iJOIN

INFSO-ICT-317941

INFSO-ICT-317941 iJOIN

Deliverable D4.3

Final definition and evaluation of network-layer algorithms and network operation and management

Editor:	Carlos J. Bernardos (UC3M)
Deliverable nature:	Public
Suggested readers:	iJOIN GA
Due date:	April30th, 2015
Delivery date:	April30th, 2015
Version:	1.0
Total number of pages:	58
Reviewed by:	GA members
Keywords:	iJOIN, RANaaS, Functional Architecture, SDN, Backhaul, Energy Optimisation
Resources consumed	15.19 PM

Abstract

This deliverable provides the final definition and evaluation of network-layer algorithms and network operation and management. This deliverable finalises the design and evaluation of applicable technologies based on the proof-of-concept work performed in WP6. It reports on the feasibility and performance of the network-layer algorithms and network operation and management.

This report includes results for each of the Candidate Technologies (CTs) obtained from simulation and/or implementation. Prior to the presentation of these results, the deliverable provides a characterisation of the iJOIN common scenarios by describing several practical realisations based on current operator policies.

The report also reviews the WP4 final functional architecture and shows how the developed approaches are harmonised within the iJOIN architecture.

List of authors

Company	Author
NEC	Peter Rost (peter.rost@neclab.eu) Fang-Chun Kuo (fangchun.kuo@neclab.eu)
UC3M	Carlos J. Bernardos (cjbc@it.uc3m.es) Antonio de la Oliva (aoliva@it.uc3m.es) Iñaki Úcar (inaki.ucar@uc3m.es) Carmen Guerrero López (carmen.guerrero@uc3m.es)
TID	Ignacio Berberana (ibfm@tid.es)
IMC	Nikolaos Sapountzis (nikolaos.sapountzis@eurecom.fr) Umer Salim(umer.salim@intel.com)
UNIS	Dongyao Wang (w.dongyao@surrey.ac.uk) Efstathios Katranaras (e.katranaras@surrey.ac.uk)

History

Modified by	Date	Version	Comments
Carlos J. Bernardos	April 30th, 2015	1.0	Final version of D4.3

Table of Contents

List of Figures

List of Tables

Abbreviations

3GPP	3rd Generation Partnership Programme
API	Application Programming Interface
BS	Base Station
CWDM	Coarse Wavelength Division Multiplexing
DMM	Distributed Mobility Management
DSL	Digital Subscriber Line
ECN	Explicit Congestion Notification
eNB	Evolved Node B
EPS	Evolved Packet System
FDD	Frequency Division Duplexing
GPRS	General Packet Radio Service
GTP	GPRS Tunnelling Protocol
IETF	Internet Engineering Task Force
iJOIN	Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks
iLGW	iJOIN Local Gateway
iNC	iJOIN Network Controller
IP	Internet Protocol
IT	Information Technology
iveC	iJOIN virtual eNodeB Controller
LOS	Line Of Sight
LTE	Long Term Evolution
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MN	Mobile Node
MPLS	Multiprotocol Label Switching
NLOS	Non Line Of Sight
OAM	Operations, Administration, Maintenance
ONF	Open Networking Foundation
OPEX	Operational Expenditures
OSPF	Open Shortest Path First
PDN	Packet Data Network
P-GW	Packet Gateway
PON	Passive Optical Networks
PQ	Phantom Queue
PtP	Point-to-Point
PtmP	Point-to-multiPoint

QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RANaaS	RAN as a Service
RRM	Radio Resource Management
SDMA	Space Division Multiple Access
SDN	Software Defined Networks
SLA	Service Level Agreement
TDD	Time Division Duplexing
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
UE	User Equipment
veNB	virtual eNodeB
WDM	Wavelength Division Multiplexing
WLAN	Wireless Local Area Network
xDSL	DSL technologies

1 Executive summary

The third deliverable produced by iJOIN WP4 provides *(i)* the evaluation of the mechanisms defined for network-layer algorithms and network operation and management, as well as *(ii)* the final definition of the WP4 architecture. Section 2 describes how the architecture is integrated in the general structure of the project. iJOIN allows having different kinds of information flows associated to different functional splits traversing the transport network. This adds a new level of complexity and requires new solutions for the support of functionalities, e.g. mobility management, load balancing, routing or congestion control. Hence, WP4 provides enablers at network layer level to support the Candidate Technologies (CTs) that are being developed in PHY and MAC/RRM layers. Section 2 also lists the main achievements of the WP4 activities, not only related to this deliverable elaboration, but also including other areas, such as contributions to standards and dissemination.

In order to set the basis for the evaluation, the deliverable first provides in Section 3 a characterisation of the iJOIN common scenarios, by considering current deployment practices from operators, and describing practical realisations of those scenarios. In this analysis, we take into account the transport network (backhaul) alternatives expected to support the iJOIN architecture, which were introduced in previous deliverables. This characterisation also serves as guidelines for the overall project on which kinds of functional splits are feasible on each of the considered scenarios.

Section 4 provides the final description of the functional architecture developed by WP4. This architecture integrates the WP4 CTs. These CTs provide solutions to technical issues posed by the iJOIN objectives that cannot be properly addressed with current architectural solutions (e.g., the existing LTE network architecture does not support in-session anchor mobility or load balancing taking into account backhaul load). They encompass five areas, namely distributed IP anchoring and mobility, network wide energy efficiency, joint path management and topology control, congestion and routing control, and network wide scheduling and load balancing. The proposed functional architecture is more than the designed CTs, as it also includes conventional functional modules that are inherited from existing systems, as well as functionality which is required for the coordination and the implementation of the interactions of the modules. The defined functional architecture is based on the software defined networking (SDN) architectural framework with a separation of the control and user plane, and with the centralisation of the control plane. Messages between the functional modules are described in order to facilitate the definition of SDN protocol (based on OpenFlow) enhancements meeting future mobile network requirements.

Section 5 serves as introduction of the WP4 evaluation, by describing the methodology followed to evaluate each of the WP4 CTs. Then, Section 6 describes the actual evaluation, by reporting the obtained results. Both simulations and proof of concept systems are used for these purposes, depending on the CT characteristics. This evaluation has been conducted following the principles describes in previous deliverable D4.2. As part of this work, a very important contribution consists in the analysis of the RANaaS placement and dimensioning, which has been also taken into account when assessing the performance of the WP4 CTs.

Finally, Section 7 concludes this deliverable.

2 Introduction

2.1 Motivation and Background

WP4 is responsible of the development of network technical solutions allowing for the **support of different functional splits between iSCs and RANaaS entities** [18] [19], as well as to **operate jointly the access and backhaul network layer** in order to optimise the system. These are the main new key concepts proposed by the project. WP4 has also proposed solutions in order to allow a more efficient operation of the network layer in terms of cost, energy consumption, mobility support, latency, etc. These activities resulted in the definition of a set of Candidate Technologies (CTs), described in D4.2 [4].

The network layer, in the context of WP4, is constituted by the set of elements that provide the connection from the termination of the radio interface at the iSC with the rest of the network, RANaaS or S/P-GW, with a certain service level. In traditional networks, the flows of information between the base station or remote radio head and the rest of the network are either the standard S1/X2 interfaces or the CPRI interface for Cloud RAN architecture. In the context of iJOIN, it is possible to have different kinds of information flows associated to different functional splits, which adds a new level of complexity and requires new solutions for the support of functionalities like mobility management, load balancing, routing, congestion control, etc.

In this sense, it is important to understand that WP4 role is not to develop technical solutions for the project key concepts indicated above (e.g., functional split of an existing distributed network-layer functionality), but to provide enablers to support the CTs that are being developed in both phy and MAC/RRM in order to achieve them.

The definition of the WP4 CTs encompasses both the specification of the algorithms used to implement the functionalities to be supported, as well as the definition of the functional entities and the messages exchanged between them. The CTs are expected to meet the requirements identified in a set of relevant scenarios defined by WP5. The specification of the CTs was finalised in the deliverable D4.2 [4].

One of the design principles adopted by WP4 for the definition of the CTs is the adoption of a software defined architecture, where the network control and forwarding functions are decoupled, enabling the network control to become directly programmable and the underlying infrastructure to be abstracted. This approach has been chosen because it is believed that it ensures that all the mechanisms are better integrated and cooperate more easily in order to achieve a joint overall optimisation of the performance. WP4 has also explored the requirements that the support of the CTs with this architecture would pose on the OpenFlow protocol, one of the solutions available for the implementation of Software Defined Networks (SDNs).

Once the architectural and specification work was finished, the potential benefits of the proposed CTs for the scenarios defined by the project were evaluated. Some of the CTs are evaluated in a test-bed that has been implemented in the context of WP6 activities. In order to carry out the evaluation activities reported in this deliverable, a first practical characterisation of the scenarios considered in iJOIN is performed. This characterisation takes as input, on the one hand, the outcome of the analysis of potential backhaul technologies that was included in D4.2 [4], and on the other hand, information on existing deployment practices from operators.

Although the specification work was mostly finished in M24, this deliverable also provides a final revision of the functional architecture, elaborating on how the WP4 architecture integrates with the iJOIN final one.

2.2 Key Contributions

The list below highlights the key contributions of this deliverable, as well as the main scientific technologies' advances, pointing out the main differences over the state of the art as well as the most significant results achieved (in terms of publications, patents, standardisation proposals, etc.).

- This document provides the final definition of the WP4 logical architecture, in collaboration with WP2 and WP3. The document updates D4.2 [4] architecture, including additional information about how WP4 architecture integrates into the iJOIN global architecture. A paper describing this architecture has been accepted for presentation in EuCNC 2015.
- This document provides an evaluation of the designed CTs. In order to do so, the evaluation methodology for each CT is carefully described in a standalone section. This includes also results of

the analysis of the placement of the RANaaS platform inside the EPC network and how to associate each RANaaS with a particular set of iJOIN Small Cells (iSC). A summary of how each of the CTs contributes to the four iJOIN KPIs has been included at the end of the evaluation.

- A paper describing the latest developments of CT4.2 on Energy Efficiency has been submitted to IEEE Globecom 2015. In this paper, an analytical framework for optimal user association that investigates the potential tradeoffs between user- and network-related performances is provided, in a more realistic setup encompassing additional key features: (i) different types of user flows, and (ii) uplink and downlink performance. The proposed framework is evaluated through extensive simulations, and we provide some qualitative and quantitative insights on the related tradeoffs.
- The standardisation work has continued, achieving significant impact in the IETF [11][12]. Within the Distributed Mobility Management Working Group (DMM WG), the two published standards (and the only ones at the time of writing of this deliverable) are co-authored by iJOIN participants. We have also presented to the IRTF SDNRG group a contribution with a proposal of SDN-based architecture [16]. We believe that this contribution could be adopted in the near future.
- Additionally, work at the ONF has continued, pushing some of the iJOIN concepts. As a result of this work, UC3M has been appointed as Research Associate of ONF.
- The Open Source DMM platform (http://www.odmm.net/) is being maintained. The SDN OpenFlow extensions developed as part of CT4.1 will be release as open source and published on this platform in the following months.
- We have implemented a framework based on SDN that allows for a much more efficient network operation and mobility management based on OpenFlow. A demo showing the SDN transport control was carried out in the Mobile World Congress 2015 in Barcelona. The software extensions developed to implement this demo are being prepared to be released as Open Source.

In addition to the list above, which focuses on the achievements related to the evaluation of WP4 CTs and cross harmonization across iJOIN, here we highlight the main results reported in D4.1[3] and D4.2 [4]:

- The paper describing a DMM solution, from which some of the technical approaches of CT4.1 are derived from, has been published in the IEEE Transactions on Mobile Computing [9].
- The paper comparing different DMM approaches, including one based on the SDN approach followed in iJOIN, has been published in the IEEE Communication Magazine (special issue on Recent Advances in Technologies for Extremely Dense Wireless Networks) [10].
- A paper describing the advantages and challenges of adopting an SDN-based approach in a mobile network was published in the IEEE Wireless Communications Magazine, Special Issue on Special Issue on “Research & Standards: Leading the Evolution of Telecom Network Architectures” [8]. This paper was selected as the feature article of the journal issue, and also as the free COMSOC article of the month.
- The standardization work was one of the main results of WP4. In addition to the standard contributions listed above (which has reached its final status as published adopted standards), we have also co-authored a document describing the applicability of congestion exposure mechanisms for EPS-based mobility based architectures within the CONEX WG [13].
- In addition to the former listed papers, we have also published WP4 papers at IEEE ICC and IEEE Globecom.
- In terms of demonstration activities in collaboration with WP6, we showed an early version of our MWC 2015 demo at EuCNC 2014.

3 Characterization of iJOIN scenarios

3.1 Introduction

The purpose of this section is to help in the specification of the values of the main parameters to be considered in the evaluation of the common scenarios by considering realistic deployment options. For each of the scenarios, it is intended to provide an indication of likely values of a set of parameters that include:

- Bandwidth;
- Latency;
- Topology;
- BH technology used.

One of the main outcomes of this document is the description of a framework for the characterisation of the transport network that would support the iJOIN network deployment, based on examples of actual networks. Based on this framework, it also provides reasonable values for each of the parameters required for the 4 common scenarios specified by WP5 for the CTs evaluation. Obviously, these values are provided just as examples of the application of the framework to the different scenarios.

In principle, it could be assumed that these values would be given by the backhaul technology selected, considering the classification that was provided in D4.2 [4]:

Table 3-1. Backhaul technologies characterisation

Number	BH technology		Latency (per hop, RTT)	Throughput	Topology	Duplexing	Multiplexing Technology
1a	Millimeter wave	60GHz Unlicensed	≤5 ms	≤800 Mbit/s	PtP (LOS)	TDD	--
1b			≤200 µsec	≤1Gbps	PtP (LOS)	FDD	--
1c		70-80GHz Light licensed	≤200 µsec	≤2.5 Gbps	PtP (LOS)	FDD	--
2a	Microwave (28-42 GHz) Licensed		≤200 µsec	≤1Gbps	PtP (LOS)	FDD	--
2b			≤10 ms	≤1Gbps	PtmP (LOS)	TDD	TDMA
3a	Sub-6 GHz Unlicensed or licensed		≤5 ms	≤500Mbps	PtP (NLoS)	TDD	--
3b			≤10 ms	≤500Mbps (shared among clients)	PtmP (NLoS)	TDD	TDMA
3c			≤5 ms	≤1 Gbps (per client)	PtmP (NLoS)	TDD	SDMA
4a	Dark Fibre		5 µs/km × 2	≤10 Gbps	PtP		--
4b	CWDM		5 µs/km × 2	≤10·N Gbps (with N≤8)	Ring		WDM
4c	Metro Optical Network		250 µs	≤1 Gbps	Mesh/Ring		Statistical Packet Multiplexing
4d	PON		≤1 ms	100M – 2.5Gbps	PtmP		TDM (DL)/ TDMA (UL)
5	xDSL		5-35 ms	10M – 100Mbps	PtP		--

Number	BH technology	Latency (per hop, RTT)	Throughput	Topology	Duplexing	Multiplexing Technology
6	1 Gigabit Ethernet	≤200 µsec	≤1Gbps	PtP		--

However, we consider that the table is only a building block to be used in the characterization of the transport infrastructure.

3.2 General framework

The general framework for the characterization of the RANaaS deployment options is based on potential deployment scenarios taking into consideration operator's practices and forecast evolutions. In this section we provide some examples of transport networks that, based on the discussions within the iJOIN consortium, may be considered as good examples of current deployment policies.

The following figure represents one potential deployment of existing network.

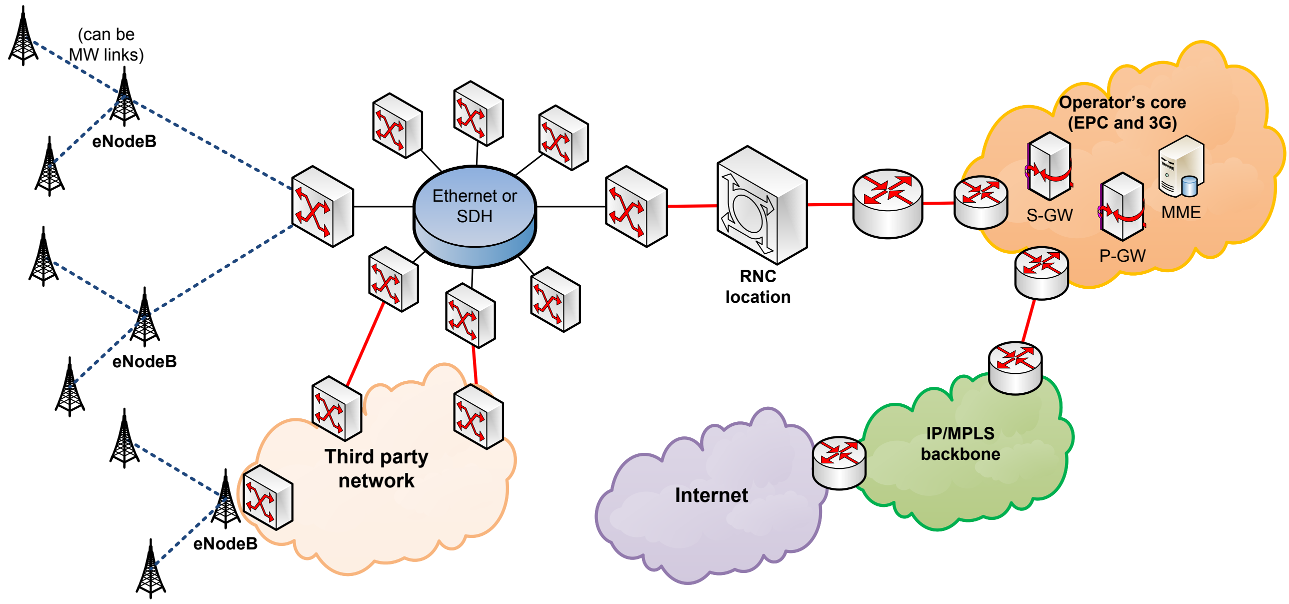


Figure 3-1: Deployment option 1 of an operator's transport network

The following figure represents another option of transport network in the context of a single RAN deployment.

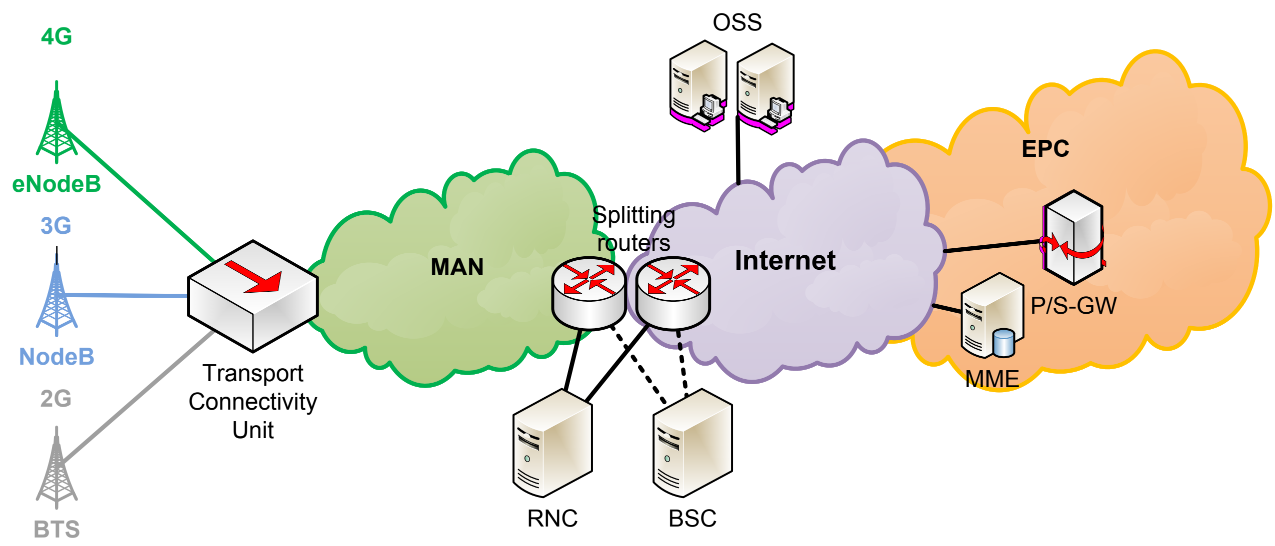


Figure 3-2: Deployment option 2 of an operator's transport network: single RAN

A third different transport architecture is shown below, where a layer 3 (IP) network interfaces with the radio nodes.

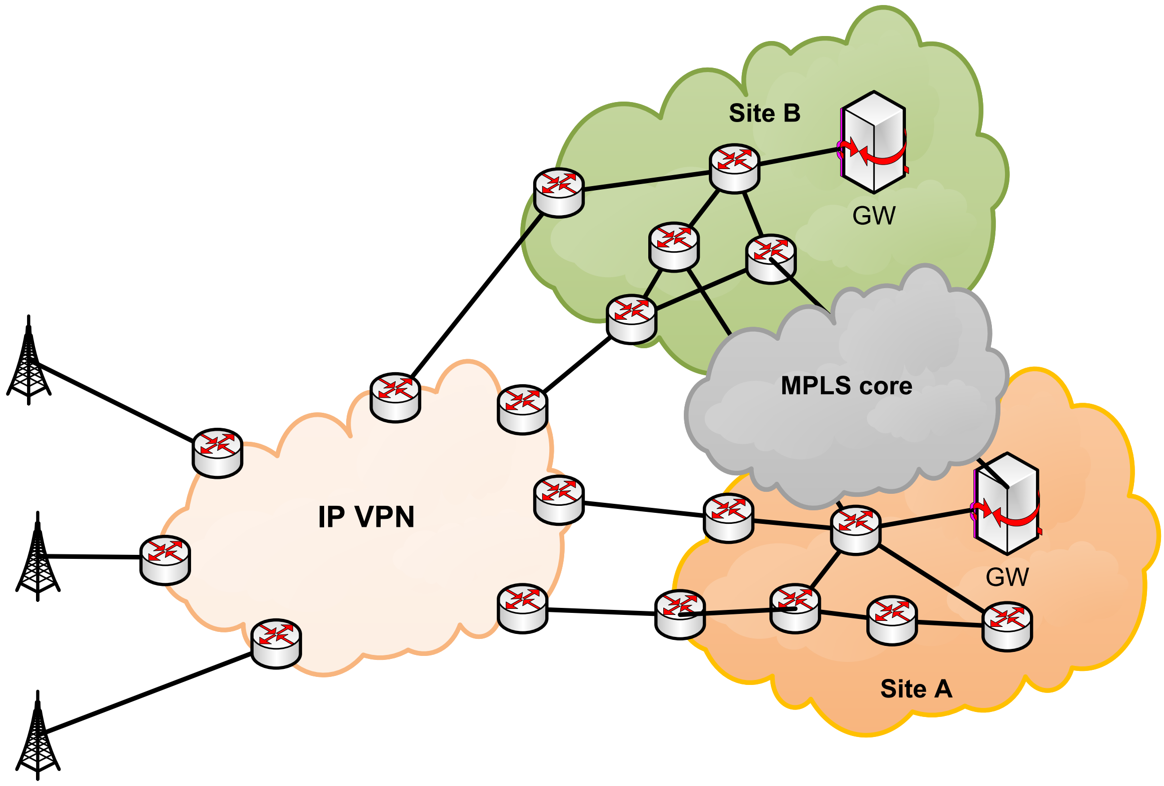


Figure 3-3: Deployment option 3 of an operator's transport network: IP network interfacing with the RAN

Based on these and other examples a characterisation of the transport network, a general framework is proposed. It should be noticed that this characterisation is a simplification that does not necessarily reflect the full complexity of any real world implementation.

