



iJOIN

INFSO-ICT-317941



INFSO-ICT-317941 iJOIN

Deliverable D3.3

Final definition and evaluation of MAC and RRM approaches for RANaaS and a joint backhaul/access design

Editor:	Andreas Maeder (NEC)
Deliverable nature:	Deliverable
Suggested readers:	iJOIN GA
Due date:	April 30 th , 2015
Delivery date:	April 30 th , 2015
Version:	1.0
Total number of pages:	86
Reviewed by:	GA members
Keywords:	Cloud RAN, RAN MAC, Backhaul MAC, RANaaS, Functional Split, iJOIN
Resources consumed	26.12

Abstract

This deliverable provides the final definition and evaluation of novel MAC and RRM approaches for RANaaS and joint RAN/backhaul design. It shows how the WP3 approaches integrate into the iJOIN system design, and how they contribute to iJOIN's key objectives. All candidate technologies are evaluated in common scenarios which reflect different network and deployment assumptions. An overall evaluation compares the potential gains from a WP3 point of view.

List of authors

Company	Author
CEA	Antonio De Domenico (antonio.de-domenico@cea.fr)
HP	Marco Di Girolamo (marco.digirolamo@hp.com)
IMC	Umer Salim (umer.salim@intel.com)
IMDEA	Jorge Ortin (jortin@it.uc3m.es)
NEC	Andreas Maeder (andreas.maeder@neclab.eu) Peter Rost (peter.rost@neclab.eu)
SCBB	Massinissa Lalam (massinissa.lalam@sagemcom.com)
TUD	Richard Fritzsche (richard.fritzsche@tu-dresden.de)
UoB	Henning Paul (paul@ant.uni-bremen.de)
UNIS	Emmanouil Pateromichelakis (e.pateromichelakis@surrey.ac.uk) Efstathios Katranaras (efstathios.katranaras@surrey.ac.uk)

History

Modified by	Date	Version	Comments
Andreas Maeder	April 30 th , 2015	1.0	Final version of D3.3

Table of Contents

List of authors.....	2
History.....	3
Table of Contents.....	4
List of Figures.....	6
List of Tables.....	8
Abbreviations.....	9
1 Executive Summary.....	12
2 Introduction.....	13
2.1 Motivation and Background.....	13
2.2 Key Contributions.....	13
3 iJOIN Architecture and Functions to support MAC and RRM approaches for RANaaS and a joint backhaul/access design.....	15
3.1 iJOIN Architecture.....	15
3.1.1 Functional Architecture.....	15
3.1.2 Logical system architecture.....	16
3.1.3 RANaaS Cloud Architecture.....	17
3.2 RAN functional split options.....	18
3.3 Joint RAN/BH Optimization.....	20
3.3.1 Background.....	20
3.3.2 Technology Approaches.....	22
3.3.2.1 Enabling RAN/BH Awareness.....	22
3.3.2.2 Joint BH/RAN Aware Centralized Connection Control.....	23
3.3.2.3 Adapting to BH constraints.....	24
3.3.3 Interface Requirements.....	26
3.3.3.1 Traffic differentiation.....	26
3.3.3.2 Required information for joint RAN/BH optimization.....	27
4 Final Description and Evaluation of the iJOIN MAC/RRM Candidate Technologies.....	28
4.1 CT 3.1: Backhaul Link Scheduling and QoS-aware Flow Forwarding.....	28
4.1.1 Final Implementation of CT.....	29
4.1.2 Evaluation of the CT.....	29
4.2 CT 3.2: Partly decentralized Mechanisms for joint RAN and Backhaul Optimization in dense Small Cell Deployments.....	33
4.2.1 Final Implementation of CT.....	33
4.2.2 Evaluation of the CT.....	33
4.3 CT 3.3: Energy-Efficient RRM at Access and Backhaul.....	36
4.3.1 Final Implementation of CT.....	36
4.3.2 Evaluation of the CT.....	38
4.4 CT 3.4: Computational Complexity and Semi-Deterministic Scheduling.....	41
4.4.1 Final Implementation of CT.....	41
4.4.2 Evaluation of the CT.....	42
4.5 CT 3.5: Cooperative RRM for Inter-Cell Interference Coordination in RANaaS.....	44
4.5.1 Final Implementation of CT.....	45
4.5.2 Evaluation of the CT.....	45
4.6 CT 3.6: Utilization and Energy Efficiency.....	47
4.6.1 Final definition of utilization efficiency metrics.....	48
4.6.2 Evaluation of RANaaS Utilization Efficiency.....	48
4.6.2.1 Analytic evaluation of the computational utilization efficiency.....	48
4.6.2.2 Characteristics of computational complexity in large scale scenarios.....	50
4.6.2.3 Resource allocation with computational complexity constraints.....	52
4.6.2.4 Scheduling with water-filling (SWF).....	53
4.6.2.5 Scheduling with complexity cut-off (SCC).....	53
4.6.2.6 Numerical results.....	54
4.6.3 Evaluation of RANaaS Energy efficiency.....	57
4.6.3.1 Proposed Measurement Methodology.....	57
4.6.3.2 Theoretical Results.....	58

4.7	CT 3.7: Radio Resource Management for Scalable Multi-Point Turbo Detection	60
4.7.1	Final Implementation of CT	60
4.7.2	Evaluation of the CT	60
4.7.2.1	Scenario	60
4.7.2.2	System Model	61
4.7.2.3	Approach	63
4.7.2.4	Final Results	63
4.8	CT 3.8: Radio Resource Management for In-Network Processing	65
4.8.1	Final Implementation of CT	66
4.8.2	Evaluation of the CT	66
4.9	CT 3.9: Hybrid local-cloud-based user scheduling for interference control	70
4.9.1	Final Implementation of CT	71
4.9.2	Evaluation of the CT	72
5	iJOIN MAC/RRM Overall Evaluation	74
5.1	CT Interoperability	74
5.2	Evaluation methodology	74
5.2.1	Metrics	75
5.2.2	Evaluation parameters	75
5.3	Evaluation Results	76
5.3.1	Area Throughput	77
5.3.1.1	Wide Area Continuous Coverage Scenario	77
5.3.1.2	Stadium	78
5.3.1.3	Square	79
5.3.1.4	Indoor	79
5.3.2	Energy Efficiency	79
5.3.3	Utilization Efficiency	80
5.4	Summary	80
6	Summary and Conclusions	82
	Acknowledgements and Disclaimer	83
	Appendix I Backhaul classification reference	84
	References	85

List of Figures

Figure 3-1: WP3 functional architecture.	16
Figure 3-2: iJOIN logical system architecture.....	17
Figure 3-3: RANaaS Cloud Architecture.	18
Figure 3-4: Functional split options and trade-offs.	19
Figure 3-5: Mapping of functional splits to backhaul technologies in WP3.	20
Figure 3-6: Flexible implementation of cooperative scheduling according to the backhaul capacity.....	22
Figure 3-7: BH link Spectral Efficiency and Maximum Delay for different number of hops.....	23
Figure 3-8: Area throughput and RAN/BH Utilization efficiency for different numbers of active users.	24
Figure 3-9: Area throughput as a function of the backhaul delay.	25
Figure 3-10: Distribution of the rate-per user for the different scheduling approaches.	26
Figure 3-11: Information and control schemes for joint RAN/BH control.	27
Figure 4-1: Backhaul link scheduling and QoS-aware flow forwarding.	29
Figure 4-2: Illustration of BH topology for $k=9$ and $k=1$	30
Figure 4-3: Average BH link throughput for different delay bounds vs. number of paths.....	30
Figure 4-4: Average delay for different delay bounds vs. number of paths.	31
Figure 4-5: Satisfaction ratio vs. number of paths for different delay bounds.	32
Figure 4-6: Perceived cell throughput for different number of paths.	32
Figure 4-7: The proposed Load & BH Aware Cell Association scheme.....	33
Figure 4-8: Area Throughput and Utilization Efficiency versus the number of active users per macro cell sector ($C_{BH}=40\text{Mbps}$).	34
Figure 4-9: Area Throughput and Utilization Efficiency versus the number of active users per macro cell sector ($C_{BH}=120\text{Mbps}$).	35
Figure 4-10: Area Throughput for the distributed approach.	36
Figure 4-11: The proposed DTX control framework.	37
Figure 4-12: Exemplary membership function to match the queue latency with the associated linguistic terms.	38
Figure 4-13: Small Cell energy per bit vs. the number of active UEs.....	39
Figure 4-14: Small Cell user latency vs. the number of active UEs.....	40
Figure 4-15: Small Cell UE Packet Error Rate Vs. the Number of Active UEs.	41
Figure 4-16: Semi-deterministic, hierarchical multi-stage scheduling.	42
Figure 4-17: Channel uncertainty as a function of the backhaul delay.	43
Figure 4-18: Area throughput as a function of the backhaul delay for a user velocity of 3 km/h.	44
Figure 4-19: Area throughput as a function of the backhaul delay for a user velocity of 30 km/h.	44
Figure 4-20: Inter-cell interference coordination between iSC.	44
Figure 4-21 Topology – wide Area scenario.	45
Figure 4-22 Area Throughput comparison.	46
Figure 4-23 User Spectral Efficiency comparison.....	46
Figure 4-24: Utilization gains in different network domains.	47
Figure 4-25: Energy efficiency evolution in mobile networks toward a sustainable 5G.....	48
Figure 4-26: Computational effort as a function of the SINR.	49
Figure 4-27: Computational utilization efficiency.	50
Figure 4-28: Distribution of per-subframe normalized computational complexities.	51
Figure 4-29: Trace of total computational complexity and number of attached UEs per cell (macro cells)...	51
Figure 4-30: Per-cell computational complexity vs. number of attached UEs.....	52
Figure 4-31: CDF of computational complexity.	55
Figure 4-32: CDF of sum-rate.	55
Figure 4-33: Average sum-rate vs. number of centralized iSCs.....	56
Figure 4-34: Mean sum-rate vs. user density.....	56
Figure 4-35: Scalable multi-point turbo detection scenario (solid lines are minimal requirements).....	61
Figure 4-36: Comparison of the user uplink throughput CDF.	64
Figure 4-37: Normalised Area Throughput gain for different thresholds.	64
Figure 4-38: Average pairing ratios (%) for different thresholds.....	65
Figure 4-39: Deployment model for CT3.8 in the Wide Area Coverage scenario.	66
Figure 4-40: Area throughput for CT3.8 with whole bandwidth allocated.	67

Figure 4-41: Area throughput for CT3.8 with equal bandwidth split among UEs of one cell.	68
Figure 4-42: Area throughput gains of INP compared to baseline.	69
Figure 4-43: Deployment model for CT3.8 in the indoor scenario.	69
Figure 4-44: Area throughput gains of INP compared to baseline in the indoor case.....	70
Figure 4-45 Schematic presentation of the architecture as studied in CT3.9	71
Figure 4-46: Message sequence chart for CT3.9.....	71
Figure 4-47: Topology of the simulated network for one random dropping of the UEs.....	72
Figure 4-48: Distribution of the area throughput.....	73

List of Tables

Table 3-1: iJOIN RRM/MAC Candidate Technologies (CTs).....	15
Table 3-2: Comparison of split options on MAC layer.....	19
Table 4-1: iJOIN RRM/MAC Candidate Technologies (CTs).....	28
Table 4-2: CT 3.1 Specific Simulation Parameters.....	29
Table 4-3: CT 3.2 simulation parameters (centralized algorithm).....	34
Table 4-4: CT 3.2 simulation parameters (distributed algorithm).....	35
Table 4-5: CT 3.3 simulation parameters.....	38
Table 4-6: CT 3.6 Simulation parameters.....	54
Table 4-7: Parameters for the energy efficiency evaluation of the baseline system.....	58
Table 4-8: Parameters for exemplary energy efficiency evaluation of the iJOIN system.....	59
Table 4-9: Node parameters.....	61
Table 4-10: Indoor Hotpot system level simulation static parameters.....	62
Table 4-11: Indoor Hotpot system level simulation dynamic parameters.....	62
Table 4-12: Area Throughput Gain for MPTD/SPTD.....	65
Table 4-13: Simulations parameters for the performance evaluation of CT3.9.....	72
Table 5-1: Compatibility of WP3 CTs.....	74
Table 5-2: CT evaluation scenario assumptions.....	76
Table 5-3: Overview of KPIs and input evaluation parameters per CT.....	76
Table 5-4: Qualitative impacts of WP3 CTs.....	77
Table 5-5: Results for area throughput in the wide area coverage scenario.....	77
Table 5-6: AT results for the stadium scenario.....	78
Table 5-7: AT results for the square scenario.....	79
Table 5-8: AT results for the indoor scenario.....	79
Table 5-9: Energy efficiency results.....	79
Table 5-10: Utilization efficiency results.....	80
Table A-6-1: Backhaul classification reference.....	84

Abbreviations

3GPP	3 rd Generation Partnership Project
ARQ	Automatic Repeat Request
AT	Area Throughput
AWGN	Additive White Gaussian Noise
BH	Backhaul
CC	Computational Complexity
CDF	Cumulative Distributed Function
CEff	Cost Efficiency
CN	Core Network
CPU	Central Processing Unit
CQI	Channel Quality Indicator
CRE	Cell Range Expansion
CS	Coordinated Scheduling
CSI	Channel State Information
CT	Candidate Technology
CWDM	Coarse Wavelength Division Multiplexing
DiCE	Distributed Consensus-based Estimation
DL	Downlink
DMRS	Demodulation Reference Signal
DSL	Digital Subscriber Line
DTX	Discontinued Transmission
EEff	Energy Efficiency
eNB	evolved NodeB
FDD	Frequency Domain Duplexing
FEC	Forward Error Correction
FFT	Fast Fourier Transform
HARQ	Hybrid Automatic Repeat Request
HetNet	Heterogeneous Network
HII	High Interference Indicator
IaaS	Infrastructure as a Service
ICIC	Inter-cell Interference Coordination
iJOIN	Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks
iNC	iJOIN Network Controller
InH	Indoor/Hotspot
INP	In-Network-Processing
IOI	Interference Overload Indication
iRPU	iJOIN virtual RAN Processing Unit
iSC	iJOIN Small Cell
iTN	iJOIN Transport Node
iveC	iJOIN Virtual eNB Controller
ITU-R	International Telecommunication Union – Radio
KPI	Key Performance Indicator
LA	Link Adaptation
LoS	Line of Sight

LTE	Long Term Evolution
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MIESM	Mutual Information based exponential SNR Mapping
MMSE	Minimum Mean Square Error
mmW	millimetre wave
MPLS	Multi-Protocol Label Switching
MPTD	Multi-Point Turbo Detection
MRS	Max-Rate Scheduling
MSC	Message Sequence Chart
MUD	Multi User Detection
NE	Nash Equilibrium
NLoS	Non Line of Sight
NP	Non-Polynomial
NRTV	Non-Real-Time Video
OAM	Operation, Administration, and Management
OFDMA	Orthogonal Frequency-Division Multiple Access
OSS	Operations Support System
PDCCP	Packet Data Convergence Protocol
PER	Packet Error Rate
PF	Proportional Fair
PHY	Physical layer
PmP	Point-to-Multipoint
PON	Passive Optical Network
PoP	Point of Presence
PtP	Point-to-Point
PUSCH	Physical Uplink Shared Channel
QCI	QoS Class Indicator
QoS	Quality of Service
RAN	Radio Access Network
RANaaS	RAN as a Service
RB	Resource Block
RLC	Radio Link Control
RRC	Radio Resource Control
RRM	Radio Resource Management
RNTP	Relative Narrowband Transmit Power
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RTT	Round Trip Time
S1-U	S1 User plane
S1-C	S1 Control plane
SCC	Scheduling with Complexity Cut-off
SINR	Signal-to-Interference-and-Noise Ratio
SNR	Signal-to-Noise Ratio
SE	Spectral Efficiency
SON	Self Organizing Networks

SPTD	Single Point Turbo Detection
SRS	Sounding Reference Signal
SWF	Scheduling with Water-Filling
TDD	Time Domain Duplexing
UE	User Equipment
UEff	Utilization Efficiency
UL	Uplink
UMi	Urban Micro
veNB	virtual eNB
X2-U	X2 User plane
X2-C	X2 Control plane
xDSL	x-Digital Subscriber Liner
VPN	Virtual Private Network
WA	Wide Area
WSRM	Weighted Sum-Rate Maximization

1 Executive Summary

This report describes the main activities carried out by Work Package 3 (WP3) which focuses on novel Medium Access Control (MAC) and Radio Resource Management (RRM) schemes in the iJOIN project. The main goal of this deliverable is to describe the final results for joint access/backhaul radio RRM and a set of novel RRM algorithms based on flexible centralization of the radio access network (RAN) functions. The evaluation of the candidate technologies show significant performance gains on MAC/RRM layer for key objectives of iJOIN for different functional split and backhaul configurations. Together with the WP3-wide evaluation on MAC/RRM layer which includes a technology integration analysis, the results demonstrate the benefits of the two main concept of iJOIN: joint access/backhaul operations and flexible functional split.

This deliverable is based on the initial definition of candidate technologies in D3.1 [1], which was the foundation for further refinements and integration in D3.2 [2] into the overall iJOIN system architecture, which is described in D 5.3 [10]. The final definition, integration and evaluation has been conducted in tight collaboration with work packages 2, 4 and 5 in order to keep coherency across network layers and system domains. In D2.3 [5], complementary parts of some candidate technologies in this deliverable can be found. In D4.3 [4], the applicability of WP3 approaches in a network-wide context is investigated. An overall evaluation of the system-wide performance can be found in D5.3 [10].

The deliverable is organized as follows: Section 2 provides an introduction of the motivation and the objectives of this deliverable. Additionally, the main contributions as well as some of main results presented in past reports are highlighted in Section 2.2.

Section 3 is devoted to the iJOIN architecture and to the upper layer RAN functional split options. In particular, Section 3.1 elaborates functional and logical aspects of the overall system architecture with relation to the upper RAN protocols and related cloud functionalities. Moreover, relevant functional split options are thoroughly investigated in terms of requirements, constraints, and centralization gains.

Section 4 presents finalized results of the WP3 candidate technologies (CT) introduced in D3.1 [1] and further developed in D3.2 [2]. This section focuses on the presentation of CT evaluation results which are used for the purpose of intra-WP and system-wide evaluations, the latter in D5.3 [10].

