronthaul user, control, and synchronization plane specification [26]. This mechanism also provides a method to extend the RoE subType field beyond the previous limit of 256 values.

After the publication of 1914.3 Standard, the IEEE 1914 working group has launched a new IEEE P1914.3a project, which is an amendment to IEEE Std 1914.3™-2018. This project is scoped to amend the base standard with the following:

- Specifications for mapping with UDP/IPv4 and UDP/IPv6 encapsulation layers
- Specification of more parameters, control messages, and mechanisms to improve OAM functions
- Specification of a management model
- Specification of a mechanism for segmenting big messages
- Extension of CPRI structure-aware mapping to the frequency domain
- Elaboration on how the rbMap function can be used to send data with different priorities
- Clarification on the relationships between all parameters of the standard

4.1.3 IEEE 802.1CM

The specification, which was previously reported as being in the sponsor ballot stage, is now finalized and can be found at [24].

4.1.4 Small Cell Forum

The specification for the MAC/PHY (option 6) functional split, network functional application protocol interface (nFAPI) [25], currently covers only LTE. In 2018 SCF started an activity to define FAPI for NR. This activity is expected to be completed in the first half of 2019. FAPI, rather than nFAPI, defines the MAC/PHY split but assumes the functions either side of the split (e.g. a DU and a CU) are in the same location. nFAPI on the other hand would add the capability for these components to be physically separated and connected via a packet switched network. SCF expects that the nFAPI for NR will follow.

4.1.5 Telecommunication Technology Association of Korea (TTA)

The TTA’s most recent specification on 5G open interface was published back in December 2017. This includes both the higher layer split (“F1 interface”) and a lower layer split (“Fx interface”). The higher layer split remains consistent with the F1 interface as specified by 3GPP. The lower layer split covers all variants of option 7 split.
TTA continue efforts to align with other industry activities. This includes options for mapping to CPRI and 1914.3 transport formats. The recent xRAN (now part of O-RAN Alliance) specification [26] can be applied to the existing TTA specification.

4.1.6 Telecom Infra Project

The Telecom Infra Project (TIP) "vRAN with non-ideal fronthaul" project initiated lab activities around April 2018 across four TIP community labs; hosted by Airtel, Cablelabs, Telecom Italia (TIM), and BT. In each of these labs, the goal is to create a multi-vendor vRAN solution against a commonly agreed high level design. The common design is shown in **Figure 9**, and adopts a functional split that is mainly option 7-3 in the downlink and 7-2 in the uplink. The 7-3 split is seen as a requirement to get the necessary compression relative to CPRI, to enable the solution to be deployable over the range of non-ideal transport [27].

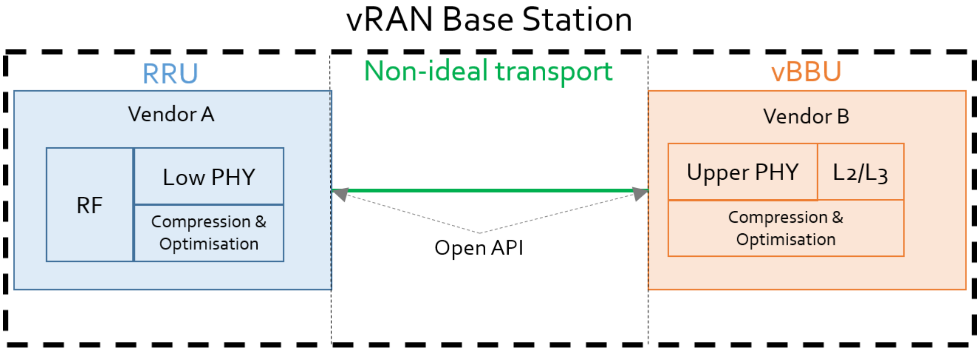

Figure 9: TIP, vRAN architecture [27]

Each of the community labs have different vendors for the RRU (Remote radio unit) and vBBU. In some labs there is more than one RRU vendor. Each lab has a different combination of vendors. Each lab, being hosted by different operators, have different non-ideal transport requirements, so the solutions are tested and optimised in a common way in each lab, but with different deployment options in mind.

The project aims to get to trials and commercial deployments of vRANs working over non-ideal fronthaul shortly. The lab activities are due to complete in early 2019, where results and next steps will be shared, with these aims in mind.

4.1.7 xRAN Alliance/O-RAN Alliance[1]

The O-RAN Alliance was formed in 2018 as a merger between xRAN Alliance and C-RAN Alliance, and aims to evolve a more open and intelligent RAN. The O-RAN Alliance is continuing work from xRAN, with workgroup WG4 "The Open Fronthaul Interfaces Workgroup", to continue the development of a LLS.