The transport network considered is composed by three segments and it is represented in Figure 3-4:

- **Last drop segment**. It is the first segment that connects the iSC to the rest of the network in iJOIN architecture, and it is assumed to be dedicated link (that may be shared by other iSCs or other services[1]). It is dimensioned according to the capacity required by each individual iSCs. The characteristics of the technologies that can be used in this segment are indicated in Table 3-1. Based on data from Telefónica Spain, average length of the last drop segment in urban/suburban environments is around 300 meters.

- **Metro aggregation network**. Its main objective is allow for the aggregation of the traffic from different iSCs/last drop segments, taking advantage of the potential statistical multiplexing, and avoiding the overprovisioning of the network. Usually it is a layer 2/2.5 (MPLS) network that connects multiple last drop segments to the core. Last drop segments connect to the aggregation network through an aggregation node (switch or router). A ring or mesh topology is used in the aggregation network, in order to provide some additional reliability.

  The capacity of the aggregation network can be estimated with a similar rule to that applied in the fixed network. In this case, the network is expected to support a certain percentage of the sum of the capacity guaranteed to the users connected in the busy hour (e.g., 10-15% of the capacity sum for 20% of the users). In the case of the mobile network, it can be estimated as a percentage of the sum

[1] This means that, in principle, there are no statistical multiplexing gains in the last drop segments, as each iSC has devoted resources.

of the traffic carried out by each cell in the busy hour. In a standard case, this traffic would be characterized by the mean bit rate of the cell, however for iJOIN it would depend on the functional split implemented. The percentage considered would depend on the number of cells aggregated (the larger the number, the lower the percentage to be considered).

- **Core transport network**. It is in charge of moving traffic to/from the aggregation network to the central processing units. It is assumed to be a layer 3 (IP) network which incorporates system redundancy and network resiliency features and extensive Layer 2 and Layer 3 IP and MPLS features to support diverse applications. This core usually relies on an optical transport layer based on WDM. The connection of the aggregation and core networks is carried out through an IP edge router. It is assumed that mobile access network elements like RNC and BSC are connected to the IP edge router, and EPC network elements are connected to the core network. The core also allows the access to the service provider core where service platforms (e.g., IMS) reside.

  Although it was the case in the past, the capacity of the core network is not assumed to be blocking factor in the capacity that can be provided to the users due to the high capacity of optical transport networks.

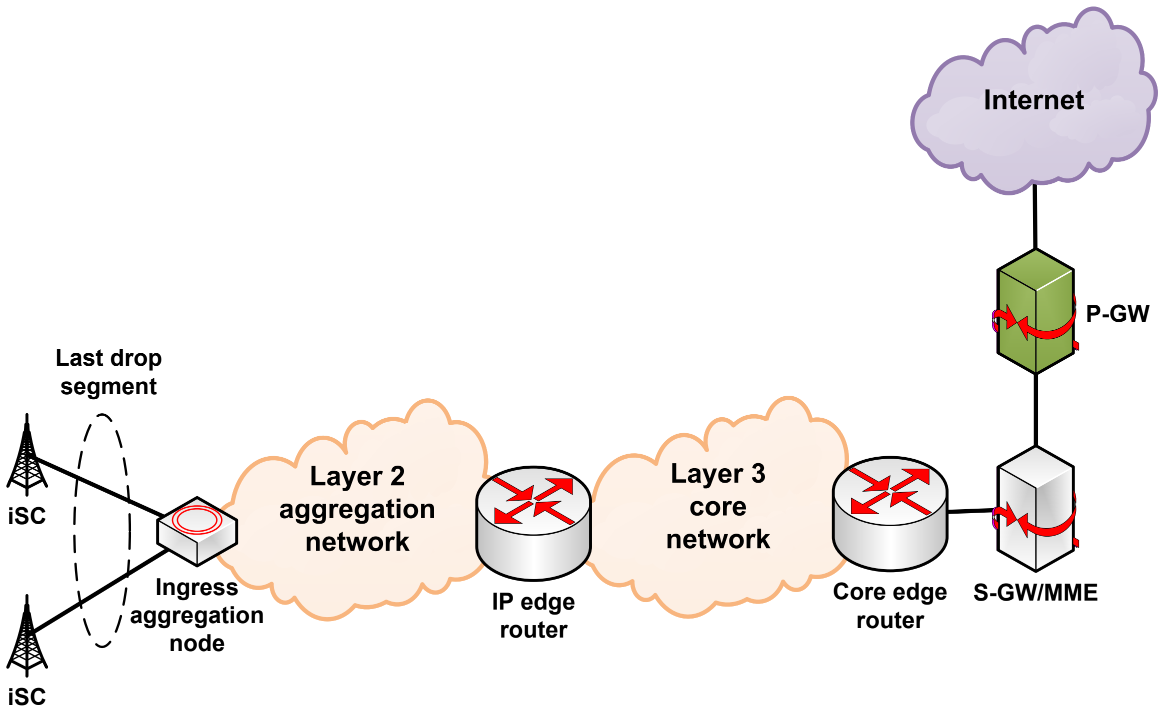


Figure 3-4: Proposed transport network

The core edge router would be a Provider Edge (PE) router in the case that the core network is not operated by the mobile network operator.

Based on this proposal, it is important to highlight the following aspects:

- Capacity estimation between the iSCs and the RANaaS infrastructure should take into account the limitations not only of the last drop segment, but also of the aggregation and core transport networks.
- Latency estimation should also take into account not only transmission and processing delays at the edge of the network, but also the contribution of the intermediate nodes of the transport network.

This classic configuration is expected to change in the future. The development of low cost long reach optical transport solutions is expected to allow for a radical transformation of this existing framework into a different one with only two segments: the last drop segment, which would be long reach. In addition, there is a trend to reduce the number of points of presence to access the core network[2], and extend the capabilities (mainly for QoS management) of MPLS to the last drop segment.

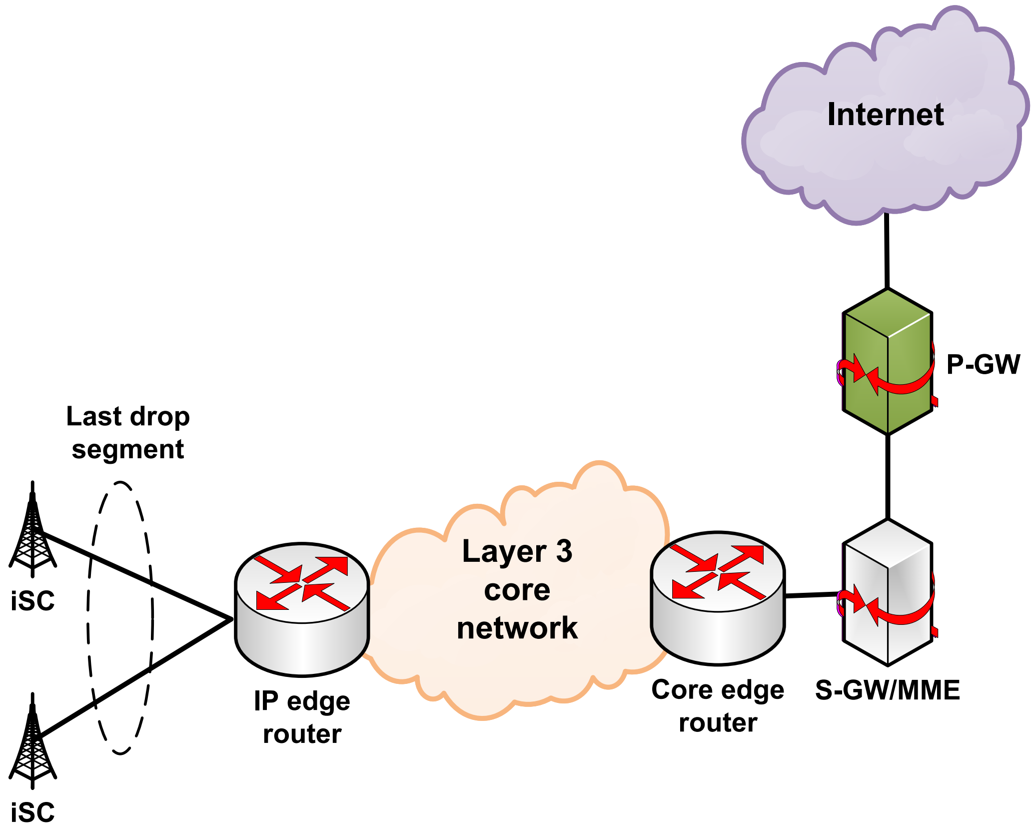


Figure 3-5: Evolved transport network

According to our analysis, the RANaaS can be connected to the iSCs through both the IP edge router and the core edge router. Another option would be to connect it through the ingress aggregation node.

3.3 Common scenario 1 – Stadium

This scenario is assumed to be characterized by the deployment of a high number of radiating elements in a much reduced volume so a very high traffic demand can be attended. The coverage should be relatively homogeneous in the seating areas and be planned taking into consideration the structure of the premises, so the area covered by each antenna is much reduced in order to maximize the capacity. Directional antennas are used to increase isolation between the cells, which may point from the pitch and upwards, or towards the pitch. Usually there will be 3-4 several rings of cells covering the seating areas, with a typical inter-cell distance in each ring of 10-30 m.

In existing deployments, the number of cells/sectors is relatively high (tens of them) as depicted in Figure 3-6. However, it can be expected that in the future this number may increase, in line with the number of Wi-Fi APs that are deployed (hundreds of them in some stadiums). Also, in the stadium there will be different types of cells, e.g., indoor/outdoor, to be used for specific tasks or with different SLAs depending on the seat.

[2] In the case of Telefónica Spain there is the expectation to reduce the number of PoPs from >700 to approximately 100. This reduction is expected to take place in conjunction with the switch-off of the PSTN.

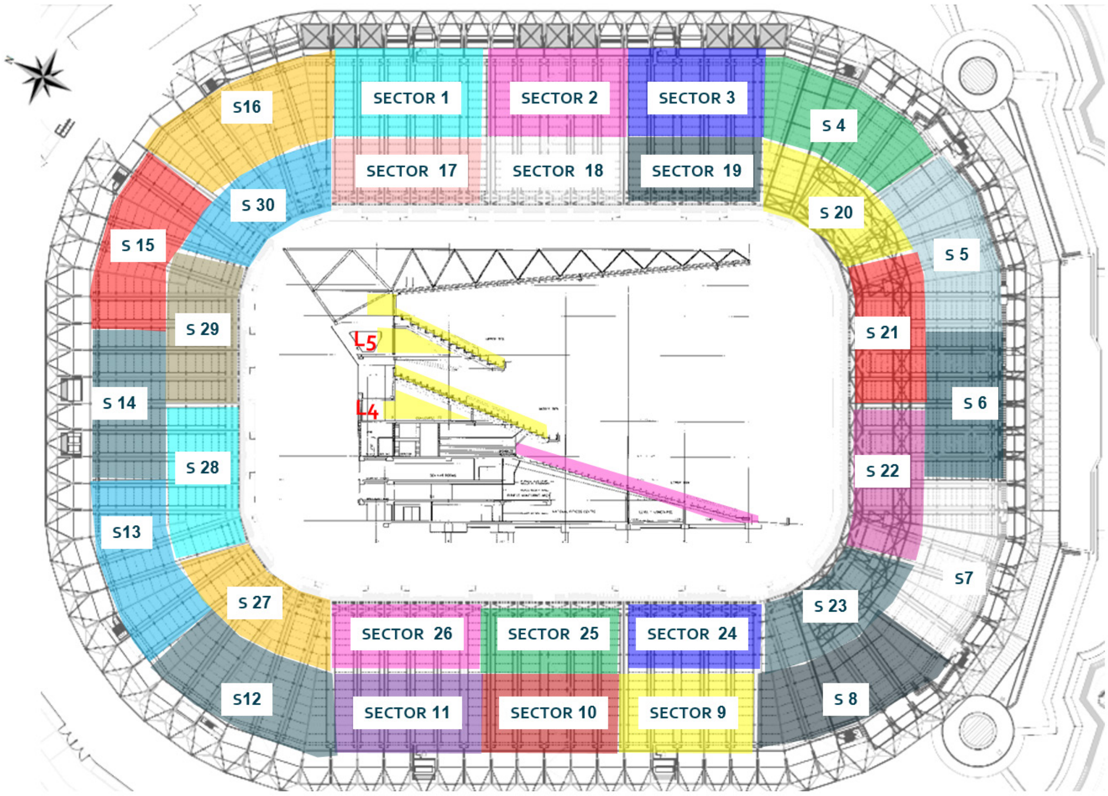


Figure 3-6: Example of sector distribution in seating area (Twickenham Stadium)

The selection of the location of the radio nodes (whether they are distributed antennas in a DAS system or remote radio heads) is a very complex task when both coverage and capacity are expected to be optimized. It usually requires the deployment of new transport infrastructure to feed them.

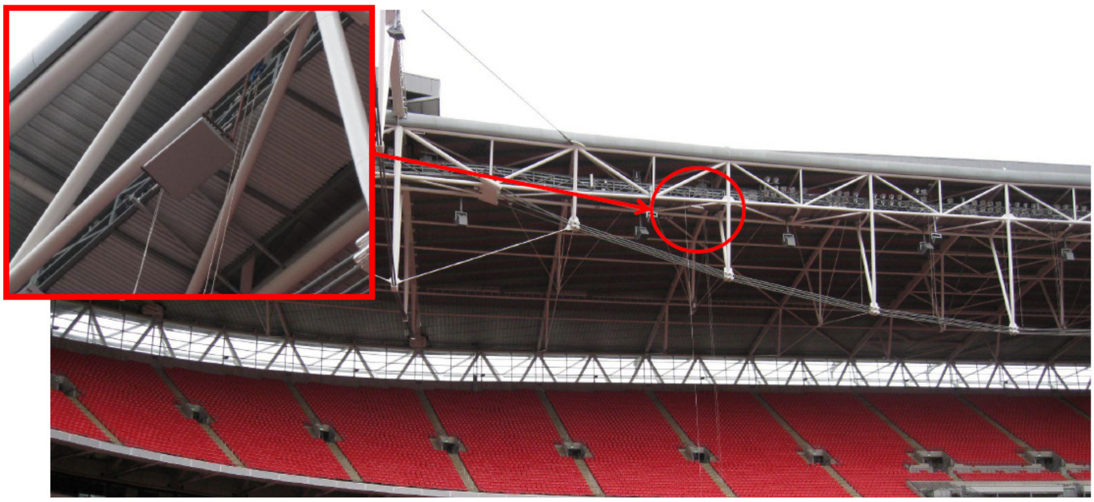


Figure 3-7: Example of detailed radio nodes location (Twickenham Stadium) (1)

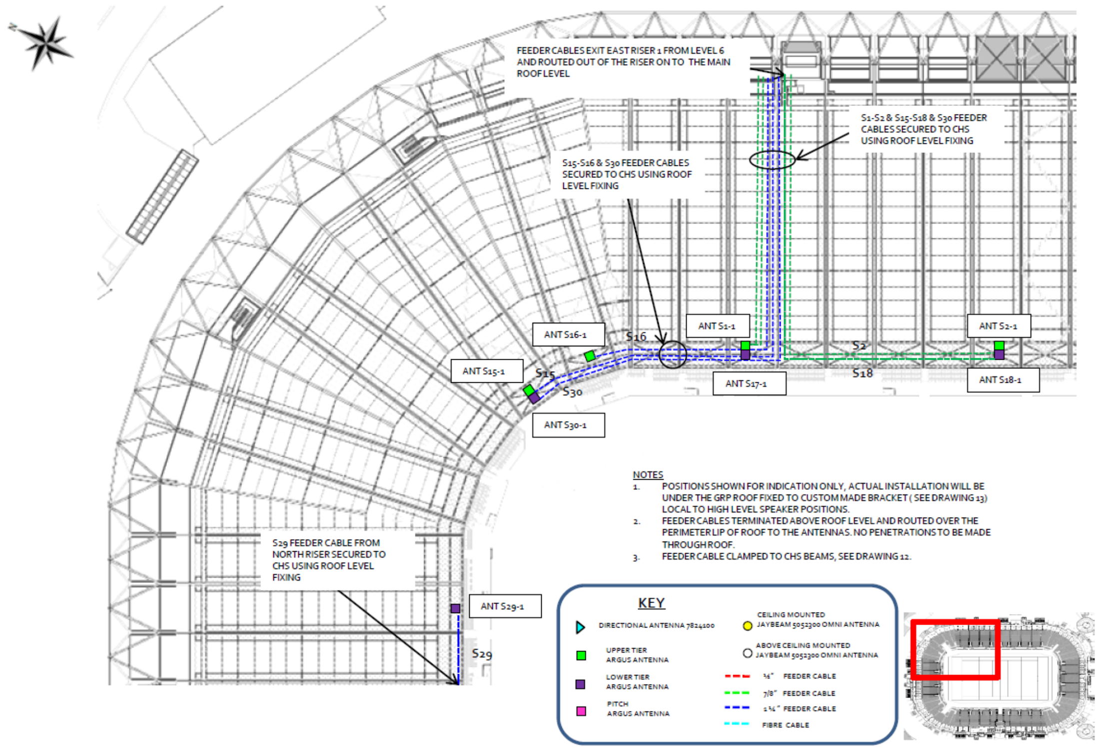


Figure 3-8: Example of detailed radio nodes location (Twickenham Stadium) (2)

In terms of the transport infrastructure deployed, in the actual solutions dedicated fibres are available for different sets of remote units.

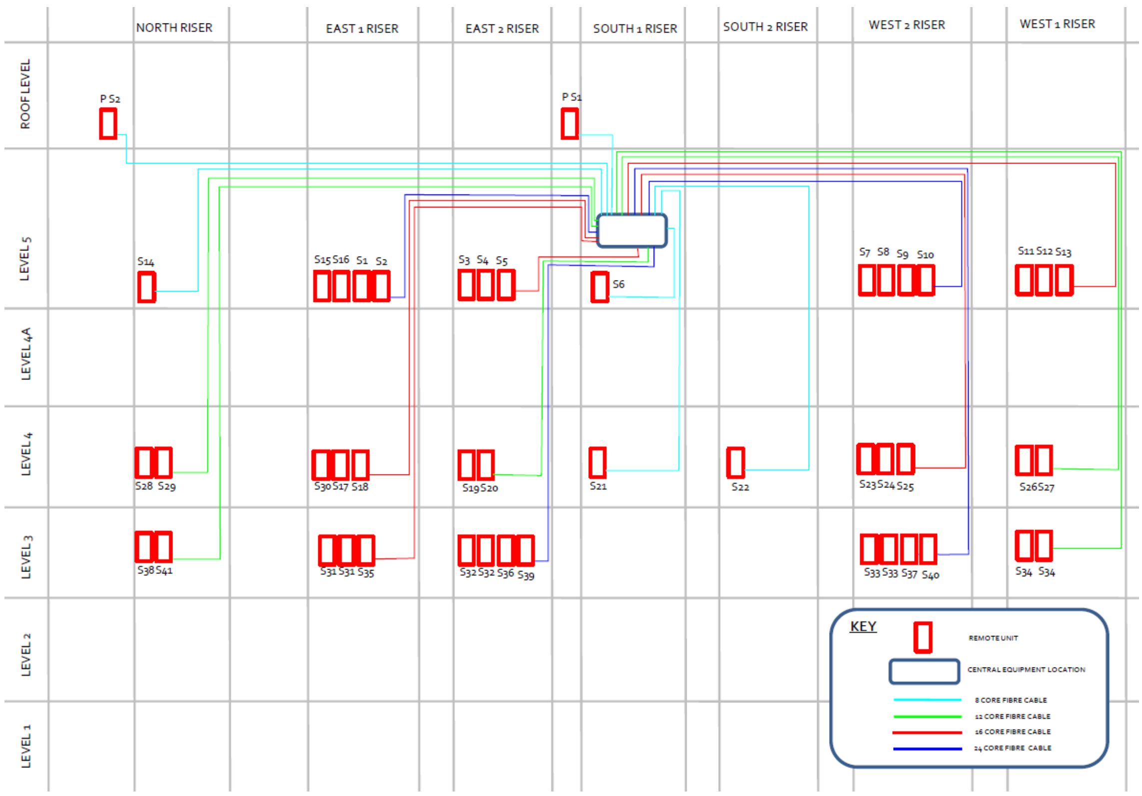


Figure 3-9: Example of interconnection of radio nodes and central processors (Twickenham Stadium)

On the other hand, there is a relatively large freedom in terms of location of the network elements and deployment of cables, except for specific areas inside the stadium. The fact that there is no need to dig trenches or other expensive construction procedures allows for the deployment of specific infrastructure that in other scenarios would not be feasible for economic reasons.

Figure 3-10: Example of cable routes inside a stadium (Twickenham Stadium)

The support of the very high traffic demands justifies the deployment of a dedicated infrastructure, both from a technical and an economic viewpoint. Also, it can be assumed that there will be no problems for the housing of the required infrastructure inside the stadium premises. However, there are some characteristics that should be taken into account:

- Passive transport infrastructure (cables, fibres, in some cases also antennas) are deployed and owned by the premises proprietor.
- This passive infrastructure should be shared by different operators and used to support different radio access technologies (possibly including Wi-Fi). For this reason, a solution based in distributed antenna systems and radio over fibre is usually implemented in this kind of scenario.

For this scenario, the following parameters and characteristics can be assumed:

- Last drop capacity: 1-10 Gbit/s, dedicated wavelength in dark fibre infrastructure.
- Last drop latency: 5-100 µs.
- Last drop transport technology: dark fibre, 10 GbE.
- Point to point topology of the last drop.
- Direct connection to the core network.
- Potential deployment of local breakout for Internet traffic.

3.4 Common scenario 2 – Square

For this scenario, the proposed RANaaS deployment is based on the standard network structure defined in Section 3.2, with the three segments indicated above: last drop, aggregation network and core network.