Section 5 discusses the intra WP3 evaluation outcomes of the CTs based on the common evaluation framework described in D5.2 [9]. In particular, Section 5.1 describes the CT interoperability, Section 5.2 presents the evaluation methodology (i.e., relevant metrics and simulation parameters), and Section 5.3 shows numerical results in terms of the key objectives of iJOIN, including area throughput, energy efficiency, and utilization efficiency.

Finally, this report is summarized and concluded in Section 6.

This report also encloses the Appendix I the outcome of the discussion for the categorization of backhaul technologies.

2 Introduction

2.1 Motivation and Background

The iJOIN project investigates the flexible centralization of Radio Access Network (RAN) functionalities on general purpose IT platforms. To realize this paradigm, named as RANaaS (RAN as a Service), in future small cell networks characterized by heterogeneous backhaul, iJOIN explores joint operation of the access and backhaul networks and flexible reconfiguration of the transport network through the software defined networking principles. Within iJOIN, WP3 develops Medium Access Control (MAC) / Radio Resource Management (RRM) solutions for the backhaul and access networks.

The deliverable D3.1 [1] has presented a state of the art on relevant MAC/RRM techniques developed in the past, introduced the main challenges for the iJOIN technologies, and provided a preliminary description of the WP3 Candidate Technologies (CTs). D3.2 [2] has introduced the logical and functional architecture from a MAC perspective; it has described in details the requirements and constraints (bandwidth, latency, protocol) associated to the possible RAN functional splits as well as the characteristics of the common backhaul technologies. Moreover, the proposed CTs have been developed and their innovations underlined. Additionally, to provide a consistent evaluation of the WP3 CTs in the iJOIN common scenarios, i.e., Square, Wide Area Continuous Coverage, Shopping Mall, and Stadium [8], a common evaluation framework has been elaborated where evaluation parameters are provided and the compatibility of the proposed CTs is investigated.

In this report, the concepts introduced in previous deliverables are elaborated in more details and finalized. In particular, the flexible selection of RAN functional split and the joint access and backhaul operations are described in terms of requirements constraints, functionalities, and potential gains. Additionally, the RAN Cloud architecture is developed and its entities and functionalities presented. The complete evaluation of the WP3 CTs is provided according to the methodology introduced in D3.2 [2]. Finally, a holistic evaluation at WP level is presented, where the interoperability of CTs is discussed, the overall impact of WP3 on the iJOIN objectives, i.e., Area Throughput, Energy Efficiency, Utilization Efficiency, and Cost Efficiency, is presented in the iJOIN scenarios.

2.2 Key Contributions

The list below highlights the key contributions of this deliverable, as well as the main scientific advances of the technologies described in the deliverable, pointing out the main differences over the state of the art as well as the most significant results achieved.

Together with WP2, WP3 have finalised the quantitative analysis (in terms of throughput, latency, and gains) of the most promising functional splits; based on this analysis, WP4 has identified the possible functional splits in each of the iJOIN common scenario [7]. Moreover, part of this evaluation has been used to provide an input to Small Cell Forum publications.

- Together with WP4, WP3 finalised the design of joint RAN/BH operations; the overall output of this collaboration has been presented in D5.3 and it has led to a joint publication accepted to EuCNC 2015 [27].
- Together with WP2 and WP4, WP3 provided input to finalise the iJOIN logical/functional architectures for WP5 [10], [14]; moreover, WP3 contributed the definition of the RANaaS Cloud architecture for WP5.
- WP3 performed a detailed and harmonized evaluation of all CTs in relevant scenarios, which has been taken as input for the iJOIN system evaluation [10].
- WP3 designed and evaluated a Cloud-RAN scheduling algorithm in CT 3.6, which improves the Cloud utilization efficiency and which provided input for the implementation on the RANaaS testbed in WP6 [11]. This algorithm has been presented in a paper submitted to IEEE Globecom 2015 [23].
- CT 3.4 has developed and assessed a partially centralized scheduler, which takes into account the impact of the backhaul latency to the link adaptation process; this solution has been presented into a journal accepted for a publication in IEEE Transaction on Wireless Communication [26].

- CT 3.3 has introduced and evaluated a backhaul-aware cell DTX controller to enhance the system energy efficiency without impairing the user performance; this algorithm has been presented into a paper submitted to IEEE Globecom 2015 [29].
- In CT 3.2 we have investigated backhaul-aware cell selection algorithm to improve the overall network Area Throughput; this algorithm has been integrated in the Joint Network-Channel Coding framework developed in WP2 (CT 2.3), and the overall study has been accepted to IEEE SPAWC [32].
- CT 3.1 has investigated the problem of joint routing and backhaul link scheduling in a dense small cell network using 60GHz multi-hop backhaul; the proposed solution has been presented into a paper accepted for publication in IEEE ICC - Workshop on Cloud-Processing in Heterogeneous Mobile Communication Networks (IWCPM) [28].

These contributions build up on results reported in deliverables D3.1 [1] and D3.2 [2]. In addition to the above results that have been produced for this deliverable, it is also worth highlighting some previous outcomes from WP3 which were reported in deliverables D3.1 and D3.2:

- A backhaul-aware solution which steers traffic across small cells to improve the overall network throughput has been developed and its results presented in IEEE Communication Magazine [13].
- An inter-cell interference coordination mechanism that improves transmission robustness and maximize the network capacity by mitigating the high levels of co-channel interference has been presented in IEEE Access [31].
- An Opportunistic Hybrid ARQ approach for the LTE uplink system that enables centralized decoding while satisfying the LTE protocol timing constraints has been published in IEEE Wireless Communication Letter [30].

3 iJOIN Architecture and Functions to support MAC and RRM approaches for RANaaS and a joint backhaul/access design

3.1 iJOIN Architecture

The final iJOIN architecture is defined in iJOIN deliverable D5.3 [10]. In WP3, functional and logical aspects of the overall system architecture have been defined with relation to the upper RAN protocol stack and related cloud functionality in the RANaaS platform.

The functional architecture described briefly in Section 3.1.1 is based on the analysis of the input and output parameters and signalling requirements of CTs and related RAN functions first conducted in D3.1 [1], and further refined in subsequent deliverables.

The concept of a flexible functional split is a key requirement of the iJOIN architecture, which has been analysed in depth in deliverable D3.2 [2], and further refined with an additional quantitative analysis of achievable performance gains in Section 3.2.

Finally, the cloud aspects of the logical architecture related to RAN functions have been investigated in Section 3.1.3.

3.1.1 Functional Architecture

This section describes the interactions between CTs related to WP3 as well as the interaction of WP3 with WP4 and WP2. The WP3 CTs are listed in Table 3-1 and are classified according to their specific functionalities and centralization requirements. In particular, CTs 3.2 and 3.3 can be characterized as SON functionalities, which enable coordinated connection control, and adapt the system parameters to changes in the cellular network, due to e.g. the network load, energy constraints, and mobility.

The other CTs are used in the centralized resource allocation framework: in particular, CT 3.1 enables optimized BH resource allocation; CTs 3.4, 3.5, and 3.9 are devoted to enhance the performance of downlink transmissions by increasing spectral efficiency, mitigating inter-cell interference, and coordinated RRM. CTs 3.7 and 3.8 increase the robustness of uplink transmissions by using inter-cell cooperation and exploiting spatial diversity. CT 3.6 has a different nature, since it provides a general scope investigation on the iJOIN Utilization and Energy Efficiency metrics. Hence, this classification it is not applicable to CT 3.6.

Table 3-1: iJOIN RRM/MAC Candidate Technologies (CTs)

CT	Topic	Abbreviation	Scope	Centralized function
3.1	Backhaul Link Scheduling and QoS-aware Flow Forwarding	BH Manager	BH RRM	Centralized Resource Allocation
3.2	Partly decentralized mechanisms for joint RAN and backhaul optimization in dense small cell deployments	Coordinated Cell Selection	SON	Centralized Connection Control
3.3	Energy-Efficient MAC/RRM at Access and Backhaul	EE RRM	SON	Centralized Connection Control
3.4	Computational Complexity and Semi-Deterministic Scheduling	SD Scheduler	DOWNLINK RRM	Centralized Resource Allocation
3.5	Cooperative RRM for Inter-Cell Interference Coordination in RANaaS	Coop. RRM	DOWNLINK RRM	Centralized Resource Allocation
3.6	Assess and Increase Utilization and Energy Efficiency	n/a	n/a	Centralized Resource Allocation (for UEff)
3.7	Radio Resource Management for Scalable Multi-Point Turbo Detection	MPTD RRM	UPLINK RRM	Centralized Resource Allocation
3.8	Radio Resource Management for In-Network-Processing	INP RRM	UPLINK RRM	Centralized Resource Allocation
3.9	Hybrid local-cloud-based user scheduling for interference control	HL Scheduler	DOWNLINK RRM	Centralized Resource Allocation

Figure 3-1 illustrates the functional interactions of the WP3 CTs (the blue box) as well as the exchange of information required between WP4 (in red) and WP2 (in green). From WP2, WP3 takes into account input and output information from the two main blocks, namely RAN-PHY Functions and BH-PHY Functions. WP3 provides to WP2 RRM and MAC information concerning the radio access and the backhaul, like scheduling maps and link adaptation parameters; WP2 forwards to WP3 estimated radio and backhaul channel information such as SINR and user data after detection and decoding.

The exchange of information between WP3 and WP4 can be divided across two iJOIN logical entities: the iJOIN Network Controller (iNC) and the iJOIN Transport Node (iTN). WP4 provides to WP3 information about the backhaul configuration and measurements such as routing information and mobility information.

In addition to the two main WP3 blocks discussed above, we identified basic functions that include standard functionalities for the BH and RAN management, which support the iJOIN RRM/MAC enablers.

Finally, we can identify in Figure 3-1 also the interaction of WP3 CTs with the iJOIN Virtual eNodeB Controller (iveC), which, according to the iJOIN architecture [10], is the logical entity that adapts the functional split configuration according to system objectives and constraints.

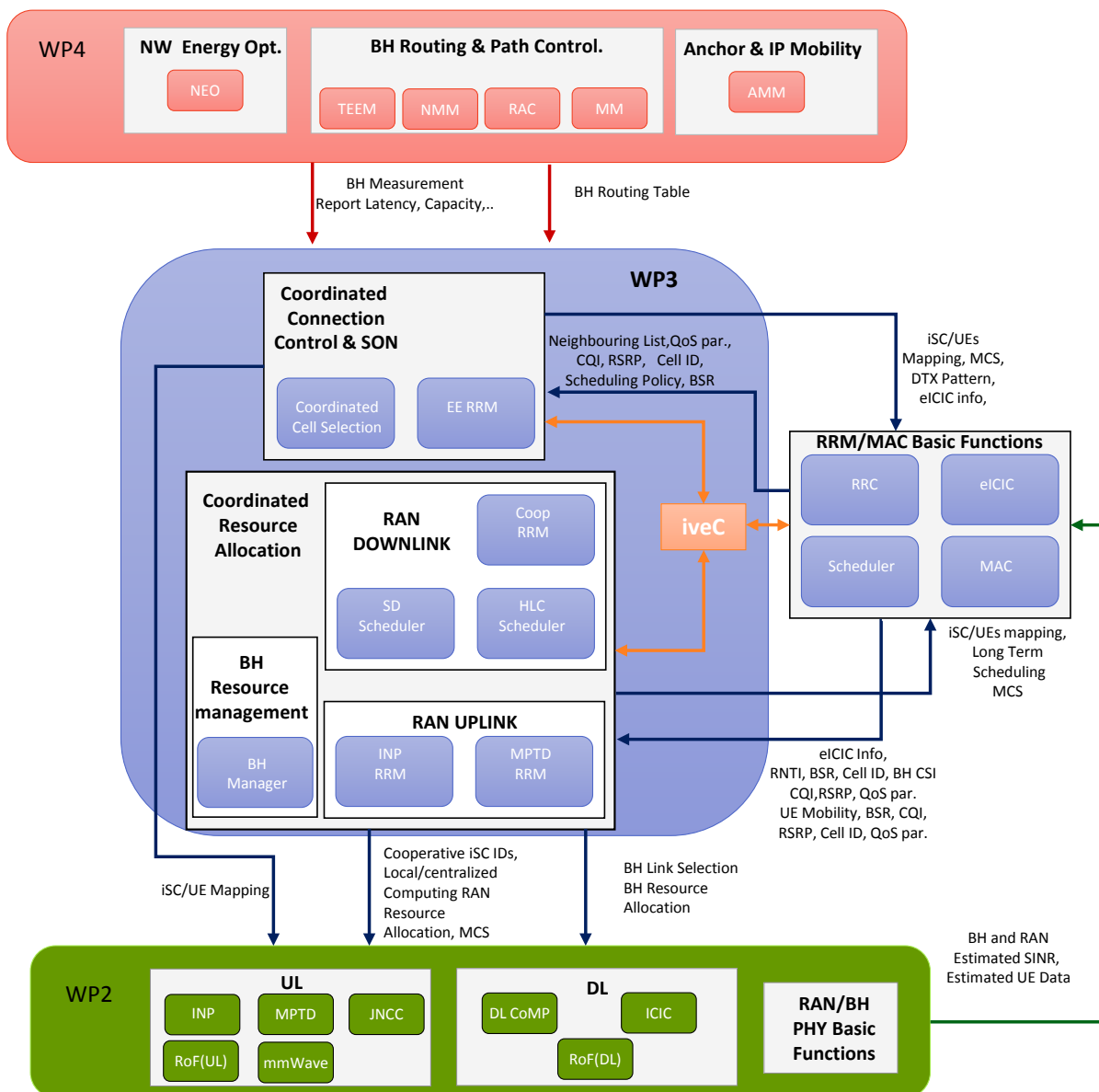


Figure 3-1: WP3 functional architecture.

3.1.2 Logical system architecture

Figure 3-2 shows the logical system architecture developed within the iJOIN project. The virtual eNB (veNB) is defined as the RANaaS instance running on a cloud platform (veNB upper domain) and one or

several iSCs (veNB lower domain). The veNB appears as classical eNB, such that it can be seamlessly integrated into the existing 3GPP LTE architecture. The core network does not need to know that the RAN functionalities are effectively split between iSCs and RANaaS. It only needs to know where to forward/get user and control planes, which by default will be the RANaaS entity. Within one veNB, the iSCs and the RANaaS platform are connected through the J1 interface, while the iSCs can exchange information directly with each other using the J2 interface. The actual split execution of the RAN functionalities between the two domains is managed by the iJOIN veNB controller. For more details on the logical system architecture, refer to [10].

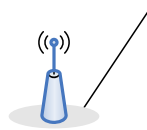


Figure 3-2: iJOIN logical system architecture.

3.1.3 RANaaS Cloud Architecture

The RANaaS module is the iJOIN architectural component where cloud computing functions of the whole platform reside. It is technologically built upon a general purpose cloud Infrastructure-as-a-Service (IaaS) baseline, whose physical resource layer has in turn been deployed on industry standard servers, considering specific extensions or enhancements where strictly needed. According with the iJOIN concept, both the cloud computing management layer and the underlying computational resources are implemented on general purpose hardware and software, to take most advantage of iJOIN's flexible functional split.

The RANaaS cloud architecture is shown in Figure 3-3. The block tagged "RANaaS" corresponds to a *RANaaS instance*, i.e. an actual implementation of a *RANaaS platform*, encompassing both the cloud management software (e.g. OpenStack modules, hypervisor, etc.) and the actual computational resources (e.g. servers, storage, physical and/or virtual network links). Inside the RANaaS framework there are the functional blocks defined by iJOIN, i.e. veNB, iJOIN veNB controller (iveC), iJOIN virtual RAN Processing Units (iRPU's), RANaaS manager.

For more detail about the RANaaS architecture, please refer to Section 5.2.1 in [10].

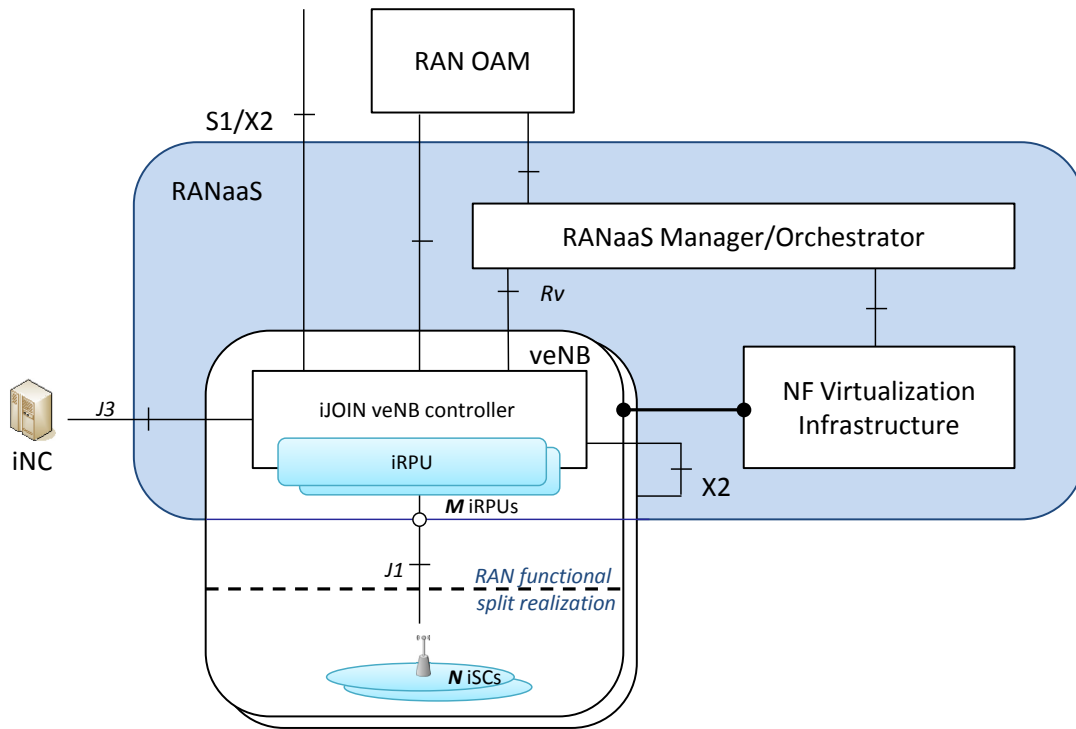


Figure 3-3: RANaaS Cloud Architecture.