Several versions of the fronthaul specification have been released over the past year. The first specification was the xRAN Fronthaul Control, User and Synchronization (CUS) Plane Specification Version 1.0 in April 2018, followed by Version 2.0 in July 2018 and Version 2.01 in January 2019. In July 2018 xRAN published xRAN Fronthaul Management Plane (MP) Specification Version 1.0 and V2.0 in January 2019. In March 2019 however, these xRAN

[1] The information provided, is a summary of public information (whitepapers and specifications) from xRAN Alliance and O-RAN Alliance, plus some additional information provided by xRAN engagement with NGMN. As a general statement the work from the O-RAN Alliance Working Group 4 continues the work of the Fronthaul Working Group of xRAN, so while much of this summary is based on xRAN, it is also true for O-RAN Alliance

specifications were replaced by O-RAN Alliance specifications, CUS Version 1.0 and MP Version 1.0 [26][28].
These specifications have been designed to allow a wide range of vendors to develop innovative, best-of-breed RRUs and BBUs (Base band unit) for various deployment scenarios, which can be easily integrated with virtualized infrastructure and management systems using standardized interfaces and data models.
There is a few notable terminology changes between xRAN and O-RAN specifications. Where xRAN has LLS between RU and lls-CU, O-RAN Alliance has the split between O-RU and the O-DU. In this document we have tried to stick with the O-RAN terminology as much as possible.

xRAN Alliance/O-RAN Alliance Fronthaul Control, User and Synchronization (CUS) Plane
The new architecture for a 7-2 functional split is shown in Figure 10. There are two categories of O-RU specified; category A and category B. The main difference lies in the placement of precoding functions for the downlink.

- Category A devices do not have precoding functions, The precoding is moved to the O-DU. The advantage here is that the O-RU design is relatively simpler. The challenge is that the fronthaul interface may have to support more traffic as it carries spatial streams rather than layers.
- Category B devices include precoding functions. The challenge here is that the O-RU design is more complex. The benefit is that the fronthaul should be relatively low as it carries layers rather than spatial streams.

The O-DU should support both categories of O-RU.

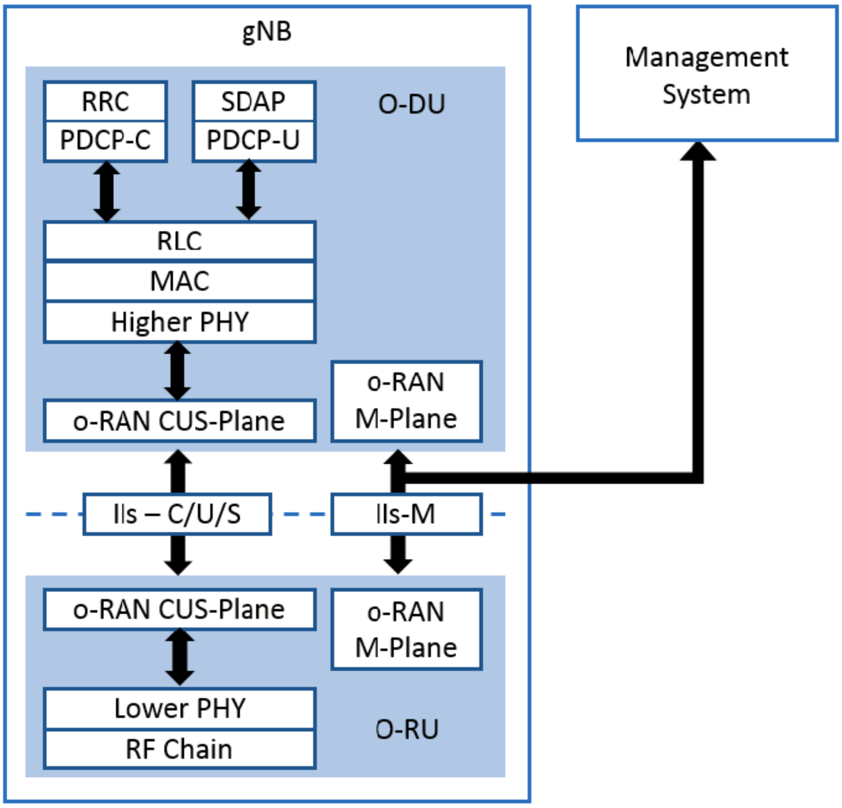

Figure 10 – xRAN/O-RAN Lower Layer Split architecture [28]