Figure 3-11 shows a physical deployment example for the square scenario. The following characteristics are considered:

- Last drop will be based on the use on wireless backhaul solutions, either microwave or millimetre wave.
- Point to pint and daisy chain topologies will be used in the last drop segment.
- Short length last drop segment (100-200 m).
- Ring/mesh is assumed for the aggregation network.

- Aggregation network capacity is considered to be the lowest value of the typical network capacity (e.g., 10 Gbit/s) and the result of the application of the dimensioning rule indicated in a previous section.

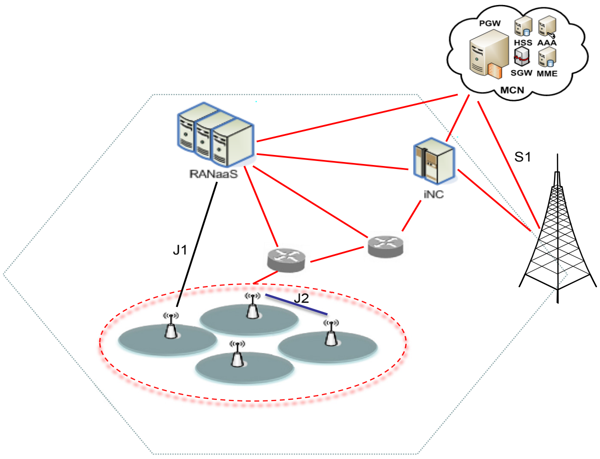


Figure 3-11: Square - Physical Deployment example

Figure 3-12 shows the small cell deployment for the square scenario according to the described characteristics.

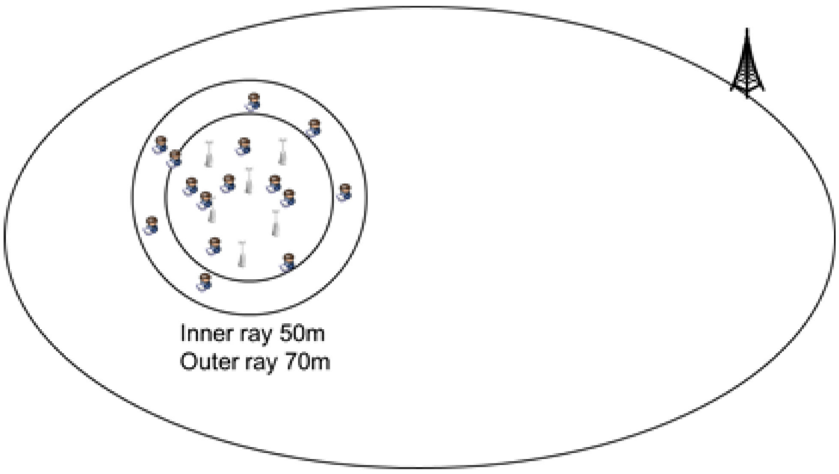


Figure 3-12: Small cell deployment in the square

It must be considered that the aggregation network in this scenario may be one of the main limiting factors in terms of bandwidth available. The reason is that the average capacity of the cells may be relatively high (LoS propagation, short distance between iSC and UE) and a multiplexing factor of 20-30% is most likely to be used.

3.5 Common scenario 3 – Wide area deployment

For this scenario, a prospective evolved transport architecture is proposed that corresponds to the expected long term evolution indicated in Section 3.2.

Figure 3-13 shows a physical deployment example for the wide area scenario. The following characteristics are considered:

- Long reach last drop, with a distance from the iSC to the core network of up to 10 km. The average distance is estimated to be 1.5 km.
- Point to point and daisy chain topologies will be used in the last drop segment.
- High last drop capacity per link, 10 Gbit/s
- No aggregation network, i.e., direct connection of the last drop to the core network.
- MPLS functionalities in the last segment (i.e., the iSC incorporates PE router functions).

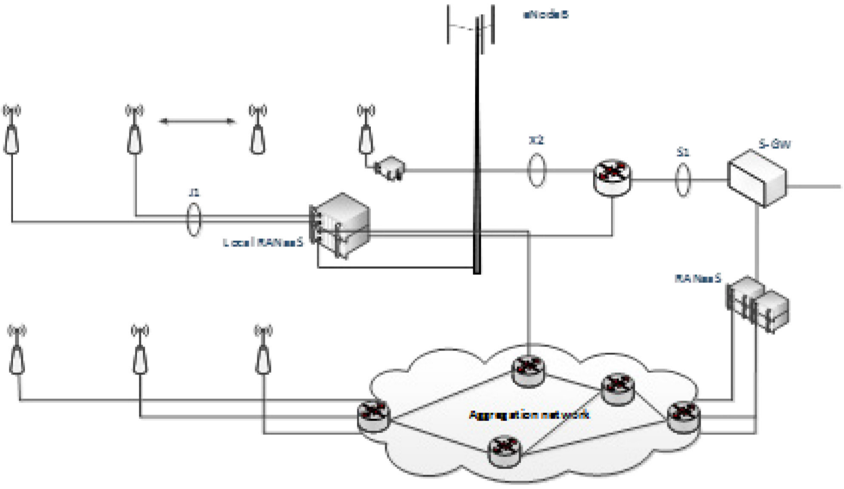


Figure 3-13: Wide Area - Physical Deployment example

Figure 3-14 shows the small cell deployment for the wide area scenario according to the described characteristics.

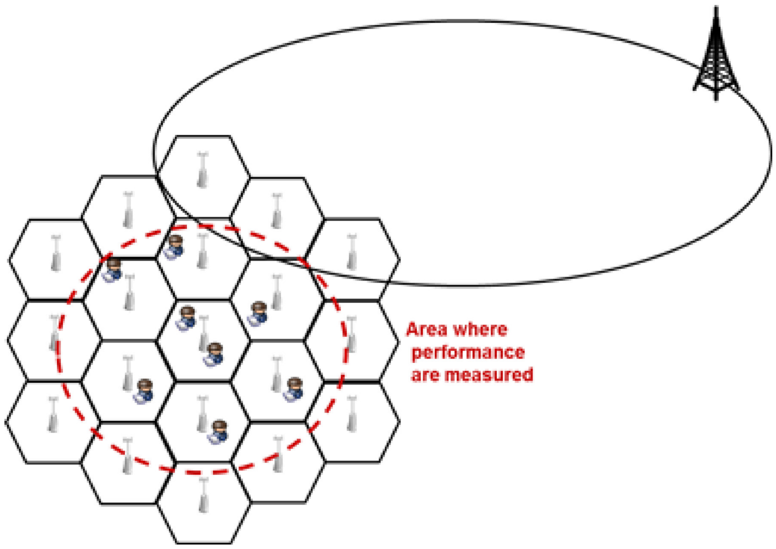


Figure 3-14: Small cell deployment for the Wide Area Coverage scenario

3.6 Common scenario 4 – Shopping Mall / Airport

In order to differentiate the shopping mall and Airport scenario from the Stadium scenario, it is proposed to assume that the solution has a lower performance for the transport infrastructure. This is in line with the fact that the traffic density that can be expected in scenario 4 would be lower than for scenario 1.

Figure 3-15 shows a physical deployment example for the shopping mall/airport scenario. The following characteristics are considered:

- Transport infrastructure may be provided by the premises owner, but it should not necessarily dedicated to the mobile network, i.e., it may be employed for other uses.
- Infrastructure should not necessarily shared by operators.
- There may be limitations in the processing infrastructure that may be deployed in the premises.

Based on these assumptions:

- Last drop capacity: 1 Gbit/s, dedicated wavelength in dark fibre infrastructure.
- Last drop latency: 5-200 µs.
- Last drop transport technology: 1 GbE
- Point to point/daisy chain topologies of the last drop.
- Direct connection to the core network.

- Potential deployment of local breakout for Internet traffic.

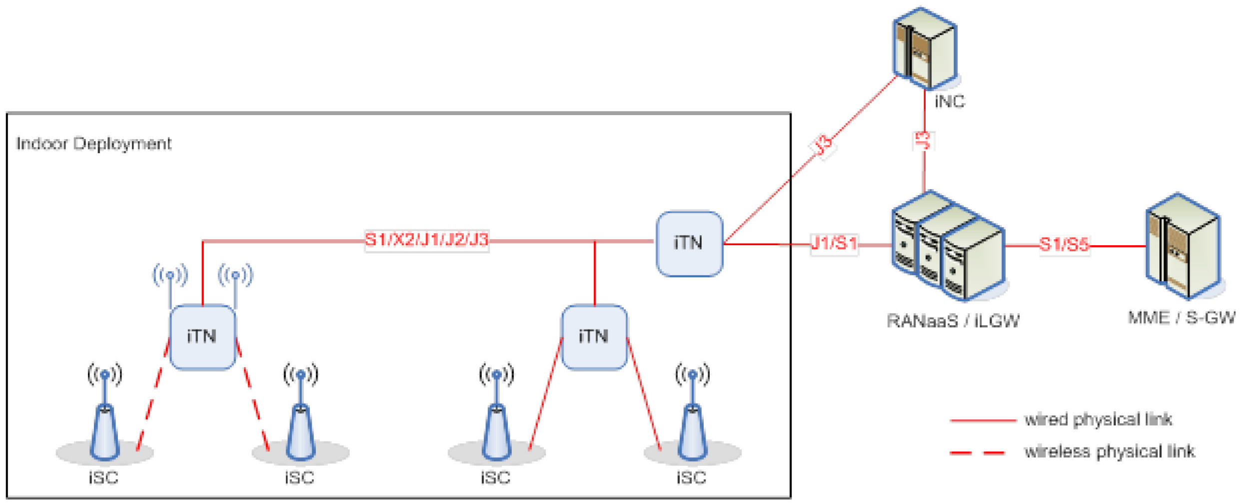


Figure 3-15: Shopping Mall/Airport: Physical deployment example

Figure 3-16 shows the small cell deployment for the shopping mall/airport scenario according to the described characteristics.

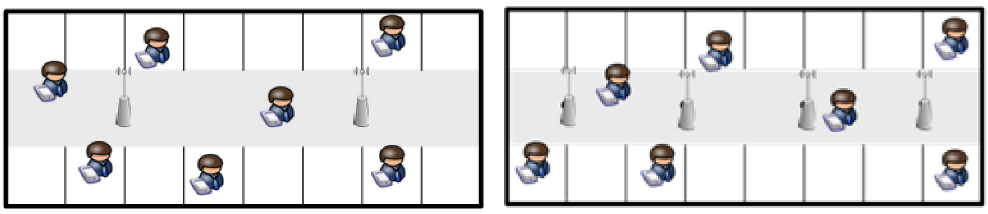


Figure 3-16: Small cell deployment in the Shopping Mall / Airport: Sparse (left) and dense (right) deployment

3.7 Common Scenario Assessment regarding functional split

In this section we provide a wrap-up regarding the feasibility of the different functional splits identified in iJOIN regarding the practical realisations of the four common scenarios described above.

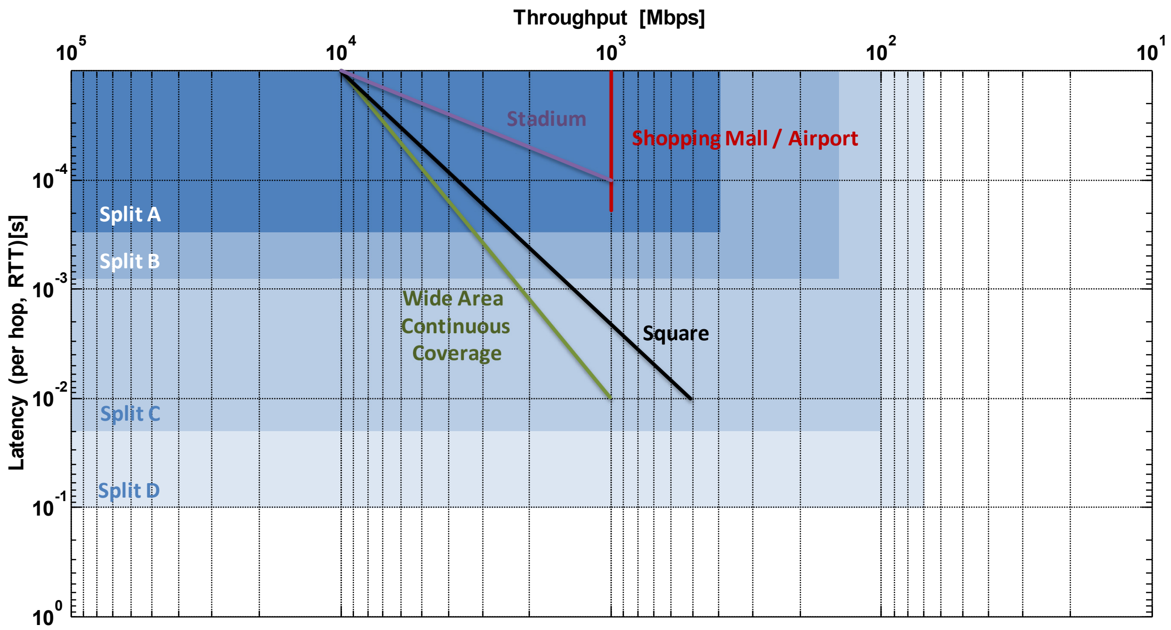


Figure 3-17: Functional split mappings to common scenarios

Figure 3-1 shows the mappings of the different functional splits regarding the common scenarios, by identifying these scenarios in the figure – which is based on latency and throughput – and highlighting the different functional splits. The take-away message of this figure is the following: the Stadium and Shopping Mall /Airport scenarios support all iJOIN functional splits, whereas the latency constraints limit the functional splits that can be selected in the Wide Area Continuous Coverage and Square scenarios.

4 Definition of the iJOIN Architecture to support Network layer approaches

4.1 WP4 in the iJOIN Architecture

The final iJOIN architecture is defined in the deliverable D5.3 [6], where the concept of the virtual eNode B (veNB), introduced in previous deliverables, is refined.

WP4 is responsible for defining the mechanisms required to allow a more efficient operation of the network layer in terms of cost, energy consumption, mobility support, latency, etc. These activities have resulted in the definition of a set of Candidate Technologies (CTs) that were described in detail in the deliverable D4.2 [4].

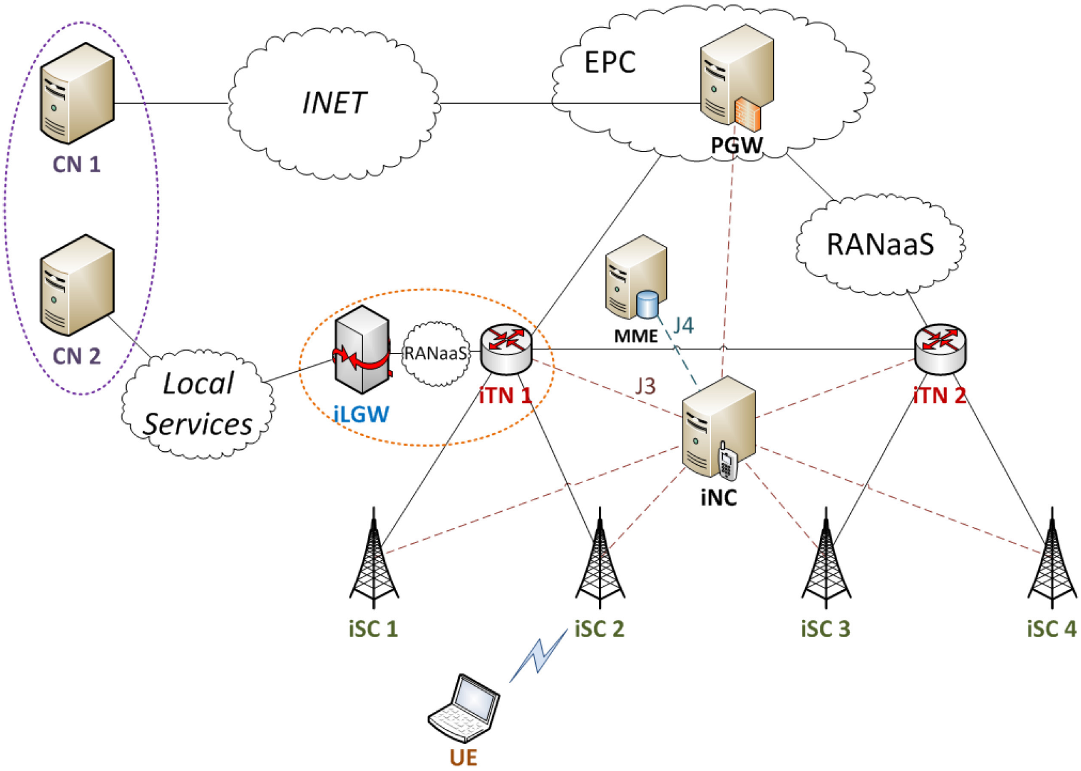


Figure 4-1: iJOIN transport network architecture

The transport network architecture is shown in Figure 4-1. The iJOIN Network Controller (iNC) is the key functional node from a WP4 point of view.

The iNC controls both the radio access and the backhaul network through an SDN protocol and it is responsible of centralized control of the transport plane of the veNB. It is in charge of configuring, monitoring and driving the operation of the rest of the RAN and backhaul network entities. In particular, we adopt the widely followed approach in the industry of using the OpenFlow protocol as SDN mechanism between the controller and the transport network entities iJOIN Transport Nodes (iTNs) as well as the iJOIN Small Cells (iSCs). In this sense, an extended OpenFlow controller located at the iNC takes care of all the protocol interactions with the rest of the network entities, which only need to support OpenFlow protocol, using the so-called Southbound protocol interface (which in iJOIN is the J3 interface). In addition to the pure forwarding computation intelligence, the iNC also hosts other functions, to support different management functions, such as mobility management (in order to perform these functions, the iNC has a J4 interface with the MME, used to exchange information about mobility events), energy efficiency or load balancing. Figure 4-2 shows the WP4 functional architecture and its relation with WP2 and WP3, as well as with other functional elements of the 3GPP architecture (like the MME).

Table 4-1 reports a summary of iNC modules described in D4.2 [4].

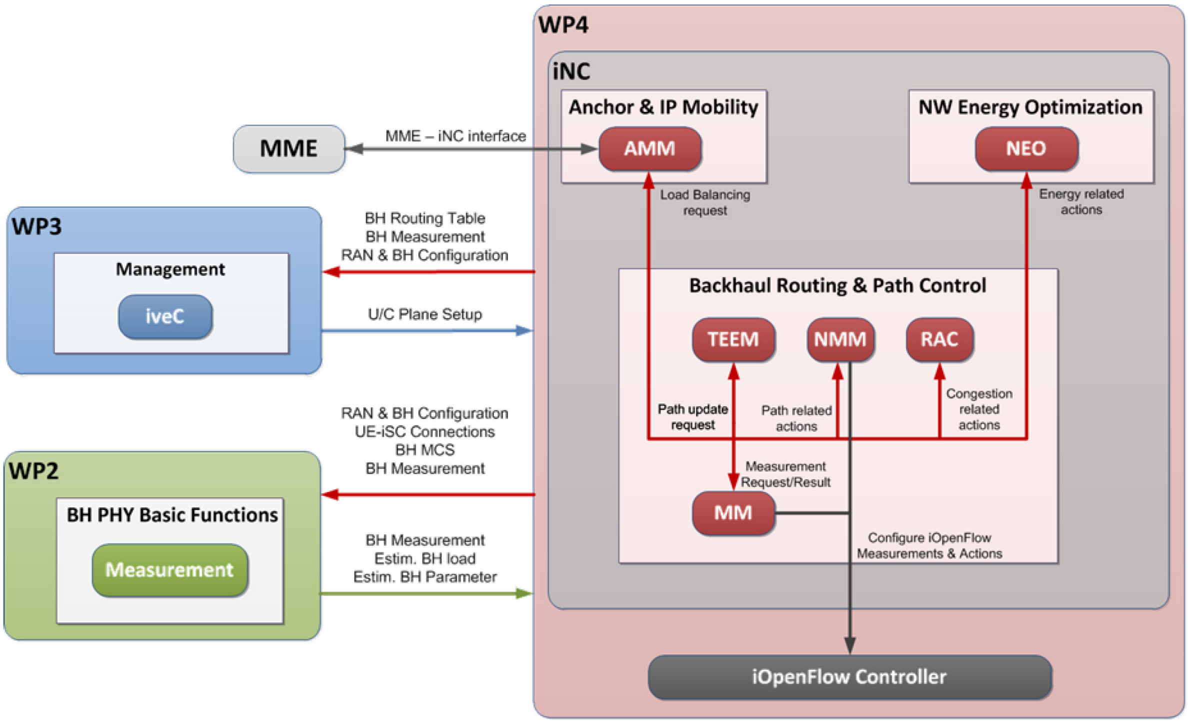


Figure 4-2: WP4 functional architecture

As introduced before, the iNC is a key network entity in the iJOIN architecture and it was designed as a composition of several modules. Each module is responsible for a specific mechanism and communicates with the other modules in order to accomplish a specific task. The operation of the iNC can be summarized as follows. Once the functional split to be executed by a set of iSCs has been determined, a routing module in the iNC is responsible of computing all the forwarding paths for each iSC-RANaaS existing flow. This computation takes into consideration the run time conditions of the network, which are obtained from continuous measurement and monitoring conducted in the backhaul. The routing module also interacts with the energy efficiency, congestion control and mobility modules, so a consistent and stable status is always maintained at the network. Note that since different functional splits may be executed at different veNBs connected to the same backhaul network, traffic differentiation might be needed at the transport level. Similarly, different RANaaS instances might be used, which can be located in different places. The distance with each of these instances is also a factor considered by the iNC when computing paths .

Table 4-1. iNC modules summary

iNC Module	Description
NMM	Network Model Module (NMM) acquires the topological and functional view of the network.
AMM	Anchor and Mobility Management (AMM) implements most of the functionality related to mobility management.
NEO	Network-wide Energy Optimisation (NEO) monitors the status and the load of the network, and runs different algorithms to optimize the overall energy consumption.
RAC	Routing and Congestion (RAC) is in charge of avoiding network congestion in the RAN/backhaul, by properly configuring the network and requesting changes on the paths used by active data flows.
TEEM	Traffic Engineering Enforcement Module (TEEM) hosts all the intelligence required to compute the best path within the backhaul to support the different traffic and network-wide requirements, providing the necessary conflict resolution functions.
MM	Measurement Module (MM) is in charge of configuring the iOpenFlow controller to perform the required measurements.
iOpenFlow Controller	The controller takes care of all the protocol interactions with the rest of the WP4 network entities which only need to support OpenFlow, using the so-called Southbound API.