3.2 RAN functional split options

A flexible functional split aims for exploiting centralization gains by adapting the assignment of RAN functions to a central entity to meet the operator requirements (e.g., in terms of cost and network performance) and to be able to adapt to the backhaul network capabilities. Figure 3-4 illustrates this concept: a higher RAN functional split in the LTE protocol stack (e.g., split options C or D) requires less backhaul capabilities in terms of latency and bandwidth than a lower one. However, the centralization gains (e.g., due to cooperative interference cancellation or joint detection schemes) are higher if lower layers are centralized as well.

Table 3-2 provides a comparison of the different split options in terms of backhaul requirements, functional impact, and expected centralization gains. The following observations can be made: with the functional split getting lower in the protocol stack:

- requirements on backhaul latency increase;
- requirements on bandwidth are unaffected (in WP3 scope);
- number of applicable CTs increase;
- centralization gains increase.

The impact on LTE depends on the individual split and cannot be directly correlated with protocol level.

By considering the compatibility analysis carried on in D3.2 [2] on WP3 CTs, we have performed a performance evaluation of the split option in terms of Energy Efficiency (EEff), Utilization Efficiency (UEff), and Area Throughput (AT): see Section 5 for more details on the definition of these metrics.

The results shown in Table 3-2 are related to the performance achieved in the wide area coverage scenario [8].

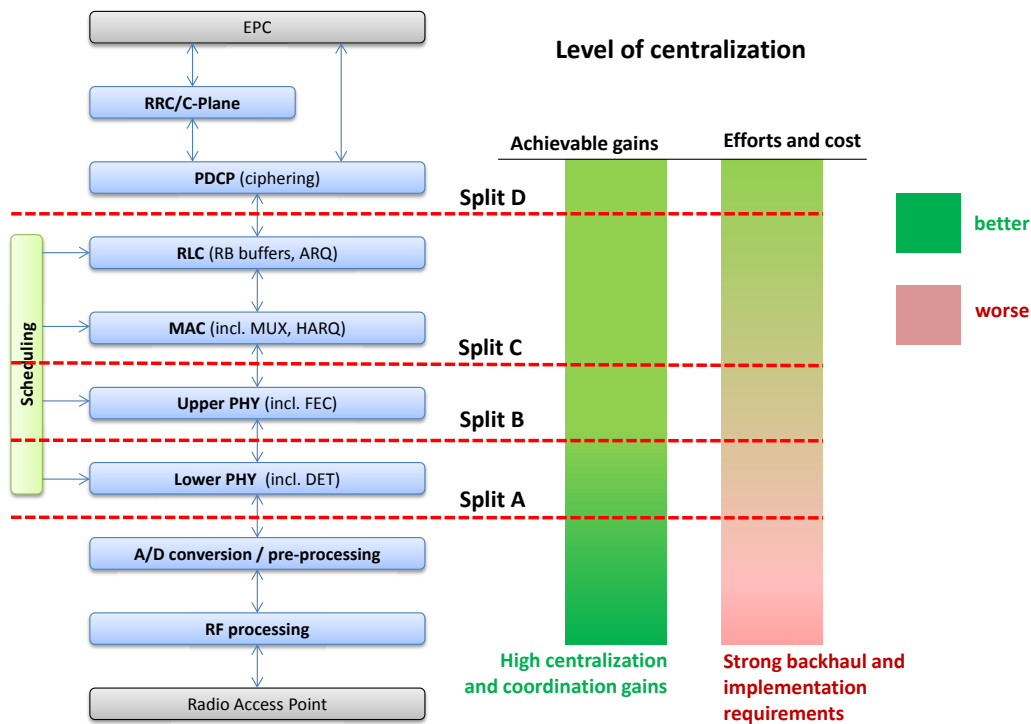


Figure 3-4: Functional split options and trade-offs.

When realizing splits D.2 or D.1, CT 3.2 and CT 3.3 can be implemented, which leads to 70% of gains in terms of EEff, due to CT 3.3 and 49% of gains in terms of UEff and AT, due to CT 3.2. These CTs are compatible with all the other CTs and obviously can also be implemented in lower layer splits. In particular, in both splits C.1 and C.2, other CTs can be realized together with CT 3.2 to further increase the AT gain. Specifically in the DL, the achieved gain is the range of [106; 249]%, depending on the selected RRM solution : CT 3.4 or CT 3.9 in split C.2 and CT 3.4, CT 3.5, or CT 3.9 in split C.1. In the same way, in the UL, according to the used RRM mechanism, the achieved gain is the range of [67; 99]%, which depends on the selected RRM solution: CT 3.7 or CT 3.8.

It can be concluded that with a low-latency backhaul, split option C.1 is preferable due to the low impact on the LTE protocol stack and the full enablement of centralization gains. If backhaul RTT values are above 3ms, split options C.2 and D.1 are preferable, depending on the optimization goals of the deployment scenario. Option D.2 is not preferable in the context of iJOIN RANaaS scenarios due to the additional burden of ciphering in the iSCs, and the potential visibility of inter-iSC handovers to the CN.

Table 3-2: Comparison of split options on MAC layer.

Split option	Lowest layer centralized	Impact on LTE	RTT requirements	Bandwidth requirements	Applicable CTs	Main centralization gains	Evaluation results (cf. Section 5.3)
D.2	RRC	Small, ciphering in iSC	Several hundred milliseconds to seconds	U-plane + C-plane overhead	CT 3.2, CT 3.3	Load balancing, energy efficiency	AT _{DL/UL} ≈+49% RAN/BH UEff≈+49% EEff≈+70%
D.1	PDCP	Small	> 50ms	U-Plane + C-Plane overhead	CT 3.2, CT 3.3	Load balancing, energy efficiency	
C.2	RLC + long-term scheduling	Split scheduling, dedicated signalling for resource allocation	Several frames (10ms each)	U-Plane + C-Plane overhead	CT 3.2, CT 3.3, CT 3.4, CT 3.7, CT 3.8, CT 3.9	Interference mitigation, cooperative schemes	EEff≈+70% RAN/BH UEff≈+49% AT _{DL} ≈+[79; 134]% AT _{UL} ≈+[79; 347]%
C.1	MAC	Small	<3 ms (HARQ)	U-Plane + C-Plane overhead	all	Interference mitigation, cooperative schemes	

As discussed, the selected functional split depends on the available transport network technologies, which determine directly the bandwidth and the latency capability to support function centralization. In previous deliverables D2.2 [4], D3.2 [2], and D4.2 [6], we have investigated the latency and data rate imposed by today backhaul technologies and evaluated the burden associated to each functional split. The result of this analysis is described in Figure 3-5, where the matching between different type of backhauls and (WP3-related) functional splits is identified. Our investigation shows how xDSL is characterized by latency constraints that can only enable split D options and at least sub 6 GHz backhaul is required to support the splits C.

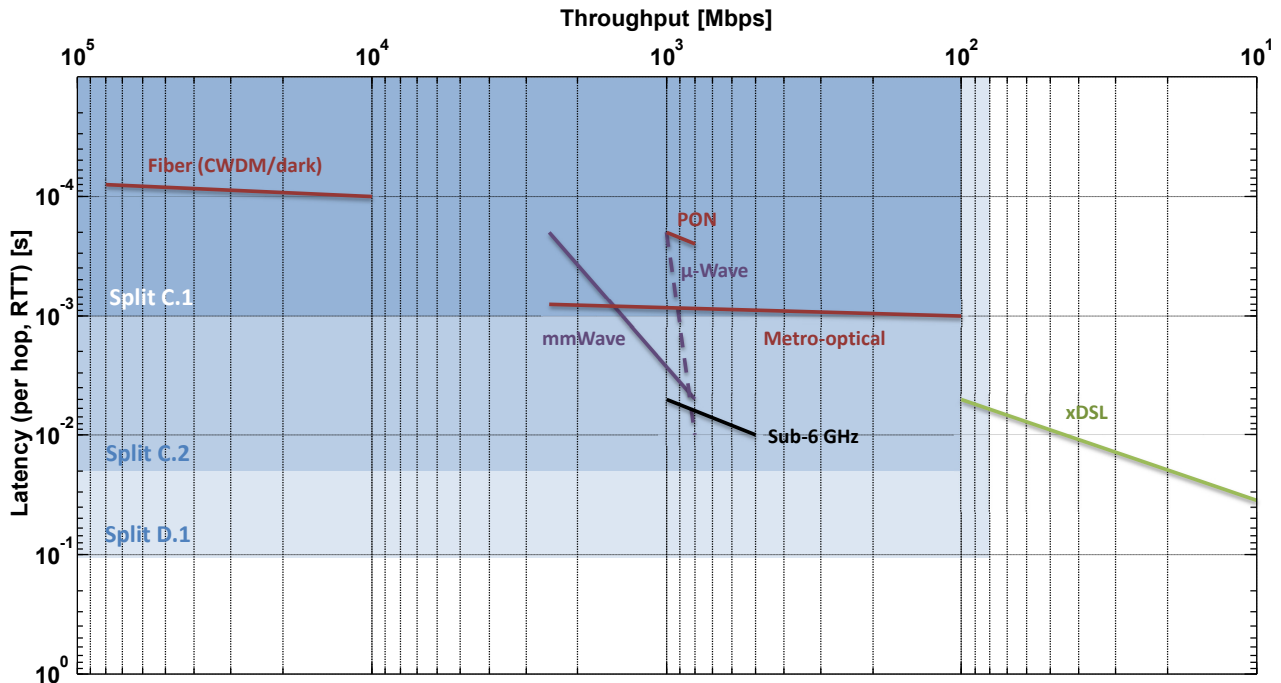


Figure 3-5: Mapping of functional splits to backhaul technologies in WP3.

3.3 Joint RAN/BH Optimization

This section provides an overview of BH/RAN optimization in dense small cell networks, assuming the iJOIN network architecture with a heterogeneous backhaul. We present some candidate MAC/RRM technologies which can be key enablers of the joint BH/RAN design. Moreover, we show the interface requirements for these technologies as a guideline for practical implementations.

3.3.1 Background

Deploying numerous small cells densely under the RANaaS framework can provide enhanced coverage and capacity in emerging wireless systems. However, in practice such deployments might require tight coordination between iSCs and RANaaS and global information from the core network, due to the profound necessity of addressing challenges like inter-cell-interference and mobility management. Thus, the efficiency of the RAN design relies heavily on the backhaul design so as to ensure that the underlying transport network can meet the requirements for fast and reliable transfer of the data to and from the end user. In this context, the sole optimization of RAN without taking into consideration the backhaul limitations might hinder the benefits of the flexible centralization at the cloud. In other words, the gains of the cloud centralization could be not so obvious in practical systems, due to the fact that the backhaul may pose some additional restrictions that should be taken into account while designing the RAN.

To this end, the joint BH / access optimization can be defined as the optimization of RAN functionalities while taking into account impact factors in RAN and BH which can influence each other. These factors can be the following:

BH capacity limitation: Backhauling a larger number of small cells will require a system with significantly greater overall backhaul capacity than the one needed for connecting few macro-cells. As discussed in D3.2 [2], the BH capacity depends on the BH technology and the topology used for the small cells connection.

- **BH topology:** The topology of the backhaul network affects the connectivity and link capacity among the PoPs and the small cells. Generally, there are two main topology types: 1) Point-to-Point and; 2) Point-to-Multipoint. Point-to-Point type links require an antenna and radio at each end. The network of small-cells and PoPs can then be backhaul connected in a star-, ring- or mesh- like manner which should be decided based on the expected capacity and latency requirements posed by the users' traffic and the centralization level. The main challenges of this topology include: a) the large number of antennas that may be required at the PoPs; b) the need of frequent re-planning or high self-organizing capabilities for the network in case new small-cells are added; c) inclusion of redundant links offering resiliency to link outages and; d) multi-hop links are implicit in such topologies leading to latency restricted performance. On the other hand, Point-to-Multipoint type links may be more efficient to pool resources across a larger, changing number of small-cells and average out any difference in traffic demand at different times of day. However, this backhaul topology, requiring high directivity antennas at PoPs, favours only a centralized type of coordination.
- **BH technology:** In the classification provided in D3.2 [2] the fiber backhaul technology can be considered either as ideal backhaul (up to 10Gbps capacity), or as non-ideal (up to 1 Gbps). Other wired technologies which can be candidate for small cell networks are DSL and cable (~100Mbps). Regarding the wireless technologies, we have three key candidates (sub-6GHz, microwave and mmW) as discussed above. Sub-6GHz can provide a good solution for urban outdoor unplanned deployments were LoS is not an option. On the other hand microwave and mmW solutions can offer high capacity if proper planning and LoS exist (up to 2 Gbps).

BH delay constraints: The backhaul latency is a key factor that also depends heavily on the BH technology to be used. The high requirement for fast transportation of data/control information to the end user can be vital for the applicability of flexible centralization, especially for real-time traffic. As discussed in D3.2 [2], the typical values for the latency can be lower than 1ms for ideal BH and in the range of 5ms-35ms for the other candidate BH technologies. Hence, designing the RAN and the functionalities to be centralized at RANaaS has to take into consideration the selection of BH technology for different scenarios.

BH availability: Backhaul solution must be available all the time and able to reach the small-cells in difficult locations. The key challenge in BH design is whether it can be available all the time to satisfy the dynamically changing small cells' requirements. This factor is more relevant to wireless BH where the capacity and delay can be dynamically affected by changes at the BH channel conditions and the potential interference.

In particular, for BH technologies that operate in lower frequencies, interference from other small cells may be an issue that can affect performance. In that case, the joint BH/Access design implies the joint scheduling of BH and access to efficiently assign resources to BH links and users to mitigate interference.

On the other hand, microwave / mmW radio can offer excellent immunity to interference, high security, and the reuse of frequency. However, clear LoS propagation is required and its range is restricted by the oxygen absorption which strongly attenuates the signals over distances. In this direction, data can be transferred via multiple low-distanced hops to ensure good backhaul link channel qualities. Hence the scheduling of BH links in a way that LoS is maintained is vital to meet the QoS requirements at all times; and these decisions shall be taken into account when selecting the optimal access configurations.

Targeting the MAC/RRM design, joint management of resources at the radio access and the heterogeneous backhaul is required to avoid bottlenecks while increasing the overall utilization efficiency. Diversity gains achieved by fast link adaptation or cooperative transmissions depend on the availability of updated channel state information, which on the other side increases overhead. To reduce this drawback, it is necessary to analyse more holistic and flexible MAC signalling that adapts to the current backhaul parameters, the centralisation requirements, and the actual access layer requirements. Moreover, reliable RRM mechanisms should be aware of the backhaul limits (such as latency and capacity) and adapt transmission parameters, accordingly (see Figure 3-6).

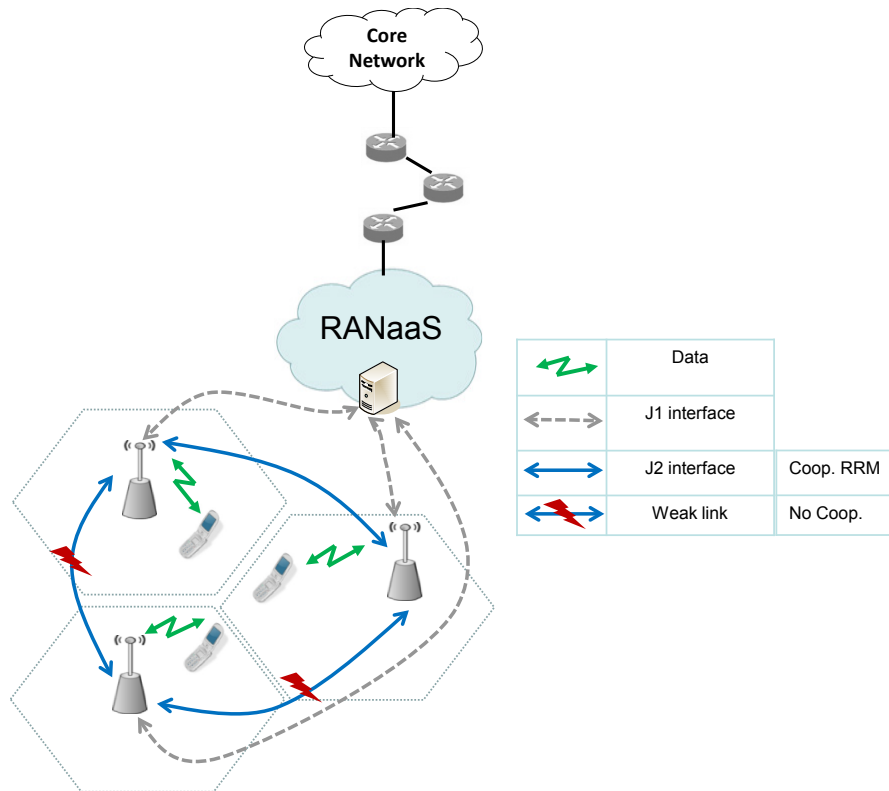


Figure 3-6: Flexible implementation of cooperative scheduling according to the backhaul capacity.

Finally, multi-hop wireless network with mmWave backhauling, which operates at 60 GHz technology, may provide up to 10 Gbps. In this scenario, cell-clustering schemes are used to improve efficiency in the utilization of access and backhaul, both in terms of cost (energy spent in the overall network) and utilization (sustainable data rate, load balance).

3.3.2 Technology Approaches

In this section, four key RRM technology approaches are outlined and some exemplary results are presented. We divide these approaches in two categories. The first category accounts for the RAN/BH awareness. In particular, information from small cells to RANaaS is exchanged in order to capture the effect of BH channel conditions and the traffic demand. The second category examines RRM schemes which adapt to BH constraints to enhance capacity with varying backhaul resources.