An important update from earlier xRAN plans, as mentioned in February 2018 [1], is that the specification focusses on a single lower layer split. The specification covers a 7-2 functional split in both downlink and uplink (previous focus was on 7-1 and 7-3 in downlink and 7-1 and 7-2 in uplink

[4]). For the downlink at least, this appears more a change of terminology, and the new specification can effectively now cover all variations of 7-1, 7-2, and 7-3:

- 7-1 split in downlink would result from the use of category A devices, which means that precoding occurs in the O-DU rather than the O-RU.
- 7-2 split in downlink would use category B devices, which means that precoding occurs in the O-RU
- 7-3 split in downlink would use category B devices and use, from the compression methods, modulation compression. The reason that this would be seen as 7-3 rather than 7-2 is that data bits are used to represent a symbol rather than a frequency domain IQ sample.

Ethernet is the selected transport. Each Ethernet payload contains an 8 byte radio transport header and the rest is radio transport payload (0-1444B). The radio transport header can be either eCPRI (mandatory) or IEEE 1914.3 (optional).
The specification provides 4 options for precoding and beamforming, which allows for support of digital beamforming, but also for digital and analog beamforming.
LAA (License assisted access) is a new feature, added since version 2.0 of the XRAN CUS specification. This allows for the listen-before-talk (LBT) operation to be managed from either the O-DU (mandatory) or the O-RU (optional)

To enable compression of UP data over the fronthaul interface, the specification provides 5 compression scheme options. The compression scheme will be determined via m-plane, the IQ bit width can be dynamically set and it is presumed that the scheduler will determine this value.
The compression schemes are:

- No compression (truncation) – This is a mandatory requirement. I and Q may be represented by 1-16 bits fixed-point values.
- Block Floating point – I and Q mantissa may be represented by 1-16 bits, with a shared 4-bit exponent. A single byte header is used.
- Block scaling – I and Q values may be represented by 1-16 bits, with a shared 8-bit scaler. A single-byte compression header is used
- µ-law compression – Combines the simple bit shift operation (dynamic range) with nonlinear piece wise approximation of µ-law compression where for implementation efficiency, µ=8 and the sign and mantissa are 1 and 2 bits respectively
- Modulation compression – DL data is encoded based on the “perfect” modulation symbols thereby allowing a high-efficiency lossless compression of DL data. This is meant to closely approximate the DL throughput efficiency of a DL 7-3 split point.

The specification also covers synchronization. Figure 11 shows some example topologies that can be supported

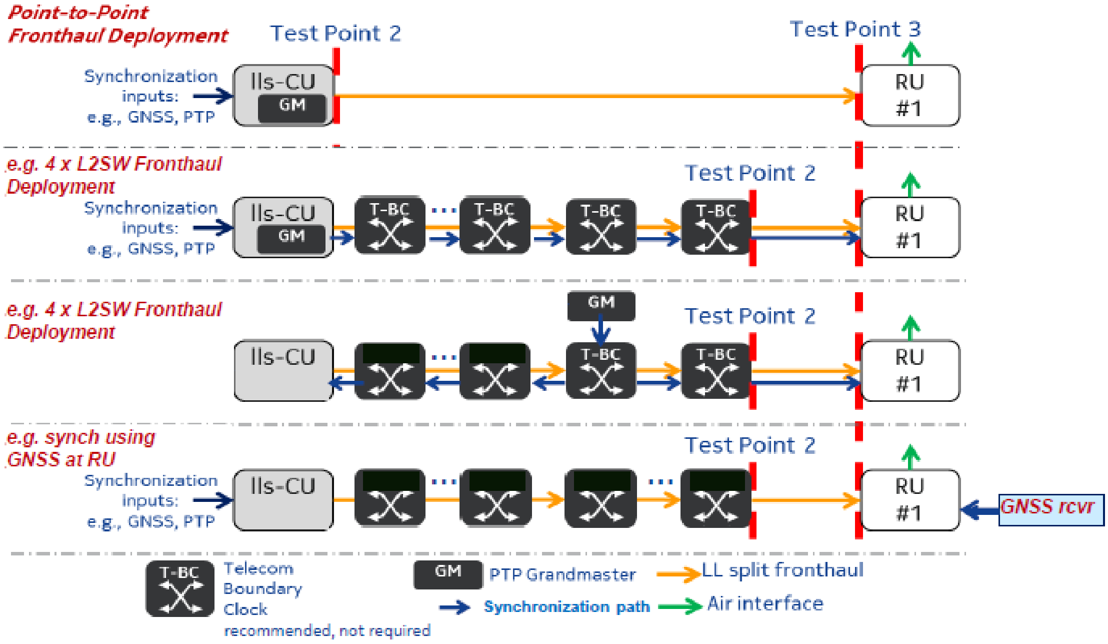

Figure 11 - Example xRAN/O-RAN synchronization topologies [29] (NOTE: xRAN terminology used. O-RAN Alliance's O-RU and the O-DU, are equivalent to xRAN's RU and lls-CU,)