As highlighted in Figure 4-2, the iNC interacts with the iJOIN virtual eNodeB Controller (iveC), which is in charge of controlling the RAN functional split[3]. The iveC is also responsible for the bootstrapping of the veNB. This encompasses several actions, such as J1 interface establishment, RAN configuration, and SON-related functionalities such as reference symbol auto-configuration. More complex aspects could also be considered, for example in the case of sharing the infrastructure among multiple tenants, as not all the iSCs might be available for a given operator (e.g., because of a lack of agreements or resources). The bootstrapping should also take care of provisioning the iSCs with proper security credentials, based on existing trust relationship (e.g., mutually trusted certificates). Note that an iSCs lacking the required security credentials would not be validated and will be left of the veNB bootstrapping process.

In addition to the bootstrapping, the iveC also regularly monitors the state of the veNB, reacting upon events such as iSCs that become unreachable or present a faulty operation. Finally, the iveC is responsible for implementing the 3GPP eNB architecture. This includes setup and termination of the eNB user and control plane interfaces (S1/X2), GTP tunnel management, specific configuration, etc.

iJOIN architecture requires some OpenFlow extensions, which are aligned with the work being carried out at the Wireless & Mobile Working Group of the Open Networking Foundation (ONF), the standardization body in charge of the OpenFlow specifications.

4.2 Compliance with ONF Architecture

The most visible of the SDN protocol stacks is the OpenFlow protocol, which is maintained and extended by the ONF. Originally this protocol was developed specifically for IEEE 802.1 switches conforming to the ONF OpenFlow Switch specification. As the benefits of the SDN paradigm have reached a wider audience, its application has been extended to more complex scenarios such as Wireless and Mobile networks [8]. Within this area of work, the ONF is actively developing new OFP extensions addressing three key scenarios: (i) Wireless backhauling, (ii) Cellular Evolved Packet Core (EPC), and (iii) Unified access and management across enterprise wireless and fixed networks.

The iJOIN SDN-architecture adopts the ONF SDN specification [17], which itself applies the generic principles of SDN, to come up with a more specific – though still open for customization – architecture made of components and interfaces. Such architecture definition is limited to functional interfaces between software components and it does not specify the internal design of the SDN controller.

ONF Architecture

Figure 4-3 shows the blocks and the functional interfaces of the ONF architecture, which comprises three planes: Data, Controller, and Application.

The Data plane comprehends several Network Entities (NEs), which expose their capabilities toward the Controller plane via a Southbound API. The Controller plane includes several cooperating modules devoted to the creation and maintenance of an abstracted resource model of the underneath network. Such model is exposed to the applications via a Northbound API where the Application plane comprises several applications/services, each of which has exclusive control of a set of exposed resources.

The Management plane spans its functionality across all planes performing the initial configuration of the network elements in the Data plane, the assignment of the SDN controller and the resources under its responsibility. In the Controller plane, the Management needs to configure the policies defining the scope of the control given to the SDN applications, to monitor the performance of the system, and to configure the parameters required by the SDN controller modules. In the Application plane, Management configures the parameters of the applications and the service level agreements. In addition to the these interactions, the Management plane exposes several functions to network operators which can easily and quickly configure and tune the network at each layer.

[3] More information about the iveC can be found in deliverables D3.3 and D5.3.

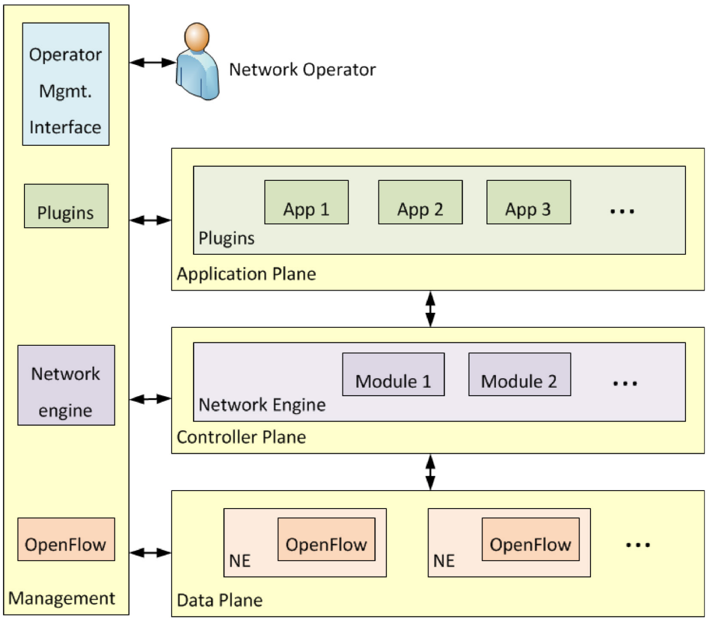


Figure 4-3: ONF SDN-based architecture

WP4 iNC modules depicted in Figure 4-2, can be implemented as application of the iOpenFlow Controller according to the ONF architecture shown in Figure 4-3. Doing so, the iJOIN architecture is compliant with the ONF SDN architecture.

4.3 RAN Functional split

Within iJOIN, four main functional split options, namely A, B, C and D, have been defined. Each functional split defines a different balance between local processing in the iSCs and central processing in the RANaaS platform. These splits are depicted in Figure 4-4 and are described in D5.2 [5].

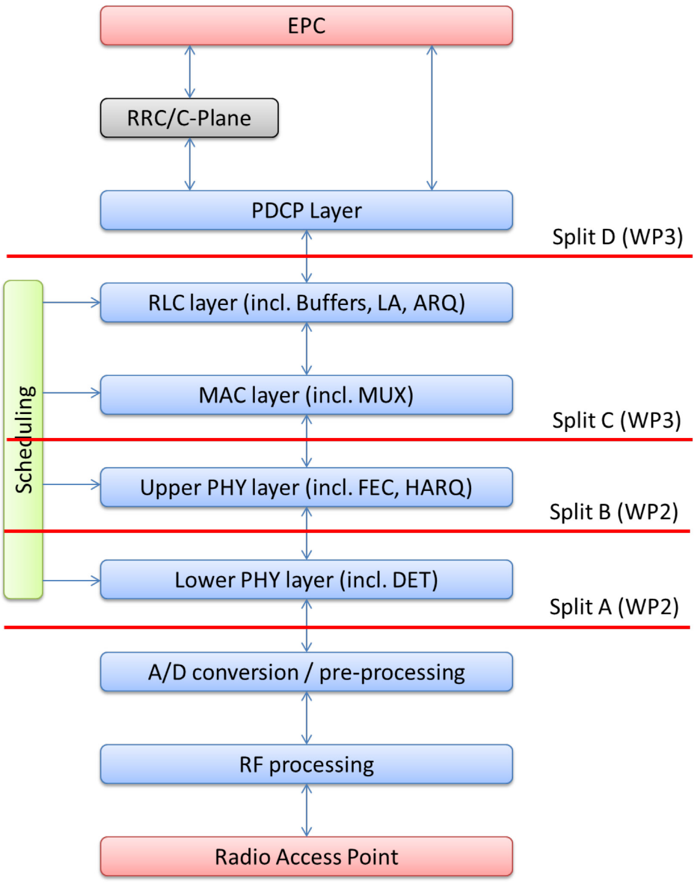


Figure 4-4: Project-wide functional split options

Note that in WP4, the different functional splits options affect only marginally the Candidate Technologies, as shown in Table 4-3, which describes the compatibility of the WP4 CTs with the iJOIN functional splits. Table 4-2 reports a summary of WP4 CTs described in D4.2 [4].

Table 4-2: WP4 CTs summary

CT	Description
CT4.1	Distributed IP Anchoring and Mobility Management selects the optimal anchor for the user via which the traffic is routed outside the mobile network.
CT4.2	Network Wide Energy Optimisation defines different algorithms to optimise the overall energy consumption while ensuring that the network wide performance is not compromised.
CT4.3	Joint Path Management and Topology Control takes care of locating and dimensioning the RANaaS data centres, and to associate each RANaaS with a particular set of iJOIN Small Cells (iSC).
CT4.4	Routing and Congestion Control Mechanisms provides an adequate technical solution for the congestion control issues in the transport nodes that support RANaaS and flexible functional split.
CT4.5	Network Wide Scheduling and Load Balancing distributes evenly the traffic in the network in order to improve the performance and pre-empt congestion by taking into account the complete end-to-end path.

Table 4-3: CT compatibility for different functional splits

CT	Split A	Split B	Split C	Split D	Aspects affected by functional split
CT4.1	x	x	x	x	No change required for different functional splits. CT4.1 gain depends on RANaaS/Gateways location.
CT4.2	x	x	x	x	No change required for different functional splits. Depending on the functional split, the energy efficiency gains might be higher or lower, as the traffic load injected in the backhaul depends on the functional split (indirect impact).
CT4.3	x	x	x	x	No change required for different functional splits. It is an enabler for functional split.
CT4.4	x	x	x	x	No change required for different functional splits. The functional split indirectly impacts this CT, as different functional splits involve different traffic loads on the backhaul, therefore affecting on the congestion control.
CT4.5	x	x	x	x	No change required for different functional splits. The functional split indirectly impacts this CT, as different functional splits involve different traffic loads on the backhaul, therefore affecting on the load balancing..

However, WP4 CTs are actually enablers of the different functional splits. Based on the backhaul topology and access technologies (e.g., fibre, wireless point-to-point, wireless point-to-multipoint, etc.), the network status, the RANaaS placement, and the users' demands, the iveC together with the iNC are capable of computing the functional split levels that are feasible. This is necessary because not every functional split is possible on a given network scenario, as the backhaul network delay and bandwidth might not be sufficient to centralize a given function on the RANaaS and it is required to execute it locally at the iSC. Note that the functional split control should also take into account the capabilities of the iSCs (e.g., CPU processing power, availability and type of antennae, etc.) and the level of connectivity among them (i.e., the expected latency and bandwidth that can be obtained between neighbouring iSCs and between the iSCs and the RANaaS).

4.4 Joint RAN/BH Optimisation

This section summarises the benefits of a joint RAN/BH optimisation in dense small cell networks from a network layer (WP4) point of view. We present some candidate Network Layer technologies which can be key enablers of the joint BH/RAN design.

Figure 4-5 shows the general information exchange among the iJOIN functional nodes for joint RAN/BH optimisation. More details can be found in D3.3 [2] and D5.3 [6]. We next summarise how the WP4 CTs make use of this information flows to benefit from and enable a joint RAN/BH optimisation.

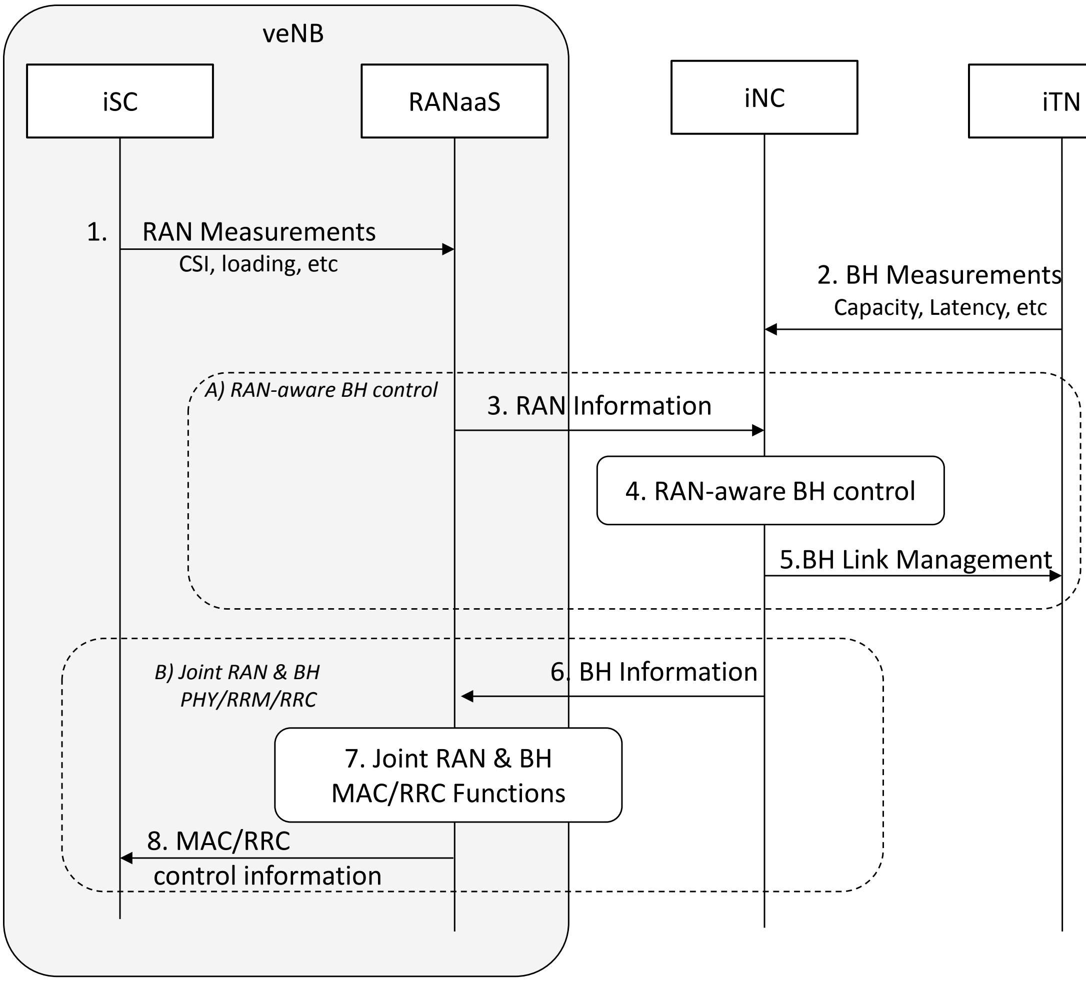


Figure 4-5: Information and control schemes for joint RAN/BH control

4.4.1 Mobility management and anchor selection

The logical centralisation enabled by the SDN-based architecture enables the mobility & anchor intelligence running on the iNC (i.e., the AMM) to be capable of dynamically and timely react to mobility events, considering not only the aspects of the radio access (which we could refer to as "legacy approach" in existing mobility architectures), but also the status and load of the backhaul network, together with the different functional splits that might be in use in the network at a given moment. This requires an active monitoring of the network topology as well as the load of the different links (iTN-iNC and iSC-iNC interactions), in addition to an up-to-date knowledge of the functional splits being used (RANaaS-iNC interaction). This introduces a signalling overhead that depends of the required freshness of the monitored information. Depending on the backhaul technologies and functional split in use, the amount of signalling and its periodicity can be dynamically adjusted. For example, retrieving the statistics associated to an OpenFlow rule produces a signalling load of 166 bytes. Similarly, 126 bytes and 66 bytes is the signalling load required for gathering the state of a single switch's port and of a single switch's queue respectively. Those values refer to the OpenFlow signalling and do not include TLS, TCP and IP headers.

4.4.2 Energy efficiency

The iJOIN network wide energy efficiency scheme fully relies on the logical view enabled by the iNC. Thanks to it, an energy reduction can be achieved by jointly considering the status at the radio access and at the backhaul. To this end, before making a switching-off decision for a cellular node one should consider the user QoS, to avoid its degradation. Some key metrics we used to model the user QoS are the following (see [15]):

- **Failure Probability**: i.e. the probability that a random user experiences poor signal quality when it needs to use the network (e.g. making a call, or sending a web request).
- **Admission control and "blocking" probabilities**, i.e. the probability that a flow that requires a certain amount of (dedicated) bandwidth, is blocked due to the lack of the available resources
- **Admission control and "service delay" for regular "best-effort" flows**, i.e. the ongoing delay for the flows that are multiplexed

4.4.3 Congestion control

With respect to the congestion control procedures, it is clear that once the traffic load growth makes the proposed mechanisms unable to provide the expected quality of service, there are three main correcting measures left (see [4]):

- Changing the route of some flows to links that are less loaded.
- Changing the location of the RANaaS.
- Changing the functional split of some of the flows that go through the congested nodes, so their associated bit rate is reduced.

The first two mechanisms can be supported with the procedures defined for other CTs, like routing, load balancing and RANaaS dynamic relocation. The third one, changing the functional split, may imply the consideration of additional requirements that hitherto have not been taken into account[4].

4.4.4 Load balancing

One crucial aspect of the network, is the one of Load Balancing. Load Balancing mechanisms distribute the load within the network elements evenly, to avoid congestion, and thus decreased QoE (e.g. if the load of an iSC exceeds 1, the expected delay goes to the infinity). To this end, we propose a centralized algorithm for load balancing that takes into account:

- Fairness issues, w.r.t. the desired degree of load balancing: as analysed in [D4.2].
- Traffic Differentiation w.r.t. the best-effort and dedicated flows and resources.

4.4.5 RANaaS placement and dynamic relocation

Another clear example of joint RAN/BH optimisation in WP4 is the RANaaS placement and dimensioning. Depending on the load of the RAN and the BH networks, different deployment of RANaaS placement and dimensioning could be adopted. Since the networks are not static, the changes on the RAN/BH may require the migration of the RANaaS or the addition of computational resources to the running RANaaS instance. For the sake of the dynamical adaption, the following information is required:

- **Capacity, latency and topology in the BH networks**: An active monitoring between iTN-iNC in the BH networks is required.
- **User traffic demands and functional split options in the RAN networks**: Traffic of each iSC should be analysed and reported to iNC. The functional split options used between iSC-RANaaS should be aware. For example, the traffic in the business area could greatly reduce in evening while the traffic in the residential area could increase. Due to the change of the traffic distribution, the deployment of RANaaS may be required to migrate accordingly.

In summary, the adaptive RANaaS relocation would be achieved by considering jointly the conditions of RAN and BH.

[4] It is not clear that the functional Split of an iSC can be changed while it is still serving users, as would be the case most likely in a congestion situation.

5 WP4 evaluation methodology

This section summarises the methodology used for the evaluation of the WP4 CTs, which is reported in Section 6. We first describe the scenarios that have been considered in the evaluation of each of the CTs, and then we enumerate and explain the assessed metrics.

5.1 Considered scenarios

5.1.1 CT4.1 - Distributed IP Anchoring and Mobility Management

With the purpose of evaluating CT4.1, we consider the Wide-Area scenario described in D5.2 [5] with the reported functional splits (A, B, and C). Such scenario defines several constraints in terms of bandwidth and latency between the iSCs and the RANaaS, but does not define any network topology. In order to evaluate our CT, we define a backhaul topology following the criteria reported in Section 3.5. As reported in D4.2 [4], CT4.1 is going to be evaluated on the SDN testbed which is described in details in D6.2 [7]. Therefore, the topology design is bounded to the testbed capabilities. Figure 5-1 depicts the designed topology which comprehends 19 iSCs, 8 iTNs, 2 iJOIN Local Gateways (iLGWs) and 1 P-GW. The iSCs are so connected: 7 iSCs are connected to the iTN 7, 5 iSCs to iTN 8 and the other 7 to iTN 9.

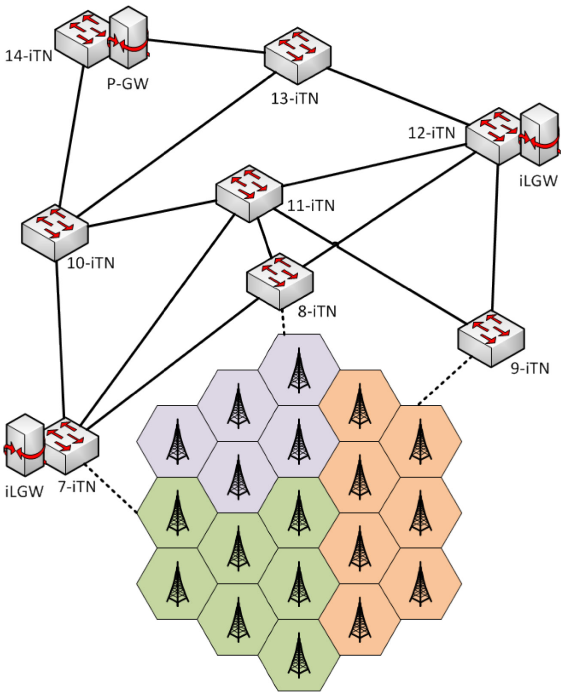


Figure 5-1: Wide-area backhaul topology

In Section 6.1 we will evaluate different combinations for link capacity and the RANaaS placement leading to multiple cases of the considered topology.

5.1.2 CT4.2 - Network Wide Energy Optimisation

The evaluation, based on simulations, considers the iJOIN common scenario Stadium case scenario (CS1), as described in D5.2 [5]. From WP4 perspective, the systems' energy efficiency mainly depends on the number of iSCs, and more precisely on their potential redundancy: for instance, if there are too many iSCs active during the night, or during a short low-traffic time period (e.g., during the lunch break in an industrial area), we can switch-off the redundant iSCs, and decrease the number of iTNs. To this end, the time-scale for the WP4 energy-efficiency scheme, is in the order of some minutes (or even hours).

In addition, the functional split affects the potential energy savings due to differentiation of the respective required rates, for example, the lower the functional split the more the exchanged signalling information between the RANaas and the iSCs, so the more the backhaul nodes needed for carrying up these (potentially large amount of) information.

For example, functional split A requires a transport capacity to the RANaas ~1.8Gbps per cell (most of the traffic should pass through the central node RANaas); whereas functional split D only 70 Mbps per bearer (the centralization here is quite low). It turns out that, Functional Split A will require to be switched-on most of the iSCs, due to the need to send information (digitized signals) independently of the user traffic volume, whereas Functional Split D will require to be switched-on a higher number of iSCs, since there is no centralization need. To this end, we chose to investigate the energy savings in Functional Split B and C, to give the insights of the aforementioned dependency. Functional Split A is not considered, since it requires almost all the cellular nodes to be switched-on.

5.1.3 CT4.3 - RANaaS location analysis

For evaluating the RANaaS deployment algorithm, the common scenarios 1-4 described in Section 3 are considered. Furthermore, simulations with large topologies are presented in order to provide a more general view. The detailed settings of the topologies for the simulations are described in the following:

Large Random Topology

Figure 5-2 shows the large random topology in our simulation. There are 500 iSCs and 50 iTNs in the network. The last drop link is point-to-point microwave or millimetre wave with 1Gbps bandwidth and 200 µs latency. The point to point topology is used in the last drop segment, while the mesh topology is assumed for the aggregation network. Two different link settings are used in the aggregation networks:

1) Metro optical network with 1Gbps bandwidth and 250 µs latency.
2) Dark fibre with 10 Gbps bandwidth and 5 µs/km latency.