3.3.2.1 Enabling RAN/BH Awareness

Joint BH/access routing and scheduling using 60GHz backhaul

This approach addresses the problem of joint path selection and backhaul link scheduling in a dense small cell network assuming 60GHz multi-hop backhaul, coordinated by a local RANaaS. The objective of this study is initially to dynamically identify BH links and paths to be scheduled per a given time window, taking into account the target global objective for the network (in terms of maximizing backhaul capacity or aggregate utility). Secondly, assuming a realistic traffic model, to identify how the incoming flows are stored in the queues and forwarded to the next hops (or destinations), based on link selections in the previous step and the QoS requirements (delay, outage or data rate) per flow. The proposed solutions framework provides tune-able operation by dynamically adjusting the number of hops from RANaaS to the destination iSCs to meet the QoS requirements, the traffic demand and enhance the RAN performance.

In Figure 3-7, we illustrate the average BH link spectral efficiency when varying the average number of hops in 19-cell case study having both LoS/NLoS mmW BH. As can be seen, more short-distanced hops with favourable BH channel conditions can provide higher capacity. On the other hand, this figure also shows the maximum delay (defined as the number of timeslots needed to serve and the last iSC) which accounts for the queuing and half duplex constraints. We observe that the higher the number of routes the lower the maximum delay. This shows that we may achieve higher throughput with more hops; however this comes at the cost of higher delays.

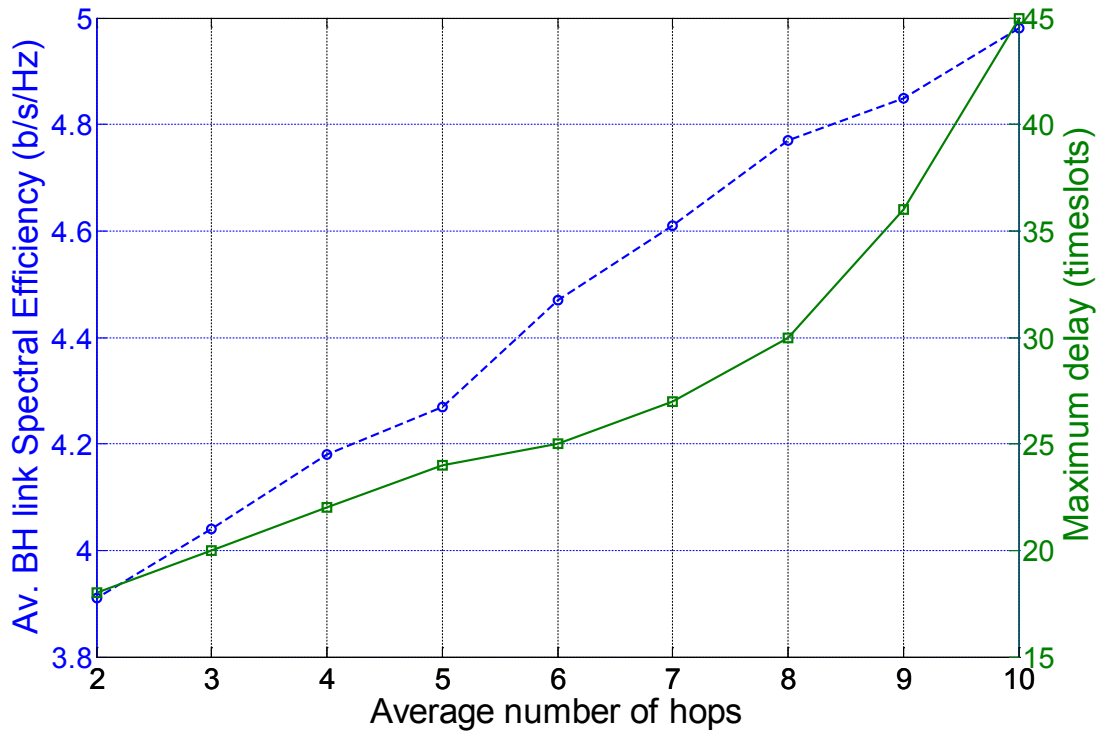


Figure 3-7: BH link Spectral Efficiency and Maximum Delay for different number of hops.

3.3.2.2 Joint BH/RAN Aware Centralized Connection Control

The connection control in future HetNets is a main challenge. First, finding the optimal association of UEs and eNBs is difficult because of the large number of possible assignments in dense small cell deployment. Second, in current technology, the association is based on the reference signal received power (RSRP) received from neighbouring eNBs; due to the downlink power unbalance in HetNets, this solution reduces the macro cell offloading and strongly limits the usage of the overall network resources. Third, the cell loads and their backhaul characteristics are taken into account to optimize the cell association. These issues affect the user performance and motivate the investigation of a joint RAN/BH aware connection control mechanism. Most of the required functions and messages to implement for this control mechanism are already defined in 3GPP LTE to steer traffic across neighbouring cells. Moreover, due to the large time scale in which the mechanism operates (seconds), it does not impose strong latency and capacity requirements on the backhaul network.

Figure 3-8 shows the average Area Throughput and joint RAN/BH utilization efficiency achieved with the Centralized Connection Control and the classic RSRP based scheme for when varying the number of active users in the network. We have considered here a tri-sectorial macro cell where three hotspots, each one composed by 4 neighbouring small cells, are deployed in the macro cell. Also, Cell Range Expansion (CRE) technique has been considered in the baseline solution to increase the macro cell offloading. Our results confirm that by efficiently using the overall RAN/BH available resources, the centralized connection control can lead up to a 32% and 42% gain with respect to the RSRP with CRE and the basic RSRP scheme.

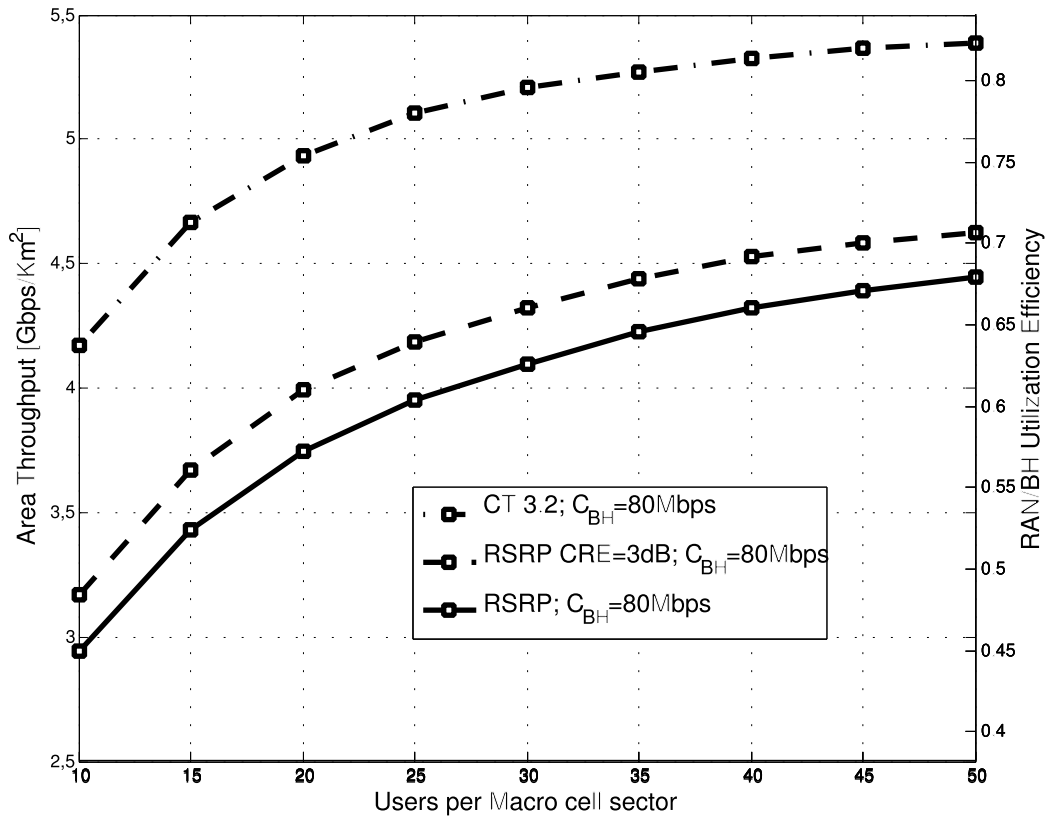


Figure 3-8: Area throughput and RAN/BH Utilization efficiency for different numbers of active users.

3.3.2.3 Adapting to BH constraints

Computational Complexity and Semi-Deterministic Scheduling

The computational resources of cloud platforms enable centralized processing of complex tasks with global knowledge, which is not available at the individual base stations. Semi-deterministic scheduling exploits these resources by shifting the computational load partially into the cloud, thus enabling the creation of a global scheduling plan for very dense small cell deployments.

For scheduling the users, only imperfect channel knowledge is available at the RANaaS as well as at iSCs. Due to the backhaul limitations in terms of capacity and/or latency, the channel uncertainty at the RANaaS is larger than the uncertainty at the iSC. In order to handle inter-cell interference the global scheduler determines whether a cell is allowed to transmit at a particular LTE subframe. Therefore, the potential throughput of each user for all possible combinations of transmitting iSCs is assessed and the best global scheduling plan selected, which is referred to as coordinated scheduling (CS). The plan is forwarded to the iSCs. Since the backhaul exchange is affected by delays, more recent channel information is available locally at the individual iSCs. Multi-stage scheduling accepts the cell specific global scheduling plan (i.e., transmitting or non-transmitting iSCs) and updates the selection of the particular user within a scheduled cell. Since local user selection has no impact on the interference received at other small cells, the local update improves the overall system performance.

Figure 3-9 illustrates the area throughput performance as a function of the backhaul delay, assuming a user velocity of 3 km/h. The baseline is represented by ‘Round Robin’ scheduling in a reuse one system (grey line). The scheduling is performed based on imperfect CSI at the RANaaS. Since centralized scheduling is affected by outdated CSI, the area throughput decreases with the delay. With the proposed multi-stage scheduling, the iSCs can update their local user selection, without changing the global scheduling plan. For higher delays, the area throughput gain can be increased (red line).

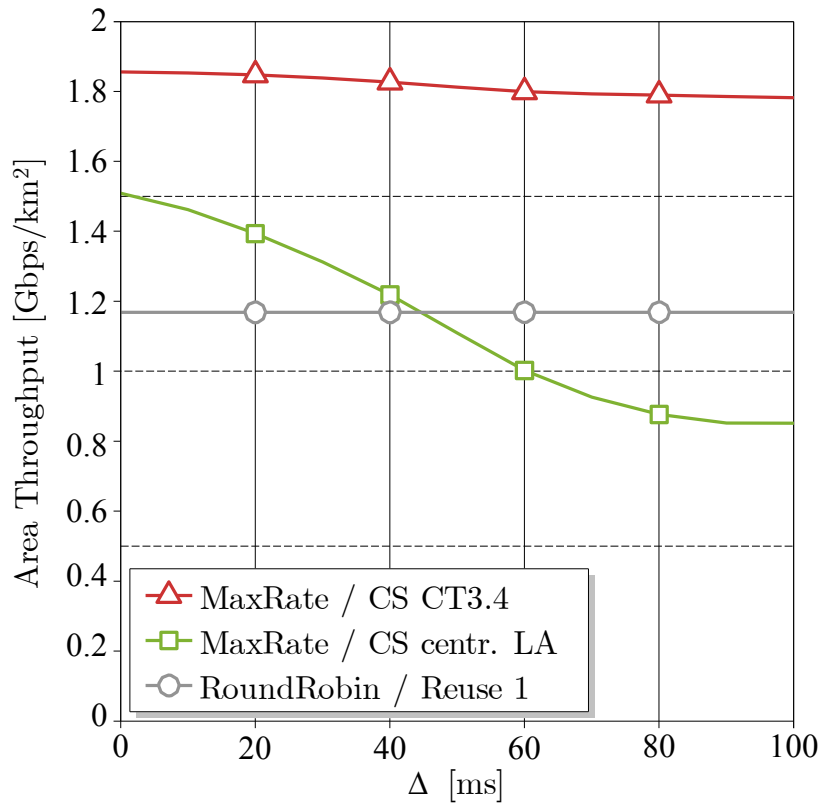


Figure 3-9: Area throughput as a function of the backhaul delay.

Hybrid local-cloud-based user scheduling for interference control

A joint centralized scheduler across multiple iSCs allows to fully exploit the multi-user diversity of the network to increase the strength of the desired signal and at the same time to reduce the interference between the adjacent iSCs. This centralized scheduling comes however at a high cost in terms of Channel State Information (CSI) sharing since all the channel coefficients (including the cross channels) of each user have to be provided to the centralized scheduler. This is in many cases unpractical in a timely manner leading the use of distributed scheduling at each iSC. In this context, we present here one possible approach to adapt to the backhaul conditions. This approach aims at taking into account a backhaul link to the RANaaS which introduces a large delay and exploiting the locally available interactions and CSI.

In that configuration, only the statistics of the channel (i.e., the topology) is available at the central controller in the RANaaS. This long term information is then used only to optimize *scheduling functions* which are applied in a distributed manner at the iSCs. The scheduling functions depend on some key parameters which are optimized in a centralized manner on the basis of the long term channel statistics, iteratively or through a learning process as they only need to be slowly updated. This approach leaves sufficient freedom at the iSCs to exploit the multi-user diversity and manage interference while enforcing coordination through the central controller.

One first approach is obtained if each iSC knows only its own direct channel [16]. The optimal distributed scheduling policy at each iSC is then to apply a threshold to decide whether or not to transmit: an iSC i transmits if the instantaneous realization of the channel gain is larger than the threshold. The value of the different thresholds can then be optimized during the centralized optimization to manage interference. Depending on the backhaul topologies, CSI or interference management information (e.g., inhibiting bits to avoid collisions) could be exchanged. The average rate per user obtained in that setting is shown in Figure 3-10. We compare the partially centralised approach with the two conventional scheduling approaches which are round robin and egoistic scheduling without coordination. The first one achieves perfect coordination but does not exploit the multi-user diversity while the second one leads to losses due to the interference. It can be seen that the partially centralized approach outperforms both approaches.

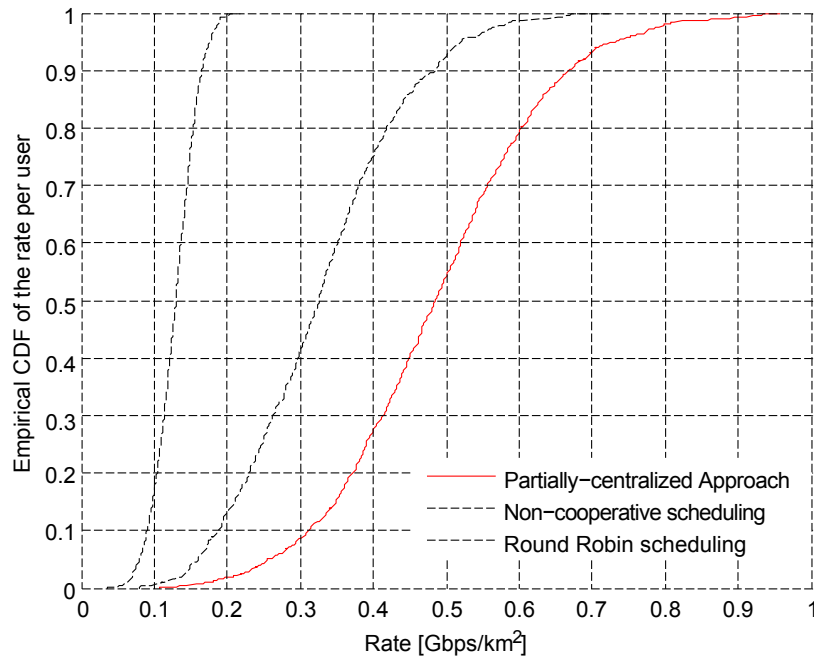


Figure 3-10: Distribution of the rate-per user for the different scheduling approaches.

3.3.3 Interface Requirements

3.3.3.1 Traffic differentiation

In the mobile backhaul, the following traffic types based on 3GPP interface definitions can be differentiated:

- S1-U traffic destined for the S-GW; note that S1-U traffic can be further differentiated according to the assigned QCI value;
- S1-C traffic destined for the MME;
- X2-U and X2-C traffic destined for other eNodeBs;
- OSS (operations support system) traffic destined for core applications that provide fault, configuration, and performance management;
- Network synchronization traffic.

All these traffic types have different requirements regarding quality of service (QoS). It can be generally stated that control plane traffic, e.g. S1-C, X2-C, and synchronization traffic, have higher requirements in terms of latency and reliability but have lower bandwidth requirements compared to user-plane traffic, e.g. S1-U and X2-U.

In today's networks, traffic differentiation for 3GPP traffic types is implemented via traffic type, e.g. control plane or user plane, and traffic class, e.g. based on QCI. Both are mapped on transport network traffic differentiation techniques, which depend on the employed transport network technology. For example, legacy ATM defines four different traffic classes which describe bandwidth requirement characteristics such as constant bit rate or variable bit rate. However, no delay requirements are specified. In LTE-Advanced, all-IP networks with layer 3 routing and VPN technologies, e.g. MPLS, or QoS and IP-aware layer 2 switching technologies, e.g. based on 802.1q/p, are expected to play a larger role due to the availability of Ethernet-capable eNodeBs in the access network and corresponding cost benefits.

While 3GPP defines a set of standardized QCI values [18], there is no standardized guideline available on how mobile network traffic is mapped to service classes on the transport layer. The problem is amplified by differences in the implementation between different vendors.

It can be concluded that neither 3GPP nor other standardization bodies offer a standardized methodology on how to map interface and protocol requirements of the mobile access network to the backhaul network. Configuration is thus a case-by-case issue which needs fine-tuning for each deployment and equipment scenario.

3.3.3.2 Required information for joint RAN/BH optimization

Figure 3-10 describes the information exchange supporting the joint RAN/BH optimization approaches described in Section 3.3.2 at the iNC and RANaaS, which in the iJOIN architecture are responsible entities for the BH and RAN management, respectively.

RAN measurements are first collected at the iSC level, then received and elaborated at the RANaaS (1), which forwards relevant information to the iNC (3). The information transferred to the iNC is typically related to the network load, QoS (PER, latency), and energy consumption. Additionally BH measurements indicating e.g., the capacity and the delay of the transport links are sent to iNC by the iTNs (2). For instance, this information can be used to change the number of hops between an iSC and the RANaaS to reduce the perceived delay (see CT 3.1). RAN-Aware BH orchestration is implemented at the iNC, which is in charge of updating the status of relevant iTN nodes (5) according to the new configuration of the transport network and of informing the RANaaS about the new status of the backhaul (6).