xRAN Alliance/O-RAN Alliance Management Plane

xRAN/O-RAN uses IETF's NETCONF/YANG standard to programmatically configure and manage the Radio Unit component in its lower layer split RAN architecture. YANG (RFC 7950) is used in the xRAN/O-RAN architecture to model the configuration and operational state of the Radio Unit, together with defining remote procedure calls (RPCs) for supporting tasks like software management, and notifications for indicating xRAN/O-RAN defined alarms.

The use of augmented IETF standard YANG models, together with xRAN/O-RAN specific models, lays the foundation for cross-domain orchestration of the RAN with other domains that have already adopted NETCONF/YANG.

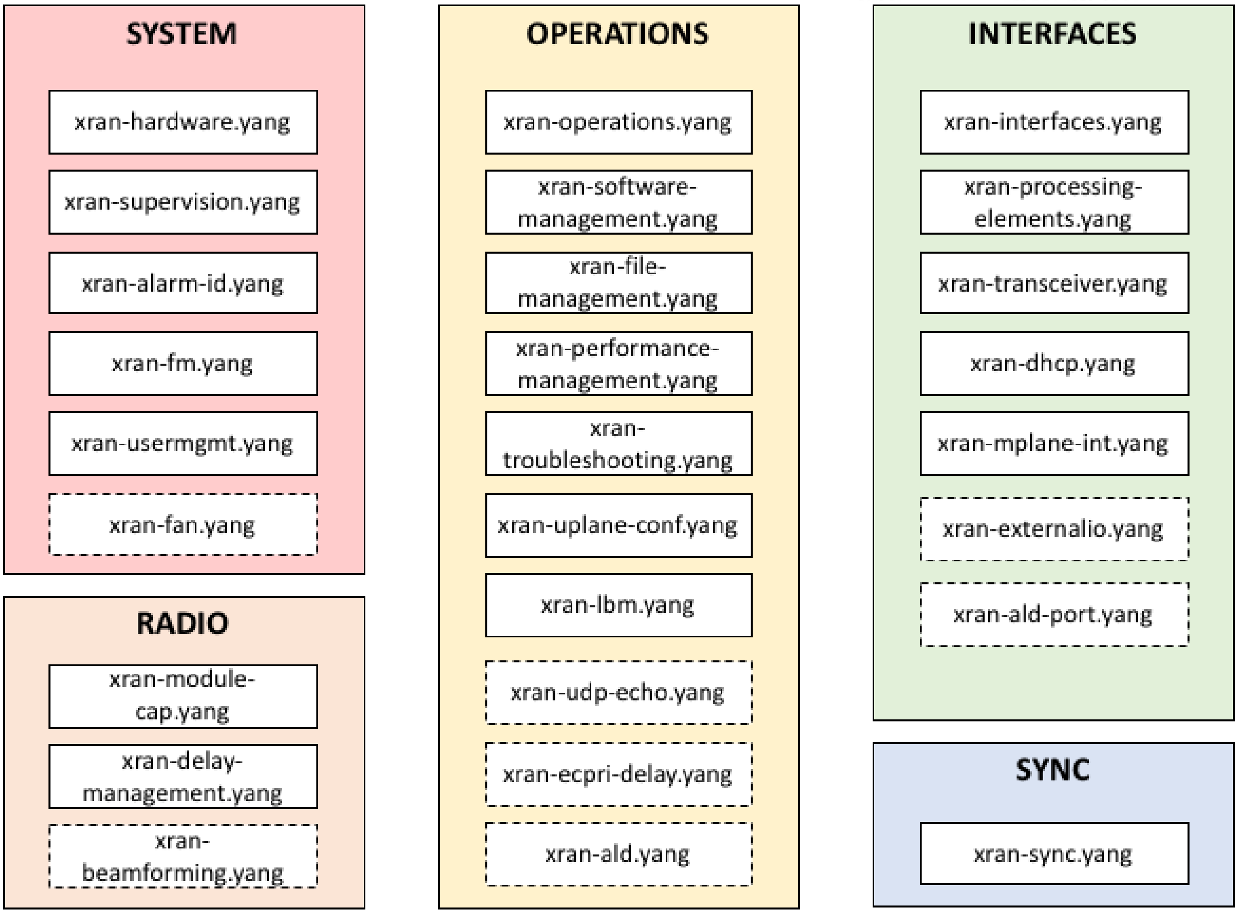

Figure 12 - YANG Models for xRAN/O-RAN Lower Layer Split [30] (NOTE: Terminology is xran version. O-RAN will have an equivalent/similar form)