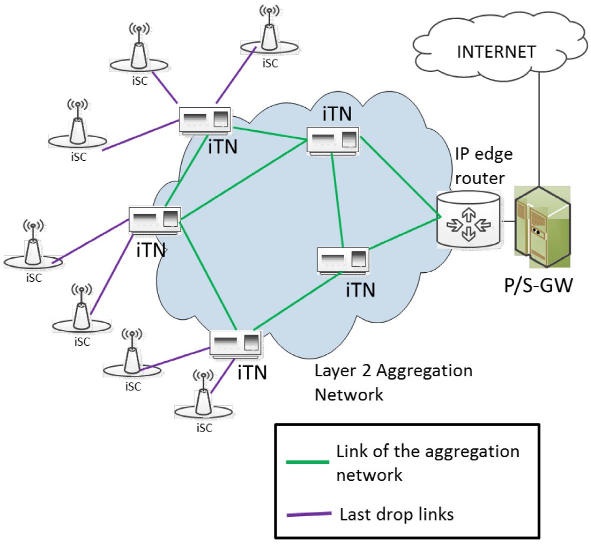


Figure 5-2: Large random topology for the evaluation of RANaaS deployment study

Common Scenario 1 – Stadium

Figure 5-3 shows the topology of common scenario 1 - Stadium for the evaluation of RANaaS deployment study. There are 100 iSCs in the networks, 1 IP edge router and 1 local breakout L-GW, collocated with the IP edge router. The link technology of the last drop is dark fibre with 10 Gbps bandwidth and 5 µs/km latency. The point-to-point topology is used for the last drop. There are direct connections between the iSCs and the core network, i.e. the iSCs are connected directly with the IP edge router.

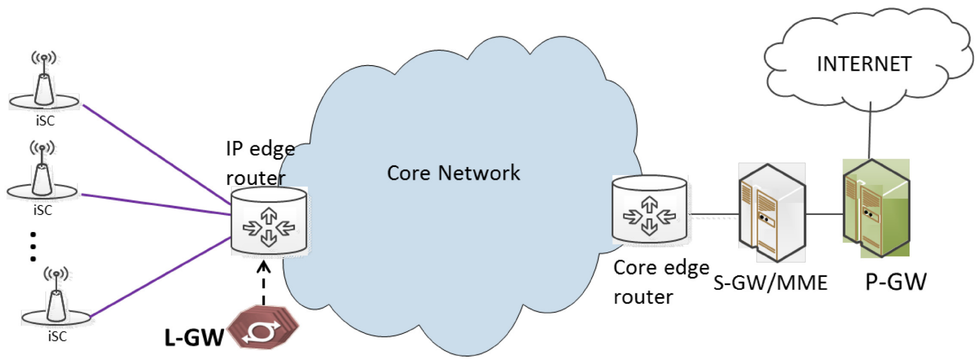


Figure 5-3: Topology of common scenario 1 - Stadium for the evaluation of RANaaS deployment study

Common Scenario 2 – Square

Figure 5-4 shows the topology of common scenario 2 - Square for the evaluation of RANaaS deployment study. There are 20 iSCs in the square and 5 iTNs in the aggregation network. The last drop link in this scenario is based on the use on wireless backhaul solutions, either point-to-point microwave or millimetre wave with 1Gbps bandwidth and 200 µs latency. The distance of the last drop segment is short, about100-200 m. The point to point and daisy chain topologies are deployed in the last drop segment. The aggregation network is deployed as a Metro optical network with 1Gbps bandwidth and 250 µs latency. A mesh topology is used in the aggregation network.

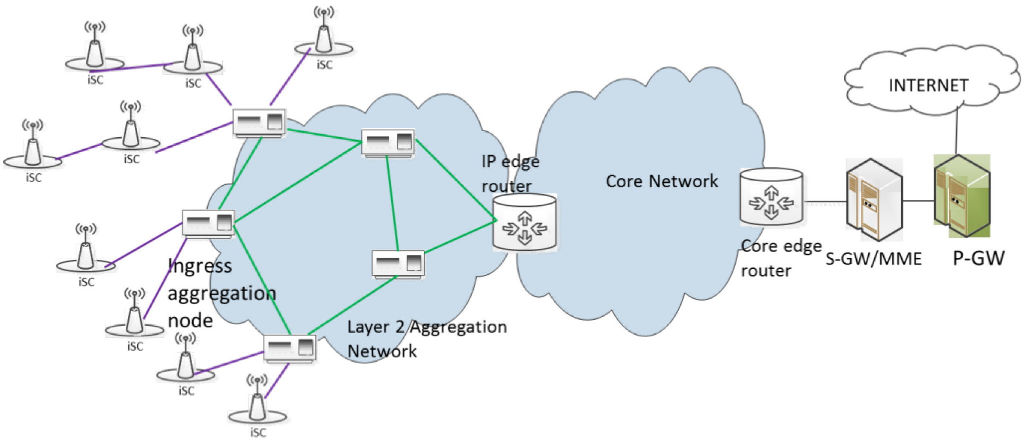


Figure 5-4: Topology of common scenario 2 - Square for the evaluation of RANaaS deployment study

Common Scenario 3 – Wide-Area Deployment

Figure 5-5 shows the topology of common scenario 3 – Wide-Area deployment for the evaluation of RANaaS deployment study. There are 50 iSCs in the area with 1 IP edge router. The distance between the iSC and the core network is average 1.5 km, up to 10 km. The point to point and daisy chain topologies are deployed in the last drop segment. The average degree of daisy chain is 2. The links in the last drop are dark fibres with 10 Gbps bandwidth and 5 µs/km latency. There are direct connections between the last drop and the core network.

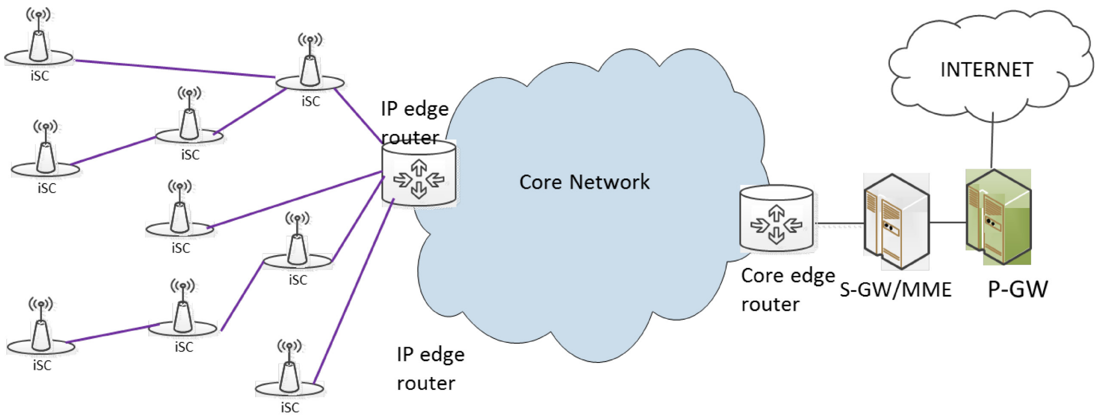


Figure 5-5: Topology of common scenario 3 – wide area deployment for the evaluation of RANaaS deployment study

Common Scenario 4 – Shopping Mall / Airport

Figure 5-6 shows the topology of common scenario 4 – Shopping Mall/Airport for the evaluation of RANaaS deployment study. There are 50 iSCs in the mall/airport, 1 IP edge router and 1 local breakout L-GW, collocated with the IP edge router. The last drop link is dark fibre with 1 Gbps bandwidth and 5 µs/km latency. The point to point/daisy chain topologies are used in the last drop segment. The average degree of daisy chain is 2. The last drop segment has the direct connection to the core network.

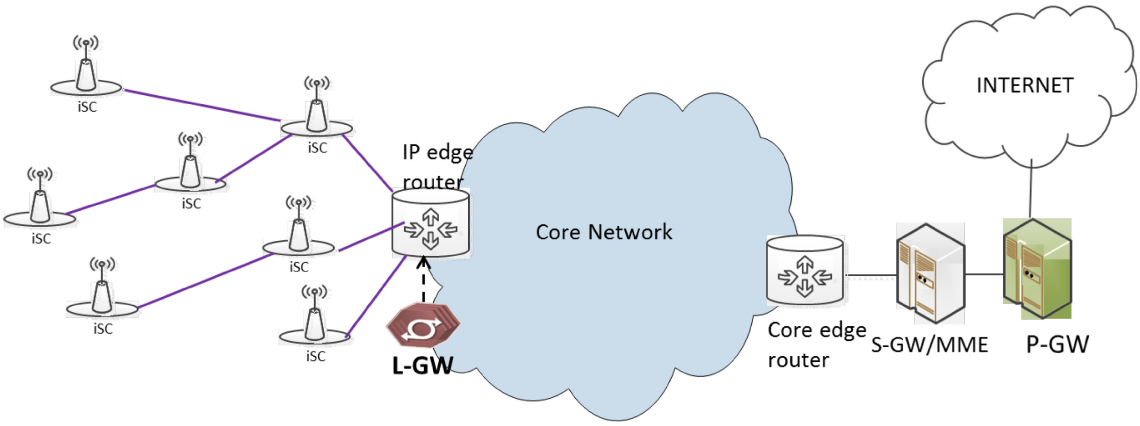


Figure 5-6: Topology of common scenario 4 – shopping mall/airport for the evaluation of RANaaS deployment study

5.1.4 CT4.4 - Routing and Congestion Control Mechanisms

This CT is considered to be applicable to the Wide-Area deployment (CS3). For the evaluation of the CT, the scenario that is being simulated is the one represented in the following figure, consisting in a variable number of iSCs (up to 8), two iTNs that implement the CT (that correspond to the IP Edge Router and Core Network Edge Router of the proposed architecture associated with CS 3), and a RANaaS entity and a EPC entity. Optionally, the use of an iNC for the control of the CT can be emulated.

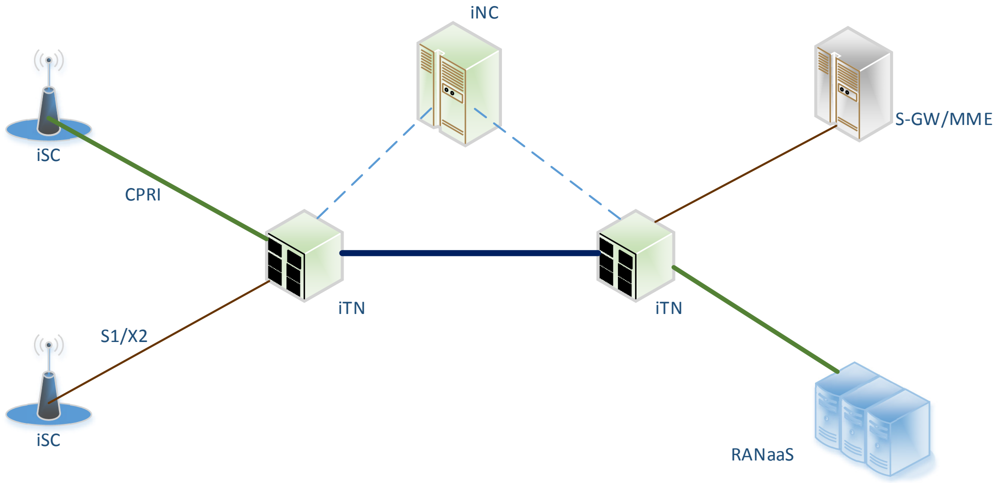


Figure 5-7: Simulation scenario for the evaluation of CT4.4

Traffic can be generated from both directions (i.e., iSCs → RANaaS/EPC and reverse). The flows that are being evaluated have been characterized in the following categories:

- Variable bit rate flows:
  - Conventional S1 flows.
  - Intra MAC layer flows (Split C.1).
  - Soft bits flows associates to intra PHY layer functional split (Split B.2).
- Constant bit rate flows:
  - High bit rates flows, corresponding to a CPRI interface (Split A.1).
  - Low bit rate flows, corresponding to a compressed CPRI interface or an intra-PHY layer functional split (Split A.2).

The characterization of these flows, especially the variable bit rate ones, is a complex issue, as the flows are the result of the aggregation of the traffic for or from different users, where this traffic may also have been originated by several sources (e.g., different applications running in the user device, some of them possibly in the background). The solution adopted is to define a set of basic traffic profiles, characterized by a packet generation rate that is consistent with the defined bit rate, where packet size is randomly selected from a set of predefined values and inter-packet arrival time is derived from the exponential distribution.

The links between the nodes may be configured with different interfaces (and associated capacities): 1 GbE, 10 GbE, 40 GbE and 100 GbE. Obviously, the inter-iTN link should have equal or larger capacity than the other links (iTN-iSC, iTN-RANaaS/EPC).

There are several parameters that are used to configure the virtual queue mechanism:

- Pacer exit rate: this parameter is used to evenly space the transmission of information blocks towards the egress port, so they are not sent in a burst. The pacer rate should be close to the link rate at the egress port.
- Virtual queue threshold: this parameter establishes when the traffic should be marked for congestion. Different thresholds can be configured for the activation of different congestion avoidance actions.
- Virtual queue link utilization factor ζ: this parameter should always be smaller than 1 and controls the rate of emptying of the virtual queue.
- Dropping packet probability.

5.1.5 CT4.5 - Network Wide Scheduling and Load Balancing

This CT is considered w.r.t. the Common Scenario 3 - wide area deployment. For the evaluation of the CT, the scenario that is being simulated is the one represented in the Figure 5-5,by considering 20 iSCs. As it will turn out, load balancing is usually coming at the price of user QoS, since these two objectives are often contrary: e.g., in a HetNet setup, associating to the cell with the highest SINR (often a macrocell) might be preferable for the user, to maximize its rate, but not to the operator who might prefer to offload traffic to available small cells (often transmitting at much lower power). There are several parameters that are used to configure our load-balancing mechanism:

- The degree of load balancing: the higher the degree the more emphasis on the load balancing, and the less in the user QoE.
- The traffic differentiation with some related "weighted" parameters, and their impact on the scheduling disciplines and iSC loads.

5.2 Evaluated metrics

5.2.1 CT4.1 - Distributed IP Anchoring and Mobility Management

As reported in D4.2 [4], Utilisation Efficiency is the metric being evaluated for CT4.1. D5.2 [5] describes several domains for Utilisation Efficiency, such as Computation Efficiency and Bandwidth/Capacity Efficiency. In CT4.1 evaluation we focus on the latter, which is defined as:

$$u_d^B(X) = \frac{B_{mean,d}(X)}{B_{cap,d}(X)},$$

where $B_{mean,d}(X)$ is the average measured data rate and $B_{cap,d}(X)$ is the corresponding outage or theoretical maximum capacity of the system. The parameter X is the capacity of the investigated link. Since CT4.1 works above the RANaaS, we consider the Utilisation Efficiency on all those link carrying the decoded signal. For instance, in case of functional split A, the links connecting the iSCs to the RANaaS are not considered, instead we consider only the link between the RANaaS and the gateways.

5.2.2 CT4.2 - Network Wide Energy Optimisation

As already discussed in D4.2 [4], we calculate the energy efficiency with respect to the three aforementioned performance metrics (failure probability, blocking probability, service delay), in order to give better qualitative and quantitative insights. To this end, we evaluate the performance of our system, with respect to *(i)* the coverage probability, *(ii)* the blocking probability, *(iii)* the service delay for the best-effort flows.

The operator should decide to fix the lower thresholds of the aforementioned dimensions, while optimizing the energy. These three fixed thresholds give different energy efficiency insights, and our aim is to understand these qualitative and quantitative insights, through extensive simulations.

5.2.3 CT4.3 - RANaaS location analysis

A RANaaS deployment algorithm based on a genetic algorithm was described in D4.2 [4]. In the design of the algorithm, both the cost of the RANaaS deployment and the network utilization are considered. For the network perspective, using RANaaS for realizing functional splitting in the RAN introduces an additional overhead in terms of volume of data to be transported. Accordingly, the shorter the distance between the iSCs and the RANaaS platform, the less the network resources are used for transmitting data between iSCs and RANaaS (i.e. there is more residual capacity in the network). However, such deployment (RANaaS platform close to iSCs) may be costly.

We aim to minimize the cost of deployment while maximize the residual capacity in the network. For this purpose, the optimization problem in the Genetic Algorithm form is formulated as follows:

$$\text{Fitness Value} = \frac{\alpha}{\text{Deployment Cost}} + \frac{\beta}{\text{Backhaul Bandwidth Usage}},$$

where α, β are the weight parameters for the deployment cost and backhaul bandwidth usage, respectively. The cost for RANaaS deployment includes the site cost and the cost of the required data processing resource (i.e. CPU cost):

$$\text{Deployment Cost} = \text{Site Cost} + \text{CPU Cost}.$$

The cost of the required data processing resource is defined as [14]:

$$\text{CPU Cost} = [(0.0834 \times N \times \gamma) + 0.007] \times 40000(\text{USD}),$$

where $\gamma = \frac{\text{Traffic Provision for } N \text{ cells}}{N \times peak}$ is the usage rate, N is the number of cells in the network. (γ =1 means BW provision for peak traffic) and the cost of each reference server unit is 40K USD. A simple estimate of the total aggregate traffic for N cells is based on the following algorithm used in the LTE guidelines [3]:

$$\text{Traffic Provision for } N \text{ cells} = max\{peak,\ N \times busy\ time\ mean\}$$

where $peak$ is the backhaul peak data rate from each iSC to the RANaaS platform. Accordingly, the usage rate is defined as

$$\gamma = \frac{\max\{peak, N \times busy\ time\ mean\}}{N \times peak}.$$

The backhaul bandwidth usage is defined as the bandwidth consumed by transmitting data between iSCs and RANaaS platform:

$$\text{Backhaul Bandwidth Usage} = \sum_{C_i} h_{c_i R_j} \times peak,$$

where $h_{c_i R_j}$ is the number of hops along the path between iSC c_i and its associated RANaaS platform R_j.

It is considered to be most efficient on network bandwidth utilization but least cost efficient to deploy the RANaaS platforms one hop away from iSCs. Therefore, this special case is used as a benchmark for measuring the cost efficiency. The cost saving rate is defined as our performance metric of cost efficiency:

$$\text{Cost saving rate} = \frac{C_i - C}{C_i},$$

where C_i is the cost of deploying RANaaS platforms one hop away from iSCs and C is the cost of the deployment to be evaluated.

For measuring the utilization efficiency, the residual capacity percentage is calculated:

$$\text{Residual capacity percentage} = \frac{B_a - B_{usage}}{B_a},$$

where B_a is the overall capacity in the backhaul network and B_{usage} is the backhaul bandwidth usage defined previously.

5.2.4 CT4.4 - Routing and Congestion Control Mechanisms

Congestion control effectiveness can be evaluated looking at a number of metrics that characterize the effectiveness of the proposed CT.

- Probability that the occupation of the egress port queue is higher than the maximum one compatible with the latency requirements of fixed bit rate flows. Values lower than 10^{-5} are expected to be achieved.

- Expected average one-way delay for end-to-end connection (iSC-RANaaS) for fix bit rate flows. A target value of 100 µs.
- Average one way delay of 1 ms to up 10 ms for variable bit rate flows (depending on the functional split is being implemented).
- Maximum jitter for fixed bit rate flows of 65 µs (this is the value usually associated with the CPRI interface used for A1 functional split).
- Packet loss probability for variable bit rate flows of 10^{-3}. It should be noticed that this packet loss is due to the dropping of packets by the dropper, not due to egress queue overflow.
- Link utilization. The achievement of the previous metrics results in an operation of the transport network that is far from full capacity utilization (depending on the mixture of flows and their characteristics). Obviously, it is one of the objectives of the CT to increase the link utilization while respecting the rest of the performance metrics.

5.2.5 CT4.5 - Network Wide Scheduling and Load Balancing

As already discussed in D4.2 [4], we have formulated a cost function that trades-off between the Load Balancing (i.e. equalization of the loads of the iSCs) and the user QoS (e.g. average system throughput) using a parameter α. These two goals are usually contradicted, that's why we chose to consider this trade-off. For example, standard SINR-based association might lead a UE to choose a nearby base station, to maximize its rate, but this iSC might already be congested on the radio channel or in the backhaul link.

To this end, we are going to show some snapshots with a sub-set of iSCs, and give some insights of the load balancing with respect to the average throughput. More precisely, we are going to show that by changing the value α, we can play around with the average user QoS and load balancing. As $\alpha \rightarrow 0$, the average throughput is maximized, whereas as $\alpha \rightarrow \infty$, load balancing is achieved. Intermediate values correspond to intermediate policies between these two goals.

6 Evaluation results

6.1 CT4.1 - Distributed IP Anchoring and Mobility Management

With the aim of evaluating CT 4.1 in the wide-area scenario, we define 3 possible configurations for different link capacity as depicted in Figure 6-1. Configuration A envisages a dark fibre link between each node in the network, including the iSC-iTN connection. Configuration B considers some link based on mmWave, while Configuration C contemplates a Sub-6GHz link for connecting the iSCs and some mmWave link between iTNs.

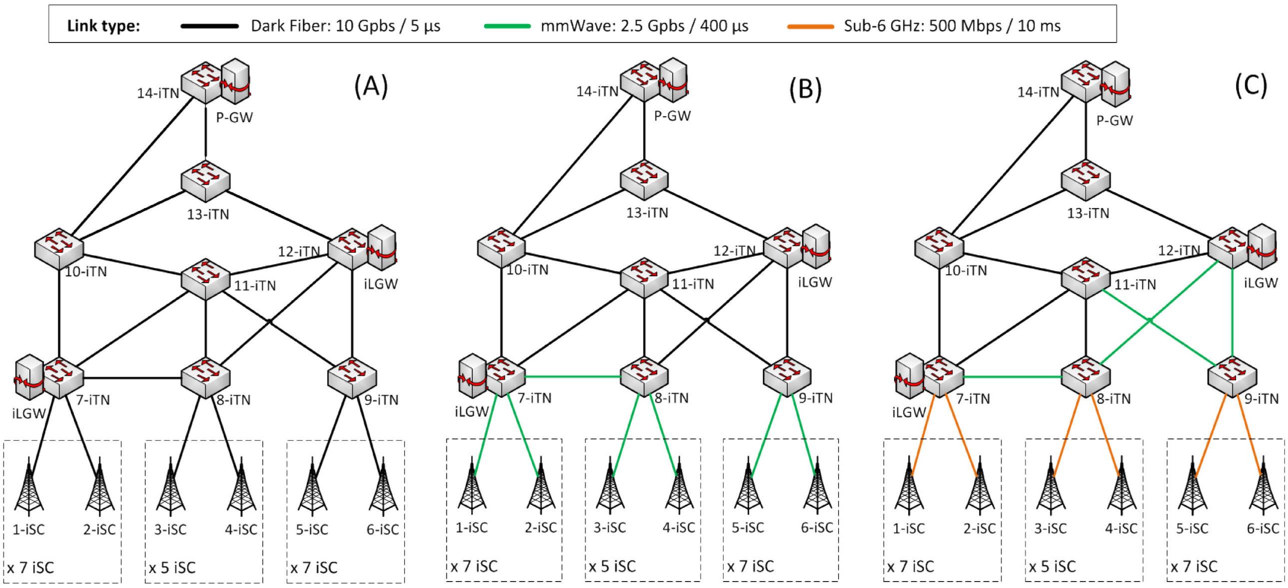


Figure 6-1: Considered scenarios for CT 4.1 evaluation

For each configuration, we consider 4 sub-scenarios where iLGWs are independently switched-on or off. That is, the P-GW is always available, while iLGW at 7-iTN and 12-iTN are enabled/disabled covering all the combinatory possibilities. For each sub-scenario and for each functional split reported in Table 6-1, we run CT4.3 RANaaS Placement algorithm in order to find RANaaS location. Table 6-2 reports the obtained RANaaS placement for each configuration, functional split and gateway combination. For instance, in case of Configuration B (Figure 6-1) with gateways available at *P-GW* and at *iLGW at 7-iTN*, the algorithm envisions two RANaaS located at nodes *7-iTN* and *11-iTN* in order to correctly implement the functional split A.2. Moreover, the algorithm shows that Configuration B does not support functional split A.1, thus is marked as *Not feasible* in Table 6-2.