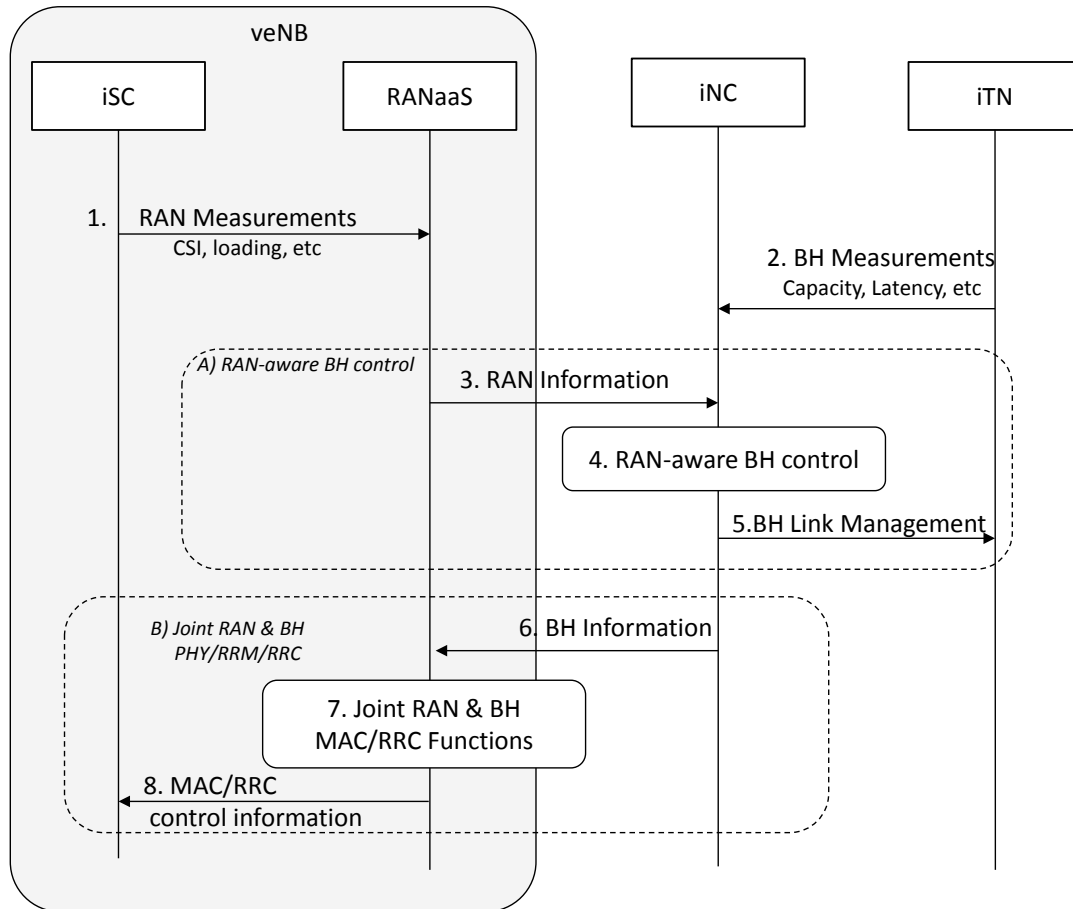


Figure 3-11: Information and control schemes for joint RAN/BH control.

The RANaaS manager is responsible for performing management and orchestration tasks in support of veNB operations (7). These RANaaS operations can be optimized by jointly using RAN/BH measurements provided by iSCs (1) and the iNC (6), respectively. RAN measurements are typically related to the quality/strength of the radio links (RSRP/RSRQ/CQI), the interference (Relative Narrowband Transmit Power Indicator (RNTP), High Interference Indicator (HII), and interference overload indication (IOI), and to the cell loads (i.e., the current radio resource usage and the composite available capacity). Moreover, the BH signalling may indicate the capacity and the delay of the transport links.

RAN measurements can be used to identify cells that suffer by excessive co-channel interference and they can be jointly used with the BH delay to setup robust ICIC mechanisms (8) (see CT3.4). Moreover, the BH capacity can be used to identify the need to distribute the load of some iSCs across the network. In this case, the veNB implements load balancing procedures by advancing the handing over of the UEs between neighbouring cells or by changing the mobility parameters (8) (see CT3.2).

4 Final Description and Evaluation of the iJOIN MAC/RRM Candidate Technologies

This section contains the final definition and the evaluation of MAC and RRM CTs (see Table 4-1) which have been developed in the iJOIN framework. Note that for most CTs, detailed descriptions are available in previous deliverables D3.1 (e.g. on required input and output parameters) and in D3.2 (e.g. on iJOIN architecture integration). As such, any information provided in this section *complements* and or *supersedes* information in previous deliverables.

Since the focus of this deliverable is on the *evaluation* of the CTs, it is recommended to read at least the corresponding section in D3.2 [2] for the full picture of a CT.

Table 4-1: iJOIN RRM/MAC Candidate Technologies (CTs)

CT	Topic	Abbreviation
3.1	Backhaul Link Scheduling and QoS-aware Flow Forwarding	BH Manager
3.2	Partly decentralized mechanisms for joint RAN and backhaul optimization in dense small cell deployments	Coordinated Cell Selection
3.3	Energy-Efficient MAC/RRM at Access and Backhaul	EE RRM
3.4	Computational Complexity and Semi-Deterministic Scheduling	SD Scheduler
3.5	Cooperative RRM for Inter-Cell Interference Coordination in RANaaS	Coop. RRM
3.6	Assess and Increase Utilization and Energy Efficiency	n/a
3.7	Radio Resource Management for Scalable Multi-Point Turbo Detection	MPTD RRM
3.8	Radio Resource Management for In-Network-Processing	INP RRM
3.9	Hybrid local-cloud-based user scheduling for interference control	HL Scheduler

4.1 CT 3.1: Backhaul Link Scheduling and QoS-aware Flow Forwarding

This CT considers a dense small cell network deployment where RANaaS operates as a coordinator and traffic aggregator for a cluster of iSCs. 60GHz multi-hop backhaul is considered to interconnect iSCs with each other and with RANaaS as illustrated in Figure 4-1. The proposed solution framework provides tuneable operations by dynamically adjusting the number of hops from RANaaS to the destination iSCs to meet the QoS requirements, the traffic demand, and enhance the RAN performance.

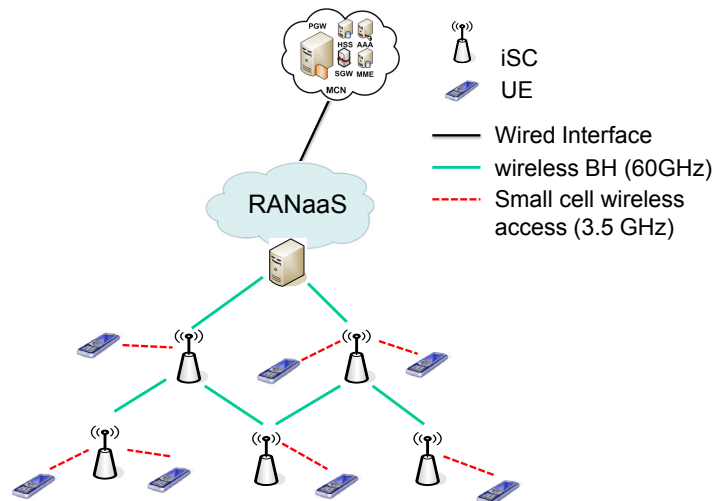


Figure 4-1: Backhaul link scheduling and QoS-aware flow forwarding.

To improve performance, the problem of joint path selection and backhaul link scheduling is tackled. BH links and paths, to be scheduled per a given time window, are initially dynamically identified, considering the optimisation of a target global objective for the network (e.g., average backhaul capacity). Then, we identify how the incoming flows are stored in the queues and forwarded to the next destination considering the link selections of the previous step and the QoS requirements (e.g., delay) per flow. For this latter step a realistic traffic model is assumed.

4.1.1 Final Implementation of CT

In D3.2 [2], we formulated the optimisation problem of the joint path selection and scheduling and proposed a framework which decouples this NP-hard combinatorial problem in two sub-problems for which we developed respective solving algorithms. Then, we presented results evaluating the average BH link spectral efficiency and the average delay (the average number of timeslots till each iSC is served for its access) versus the number of routes considered as well as the trade-off between maximum delay and average link throughput.

Compared to D3.2 [2], we enhance our simulation results in order to obtain realistic performance evaluations. Regarding the flow forwarding problem we consider a realistic traffic model. Moreover, we transform our evaluations on backhaul SE into backhaul throughput by implementing simulations for the wide area scenario as described in D5.2 [9]. Finally, we consider additional metrics in order to holistically evaluate the performance of the proposed solution. More specifically, we introduce the satisfaction ratio metric, reflecting how well the solution meets the target delay threshold, i.e., the ratio of the successful snapshots over the total number of snapshots. Furthermore, we utilize perceived cell throughput, defined as the aggregated cell throughput divided by the number of timeslots needed for the data to reach a specific cell.

4.1.2 Evaluation of the CT

To evaluate our work, Monte Carlo simulations for a 9-cell deployment are performed in a wide area scenario as described in D5.2 [9]. For the BH modelling, we used the parameters as shown in Section 4.1 of D3.2 [2]. In particular, Table 4-2 provides a summary of the simulation parameters.

Table 4-2: CT 3.1 Specific Simulation Parameters.

Parameter	Value
iSCs	9
ISD	20m
Users	Poisson arrivals per cell ($\lambda=2.5$)
Traffic	Random traffic demand per user (10-50Mbits)
Radio access channel	ITU UMi [TR 36.814]]
Carrier	2GHz (access), 60GHz (BH)

Bandwidth	10MHz (access), 100MHz (BH)
Snapshots	5000

The metrics used for the evaluations are the average BH link throughput, the average delay from the CU to reach each destination iSC, the satisfaction ratio, and the perceived backhaul throughput per cell.

The implementation of this scheme comprises two stages. The first stage is the extraction of results for the path selection problem. Here, we adjust the number of paths (k), so as to find the optimal path selection in different cases. In Figure 4-2 two extreme cases are shown. The first one is the single-hop case ($k = 9$, Figure 4-2-left), and the second one is the 1-path case with 9 hops ($k = 1$, Figure 4-2-right).

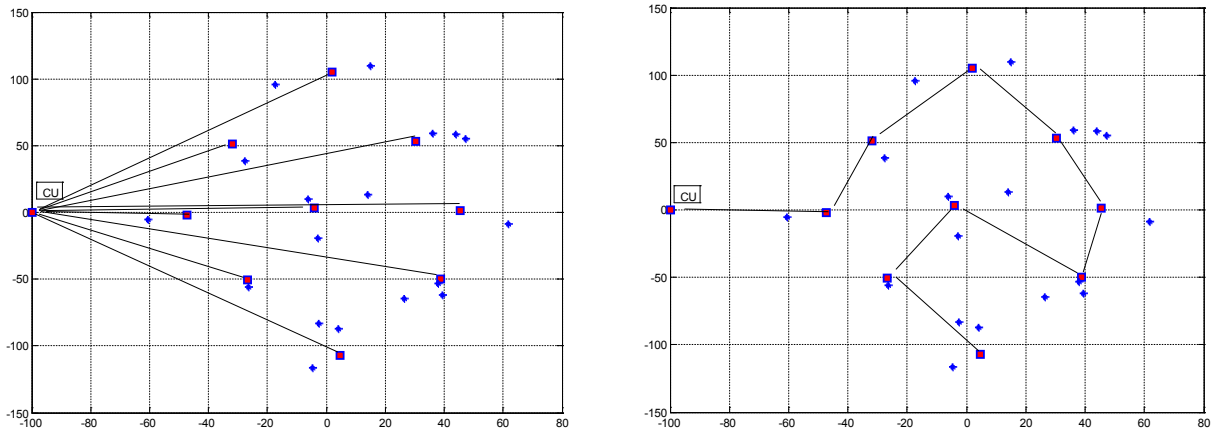


Figure 4-2: Illustration of BH topology for $k=9$ and $k=1$.

The path selection algorithm was analysed for all the possible number of paths. As shown in Figure 4-3, the average BH link spectral efficiency drops when we increase k . This is due to the fact that the higher the number of paths, the lower the number of hops. In other words, long-distanced links with NLoS will impact the performance. On the other hand, in low- k regime with more hops, short-distanced LoS links increase the throughput performance.

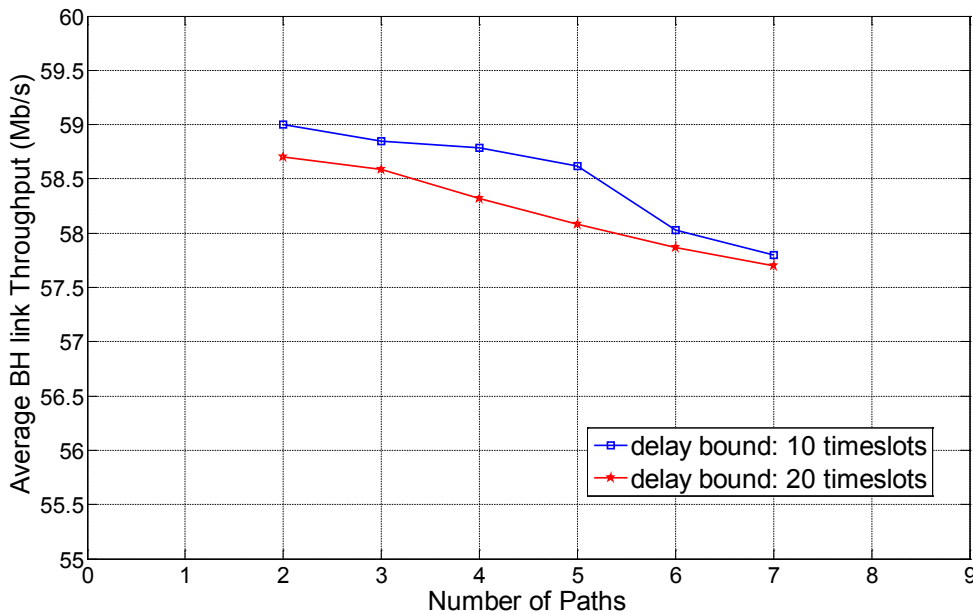


Figure 4-3: Average BH link throughput for different delay bounds vs. number of paths.

Another key observation is on the effect of the delay bound (i.e. the maximum allowed number of timeslots until the last iSC receives the last packet) in different cases. In particular, we used two thresholds of 10 and

20 timeslots. For 10-timeslot case, we observe that the BH link throughput has improved; however, as we discuss in Figure 4-5, this comes at the cost of lower satisfaction ratio.

In the second stage of our solution, the backpressure scheduling is used. As can be seen in Figure 4-4, we evaluate the average delay (the average number of timeslots till each iSC serves the required traffic) in this scenario. As shown, the higher the number of paths, the lower will be the delay. This is in line with our analysis of Figure 4-3. In other words, we may achieve higher throughput with more paths; however, this comes at the price of higher levels of delay. Another important observation comes from the impact of delay bound. As shown, by increasing the number of paths, the average delay with lower threshold bound (10 timeslots) decreases steadily. In the higher threshold (20 time-slots), also, the average delay drops; however, the pace of drop is relatively slower compared to the low threshold due to the fact that links with higher delays are admitted as feasible solutions.

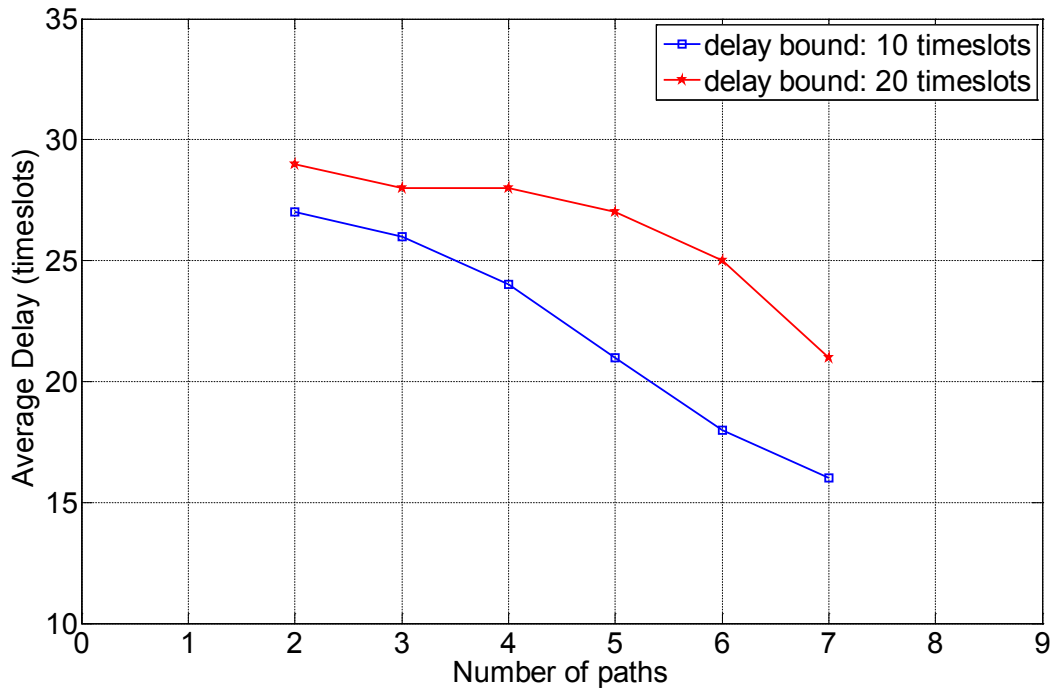


Figure 4-4: Average delay for different delay bounds vs. number of paths.

Combining the analysis of Figure 4-3 and Figure 4-4, both BH link throughput and delay are improved for lower delay threshold. Nevertheless, setting the threshold should be carefully monitored as shown in Figure 4-5. This is mainly due to the fact that lower threshold can potentially cause lower level of satisfaction ratio, making the routing optimization infeasible in some instances. In particular, the satisfaction ratio can severely degrade for higher numbers of path in lower bound case.