xRAN Alliance/O-RAN Alliance Management Plane Architecture

The xRAN/O-RAN architecture supports two deployment models, a “hierarchical model”, where the O-RU is managed entirely by one O-DU using a NETCONF based M-Plane interface, and a “hybrid model”, where the O-RU has one or more direct logical interface(s) between management system(s) and O-RU in addition to a logical interface between O-DU and the O-RU. This hybrid model allows functions like O-RU software management, performance management, configuration management and fault management functions to be offloaded from the O-DU(s).

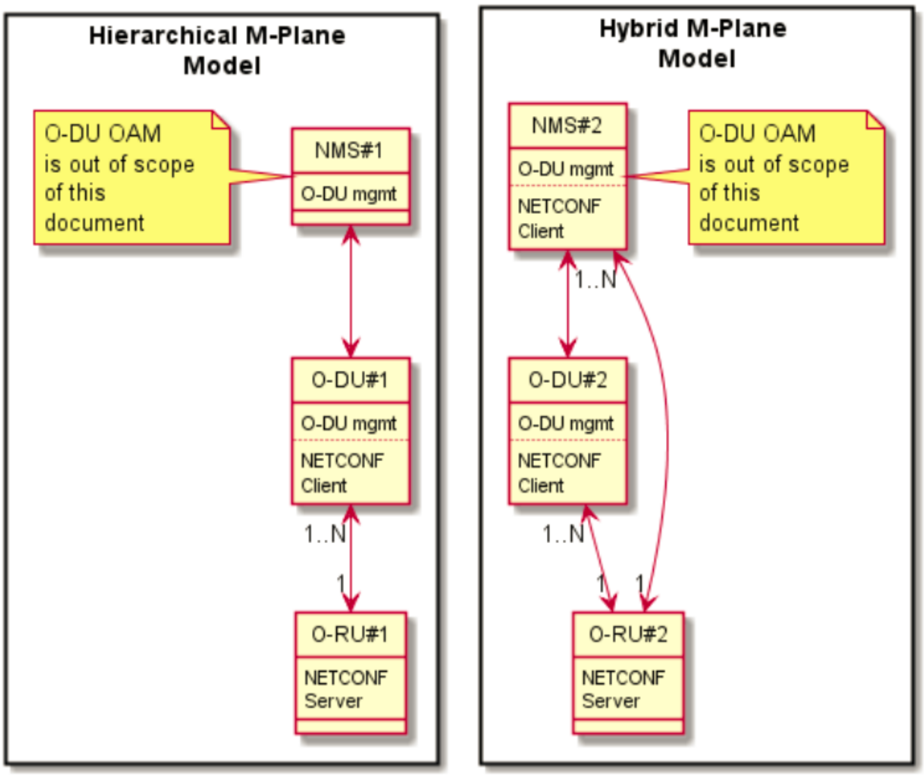

Figure 13 - xRAN/O-RAN Management Plane Architecture [28]

xRAN Alliance/O-RAN Alliance **Transport Aspects**
All xRAN/O-RAN Radio Units are required to support IP in order to support the NETCONF-based management plane, with IPv4 support being mandatory and IPv6 support optional. The management plane has been designed to support a variety of network and IP addressing deployments, including operation across NATs.

xRAN/O-RAN specifies how transport configurations can be used for transporting the control and user-plane communications, with transport options being defined for operation in a bridged-domain, required to be supported by all Radio Units, and optional support for operation across a routed domain.

When operating in a bridged domain, all Radio Units are required to support flow separation based on different VLANs. Optionally, Radio Units may also support flow separation based on different (alias) MAC addresses. When operating in a routed domain, flow separation is based on UDP-ports and IP addresses.

In order to verify end-to-end transport connectivity between the O-DU and O-RU, xRAN/O-RAN specifies the use of transport connectivity checking procedures. When the C/U plane operates in a bridged-domain, Ethernet Loop-back Protocol (LB/LBM) as defined by IEEE 802.1Q (amendment 802.1ag) is used for connectivity checking purposes and when the C/U plane operates over IP, UDP echo (RFC 862) is used for connectivity checking.