Table 6-1: Considered functional split options

Split option	Lowest functionality centralized	BH Latency required	Required BH Bandwidth
A.1	FFT/IFFT	5 µs	1.8 Gbps
A.2	Subcarrier mapping	1 ms	470 Mbps
B.2	FEC	3 ms	100 Mbps
C.1	MAC	3 ms	70 Mbps
D.1	PDCP	~100 ms	70 Mbps

With the purpose of evaluating the CT4.1, we configured the SDN test-bed topology accordingly to Table 6-1. The functional splits are not directly implementable on the SDN test-bed, but this is not an inconvenient since we are only interested in RANaaS and gateways location. In order to assess the wide-area scenario, we employ one UE which connects to the test-bed and generates the amount of traffic reported in D5.2 (10-15 Mbps in peak hour). Since the test-bed does not envisage 19 iSCs but only 6, we proceed in the following way: firstly we attach the UE to an iSC and we generate the considered amount of traffic during 1 minute, and, secondly we collect the amount of traffic passed through of each link during this interval. Once we collected the results, we attach the UE to a different iSC and we repeat the measurement.

Table 6-2: RANaaS placement in backhaul topology for different configurations, gateway availability and functional splits

Configuration	Gateways	RANaaS location (refers to node name in Figure 6-1)				
		Split A.1	Split A.2	Split B.2	Split C.1	Split D.1
A	P-GW	7-iTN 8-iTN 9-iTN	10-iTN 12-iTN	14-iTN	14-iTN	14-iTN
	P-GW iLGW at 7-iTN	7-iTN 8-iTN 9-iTN	7-iTN	7-iTN	7-iTN	7-iTN
	P-GW iLGW at 12-iTN	7-iTN 8-iTN 9-iTN	12-iTN	12-iTN	12-iTN	12-iTN
	P-GW iLGW at 7-iTN iLGW at 12-iTN	7-iTN 8-iTN 9-iTN	7-iTN 12-iTN	7-iTN 12-iTN	7-iTN 12-iTN	7-iTN 12-iTN
B	P-GW	*Not feasible*	10-iTN 12-iTN	14-iTN	14-iTN	14-iTN
	P-GW iLGW at 7-iTN	*Not feasible*	7-iTN 11-iTN	7-iTN	7-iTN	7-iTN
	P-GW iLGW at 12-iTN	*Not feasible*	12-iTN	12-iTN	12-iTN	12-iTN
	P-GW iLGW at 7-iTN iLGW at 12-iTN	*Not feasible*	7-iTN 12-iTN	7-iTN 12-iTN	7-iTN 12-iTN	7-iTN 12-iTN
C	P-GW	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	14-iTN
	P-GW iLGW at 7-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	7-iTN
	P-GW iLGW at 12-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	12-iTN
	P-GW iLGW at 7-iTN iLGW at 12-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	7-iTN 12-iTN

We assume that the UEs' traffic is independent, that is, there is no correlation between the traffic generated by two different UEs. Doing so, we can find the total link occupation by summing the partial measurement. Figure 6-2, Figure 6-3, and Figure 6-4 show the cumulative link occupation for Configuration A, B, and C, respectively. The red line illustrates the link occupation when the P-GW is the only gateway available. The green line represents when the P-GW and the iLGW located at 7-iTN are present. Similarly, the blue line shows when the P-GW and the iLGW located at 12-iTN are available. Finally, the purple line shows the link occupation when all the gateways are available. The red line is the baseline considered for the evaluation. One of the aim of CT4.1 is to allow multiple gateways in the network with the purpose of offloading the network as soon as possible, hence, the lower the value in the figures, the better.

As it can be noticed by Figure 6-2, Figure 6-3, and Figure 6-4, any combination of functional split and iLGWs always outperforms the baseline, that is there is no iLGW available. Moreover, we observe how the results are not affected by the functional split itself, but are affected by the RANaaS location. For instance, if we analyse the results for functional split A.2, B.2, C.1 and B.1 in case of Configuration A, we observe that the results are the same for those functional splits. This is due to the fact that the RANaaS locations are the same for each case, hence, CT4.1 algorithm for anchors selection chooses always the same gateways in all the cases since they are the closest ones to the RANaaS. The results show that the highest network offloading happens when both iLGWs are available in the network. In addition, in case of having only one iLGW available, we observe that placing the iLGW closer to the UEs (iLGW at 7-iTN) allows a higher network offloading compared to the case when the iLGW is placed farther (iLGW at 12-iTN).

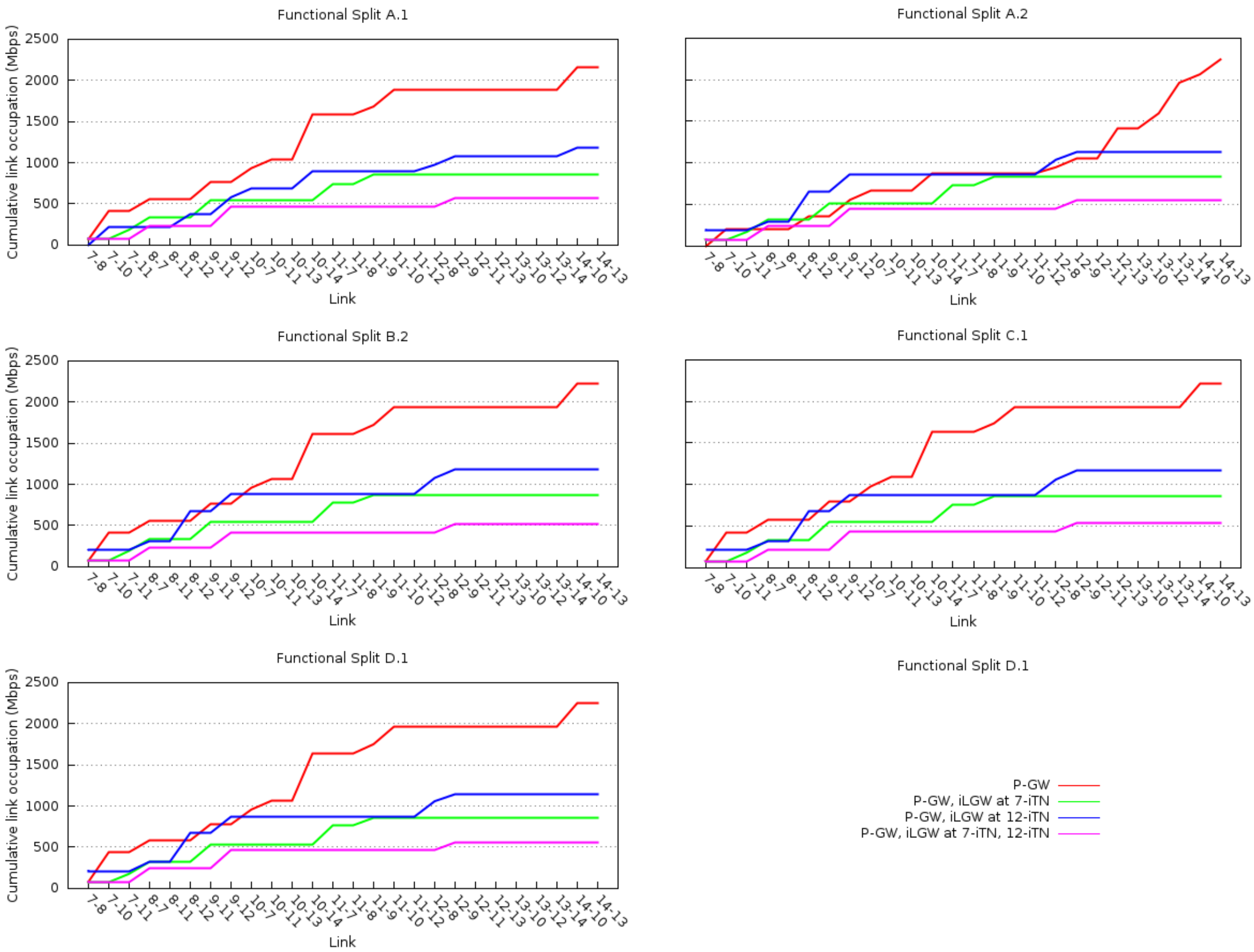


Figure 6-2: Configuration A, Cumulative link occupation

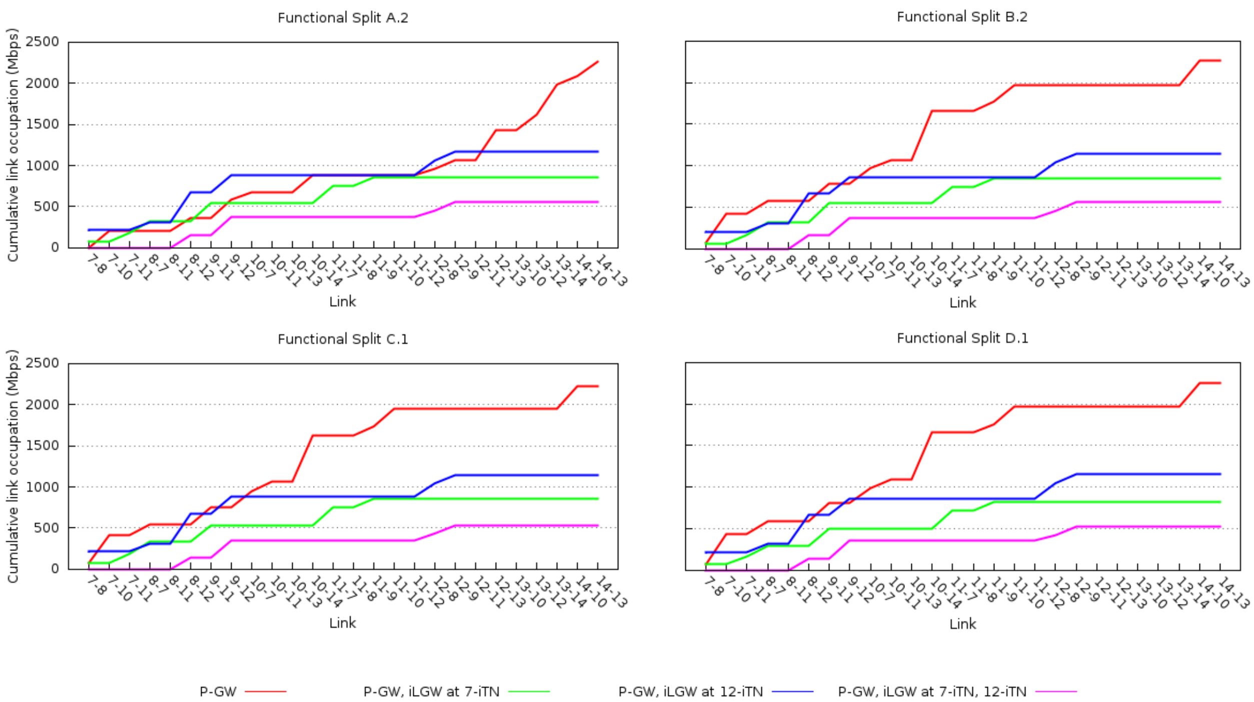


Figure 6-3: Configuration B, Cumulative link occupation

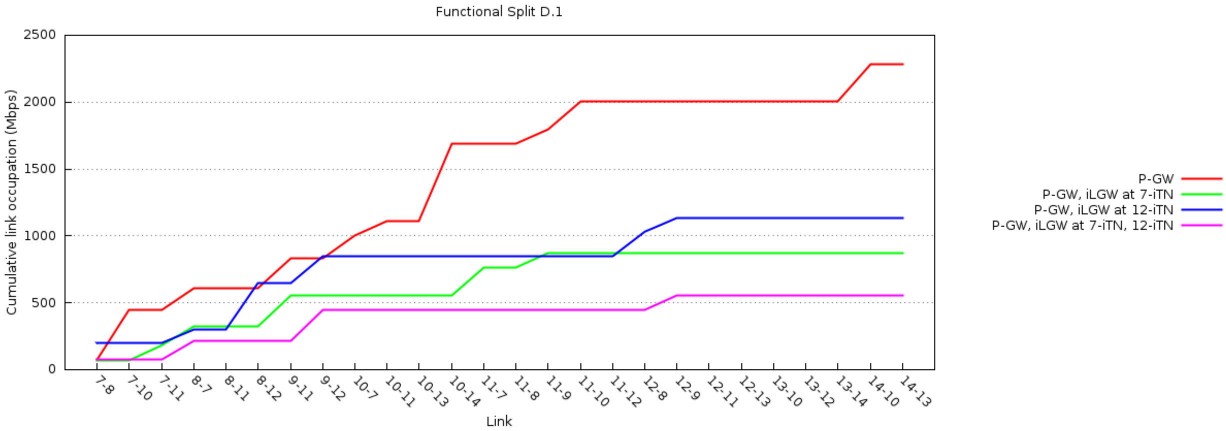


Figure 6-4: Configuration C, Cumulative link occupation

Table 6-3 reports the measured utilisation efficiency gain for each configuration. We consider as utilisation efficiency gain the percentage of the total traffic offloaded from the network against the baseline. As it can be noticed, in case of employing both iLGWs we offload 75% of traffic. This result is valid for each assessed configuration. In case of employing an iLGW at 12-iTN we achieve the minimum network offloading of 46%, while we achieve a higher offloading in case of iLGW at 7-iTN (60%).

Table 6-3: Measured Utilisation Efficiency gain for different configurations

Config.	Gateways	Split A.1 (Mbps/gain)	Split A.2 (Mbps/gain)	Split B.2 (Mbps/gain)	Split C.1 (Mbps/gain)	Split D.1 (Mbps/gain)
A	P-GW	2162 / -	2244 / -	2231 / -	2213 / -	2249 / -
	P-GW iLGW at 7-iTN	851 / 60%	832 / 63%	876 / 61%	864 / 61%	859 / 62%
	P-GW iLGW at 12-iTN	1184 / 46%	1137 / 49%	1187 / 46%	1166 / 47%	1150 / 48%
	P-GW iLGW at 7-iTN iLGW at 12-iTN	570 / 74%	552 / 75%	523 / 76%	540 / 75%	561 / 75%
B	P-GW	*Not feasible*	2263 / -	2262 / -	2230 / -	2257 / -
	P-GW iLGW at 7-iTN	*Not feasible*	858 / 62%	844 / 62%	853 / 61%	827 / 63%
	P-GW iLGW at 12-iTN	*Not feasible*	1171 / 48%	1145 / 49%	1146 / 49%	1159 / 48%
	P-GW iLGW at 7-iTN iLGW at 12-iTN	*Not feasible*	558 / 75%	569 / 74%	536 / 76%	525 / 76%
C	P-GW	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	2287 / -
	P-GW iLGW at 7-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	871 / 62%
	P-GW iLGW at 12-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	1136 / 50%
	P-GW iLGW at 7-iTN iLGW at 12-iTN	*Not feasible*	*Not feasible*	*Not feasible*	*Not feasible*	554 / 75%

Concluding, we showed how CT 4.1 can efficiently operate on any functional split, and how the functional split itself does not affect the CT's results. Moreover, the RANaaS placement is the only factor that affects the results. Therefore, we believe that whenever a RANaaS is deployed, the RANaaS should be provided with iLGWs functionalities in order to offload as much as possible the network.

6.2 CT4.2 - Network Wide Energy Optimisation

iJOIN provides flexible energy efficiency with respect to different scenarios and parameters. In WP4, we present the power/energy savings that can be achieved by switching-off iSCs and iTNs.

Thus, in the following we present the results obtained, regarding the energy efficiency, with respect to the Stadium scenario, in the three aforementioned dimensions: *(i)* the failure probability, *(ii)* the blocking probability, and *(iii)* the service delay for the best-effort flows. As already discussed previously, we are going to consider the energy efficiency along with:

- **Failure Probability;**
- **Admission control and "blocking" probabilities**;
- **Admission control and "service delay" for regular "best-effort" flows.**

To this end p_f, p_{block}, D_{max} are the upper thresholds specified from the operator, and influence the energy efficiency. More precisely, p_f is the maximum threshold as regards the failure probability a random user will experience, given that his origin iSC is switched-off and he is handed over in a neighbour iSC. p_{block} is the average blocking probability a user will experience for his dedicated flows, given that his iSC will be more "loaded" due to the increase of hand overs caused from switching-offs of neighbouring iSCs. Similarly D_{max} is the average service delay a user will experience for his best-effort flows, given that his iSC will be more "loaded" due to the increase of hand overs caused from switching-offs of neighbouring iSCs.

Indeed, the less strict the operator is (i.e. the higher the thresholds he defines) the more the iSCs we can safely switch-off. To put in another words, as these thresholds increase, the user QoS is decreased. It turns out that this has as an impact, that the more the iTNs we can also safely switch-off, as well. The following six figures show this behaviour that is consistent with our model and predictions.

In Figure 6-11 we show the functional split options between iSC and RANaaS: the lower the *functional split* the higher the overhead and the more stringent backhaul requirements are, so the less the energy savings. Hence, the different backhaul rate requirements in different functional splits, as presented in Figure 6-11, are affecting directly the energy efficiency. We are going to consider the functional B2 and C1.

In the following six figures we show the energy savings with respect to the different functional splits B2 and Functional Split C1. As expected, as the thresholds are increased, the energy savings are increased both in the access and backhaul network. Also, functional split C1 has always enhanced energy savings, due to the fact that are less bandwidth hungry.

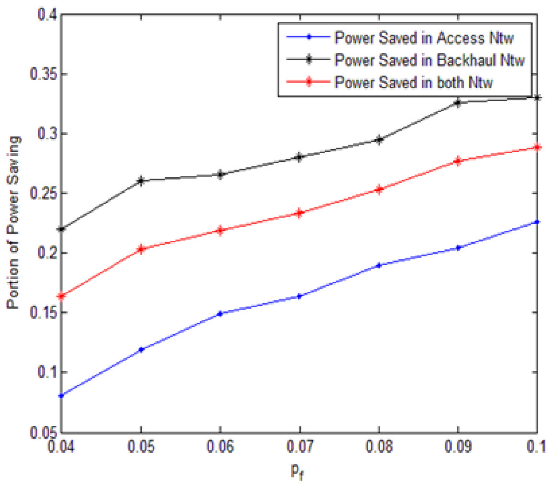


Figure 6-5: Energy Efficiency wrt to the failure probability (functional split B2)

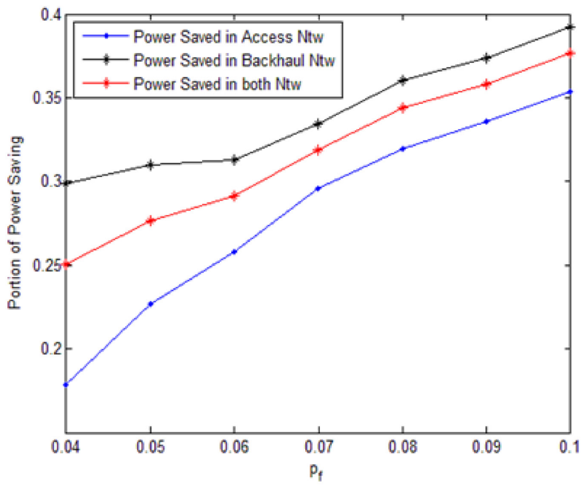


Figure 6-6: Energy Efficiency wrt to the failure probability (functional split C1)

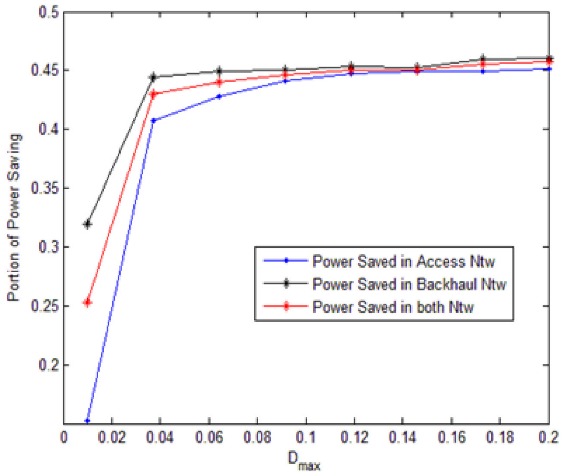


Figure 6-7: Energy-Efficiency wrt to the service delay (functional split B2)

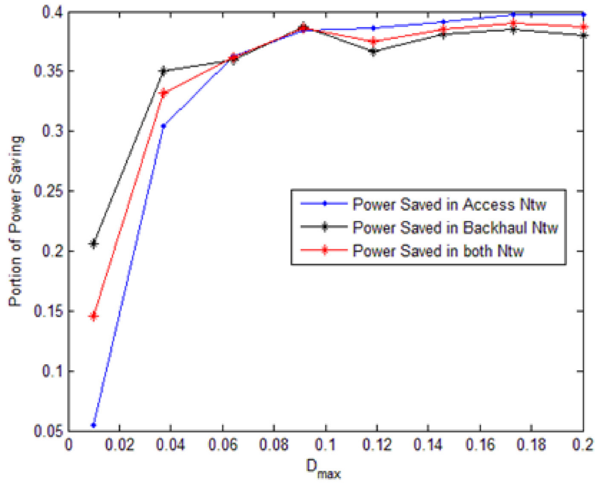


Figure 6-8: Energy Efficiency wrt to the service delay (functional split C1)

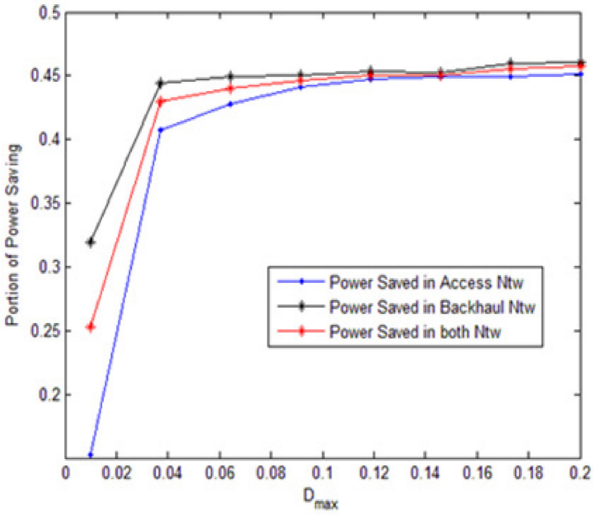


Figure 6-9: Energy Efficiency wrt to the blocking probability (functional split B2)

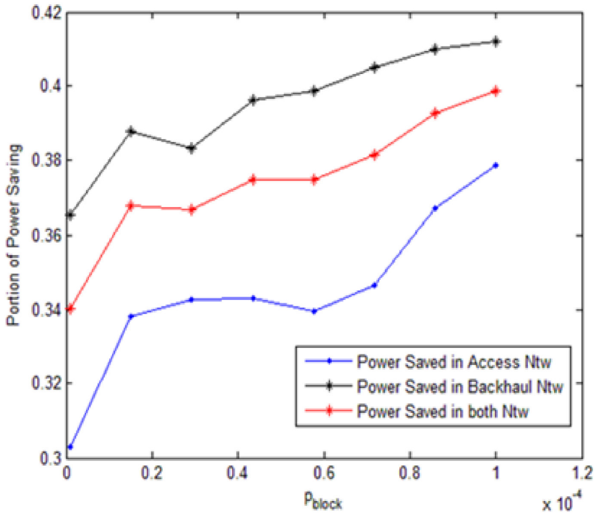


Figure 6-10: Energy Efficiency wrt to the blocking probability (functional split C1)

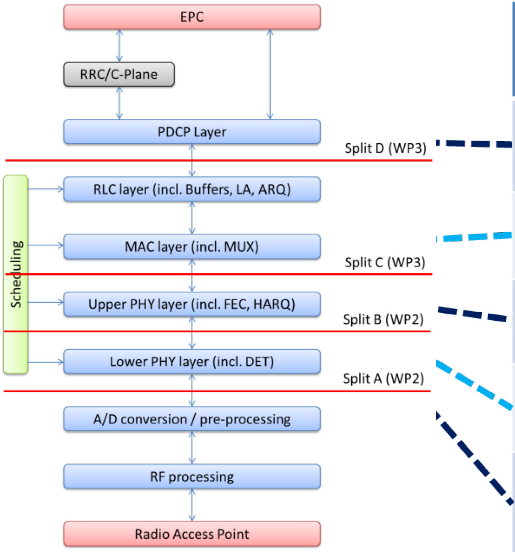
To better illustrate the energy efficiency, we introduce the Table 6-4. In this table, we assume that each metric can take values in a given interval. So, we assume that p_f=[0.04, 0.1], p_{block}=[10^{-5}, 10^{-3}], D_{max} =[0.01, 0.2] sec, so we depict the maximum energy savings we can achieve in our energy-efficient joint RAN\BH algorithm.