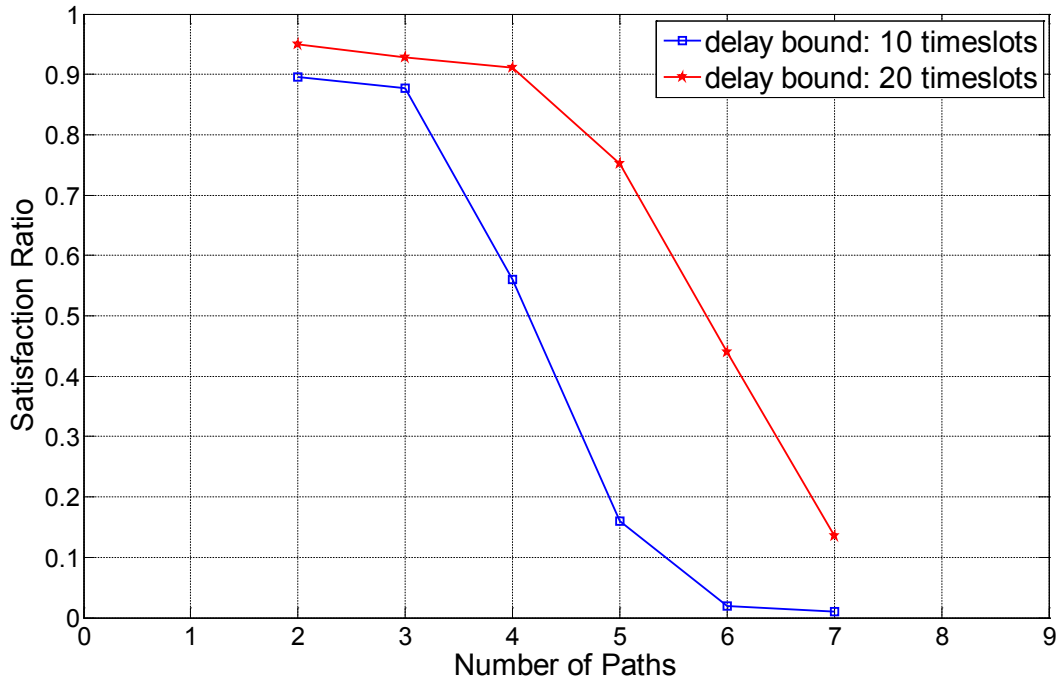


Figure 4-5: Satisfaction ratio vs. number of paths for different delay bounds.

Finally, Figure 4-6 shows how the BH path selection and scheduling affects the small cell performance in terms of CDF of perceived backhaul throughput per cell for 2 to 5 paths in high delay bound. It is observed that the higher the number of paths (less hops) the higher will be the perceived cell throughput, in particular in the tail end of the CDF curve. However, in the single-hop case there will always be a very high chance that several RANaaS-iSC links will be NLoS links and the fact that there is no alternative path to be chosen for the transmission will degrade the average system performance. As can be seen in Figure 4-6 where we also provide the single-hop case (with no delay bound), due to the fact that NLoS links can be chosen for the transmission, the perceived throughput is far worse compared to the proposed solution.

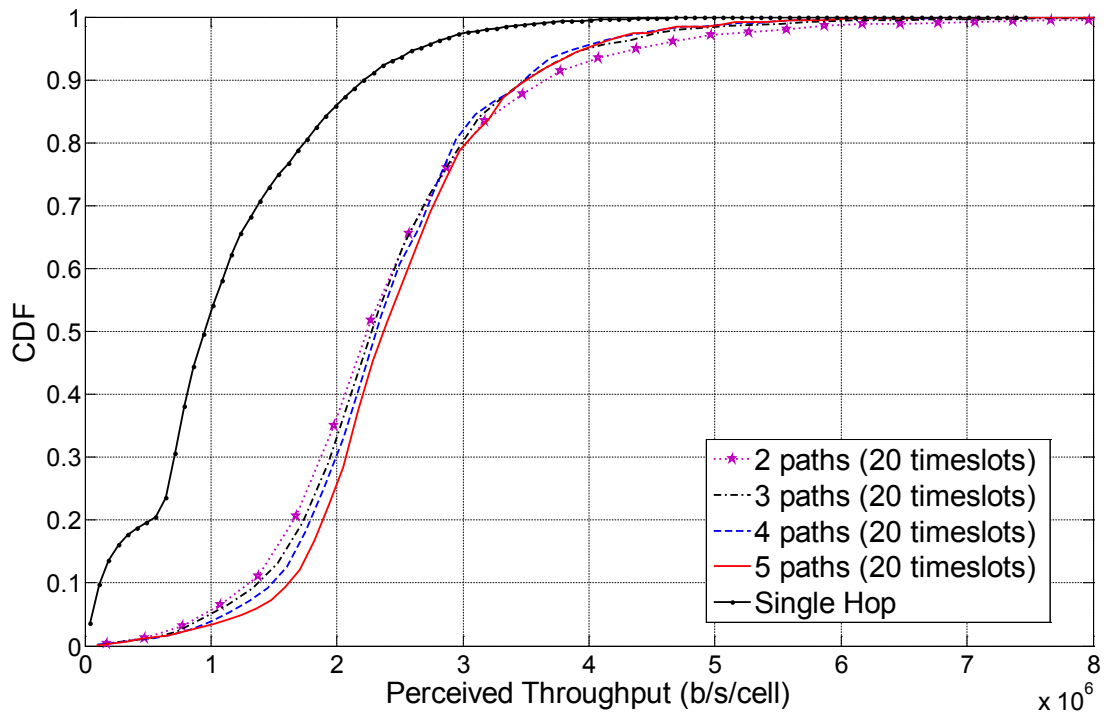


Figure 4-6: Perceived cell throughput for different number of paths.

4.2 CT 3.2: Partly decentralized Mechanisms for joint RAN and Backhaul Optimization in dense Small Cell Deployments

This CT investigates a scenario where iSCs are densely deployed to satisfy the demand of high data rate services of future wireless networks. However, wireless backhaul characterized by limited capacity affects the maximum cell throughput. Therefore, we have introduced a backhaul aware cell selection mechanism that iteratively enables network-wide load balancing and improves the overall network area throughput [2]. Our results confirm the benefits of our proposal both in terms of throughput and joint RAN/BH utilization efficiency.

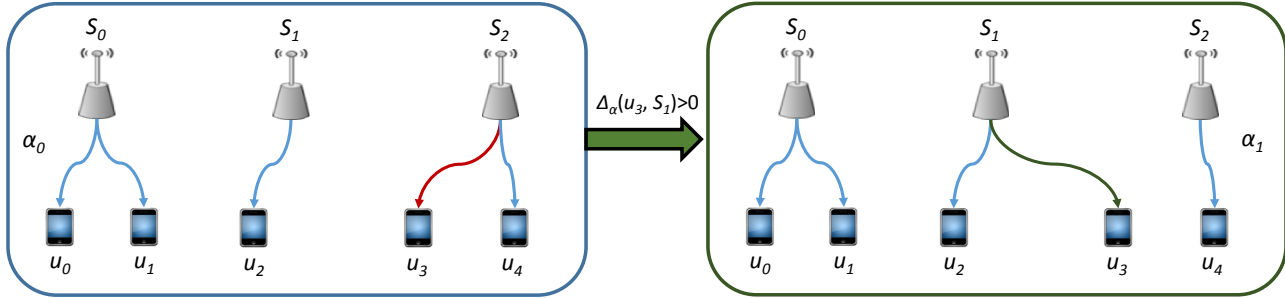


Figure 4-7: The proposed Load & BH Aware Cell Association scheme.

4.2.1 Final Implementation of CT

The proposed cell association framework is composed of two solutions: a centralized and a distributed one. These algorithms have been described in details in D3.2 [2]; here, we will demonstrate their advantages in terms of Area Throughput and Utilization Efficiency.

4.2.2 Evaluation of the CT

We consider a mobile wireless cellular network in which user terminals and eNBs implement an OFDMA air interface based on 3GPP/LTE downlink (DL) specifications. Coherently with the study on small cell enhancement, which is currently under investigation in 3GPP [12], our research focuses on HetNets where small cells operate in a dedicated carrier with respect to the macro cell. We also consider that the RANaaS [13] orchestrates the connection control functionalities (cell selection, mobility, etc.).

Here we assess our solution in the iJOIN square use case scenario where a cluster of four small cells is deployed in each macro cell sector (see Table 6-4 in [2]) and we compare the performance of our algorithm with the reference RSRP approach. Furthermore, to improve the macro-cell offloading we also evaluate the impact of a cell range expansion technique.

Our centralized connection control scheme aims at optimizing the overall network area throughput (R_A)

$$R_A = \frac{C}{A} \text{ [Gbps/km}^2\text{]}, \quad (4.1)$$

where C and A represent the network average throughput and the HetNet area (which coincides here with the macro cell area), respectively.

However, the load balancing functionalities has also a positive impact in terms of both RAN and BH utilization efficiency, which can be computed as

- RAN Utilization efficiency:

$$u^{RAN} = \frac{C}{C_{RAN}^T}, \quad (4.2)$$

- BH Utilization efficiency:

$$u^{BH} = \frac{C}{C_{BH}^T}, \quad (4.3)$$

Where $C_{RAN}^T = N_{iSC} \cdot C_{RAN}$, $C_{BH}^T = N_{iSC} \cdot C_{BH}$, N_{iSC} is the number of deployed small cells, C_{BH} is the BH capacity at each small cell, and C_{RAN} is the maximum cell capacity; then the joint RAN/BH utilization efficiency can be expressed as

$$u^{RAN/BH} = \frac{C}{2 \cdot C_{BH}^T} + \frac{C}{2 \cdot C_{RAN}^T}, \tag{4.4}$$

Table 4-3: CT 3.2 simulation parameters (centralized algorithm).

Parameter	Value
Macro ISD	500 m
N_{iSC} per sector	4
Users per sector	10-50
Traffic	Full buffer
Radio access channel	ITU UMi [18]
Carrier	2GHz (macro), 3.5GHz (small cells)
Bandwidth (RAN)	10MHz
Max Spectral Efficiency (RAN)	12 bps/Hz
C_{BH}	40-120 Mbps

Relevant simulation parameters are summarized in Table 4-3.

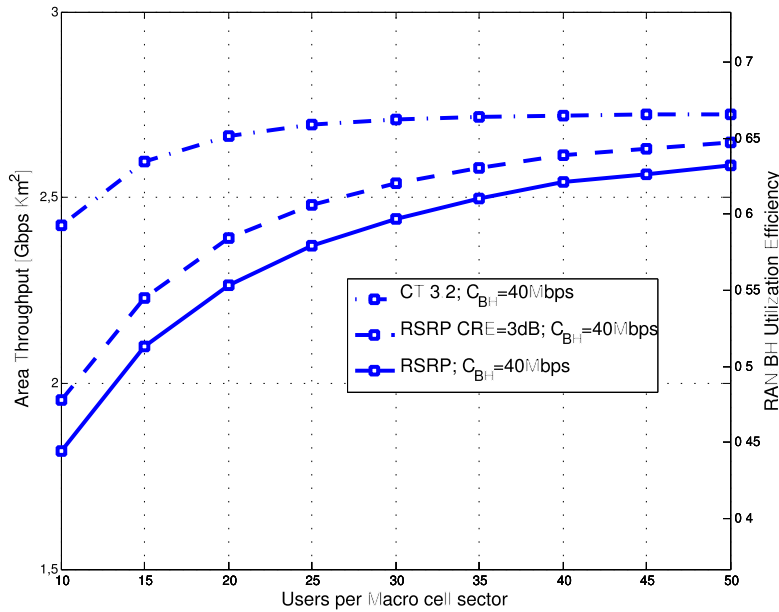


Figure 4-8: Area Throughput and Utilization Efficiency versus the number of active users per macro cell sector ($C_{BH}=40$ Mbps).

Figure 4-8 shows the performance gains of our solution when the BH capacity per cell is equal to 40Mbps (very poor BH). As expected, CRE leads to some improvements with respect to the basic RSRP scheme due to the enhanced macro cell offloading. However, in lightly loaded scenarios, our solution gains up to 24% and 33% with respect to the RSRP with CRE and the basic RSRP solution, both in terms of Area Throughput and joint RAN/BH utilization efficiency. Nevertheless, when the number of active users increases, the RAN capacity is limited by the BH capacity constraints and the advantages of a centralized connection control solution are strongly reduced. Actually, the area throughput floor indicates that in this region, the user throughput has to be reduced to satisfy a large number of active users. We can conclude that in this scenario, the achievable RAN capacity and utilization efficiency is largely affected by the BH capacity.

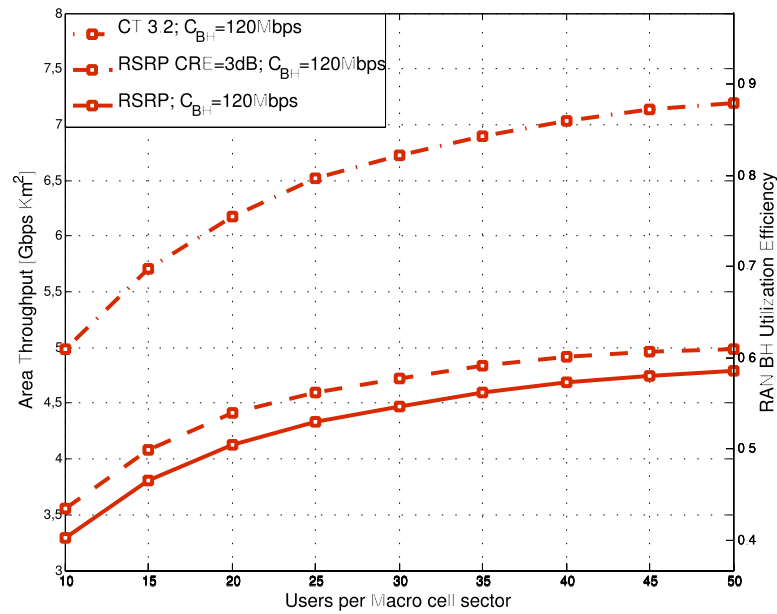


Figure 4-9: Area Throughput and Utilization Efficiency versus the number of active users per macro cell sector ($C_{BH}=120\text{Mbps}$).

Figure 4-9 shows the performance gains of our solution when the BH capacity per cell is equal to 120Mbps. In this case, the BH capacity does not constraint the RAN performance since $C_{BH} = C_{RAN}$; hence, there is a large room for improving the network performance by steering traffic across the network cells. In particular, results confirm that centralizing the connection control functionalities can lead to larger gain with respect to the CRE solution. We can observe that with the RSRP solutions, the network performance is constrained by C_{RAN} and that the performance slope is limited when increasing the number of active users. On the contrary, by distributing traffic across the network, the proposed solution take advantage of the overall available resources and leads up to 40% and 50% of gains with respect to the RSRP with CRE and the basic RSRP scheme.

The previous results correspond to the centralized approach explained in D3.2 [2]. In that document, a second algorithm was proposed as well. This algorithm allows a distributed implementation and relies on a game-theoretic approach of the cell selection problem. Some preliminary results for a cluster of iSC were already given in D3.2 [2]. Here, we provide further results for the wide-area continuous coverage scenario described in D5.2. The performance metric is the RAN Utilization efficiency as defined above. Table 4-3 presents the main simulation parameters.

Table 4-4: CT 3.2 simulation parameters (distributed algorithm).

Parameter	Value
ISD	50 m
Number of iSC	19
Users per iSC (avg)	3
Traffic	Full buffer
Small cell dropping	Regular on Hexagonal Grid
Radio access channel	ITU UMi [TR 36.814]
UE dropping in a cluster	Random dropping
Bandwidth (RAN)	10MHz
Minimum distance	UE-iSC 5 m
Max Spectral Efficiency (RAN)	5.55 bps/Hz
C_{BH}	50-100 Mbps

Figure 4-10 shows the CDF of the area throughput for the Rate Potential Game, which was the best option to maximize the network capacity as shown in D3.2 [2]. As it can be seen, the proposed solution improves the total capacity of the network compared to the capacity obtained when the user always connects to the closest iSC (benchmark scenario). The capacity increase of the proposed algorithm is about a 5% on average with

respect to the benchmark scenario. In general, this value will depend on the diversity of possible backhaul capacities for the small cells.

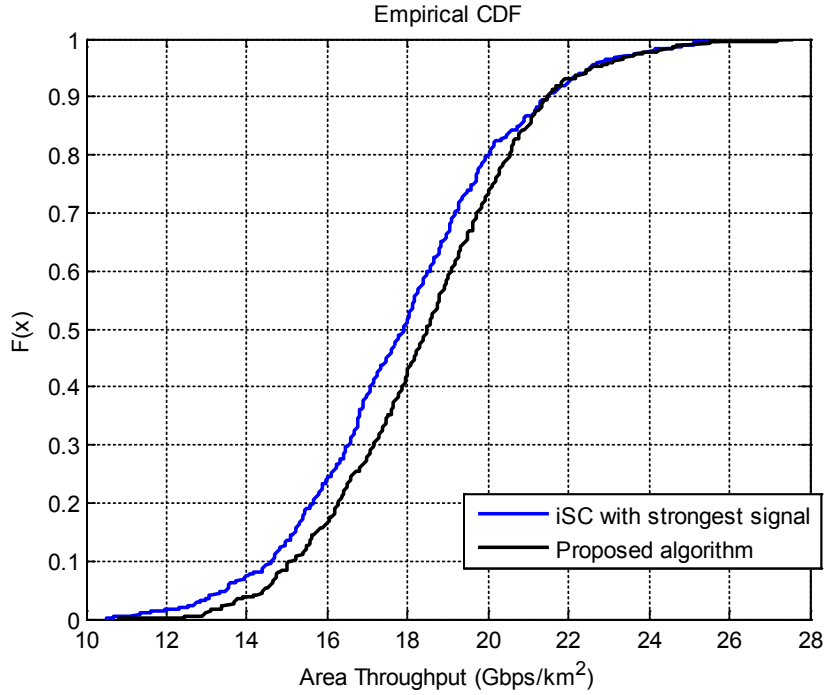


Figure 4-10: Area Throughput for the distributed approach.