4.2 Dimensioning - Capacity-based analysis

4.2.1 LLS scalability for Massive MIMO

Legacy CPRI is a stream-based interface scaling with the number of transmit/receive chains, with a constant throughput. It makes it impractical to support configurations with many antennas, like massive MIMO. In the specification work of the LLS conducted by the xRAN/O-RAN Alliance [26][28], particular attention has been paid to the efficient support of massive MIMO.
The throughput at the interface scales with the cell load, and the maximum number of layers, regardless of the number of transmit/receive chains.
The xRAN/O-RAN specifications defines several control plane and user plane formats, making a generic throughput assessment difficult. The rest of the section is an approximate assessment of the DL peak throughput for a representative case in NR with massive MIMO. The assumptions are listed below:

- Carrier size: 100 MHz, 273 PRB, Sub Carrier Spacing =30kHz
- Class B O-RU (precoding done in the O-RU)
- Maximum of 16 layers
- Maximum of 16 FDM users per slot or sub-bands. This parameter is used to derive the number of segments per slot
- Modulation compression method assumed: the symbol index is transmitted: 8 bits per symbol to support up to 256 QAM
- Precoding coefficients sets are sent on the interface every slot. There is one set per segment. Each complex precoding coefficient is a complex number coded on 2*8 bits.
- Maximum size of the Ethernet payload: 1500 bytes

4.2.2 User plane

Each packet includes several headers: Ethernet, Radio transport (e.g., eCPRI header) as well as an application header (defined in the xRAN/O-RAN specification). There are up to 16 UEs co-scheduled per slot, each supported by a dedicated segment. Due to the payload size limitation, several packets are needed to transport the user plane data of one symbol. Based on this assumption, the UP peak rate can be derived as shown in the table below:

Ethernet overhead	42	byte
Radio transport header	8	byte
Application header	4	byte
Per segment header UP	4	byte
Max ethernet payload	1500	byte
Max. number of layers	16	
TTI length	0.5	ms
Max number of PRB	273	
Max number users per slot	16	
Bits/symbol	8	bit

Section size	208	Bytes
# sections in a Eth. Packet	7	
#sections / symbol	256	
Nber of symbols/s	28000	

# packets/second	1E+06	
Packet size	1510.0	
Peak total throughput UP	**12.5**	Gb/s

It is important to note the gain with respect to the conventional CPRI, which scales with the number of TRX.

Sample rate	3.7	Gb/s
16/15 ctrl channel	3.9	Gb/s
66/64 coding	4.1	Gb/s
Total for 64 TRX	**259.5**	Gb/s

The UP throughput gain is above 20. With compressed CPRI, the ratio can be reduced to 10 approximately, but the advantage of the xRAN/O-RAN -defined interface remains nevertheless very significant.

A similar methodology could be used to derive the UL UP throughput. However, the same kind of modulation compression is not possible for uplink, so transport would consist of I/Q samples. The number of bits required to transport the I/Q samples could vary, let us consider 2*8 bits here. In DL we consider up to 16 layers, whereas for UL the maximum number of layers supported in products is generally lower (up to 8). When we combine these two effects, it appears that UL throughput would be roughly balanced with the DL.

4.2.3 Control plane

The dominant flow for the control plane consists of the precoding coefficients. In a similar way to the user plane, the control plane information is organized in segments. The segments related to one precoding set needs to be fragmented over several packets due to the payload size limitation. The control plane throughput can be derived as indicated in the table below:

Number of sets of precoding coefficients	256	
Precoding set size	128	byte
Precoding set size with overhead	140	byte
Number of sets/packet	10	
Number of packets/TTI	26	
Packet size	1454	byte
CP throughput	**0.6**	**Gb/s**

4.2.4 Considerations on peak versus average rate

It is well known in mobile communications that the average cell throughput is far below the peak. Sizing transport at peak would lead to a significant overestimation of the transport resources. It seems therefore relevant to estimate the throughput distribution in view of assessing the possible statistical multiplexing gain achievable when several cells are multiplexed on the same transport links. An accurate assessment of the required transport capacity would require simulations. At first glance the dimensioning methods used for the HLS can be applied for the LLS. The jitter created by queuing may be the limiting factor in some cases, and would require a more detailed analysis.