Table 6-4: Energy Efficiency gain

Functional Split	Energy Efficiency
B2	15-30%
C1	35-45%

6.3 CT4.3 - RANaaS location analysis

iJOIN provides the flexible centralisation of RAN functionalities. Figure 6-11 shows the functional split options between iSC and RANaaS. In general, the lower we place the *functional split* within the protocol stack, the higher the overhead and the more stringent backhaul requirements are. For example, in option C.1, MAC function is centralized in RANaaS platform and the required backhaul data rate is 70 Mbps. On the other hand, option A.2 is soft-bit forwarding and it requires 470 Mbps backhaul data rate. In our evaluation, the functional split options A.1, A.2, B.2, C.1 and D.1 are considered. Figure 6-11 depicts the required bandwidth and the latency under different options.



Split option	Lowest functionality centralized	BH Latency requirements	Required BH Bandwidth
D.1	PDCP	~100ms	70Mbps
C.1	MAC	3 ms (HARQ)	70 Mbps
B.2	FEC	3 ms (HARQ)	100 Mbps
A.2	Subcarrier mapping	1 ms	470 Mbps
A.1	FFT/IFFT	5 µs (CPRI)	1.8 Gbps

Figure 6-11: The required bandwidth and latency under different functional split options

As mentioned in Section 6.2, the fitness value in the GA algorithm is defined as:

$$\text{Fitness Value} = \frac{\alpha}{\text{Deployment Cost}} + \frac{\beta}{\text{Backhaul Bandwidth Usage}}$$

$$\text{Deployment Cost} = \text{Site Cost} + [(0.0834 \ \times N \times \gamma) + 0.007] \times 40000(\text{USD})$$

$$\text{Bankhaul Bandwidth Usage} = \textstyle\sum_{C_i} h_{c_i R_j} \times peak,$$

where α, β are the weight parameters for the deployment cost and backhaul bandwidth usage, respectively, $\gamma = \frac{\max\{peak, N \times busy\ time\ mean\}}{N \times peak}$ is the usage rate, N is the number of cells in the network, $h_{c_i R_j}$ is the number of hops along the path between iSC c_i and its associated RANaaS platform R_j. In our simulation, the site cost is 50000 USD. The peak traffic from the cell to the RANaaS is the required backhaul bandwidth in each functional split option. The busy time mean is half of the peak traffic, i.e.

$$Busy\ time\ mean = Peak * 0.5.$$

For the settings of the weight parameters α, β, three combinations are used in the simulations, namely cost-dominating setting, capacity-dominating setting and balanced setting:

Table 6-5: RANaaS location parameters

Combination	Cost Weight α	Capacity Weight β
Cost-Dominating Setting	10^6	1
Balanced Setting	160	1
Capacity-Dominating Setting	1	10^6

When $\alpha = 1$, $\beta = 10^6$ (capacity-dominating setting), the effect of backhaul bandwidth usage dominates. Therefore, the GA algorithm tends to find the solution with the maximum residual capacity in the network. On the other hand, when $\alpha = 10^6$, $\beta = 1$ (cost-dominating setting), the effect of cost dominates. Accordingly, the solution with the minimum cost is selected by the algorithm. In the balanced setting, both the cost and the capacity effects are considered so that the goal of minimizing the cost of deployment while maximizing the residual capacity in the network could be achieved.

Large Random Topology

In this scenario, there is no feasible solution for the functional split A.1 since the required backhaul bandwidth in option A.1 is 1.8 Gbps, which is more than the capacity of the last drop link (1 Gbps) in the scenario. As mentioned in Section 6.1.1, two different link settings are used in the aggregation networks:

- Case 1: Metro optical network with 1Gbps bandwidth and 250 µs latency
- Case 2: Dark fibre with 10 Gbps bandwidth and 5 µs/km latency

Table 6-6 shows the simulation results of the large random topologies with the setting of the case 1 for the aggregation network. As shown in Table 6-6, the options C.1 and D.1 require a lower number of RANaaS platforms deployed in the network. The locations of the RANaaS platforms are about 2.6 hops away from iSCs when the balanced setting is applied. On the other hand, the functional split option A.2 requires more RANaaS platforms deployed in the network and the locations of the RANaaS platforms are very close to the iSCs. This is because the option A.2 requires more bandwidth and shorter latency for transmitting the data from iSCs to the RANaaS. Consequently, the RANaaS platforms need to be deployed close to iSCs for shorter transmitting time and more residual capacity in the backhaul network. Furthermore, when more iSCs are aggregated in few RANaaS platforms, additional savings are achieved because of the statistical multiplexing gain of the aggregation traffic and the reduced site cost.

Table 6-7 shows the simulation results of the large random topologies with the setting of the case 2 for the aggregation network. As shown in Table 6-7, the options C.1 and D.1 in the case 2 require a lower number of RANaaS platforms when compared to that of the options C.1 and D.1 in the case 1 in the Table 6-7. This is because the capacity of the backhaul network increases in the case 2 so that the RANaaS platforms could be deployed more hops away from the iSCs with reasonable additional bandwidth consumed in the backhaul network. Therefore, savings increase. Similarly, the functional split options B.2 and A.2 in the case 2 required less number of RANaaS platforms, when compared to that of the options B.2 and A.2 in the case 1 in the Table 6-7.

Table 6-6: Simulation results of the RANaaS deployment in the large random topology with the setting of metro optical network for the aggregation network

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	50	94%	0%
	Balanced Setting	**2.6**	**8.6**	**84.5%**	**80.1%**
	Cost-Dominating setting	3.1	5	81.5%	87.5%
B.2	Capacity-Dominating setting	1	50	91.5%	0%
	Balanced Setting	**1.6**	**25.9**	**86.3%**	**43.4%**
	Cost-Dominating setting	2.1	14.9	82.1%	68.9%
A.2	Capacity-Dominating setting	1	50	60%	0%
	Balanced Setting	**1.2**	**40.1**	**52%**	**19.7%**
	Cost-Dominating setting	1.3	36	48%	28.1%

Table 6-7: Simulation results of the RANaaS deployment in the large random topology with the setting of dark fibre for the aggregation network

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	50	97.4%	0%
	Balanced Setting	**2.8**	**6.9**	**92.8%**	**83.9%**
	Cost-Dominating setting	3.7	2.9	90.5%	92%
B.2	Capacity-Dominating setting	1	50	96.4%	0%
	Balanced Setting	**2**	**16.7**	**92.7%**	**65.6%**
	Cost-Dominating setting	2.5	9.6	90.9%	78.8%
A.2	Capacity-Dominating setting	1	50	82.8%	0%
	Balanced Setting	**1.6**	**25.9**	**72.5%**	**48.4%**
	Cost-Dominating setting	1.8	20.8	69.1%	57.9%

Common Scenario 1 – Stadium

Table 6-8 shows the simulation results of the RANaaS deployment in the common scenario 1 - Stadium. As shown in Table 6-8, the number of the RANaaS platform is one in every functional split options and any weight parameter settings. The reason for that is the limitation of the topology in this scenario, where the iSCs are only one hop away from the IP edge router and there is only one IP edge router for the whole stadium backhaul network. Furthermore, as (by design) the local breakout gateway iLGW is collocated with the IP edge router, the algorithm always results in the allocation of the RANaaS platform also collocated with the IP edge router. Therefore, in this scenario, the RANaaS platform is deployed together with the IP edge router in all the functional split options.

Table 6-8: Simulation results of the RANaaS deployment in the common scenario 1 – Stadium

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	1	99%	0%
	Balanced Setting	**1**	**1**	**99%**	**0%**
	Cost-Dominating setting	1	1	99%	0%
B.2	Capacity-Dominating setting	1	1	99%	0%
	Balanced Setting	**1**	**1**	**99%**	**0%**
	Cost-Dominating setting	1	1	99%	0%
A.2	Capacity-Dominating setting	1	1	95.3%	0%
	Balanced Setting	**1**	**1**	**95.3%**	**0%**
	Cost-Dominating setting	1	1	95.3%	0%
A.1	Capacity-Dominating setting	1	1	82%	0%
	Balanced Setting	**1**	**1**	**82%**	**0%**
	Cost-Dominating setting	1	1	82%	0%

Common Scenario 2 – Square

Table 6-9 shows the simulation results of the RANaaS deployment in the common scenario 2 - square. In this scenario, there is no feasible solution for the functional split A.1 since the required backhaul bandwidth in option A.1 is 1.8 Gbps, which is more than the capacity of the last drop link (1 Gbps) in the scenario. As shown in Table 6-9, the functional split options B.2, C.1 and D.1 require one RANaaS platforms (collocated with the IP edge router). On the other hand, the option A.2 requires 10.1 RANaaS platforms because the option A.2 requires more bandwidth and shorter latency for transmitting the data from iSCs to the RANaaS. However, the capacity of the last drop links is only 1 Gbps. Accordingly, the RANaaS platforms need to be deployed close to iSCs for shorter transmitting time and more residual capacity in the backhaul network. In summary, a RANaaS platform is deployed in the IP edge router in the square scenario when the functional split options B.2, C.1 and D.1 are adopted, while a number of RANaaS platforms are deployed one hop away from iSCs when the functional split option A.2 is used.

Table 6-9: Simulation results of the RANaaS deployment in the common scenario 2 – Square

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	10.1	99%	0%
	Balanced Setting	**3.9**	**1**	**95%**	**44.2%**
	Cost-Dominating setting	3.9	1	95%	44.2%
B.2	Capacity-Dominating setting	1	10.1	99%	0
	Balanced Setting	**3.9**	**1**	**94%**	**32.6%**
	Cost-Dominating setting	3.9	1	94%	32.6%
A.2	Capacity-Dominating setting	1	10.1	99%	0%
	Balanced Setting	**1**	**10.1**	**99%**	**0%**
	Cost-Dominating setting	1	10.1	99%	0%

Common Scenario 3 – Wide Area Deployment

Table 6-10 shows the simulation results of the RANaaS deployment in the common scenario 3 – wide area deployment. In this scenario, there is no feasible solution for the functional split A.1 since the required latency in option A.1 is 5 μs and the latency of the last drop link is 5μs/km. As shown in Table 6-10, the functional split options A.2, B.2, C.1 and D.1 require one RANaaS platforms (collocated with the IP edge router). Because of the sufficient capacity in the backhaul network for the functional split options A.2, B.2, C.1 and D.1, the algorithm tends to deploy one RANaaS platform for saving the deployment cost. Therefore, one RANaaS platform is deployed in the wide area deployment scenario when the functional split options A.2, B.2, C.1 and D.1are used.

Table 6-10: Simulation results of the RANaaS deployment in the common scenario 3 – Wide Area deployment

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	26.5	99%	0%
	Balanced Setting	**2**	**1**	**98%**	**62.1%**
	Cost-Dominating setting	2	1	98%	62.1%
B.2	Capacity-Dominating setting	1	26.5	99%	0%
	Balanced Setting	**2**	**1**	**98%**	**48.4%**
	Cost-Dominating setting	2	1	98%	48.4%
A.2	Capacity-Dominating setting	1	26.5	99%	0%
	Balanced Setting	**2**	**1**	**96%**	**35.2%**
	Cost-Dominating setting	2	1	96%	35.2%

Common Scenario 4 – Shopping Mall/Airport

Table 6-11 shows the simulation results of the RANaaS deployment in the common scenario 4 – shopping mall / airport. In this scenario, there is no feasible solution for the functional split A.1 since the required backhaul bandwidth in option A.1 is 1.8 Gbps, which is more than the capacity of the last drop link (1 Gbps) in the scenario. As shown in Table 6-11, the functional split options B.2, C.1 and D.1 require one RANaaS platforms (collocated with the IP edge router). As shown in Table 6-11, the functional split options B.2, C.1 and D.1 require one RANaaS platforms (collocated with the IP edge router). On the other hand, the option A.2 requires 25.4 RANaaS platforms because the option A.2 requires more bandwidth and shorter latency for transmitting the data from iSCs to the RANaaS. However, the capacity of the last drop links is only 1 Gbps. Accordingly, the RANaaS platforms need to be deployed close to iSCs for shorter transmitting time and more residual capacity in the backhaul network. In summary, a RANaaS platform is deployed in the IP edge router in the shopping mall/airport scenario when the functional split options B.2, C.1 and D.1 are adopted, while a number of RANaaS platforms are deployed one hop away from iSCs when the functional split option A.2 is used.

Table 6-11: Simulation results of the RANaaS deployment in the common scenario 4 – Shopping Mall/Airport

Functional Split option		# of hops between iSCs and RANaaS	# of RANaaS	Residual capacity percentage	Cost saving rate
C.1, D.1	Capacity-Dominating setting	1	25.4	99%	0%
	Balanced Setting	**2**	**1**	**98%**	**70.3%**
	Cost-Dominating setting	2	1	98%	70.3%
B.2	Capacity-Dominating setting	1	25.4	99%	0%
	Balanced Setting	**2**	**1**	**97%**	**63.2%**
	Cost-Dominating setting	2	1	97%	63.2%
A.2	Capacity-Dominating setting	1	25.4	98%	0%
	Balanced Setting	**1**	**25.4**	**98%**	**0%**
	Cost-Dominating setting	1	25.4	98%	0%

6.4 CT4.4 - Routing and Congestion Control Mechanisms

The evaluation has been carried out for two sets of configurations. In each configuration the CT has been evaluated in the nodes where congestion occurs for the uplink and downlink traffic.

6.4.1 Configuration 1 – Uplink traffic

In this configuration, the effectiveness of the proposed CT is tested in the iTN that links with the iSCs through its ingress ports, with only a single egress port. An example of a potential implementation of this configuration is represented in the following figure:

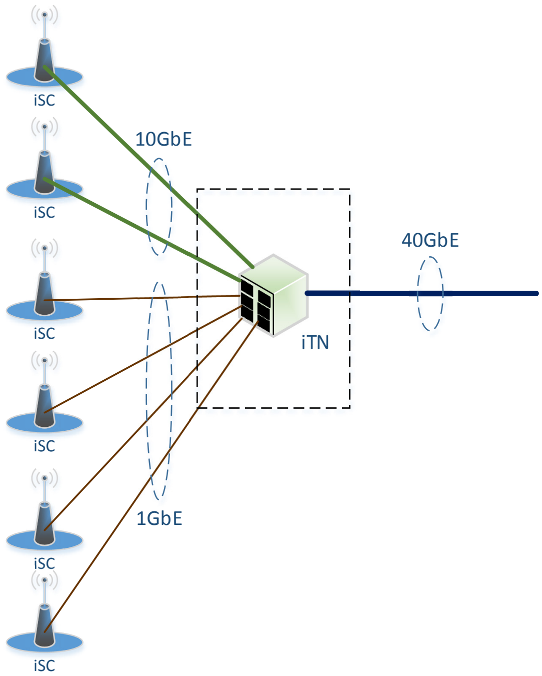


Figure 6-12: Uplink traffic evaluation scenario

In this configuration, it is a design requirement that the capacity of the link in the egress port must be higher than the sum of the maximum capacities of the links in the ingress ports.

For the evaluation of this configuration, the two fixed bit rate flows are configured as implementing A.1 functional split with a CPRI interface (which would correspond to the worst case scenario). Each iSC generates samples with a bit rate of 1.722 Gbit/s, which corresponds to a 20 MHz channel, 2 antennas, 14 bits per sample and oversampling of 1. As the CPRI frame size contains 480 bits of samples, frames are generated once every 278 ns (each frame contains 512 bits, as there is a frame header of 32 bits, so the effective bit rate is 1.837 Gbit/s).

The variable bit rate flows (up to 6) are configured so their sum provides an average specified value, which is denominated as traffic intensity. In the simulation framework the traffic intensity of the variable bit rates flows is kept constant for a certain period of time (till the average throughput converges to the intended value with a 5% tolerance), then it is increased to the next level. It should be noticed that the instantaneous bit rate should not be necessarily equal to the expected traffic intensity.

The following graph represents the probability that a CPRI frame finds the egress port queue filled up to a level which corresponds to the spare latency that can be consumed in order to meet the end-to-end latency (minus a programmable back-off), as a function of the variable rate traffic intensity.

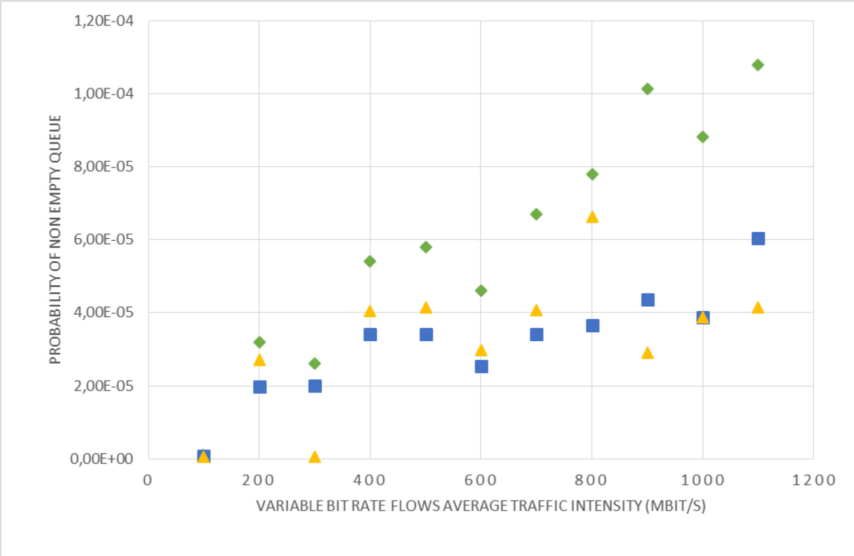


Figure 6-13: Uplink probability of non-empty egress port queue as a function of the variable bit rate traffic intensity

The results represented correspond to three different strategies implemented:

- For the green dots only the ECN (packet marking) is used to signal congestion. As a result of the notification, the bit rate is reduced after a delay associated with the round trip delay to the policy enforcement entity.
- For the yellow triangles the strategy is to drop packets. This originates a reduction of the corresponding variable bit rate flow due to conventional TCP based mechanisms that becomes apparent after a certain delay.
- Blue dots represent the case where dropping is combined with the deactivation of TCP congestion control mechanisms.

The CT is able to keep a very low probability value of finding an empty queue, as intended. Compared with the results obtained without application of the CT (only end-to-end congestion control applied), the values obtained in this case are always higher than 10^{-3}, even for low intensity traffic levels. It should be noticed that for the traffic intensities tested, the mechanism configuration parameters (α, packet drop probability) have not been changed. Also, it is feasible to have egress link's utilization levels of 25% with acceptable quality of service if a proper selection of the configuration parameters' values has been performed.

For a given traffic intensity, it is possible to change the behaviour of the mechanism using different values for its configuration parameters. In the following figure the results from different sets of values of the CT parameters fora variable bit rate flows traffic intensity of 1,5 Gbit/s

- Points marked as "+" correspond to a policy that drops packets but does not activate ECN.
- Points marked as "o" implement a policy with a double threshold in the virtual queue, a lower one that activates ECN and a second one that activates packet dropping.
- Points marked as "x" corresponds to a congestion control based only in ECN.