4.3 CT 3.3: Energy-Efficient RRM at Access and Backhaul

4.3.1 Final Implementation of CT

In this deliverable, we extend the solution proposed in D3.2 [2] through fuzzy Q-Learning, which allows for simplifying the characterization of the network status and finding an optimal solution to control the DTX pattern in a realistic environment.

In this CT, the RANaaS receives data from the core network through the S1 interface and stores it in dedicated buffers. When required, the RANaaS, activates a given iSC and forwards to it part of the associated traffic through the J1 interface. Thereafter, the activated iSC will autonomously manage available radio resources to efficiently transmit the data received from the RANaaS according to a first-in-first-out policy.

In our model the RANaaS is equipped with N buffers of size M packets, each one dedicated to a specific iSC. Let S be a finite set referred to as the small cell state set and defined as $S = Q \times R \times UE$, where Q , R , and UE are the composite state sets, which describe the buffer state, the cell capacity, and the number of active UEs per iSC. In particular, at each time step, the small cell queue is described through the couple $\mathbf{q}_t = (\mathbf{l}_t, \mathbf{n}_t) \in Q$, where \mathbf{n}_t indicates the size of each packet presents in the queue and \mathbf{l}_t the packet time-to-live. Moreover, $r_t \in R$ and $u_t \in UE$ describe the cell capacity and the number of served UEs, respectively. Accordingly, each small cell state can be represented through the state vector $\mathbf{s}_t = (\mathbf{q}_t, r_t, u_t)$. The RANaaS observes the current state of each small cell and selects an action $a_t \in A, \{0, 1\}$, where $a_t = 0$ corresponds to keep the small cell and the associated BH link idle, while $a_t = 1$ indicates that the small cell is activated and $\min(r_t, \sum \mathbf{n}_t)$ are transferred from the RANaaS to the small cell. A cost is associated to each state-action pair given by

$$C(\mathbf{s}_t, a_t) = P(\mathbf{s}_t, a_t) + \beta d(\mathbf{s}_t, a_t), \quad (4.5)$$

where $P(\mathbf{s}_t, a_t)$ describes the aggregated BH and small cell power consumption [15], $d(\mathbf{s}_t, a_t)$ indicates the number of bits dropped due to the latency constraints, and β is a constraint that prioritizes amongst QoS and power consumption.

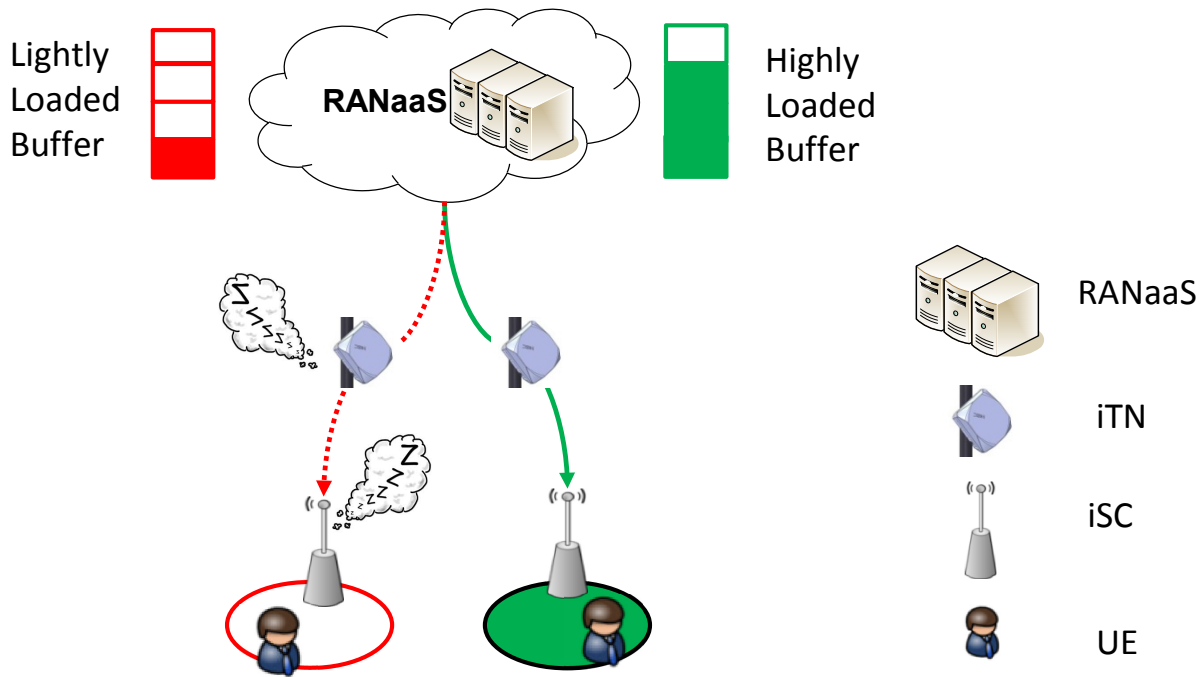


Figure 4-11: The proposed DTX control framework.

Fuzzy Q-Learning

The purpose of fuzzy logic is to realistically represent those environments that cannot be efficiently described through mathematical models.

In our model, the small cell state vector is represented by $L = 4$ linguistic variables. Then by denoting as $\bar{S} = \{\{\bar{s}_1, \dots, \bar{s}_n\}\}$ the set of fuzzy state vectors composed by the linguistic variables the fuzzy inference rule is

IF input state vector \mathbf{s} is \bar{s}_i
 Then action a_0 with $q(\bar{s}_i, a_0)$
 or action a_1 with $q(\bar{s}_i, a_1)$,

where $a_j \in A$ is the j -th action candidate which is possible to choose for state \bar{s}_i , and $q(\bar{s}_i, a_j)$ is the fuzzy Q-value for the state-action pair (\bar{s}_i, a_j) . The FQL has two outputs: one corresponds to the inferred action and the other represents the Q-value for the state-action pair (\mathbf{s}, a) :

$$a = \frac{\sum_i^n w_i \cdot a_i}{\sum_i^n w_i} \quad (4.6)$$

$$Q(\mathbf{s}, a) = \frac{\sum_i^n w_i \cdot q(\bar{s}_i, a_i)}{\sum_i^n w_i}, \quad (4.7)$$

where w_i represents the truth value (i.e., the output of fuzzy-AND operator) of the rule representation of FQL for \bar{s}_i , and a_i is the action selected for state \bar{s}_i . Q-values have to be updated after the action selection process according to $q(\bar{s}_i, a_i) = q(\bar{s}_i, a_i) + \alpha \cdot \Delta q(\bar{s}_i, a_i)$, where α is the learning rate, $\Delta q(\bar{s}_i, a_i) = (C(\mathbf{s}, a) + \delta Q(\mathbf{y}, a') - Q(\mathbf{s}, a)) \cdot \frac{w_i}{\sum_i^n w_i}$, $C(\mathbf{s}, a)$ represents the cost obtained applying action a in state vector \mathbf{s} , and $Q(\mathbf{y}, a')$ is the next-state optimal Q-value defined as:

$$Q(\mathbf{y}, a') = \frac{\sum_i^n w_i \cdot q(\bar{y}_i, a_i^*)}{\sum_i^n w_i}, \quad (4.8)$$

and $a_i^* = \underset{a_j^* \in A}{\operatorname{argmin}} q(\bar{y}_i, a_j^*)$ is the optimal action for the next state \bar{y}_i , after the execution of action a_i in the fuzzy state \bar{s}_i .

The proposed FQL algorithm consists of a four layer Fuzzy Inference System (FIS), where the first layer has as input the four linguistic variables defined by the term sets:

$$\begin{aligned}
T(\mathbf{l}) &= \{Very\ High\ Urgent, Very\ Urgent, Urgent, Almost\ urgent, Not\ Urgent\}, \\
T(\mathbf{n}) &= \{Very\ Highly\ Loaded, Higly\ Loaded, Mid\ Loaded, Lightly\ Loaded, UnLoaded\}, \\
T(r) &= \{Low\ Capacity, Mid\ Capacity, High\ Capacity\}, \text{ and} \\
T(u) &= \{Few\ UEs, Some\ UEs, Many\ UEs\}.
\end{aligned}$$

In this layer, there are $z = |T(\mathbf{l})| + |T(\mathbf{n})| + |T(r)| + |T(u)| = 16$ output nodes, each one describing (via a trapezoidal membership function $\mu(x)$, see (Figure 4-12) to which degree the small cell state variables belong to the appropriate fuzzy sets (fuzzification process). Let denote as $O_{1,k}$ with $1 \leq k \leq z$ the output of this layer.

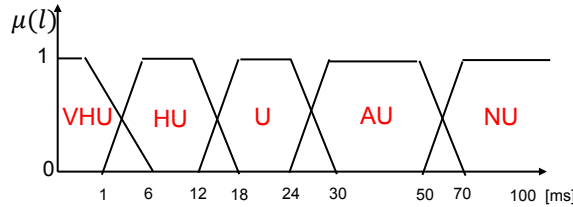


Figure 4-12: Exemplary membership function to match the queue latency with the associated linguistic terms.

The second Layer is named as the rule nodes layer and is composed by $n = |T(\mathbf{l})| \cdot |T(\mathbf{n})| \cdot |T(r)| \cdot |T(u)| = 225$ nodes. Each node gives as output the truth value of the i -th fuzzy rule $O_{2,i}$ with $1 \leq i \leq n$ which is the product of four membership values corresponding to the term set inputs.

The third layer is also composed of $n = 225$ nodes named as action-selection nodes. The selection of the per node action a is based on the ϵ -greedy action selection policy. Each node $1 \leq i \leq n$ generates two output values as follows:

$$O_{3,i}^A = \frac{O_{2,i} \cdot a_i}{\sum_i^n O_{2,i}} \quad (4.9)$$

$$O_{3,i}^Q = \frac{O_{2,i} \cdot q(\bar{s}_i, a_i)}{\sum_i^n O_{2,i}} \quad (4.10)$$

The fourth layer is has two output nodes, action node O_4^A and Q-value node O_4^Q , which represent the results of the defuzzification process.

$$O_4^A = \sum_i^n O_{3,i}^A \quad (4.11)$$

$$O_4^Q = \sum_i^n O_{3,i}^Q \quad (4.12)$$

4.3.2 Evaluation of the CT

We consider a mobile wireless cellular network in which user terminals and eNBs implement an OFDMA air interface based on 3GPP/LTE downlink (DL) specifications. Coherently with the study on small cell enhancement, which is currently under investigation in 3GPP [12], our research focuses on HetNets where small cells operate in a dedicated carrier with respect to the macro cell.

Here we assess our solution in the iJOIN square use case scenario where a cluster of four small cells is deployed in each macro cell sector (see Table 6-4 in D3.2 [2]). Near Real Time Video traffic (average user rate and latency constraints equal to 64 Kbps and 100 ms, respectively) is simulated for the UEs located in the central macro cell, while other cells generate only additive interference. Moreover, since we focus on a small cell controller, we present here only results related to small cells. Results are averaged over 50 independent runs that simulate 10 seconds of network activity. At the beginning of a run, the small cell hotspots and the UEs are randomly deployed in the macro cell area.

The most relevant simulation parameters are summarized in Table 4-5.

Table 4-5: CT 3.3 simulation parameters.

Parameter	Value
Macro ISD	500 m
N_{iSC} per sector	4
Users per sector	20-45
Traffic	NRTV 64Kbps
Radio access channel	ITU UMi [18]
Carrier	2GHz (macro), 3.5GHz (small cells)
Bandwidth (RAN)	10MHz

Figure 4-13 shows the average small cell energy per bit [j/bit] with respect to the number of active UEs in the macro-cell. The proposed Fuzzy Q-Learning scheme is compared with

- 1) the baseline solution where small cells do not have sleep mode (No DTX),
- 2) the classic DTX where small cells enter in sleep mode when there is no data to transmit (RAN DTX), and
- 3) the joint RAN/BH DTX where both small cells and BH nodes enter in sleep mode in absence of data. The parameters used for the power model are described in D5.2 [9].

First, it is worth to underline that the small cell energy per bit decreases when increasing the number of active UEs due to the rising of the network throughput. The small cell network energy consumption is slightly affected by the network load, especially if the system does not use any energy saving scheme [2]. Second, we can observe as the classic DTX solution leads to limited energy saving since the overall power consumption is largely due to the BH contribution. In fact, the joint RAN/BH DTX gains up to the 48 % with respect to the classic DTX scheme. However, the proposed solution, by jointly exploiting the energy-delay trade-off and Fuzzy Q-Learning optimization achieves up to 60% and 22% of energy saving with respect to the RAN DTX and joint RAN/BH schemes, respectively.

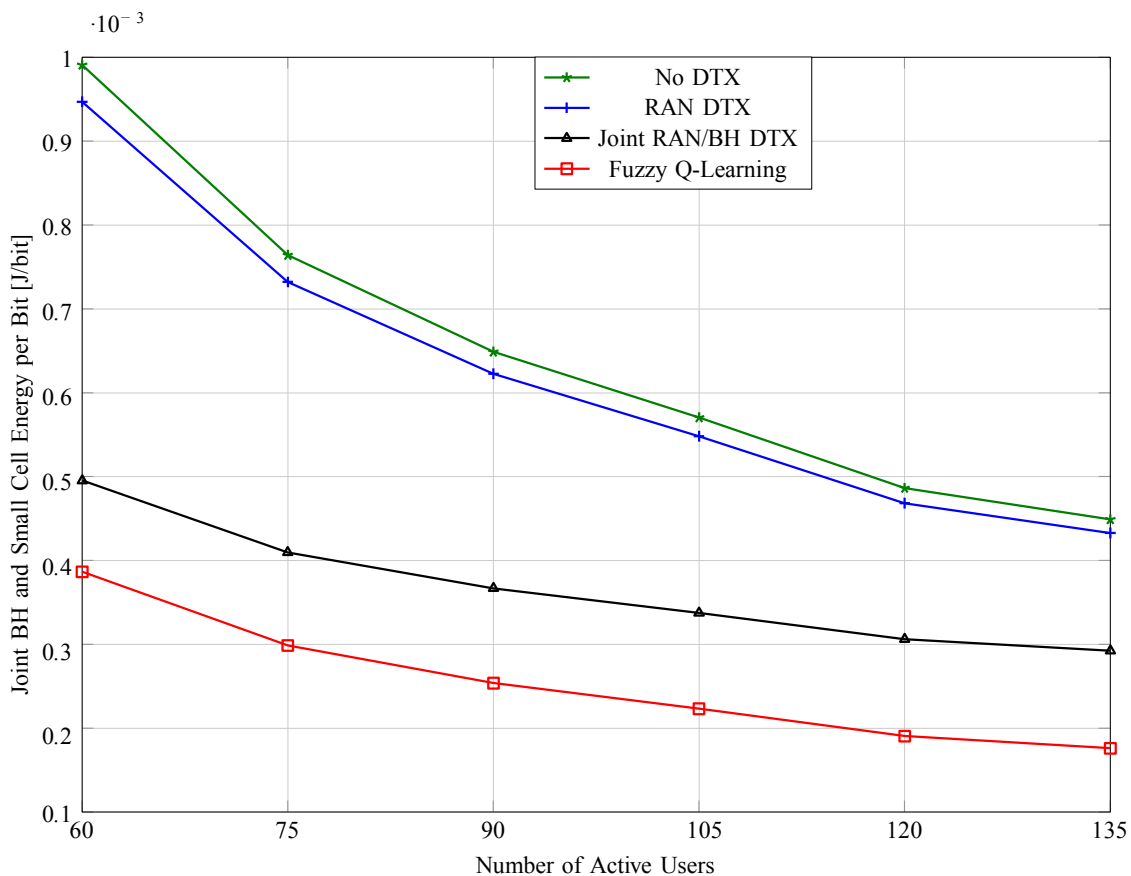


Figure 4-13: Small Cell energy per bit vs. the number of active UEs.

Figure 4-14 shows the average perceived latency at the small cell UEs with and without the proposed energy saving scheme. Note that DTX has no impact on the latency; accordingly we compare our technique only with the baseline solution without DTX. We can observe that the average latency associated to both schemes is far lower than the latency constraint in NRTV traffic (100 ms). Then, we can underline that the Fuzzy Q-Learning scheme has three major effects: it avoids packet losses due to the latency constraints, it adapts the small cell control to the overall delay required by the HARQ retransmissions (i.e., 24 ms), which leads to an average latency close to 70 ms, and it increases the small cell duty cycle when increasing the network load. In fact, in highly loaded scenarios, the maximum amount of data that a small cell can transmit at once is reached faster, which reduces the user perceived latency.

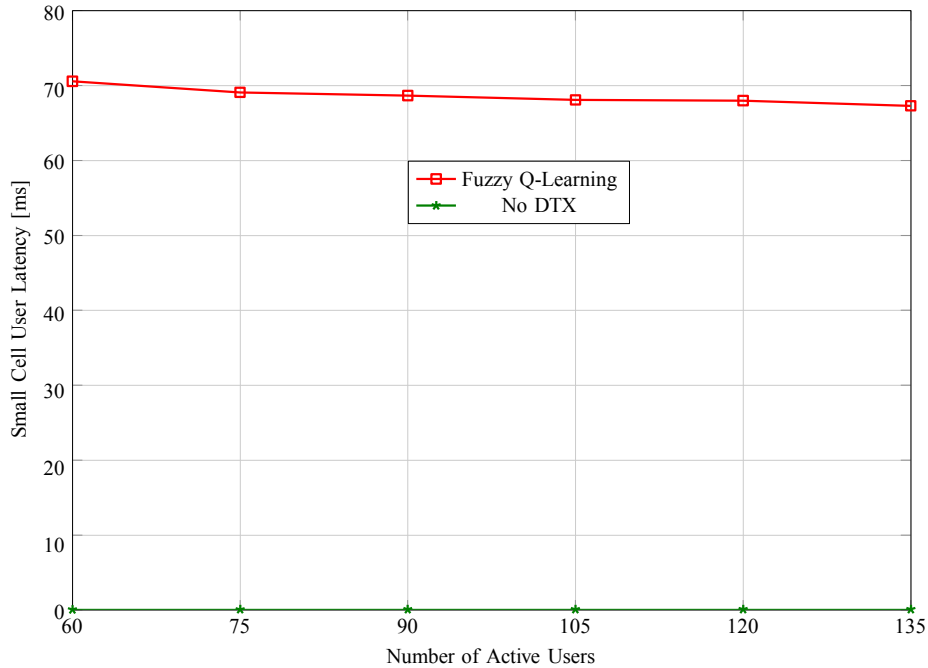


Figure 4-14: Small Cell user latency vs. the number of active UEs.