5 CONCLUSIONS

Over the past year, there has been significant progress in defining the 5G RAN architecture. In particular, for the high layer split (HLS), many of the interfaces have already been defined by 3GPP, and there are a wide range of transport solutions already available that could support a HLS in early 5G deployments. In many ways the transport dimensioning is similar to considerations for backhaul, a possible rule of thumb for a 3-sector site could be to provision at least 1 sector with peak rates plus the other two sectors with average data rates. Indicative transport dimensioning numbers are provided, which suggest that, to satisfy the requirements for 5G services, a typical transport interface may have to support 10Gbit/s and possibly even 25Gbit/s.

Detailed discussion on security options for HLS has been provided, for the various interfaces used in HLS solutions. IPSec and SCRP DTLS have been reviewed and their transport impact summarized.

Although, not as mature as the HLS, the low layer split (LLS) has seen significant industry progress recently. In [1] it was highlighted that there was a risk of fragmentation of the industry due to so many industry groups potentially creating competing solutions. So far this fragmentation has been avoided reasonably well, with various industry groups working together. However, there are still multiple options for LLS being developed in parallel. Although these options provide different split options to support different use cases, there is still a need to encourage industry groups to ensure fragmentation is avoided wherever possible, and related groups should continue to collaborate and look for ways to align with one another. Transport dimensioning results in this document also provide some indications that the throughput requirements for LLS, using the recent xRAN/O-RAN specification, can be a significant improvement in comparison to compressed CPRI. This is representative of the impact we could expect when NR uses massive MIMO.

NGMN is still monitoring the developments across the industry and will aim to provide further insight into the requirements and deployment considerations for 5G RAN, in particular the interfaces, their radio performance benefits and the transport options that can support them.

ABBREVIATIONS

Co-DBA	Cooperative Dynamic Bandwidth Allocation
CP	Control Plane
CU	Centralized Unit
CU-CP	Centralized Unit-Control Plane
CUS	Control, User and Synchronization
CU-UP	Centralized Unit-User Plane
CPRI	Common Public Radio Interface
C-RAN	Centralized/Cloud RAN
CWDM	Coarse WDM
DDoS	Distributed Denial of Service
DoS	Denial of Service
D-RAN	Distributed RAN
DTLS	Datagram Transport Layer Security
DU	Distributed Unit
DWDM	Dense WDM
FAPI	Functional Application Programming Interface
FSO	Free Space Optics
FTTP	Fibre to the Premise
gNB	5G NodeB
GPON	Gigabit-capable Passive Optical Networks
HARQ	Hybrid Automatic Repeat reQuest
HLS	High Layer Split
LLS	Low Layer Split
LOS	Line Of Sight
LTE	Long Term Evolution
MAC	Medium Access Control
MIMO	Multiple Input Multiple Output
mmW	Millimetre Wave
MW	Microwave
nFAPI	Network Functional Application Programming Interface
NGFI	Next Generation Fronthaul Interface
NGMN	Next Generation Mobile Networks
NG-PON2	Next Generation Passive Optical Network stage 2
NR	New Radio
OAM	Operations Administration and Maintenance
OAM	Orbital Angular Momentum (As mentioned in section 2.6)
OBSAI	Open Base Station Architecture Initiative
OIF	Optical Internetworking Forum
OLT	Optical Line Terminal
O-DU	O-RAN Distributed Unit
O-RU	O-RAN Radio Unit
ORI	Open Radio Equipment Interface
OTN	Optical Transport Network
PDCP	Packet Data Convergence Protocol
PDCP-C	PDCP Control plane
PDCP-U	PDCP User plane
PHY	Physical layer
RAN	Radio Access Network
RLC	Radio Link Control

RoE	Radio over Ethernet
RRC	Radio Resource Control
RRH	Remote Radio Head
RRU	Remote Radio Unit
RU	Radio Unit
SCF	Small Cell Forum
SCTP	Stream Control Transmission Protocol
SDAP	Service data adaptation protocol
TDM-PON	Time-Division Multiplexing Passive Optical Network
TIP	Telecom infra project
UP	User plane
vBBU	Virtualized BaseBand Unit
WDM	Wavelenth-Division Multiplexing
XG-PON	10 Gigabit PON
XGS-PON	10-Gigabit-capable Symmetric Passive Optical Network
XPIC	Cross Polarisation for Interference Cancellation

REFERENCES

[1] NGMN White Paper, “NGMN Overview on 5G RAN Functional Decomposition,” v1.0, Feb. 2018.

[2] 3GPP TR 38.801, V14.0.0, "Study on new radio access technology: Radio access architecture and interfaces”, Mar. 2017.