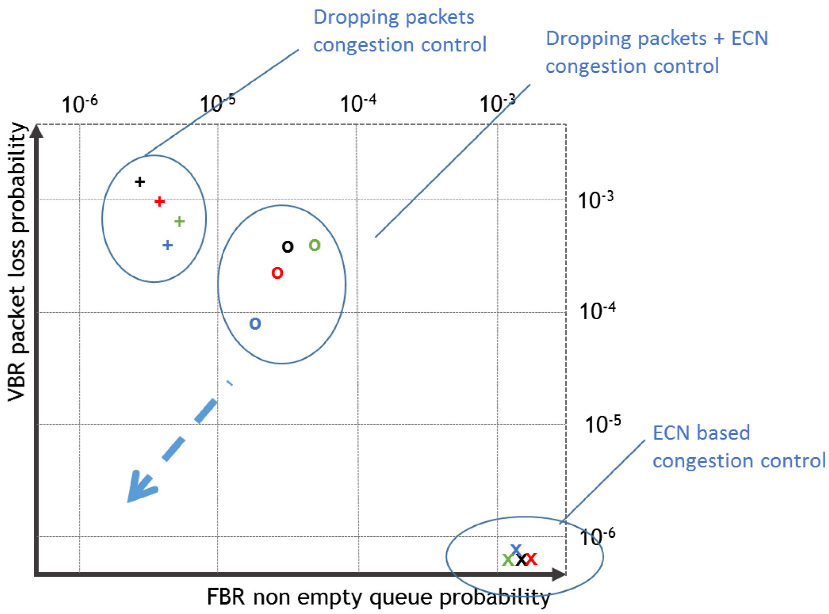


Figure 6-14: Uplink non empty egress port queue probability vs variable bit rate flows' packet loss probability

Several conclusions can be extracted from the results obtained:

- ECN based congestion control does not benefit from any change in the CT configuration parameters in order to reduce the probability of not empty queue for the fixed rate flows. This is due to the fact that the mechanism is not fast enough.
- It is possible to improve the variable bit rate dropping rate acting on the CT configuration parameters without sacrificing the probability of finding the queue empty by the fixed rate flows. Specifically, the use of additional information, like the occupation of other queues in the network (that would be available for a centralized implementation of the CT) may help to avoid dropping an excessive number of packets in successive iTNs. Also, it has been verified that a careful selection of the dropping probability may also be helpful[5].

6.4.2 Configuration 2 – Downlink traffic

In this configuration, traffic flows from RANaaS and EPC are switched to a single egress port, as represented in the following

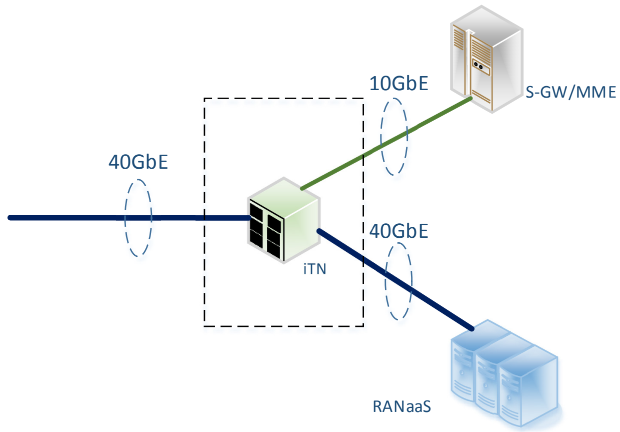


Figure 6-15: Downlink traffic evaluation scenario

[5] The drop probability, p, can be calculated based on the data packet size, which divided by inter data packet inter arrival time provides the actual bit rate v. This bit rate can be compared with the bit rate that would be required to keep the queue occupation below the established threshold, v'. In this way it is possible to calculate the reduction factor γ, such that $v' = \gamma \cdot v$. It is easy to calculate that for reducing the bit rate for v to v' it is necessary to drop 1 every n packets, with n being equal to $1/(1 - \gamma)$. Then the dropping probability is adjusted to $(1 - \gamma)$. Applying filtering to this dropping probability has shown to improve the performance of the CT.

The main differences with the previous scenarios are:

- The average bit rate of the variable bit rate flows is higher than that in the uplink.
- The constant and variable bit rate flows that are delivered from the RANaaS arrive to the iTN sequentially, so there is no need for flow pacer.

The following graph represents the results obtained in the simulation of this configuration when

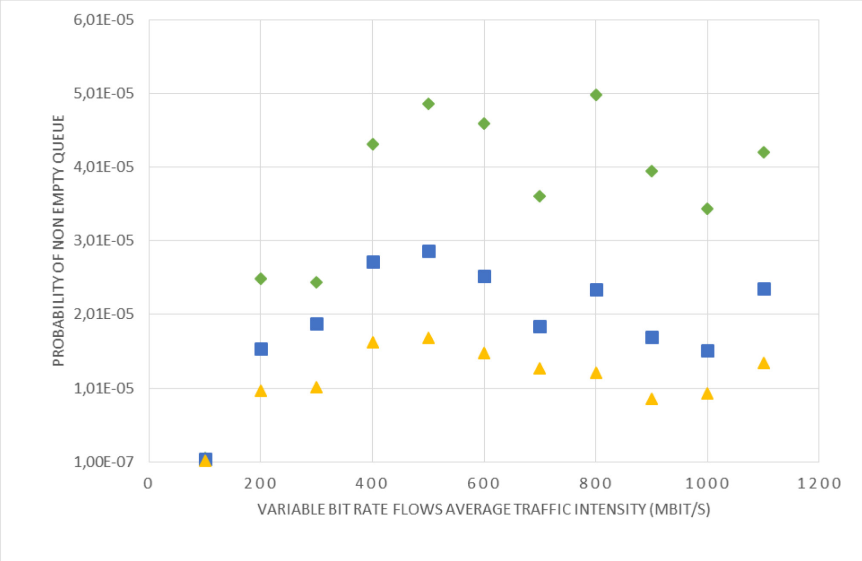


Figure 6-16: Downlink probability of non-empty egress port queue as a function of the variable bit rate traffic intensity

The results in this case show a better performance (a lower probability of finding the egress port queue non empty) than in the previous one for the same traffic intensity. This is due to the fact that the link between the EPC and the iTN acts like a pacer that precludes potential incast congestion situations (that happen when data is sent to a same receiver in parallel by multiple transmitter) between variable bit rate flows.

These results suggest that an adequate design of the flow pacer element of the CT can also help to improve results for configuration 1.

It should be noticed that the reliability of the results for this configuration is lower than for configuration 1, due to the lower probabilities, that require exceedingly large simulation times.

6.4.3 Conclusions from the evaluation

There are several conclusions obtained from the evaluation process:

- There is the need for a new procedure to avoid congestion issues when the same transport infrastructure is employed for supporting different functional splits. Performance clearly degrades when only window-based end-to-end flow control, like the one provided by TCP or its variants, is used for avoiding congestion. This is mainly because these mechanisms are not fast enough (nor strong enough) to overcome the congestion situations before significant degradation happens.
- The proposed CT improves the performance of some of the indicators, although there is always a trade-off: improving some KPI (e.g., reducing latency) may be realized at the expense of other (like link occupation).
- The use of ECN, that allows end-to-end notification of the network congestion without dropping packets, is not effective due to the latency of the mechanism. Random dropping of packets is much more efficient in preserving the required QoS for constant bit rate flows. However, dropping has a double effect: directly reducing egress port queue occupation by eliminating a packet, and reducing the intensity of the packet flow, due to the action of the end-to-end congestion control protocol like TCP. In some cases, this causes a loss of performance.
- The mechanism allows for the configuration of multiple parameters that can be used for further optimising the performance – based not only in local information, but also in global KPIs. In this

sense, it seems clear that a programmable implementation, based on the separate control plane integrated in the iNC, can achieve an improved performance.

- On top of those that have been used in the current evaluation, additional parameters can be defined in order to further optimise the performance. E.g., there may be thresholds associated with the activation of the different congestion correction mechanisms proposed (marking and dropping). In this sense, a lower threshold to activate congestion notification procedures and a larger one for activating packet dropping ones can be defined. There may be also different thresholds associated with flows with different priorities, i.e., lower threshold values for those flows that have lower priority, so they are more likely to be marked or discarded.
- Performance of other congestion control mechanisms designed for data centres (e.g., XCP) has not been simulated in detail, but there are indications that they are not effective for the traffic profile that characterise the iJOIN network. These mechanisms assume that flows with a high bit rate are delay tolerant, while short messages, associated with control functionalities, should be delivered without significant delay.

One important point in terms of feasibility of the solution proposed is that iTN that should be able to switch both CPRI flows and other flows is expected to be an Ethernet switch. However, as of today, Ethernet standard is not adequate for supporting CPRI. However, we consider that the new Ethernet capabilities that are being developed by IEEE 802.1 may allow it in the short term. The most significant ones are:

- 802.1ASbt Precise Timing Protocol Gen 2 (gPTP Gen 2): Improve performance, support redundancy and link aggregation and other media.
- 802.1Qbu Preemption (collaborating w/ 802.3br Interspersing Express Traffic): Allow time sensitive frames to preempt other frames.
- 802.1Qbv Time Aware Shaper (TAS) ─ Scheduled Traffic: Adds windows where non-scheduled traffic is blocked to ensure lowest latency.
- 802.1Qca Shortest Path Control & Reservations: Uses IS-IS to find all paths through a network – for redundancy.
- 802.1CB Frame Replication & Elimination: Bridges in a Ring automatically Replicate & Eliminate Duplicate frames.
- 802.1Qcc Stream Reservation Protocol Gen 1.1 (SRP Enhancements and Performance Improvements): Bandwidth and latency reservations to avoid time-sensitive queues to overflow and drop packets.

6.5 CT4.5 - Network Wide Scheduling and Load Balancing

In Figure 6-17, we depict the heterogeneous arrival rates of a considered area of 1000m X 1000m: with blue we depict low arrival rates, whereas with red high. Similarly, in the left picture of Figure 6-18, we depict the optimal user-associations, with respect to the maximization of the user QoS: each user is attached to the BS (BSs are illustrated with black dots) with higher SINR. However, in the right picture of Figure 6-18, we depict the optimal user-associations, with respect to load-balancing: the BSs that correspond to the high arrival rates area, are shrink since they will have to serve more “bandwidth-hungry” users.

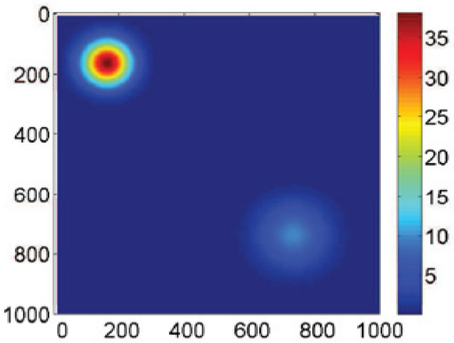


Figure 6-17: Snapshots of cell coverages

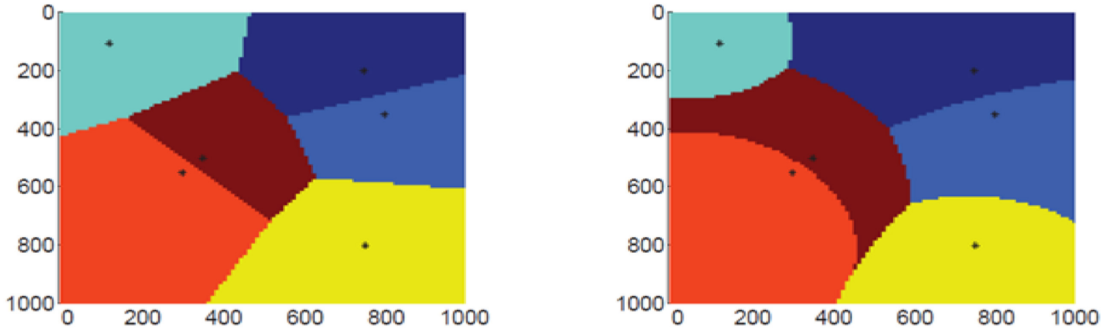


Figure 6-18: Heterogeneous Arrival Rates

Finally, in Figure 6-19 and Figure 6-20, we illustrate the quantitative results of the trade-off between user perspective (user QoS) and network perspective (Load Balancing): if α=0, the emphasis is on the user perspective (spectral efficiency), thus the average throughput, and average number of “dedicated” slots is maximized. However, as α is increased, the emphasis is moving on the network-perspective: the equalization of the loads is considered. Regarding the network-related performance goal, which is that of load balancing, we chose to use the Mean Squared Error (MSE) between the utilization of different BSs:

$$\mathrm{MSE}_1 = \frac{1}{2}\sum_i \sum_j \left(\rho_{i,1} - \rho_{j,1}\right)^2.$$

Similarly for MSE_2.Indeed, as $\alpha \rightarrow \infty$, we notice that the equalization (load balancing) of the nodes is achieved.

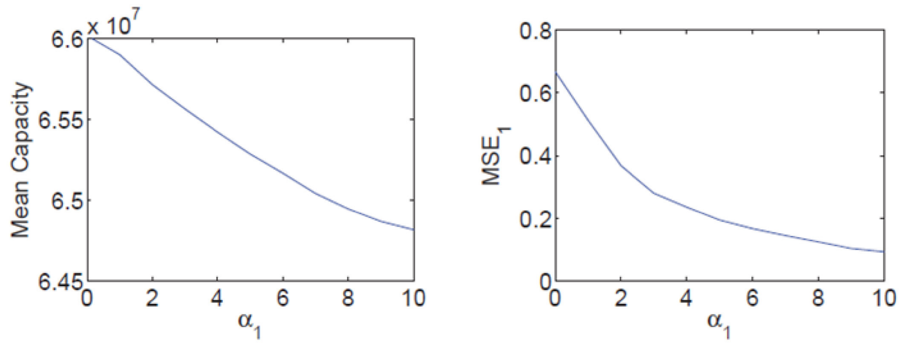


Figure 6-19: Trade-off: User QoS vs Load Balancing (best-effort flows)

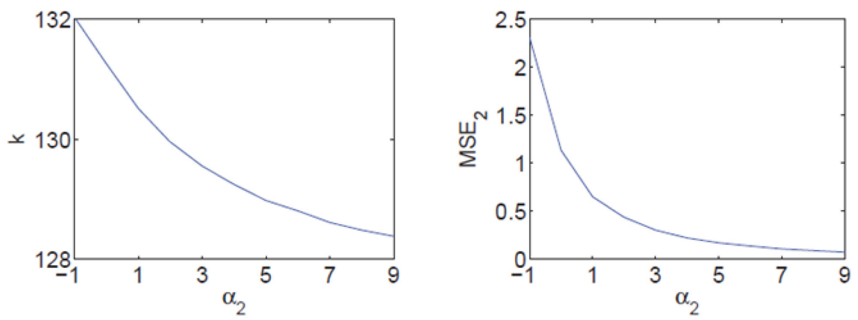


Figure 6-20: Trade-off: user QoS VS Load Balancing (dedicated flows)

To conclude, we notice that indeed α is the degree of load balancing: the higher the α the higher the load balancing, since the imbalanced load (reflected from MSE_1 and MSE_2) are alleviated. Also, this comes at the price of spectral efficiency, since the average throughout is decreased, as can be seen from the above figures.

6.6 Analysis of evaluation results

This subsection provides a global, but summarised, analysis of the obtained evaluation results, by relating them to the iJOIN KPIs.

WP4 mechanisms mainly contribute to the Energy efficiency KPI, by CT 4.2 selectively switching on/off iSCs and iTNs while keeping a given level of service. Obtained results show average savings of around 30-35%, which could even be improved when CT 4.2 operates in conjunction with the energy efficiency mechanisms defined by WP3, as reported in D5.3.

Another iJOIN KPI addressed by WP4 mechanisms is Utilisation Efficiency. In particular, CT 4.1, CT 4.4 and CT 4.5 are the main contributors to the KPI.

CT 4.1 improves the utilisation efficiency by offloading traffic from the backhaul and the core as compared to the baseline scenarios. The improvements achieved can be of around 75% depending on the scenario.

CT 4.4 is expected to contribute to the utilisation efficiency iJOIN KPI. Although one of the results of the CT is to ensure levels of utilization of the iTN switching capacity low enough to guarantee the fulfilment of other performance indicators when coexisting fixed and variable bit rate flows, it should be noticed that the alternative would be to have a duplicated transport infrastructure, with specific nodes and links for each kind of nodes. The support of the Ct also increases the flexibility of the operator when deploying new iSCs and adapting to changes in the traffic demand.

On the other hand, in case a single kind of flow (either fixed rate or variable rate) is employed the behaviour of the CT is expected not to have any negative impact – i.e., the behaviour would be very similar to that of a dedicated switching element. The proposed mechanism is expected to have a reduced impact in terms of cost for the iTNs, especially if the control is implemented in the iNC.

CT 4.2 improves the energy-efficiency in the timescale of the minute or hour, by considering jointly the access and backhaul network. We underlined the dependency between the energy efficiency we can achieve in the access network and in the backhaul, and we saw that the improvements achieved can be of around 40% depending on the scenario and how much "strict" the operator wants to be.

CT4.5 improves the load balancing, and we showed how this can be achieved at the price of average user throughput. Our model can be flexibly adapted according to a desired policy, since we can be more or less aggressive in the load balancing, by controlling the aforementioned parameter α.

In terms of cost efficiency, the main contributions of WP4 comes from the RANaaS location analysis, which allows to minimise the costs in terms of RANaaS required infrastructure, while ensuring a given set of constraints. The flexible functional split provided by the RANaaS concept ultimately enables the cost efficiency achieved by iJOIN. CT 4.1 also contributes, to a minor extent, to reducing costs, due to the traffic offload it enables. As commented above, CT4.4 can also improve this KPI, as it allows for avoiding the need of a duplicated transport infrastructure.

WP4 does not directly address the Area throughput KPI, as this one depends, fundamentally, on the spectral efficiency of the radio interface, which is beyond the scope of WP4.

7 Summary and Conclusions

This deliverable completes WP4 work, by providing a final update on the WP4 functional architecture (and its interaction with WP2 and WP3) and reporting on the evaluation results obtained at the final testing campaign.

In addition, we have characterised the iJOIN common scenarios taking also as input the analysis of the backhaul technologies that was reported in D4.2. This characterisation is intended to come up with practical realisation of the common scenarios defined in D5.1, with two specific purposes: *(i)* serve as basis for the WP4 evaluation, and *(ii)* provide a tangible set of scenarios where the feasibility of the different functional splits could be assessed. This characterisation has been performed considering current and planned deployment practices from operators.

The updated functional architecture of WP4 has been put into context within the overall iJOIN architecture, by showing the high level interactions with WP2 and WP3, and also analysing how WP4 contributes to and is affected by the functional split, as well as the joint RAN/BH design. These two are the main pillars of the iJOIN project, together with the SDN-based transport control, which ultimately is an enabler of both of them.

In this deliverable, a detailed evaluation of the WP4 mechanisms has been provided, building on top of the methodologies that were identified in D4.2, and providing additional information on both the methodology finally followed and the parameters evaluated.

From a dissemination point of view, several first-class publications including WP4 work have been published since D4.2, including a paper in IEEE Communications Magazine. Our active participation and contribution to standardisation activities have also continued, resulting in the final publication of a new RFC at the IETF, together with the acceptance of UC3M as Research Associate of ONF (only a few selected academic institutions are part of this group). Our Open Source contributions keep being maintained, which plans already scheduled for the release of SDN OpenFlow extensions developed with CT 4.1. Last, but not least, the SDN-based transport network control was shown in the iJOIN demo at the Mobile World Congress 2015.

Acknowledgements and Disclaimer

This work was partially funded by the European Commission within the 7th Framework Program in the context of the ICT project iJOIN (Grant Agreement No. 317941).

References

[1] S.E. Nai, T. Quek and M. Debbah, "shadowing time-scale admission and power control for small cell networks", in Wireless Personal Multimedia Communications, Sept 2012, pp.628-632.

[2] iJOIN INFSO-ICT-317941, "D3.3 - Final definition and evaluation of MAC and RRM approaches for RANaaS and a joint backhaul/access design," April 2015. [Online].

[3] iJOIN INFSO-ICT-317941, "D4.1 - Report on SotA and requirements for network-layer algorithms and network operation and management," November 2013. [Online]. Available: http://www.ict-ijoin.eu/wp-content/uploads/2014/01/D4.1.pdf. [Accessed April 2015].

[4] iJOIN INFSO-ICT-317941, "D4.2 - Network-layer algorithms and network operation and management: candidate technologies specification," November 2014. [Online]. Available: http://www.ict-ijoin.eu/wp-content/uploads/2012/10/D4.2.pdf. [Accessed April 2015].

[5] iJOIN INFSO-ICT-317941, "D5.2 - Final Definition of iJOIN Requirements and Scenarios," November 2014. [Online]. Available: http://www.ict-ijoin.eu/wp-content/uploads/2012/10/D5.2.pdf. [Accessed April 2015].

[6] iJOIN INFSO-ICT-317941, "D5.3 - Final definition of iJOIN architecture," April 2015. [Online].

[7] iJOIN INFSO-ICT-317941, "D6.2 - Final proof-of-concept results for selected candidate algorithms," April 2015. [Online].

[8] Carlos J. Bernardos, Antonio de la Oliva, Pablo Serrano, Albert Banchs, Luis M. Contreras, Hao Jin, Juan Carlos Zúñiga, "An Architecture for Software Defined Wireless Networking", IEEE Wireless Communications, Special Issue on Special Issue on "Research & Standards: Leading the Evolution of Telecom Network Architectures".

[9] Giust, F.; Bernardos, C.J.; de la Oliva, A., "Analytic Evaluation and Experimental Validation of a Network-based IPv6 Distributed Mobility Management Solution," Mobile Computing, IEEE Transactions on, vol.PP, no.99, pp.1,1, February 2014.

[10] Giust, F.; Cominardi, L.; Bernardos, C., "Distributed mobility management for future 5G networks: overview and analysis of existing approaches," Communications Magazine, IEEE , vol.53, no.1, pp.142,149, January 2015, doi: 10.1109/MCOM.2015.7010527.

[11] H. Chan et al., "Requirements for Distributed Mobility Management", RFC 7333, August 2014.

[12] D. Liu et al., "Distributed Mobility Management: Current practices and gap analysis", RFC 7429, January 2015.

[13] D. Kutscher, R. Winter, S. Krishnan, Y. Zhang, C. J. Bernardos, "Mobile Communication Congestion Exposure Scenario", IETF Internet Draft: draft-ietf-conex-mobile-04, (work-in-progress), September 2014.

[14] V. Suryaprakash, P. Rost, and G. Fettweis, "Are heterogeneous cloud-based radio access networks cost effective," accepted for publication on IEEE Journal on Selected Areas in Communications.

[15] N. Sapountzis, S. Sarantidis, T. Spyropoulos, N. Nikaein, U. Salim, "Reducing the energy consumption of small cell networks subject to user QoE constraints", Globecom 2014.

[16] Luis M. Contreras, Carlos J. Bernardos, Diego López, M. Boucadair, "Cooperating Layered Architecture for SDN", draft-contreras-sdnrg-layered-sdn-02, Work-in-progress, February 2015.

[17] SDN architecture, Open Networking Fundation Std. 1.0, June. 2014.

[18] P. Rost, C.J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, *Cloud Technologies for Flexible 5G Radio Access Networks,* IEEE Communications Magazine, Vol. 42, No. 5, pp. 68-76, May 2014.

[19] D. Wübben, P. Rost, J. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, G. Fettweis, *Benefits and Impact of Cloud Computing on 5G Signal Processing,* IEEE Signal Processing Magazine, Vol. 31, No. 6, pp. 35-44, Nov. 2014.