[3] Wang, J; Jia, Z; Campos, LA; Knittle, C, “Delta-Sigma Modulation for Next Generation Fronthaul Interface,” in CABLE-TEC EXPO, Oct. 2018.

[4] XRAN Alliance White Paper, “Fronthaul Working Group,” Oct. 2017

[5] NGMN White Paper, “Backhaul and Fronthaul Evolution,” v1.01, Mar. 2015.

[6] P. Chanclou, H. Suzuki, J. Wang, Y. Ma, M. Boldi, K. Tanaka, S. Hong, C. Rodrigues, L. Anet Neto, and J. Ming, "How Does Passive Optical Network Tackle Radio Access Network Evolution?," J. Opt. Commun. Netw. 9, 1030-1040 (2017)

[7] Supplement G.sup.5GP to ITU-T G-series of Recommendations - 5G Wireless Fronthaul Requirements in a PON Context

[8] ETSI GR mWT 012 v1.1.1 “5G Wireless Backhaul/X-Haul”, ETSI group report, Nov. 2018.

[9] ITU-T TR, SSTR-TN5G “Transport Network Support of IMT-2020/5G,” Oct. 2018.

[10] Microwave Journal, “NEC Successfully Demos Real-Time Digital OAM Mode Multiplexing Transmission in 80 GHz-Band” [online], Dec. 2018. Available from: http://www.microwavejournal.com/articles/31556-nec-successfully-demos-real-time-digital-oam-mode-multiplexing-transmission-in-80-ghz-band

[11] ICNIRP, “ICNIRP GUIDELINES FOR LIMITING EXPOSURE TO TIME-VARYING ELECTRIC, MAGNETIC AND ELECTROMAGNETIC FIELDS (UP TO 300 GHZ)," 1998

[12] IEEE, “IEEE P802.3ca 50G-EPON Task Force” [online], Available from: http://www.ieee802.org/3/ca/index.shtml

[13] ITU Work Item “G.hsp.50Gpmd” [online], Available from: https://www.itu.int/itu-t/workprog/wp_item.aspx?isn=14550

[14] PMC-Sierra white paper, “The Importance of dynamic Bandwidth Allocation in GPON Networks,” 2008.

[15] ITU-T Series G Supplement 66 “5G Wireless Fronthaul Requirements in a PON Context”, Oct. 2018

[16] M. Santa, C. Antony, M. Power, A. Jain, P. Ossieur, G. Talli, and P. Townsend, "25Gb/s PAM4 Burst-Mode System for Upstream Transmission in Passive Optical Networks," in Optical Fiber Communication Conference, March 2017.
[17] O-RAN Alliance White Paper, "O-RAN: Towards an Open and Smart RAN", Oct. 2018.
[18] NGMN White Paper, "Guidelines for LTE Backhaul Traffic Estimation," July 2011.
[19] 3GPP TS 33.823 "Packet Data Convergence Protocol (PDCP) specification," Release 15, v15.3.0, Sep. 2018.
[20] 3GPP TS 33.501 "Security architecture and procedures for 5G system," Release 15, v15.2.0, Sep. 2018.
[21] 3GPP TS 38.473 "NG RAN; F1 application protocol (F1AP)," Release 15, v15.3.0, Sep. 2018.
[22] 3GPP TS 33.310 "Network Domain Security (NDS); Authentication Framework (AF)," Release 16, Jun. 2018.
[23] A. Hussein, I. H. Elhajj, A. Chehab, and A. Kayssi, "Securing Diameter: Comparing TLS, DTLS, and IPSec," IEEE International Multidisciplinary Conference on Engineering Technology (IMCET 2016), 2016.
[24] IEEE 802.1CM "IEEE Standard for Local and metropolitan area networks — Time-Sensitive Networking for Fronthaul", [online], Available from: https://1.ieee802.org/tsn/802-1cm-2018/
[25] Small Cell Forum, SCF082, "nFAPI and FAPI specification", Oct. 2016.
[26] O-RAN Alliance ORAN-WG4.CUS.0-v01.00, "O-RAN Fronthaul Working Group Control, User and Synchronization Plane Specification", Mar. 2019.
[27] Telecom Infra Project White Paper, "Creating an Ecosystem for vRANs Supporting Non-Ideal Fronthaul," Jan. 2019.
[28] O-RAN Alliance XRAN-FH.MP.0-v01.00, "O-RAN Alliance Working Group 4 Management Plane Specification", Mar. 2019.
[29] XRAN Alliance, "xRAN presentation for NGMN", April 2018.
[30] Cisco, "xRAN M-Plane Update. Presentation to NGMN" June 2018.