Finally, Figure 4-15 describes the average network Packet Error Rate (PER) achieved by the proposed scheme and the baseline solution without DTX. Note that, as for the latency, the classic DTX does not impact the system PER. Simulation results show that both the compared solutions ensure a PER below the PER target (10^{-3}). However, the proposed solution results in much higher PER than the baseline, especially in highly loaded scenarios. These losses are due to two concurrent consequences of the proposed scheme: first, it affects the link adaptation reliability, since active UEs are not able to measure the interference that can be generated by neighbouring small cells in sleep mode. Second, when two neighbour iSCs are activated in the same LTE subframe, they may perceive spikes of interference on most of the bandwidth (due to the accumulated traffic load).

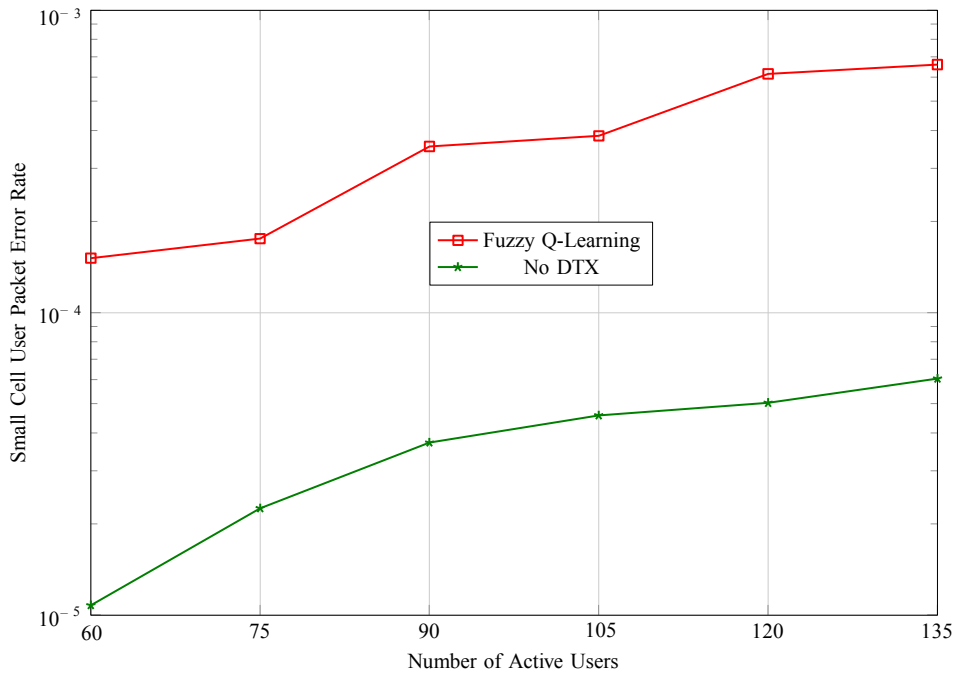


Figure 4-15: Small Cell UE Packet Error Rate Vs. the Number of Active UEs.

4.4 CT 3.4: Computational Complexity and Semi-Deterministic Scheduling

The computational resources of cloud platforms enable centralized processing of complex tasks with global knowledge, which is not available at the individual base stations. Semi-deterministic scheduling exploits these resources by shifting the computational load partially into the cloud, thus enabling the creation of a global scheduling plan for very dense small cell deployments. This is necessary to combat the severe inter-cell interference caused by short inter-site distances in such scenarios.

Previously, we presented novel robust link adaptation (LA) and scheduling algorithms, which are designed for scenarios where channel state information (CSI) of the users is only imperfectly available at the network [2], [17], [26]. In particular, the exchange of CSI between iSCs and RANaaS via latency affected backhaul connections is in the focus of this work. In such a setup, centralized processing has to deal with outdated channel information. We could show that the proposed robust algorithms outperform non-robust methods and are able to guarantee target outage probability constraints. Those schemes have been designed under the assumption that interference from other cells can be considered Gaussian.

4.4.1 Final Implementation of CT

In this document we focus on the multi-cell aspect of the proposed scheduling scheme considering common scenarios of iJOIN [9]. In this regard, hierarchical multi-stage scheduling is of special interest. Therefore, transmission resources of users, which are located within the cooperation cluster, are assigned by coordinated scheduling (CS) at the RANaaS. This functional entity refers to the global scheduler (see Figure 4-16). It benefits from flexible computational resources. However, processed channel information is outdated due to backhaul delays. In order to compensate the effect of outdated CSI, local schedulers are implemented at the iSCs, which offers the opportunity to update the global scheduling decision. We analyse this procedure in the following.

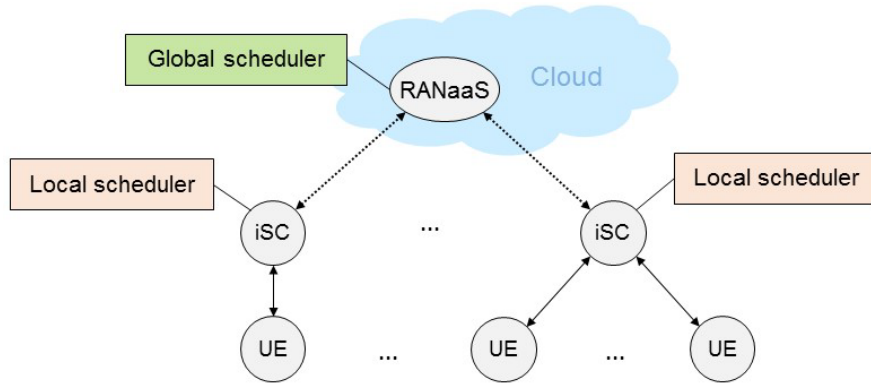


Figure 4-16: Semi-deterministic, hierarchical multi-stage scheduling.

4.4.2 Evaluation of the CT

This section presents more details on the actual implementation of the CT w.r.t. the common scenarios in iJOIN [9]. We focus on two common scenarios (CSs), *CS2 - Square* as well as *CS3 - Wide-area continuous coverage*. Furthermore, we use functional split option C.

For the basic model describing assumptions regarding channel characteristics, CSI impairments, achievable data rates and throughput, we refer to the [2] as well as to [17], [25], and [26]. Beyond that, we extend the system model to multi-cell setups.

Considering M cells (iSC or macro cell base station), K users are dropped within the cooperation area. For a detailed description regarding cell and user locations, see [2], [9]. The channel between cell m and user k present at time slot n is denoted by $h_{k,m}[n] \sim N_{\mathbb{C}}(0, \lambda_{k,m})$, where $\lambda_{k,m}$ denotes the mean channel gain. Each user k is assigned to the base station A_k with the largest mean channel gain. Only base station A_k is allowed to transmit the requested data to user k . The other base stations can basically transmit data to their own users at the same time slot n , which causes inter-cell interference and leads to the achievable data rate

$$R_k[n] = \log_2 \left(1 + \frac{\rho_m |h_{k,m}[n]|^2}{\sigma^2 + \sum_{l \in M_n, l \neq m} \rho_l |h_{k,l}[n]|^2} \right), \quad (4.13)$$

where M_n denotes the set of base stations which are transmitting at time slot n . Furthermore, ρ_m is the transmit power of base station m and σ^2 is the noise power at the receiver.

Hierarchical Scheduling Approach

Based on the achievable rates in (4.13) the global scheduler can select the base stations, which are transmitting in time slot n . This selection defines the set M_n . Beyond that, also the users, which are served by the selected base stations, are chosen by the scheduler. Both decisions are coupled. The set of all users, which are served at time slot n is denoted by K_n . The overall achievable rate at time slot n is then given by

$$R_k[n] = \sum_{l \in K_n} R_k[n] \quad (4.14)$$

Knowledge of the channels between all base stations and all users is only available at the global scheduler. Consequently, coordinated scheduling can only be performed at the central entity, the RANaaS. However, the selection of the served user can be performed locally at the iSCs. Therefore, we propose the following hierarchical scheduling method:

- i. The global scheduler decides which base stations are transmitting at a particular time slot n .
- ii. The local scheduler selects the users which are served at time slot n , considering the global scheduling pattern.

As decision objective function we assume the maximization of the overall achievable data rate $R[n]$.

Performance Results

We evaluated the proposed scheduling schemes based on simulations w.r.t. the Common Scenarios 2 and 3 (see [9]). In this section we focus on *CS3 - Wide-area continuous coverage*.

Figure 4-17 shows the average channel uncertainty of the channel from the serving cell (the best mean channel gain) and from the worst cell (the worst mean channel gain), assuming a user velocity of 3 km/h (about 40 ms coherence time). The plot indicated substantial differences in the CSI quality. Consequently, without any backhaul delay the CSI from the cell with the worst mean channel gain has similar quality as the CSI of the serving cell when a backhaul delay of 100 ms (about twice the coherence time) is present.

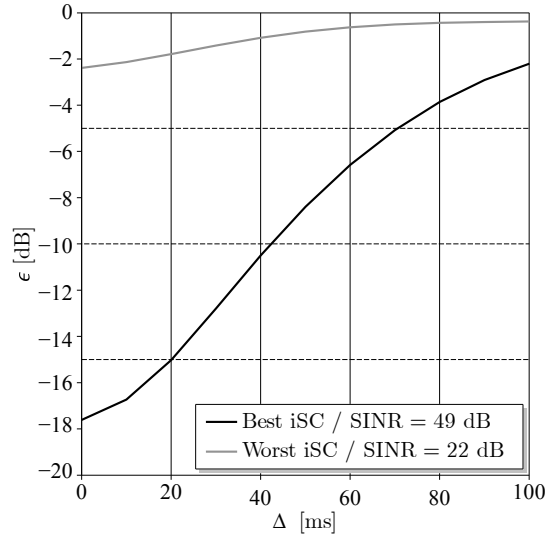


Figure 4-17: Channel uncertainty as a function of the backhaul delay.

Figure 4-18 illustrates the area throughput performance as a function of the backhaul delay. The base-line is represented by Round Robin scheduling in a reuse one system (grey line), which refers to common LTE systems.

The proposed scheduling scheme is based on max rate scheduling (opportunistic scheduling). It is evaluated for a reuse one system as well as for coordinated scheduling (CS), as discussed in the previous section. The scheduling is performed based on imperfect CSI at the central entity, represented by the RANaaS according to the iJOIN architecture. The centralized scheduling decisions are affected by outdated CSI due to backhaul latency. Consequently, the area throughput decreases with the delay.

With the proposed multi-stage scheduling, the iSCs can update their local user selection, without changing the global scheduling pattern (transmitting and non-transmitting iSCs). For higher delays, the area throughput loss can be reduced (red line), since local updates are based on more accurate CSI, than available at the RANaaS.

In addition, Figure 4-19 shows the area throughput for a user velocity of 30 km/h. It can be observed that the behaviour is basically similar to the one for a user velocity of 3 km/h but shifted towards smaller latencies. Consequently, for larger user velocities smaller backhaul delays are required for keeping the performance constant.

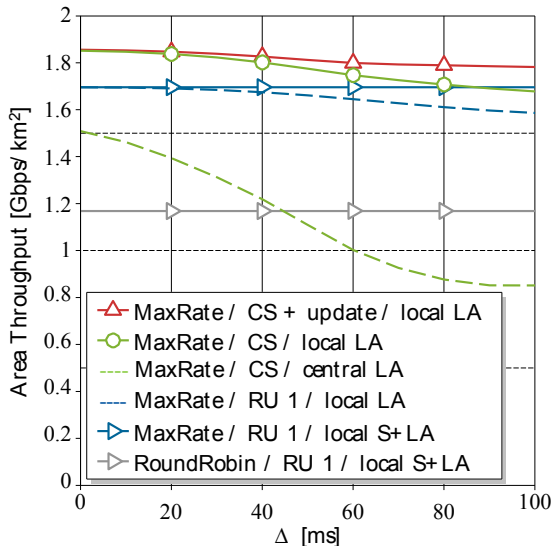


Figure 4-18: Area throughput as a function of the backhaul delay for a user velocity of 3 km/h.

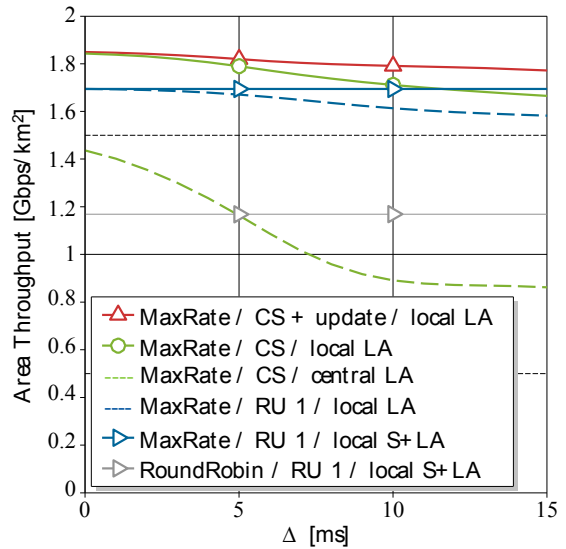


Figure 4-19: Area throughput as a function of the backhaul delay for a user velocity of 30 km/h.

4.5 CT 3.5: Cooperative RRM for Inter-Cell Interference Coordination in RANaaS

In this CT, we consider that dense deployed iSCs are inter-connected through a backhaul, which is characterized by limited capacity and non-negligible latency. The backhaul network enables cooperation among iSCs in order to employ ICIC mechanisms which can improve transmission robustness and maximize the network capacity by mitigating the high levels of co-channel interference arising in such a dense deployment (better dynamic reuse across the cells in a networked small-cell environment). Moreover, the cellular network can exploit the iJOIN RANaaS architecture to flexibly implement ICIC functionalities either in a centralized or a distributed fashion. The system model is illustrated at the following figure.

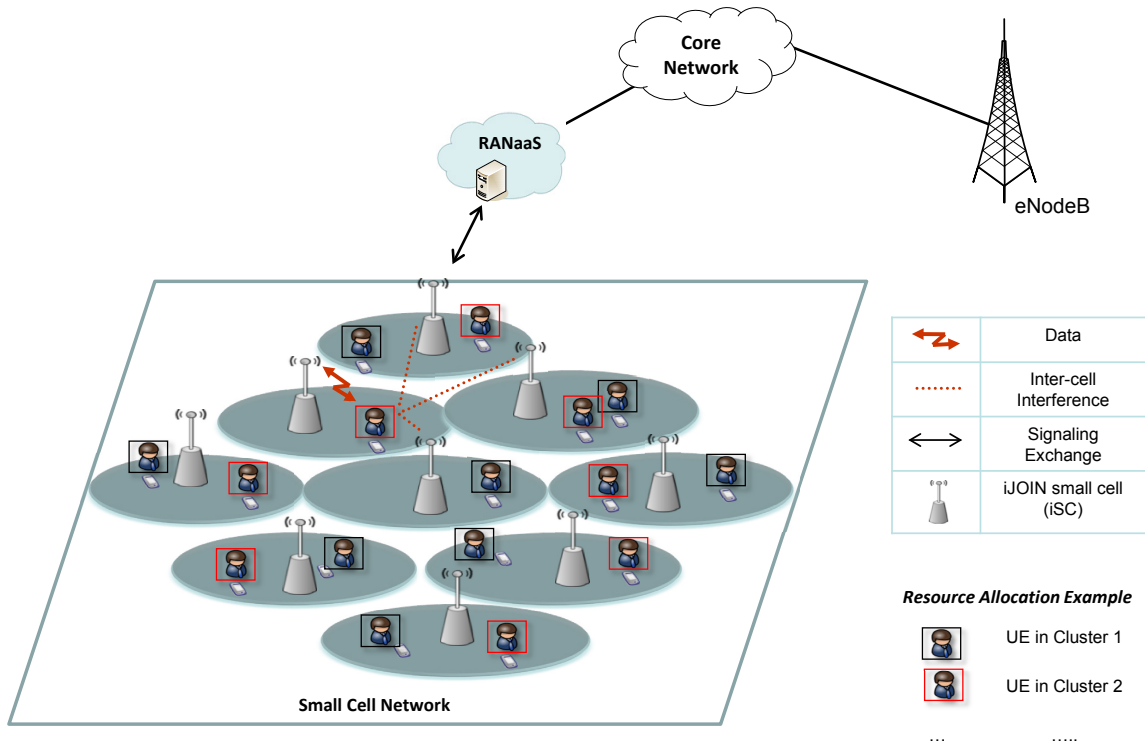


Figure 4-20: Inter-cell interference coordination between iSC.

4.5.1 Final Implementation of CT

In D3.2 [2], we formulated the non-convex weighted sum-rate optimisation problem of the resource allocation (subcarrier and power control) considering also fairness constraints and proposed a graph-based framework which efficiently solves this problem. This framework involved a locally-centralized graph-based Inter-cell Interference Coordination (ICIC) via user partitioning across different clusters and the formulation of the resource allocation policy as weighted sum rate maximization (WSRM) to optimize system performance in terms of both throughput and fairness. Then, we presented results evaluating the average cell spectral efficiency for an indoor dense small cell deployment.

Compared to [2], here we provide an evaluation study for the Wide Area scenario as defined in D5.2 [9]. To this end, we obtain realistic performance evaluations regarding the area throughput gain of the CT compared to the single-cell Proportional Fair (PF) scheduling case. In addition, we introduce the per-user spectral efficiency metric which captures how the user's performance is affected by the proposed scheme.

4.5.2 Evaluation of the CT

In addition to the description and evaluation of this CT in D3.2 [2] for an indoor dense small cell deployment, here we provide an evaluation study for the Wide Area scenario as defined in D5.2 [9]. For the evaluation, Monte Carlo system level simulations were performed for the topology of 19 iSCs.

As can be seen in the following figure the evaluation was done for the interior 7 iSCs to ensure similar interference conditions on all cells. In particular 4 users were randomly dropped each snapshot (5000 snapshots were tested) at the area covered by each of the 7 iSCs (regular deployment).

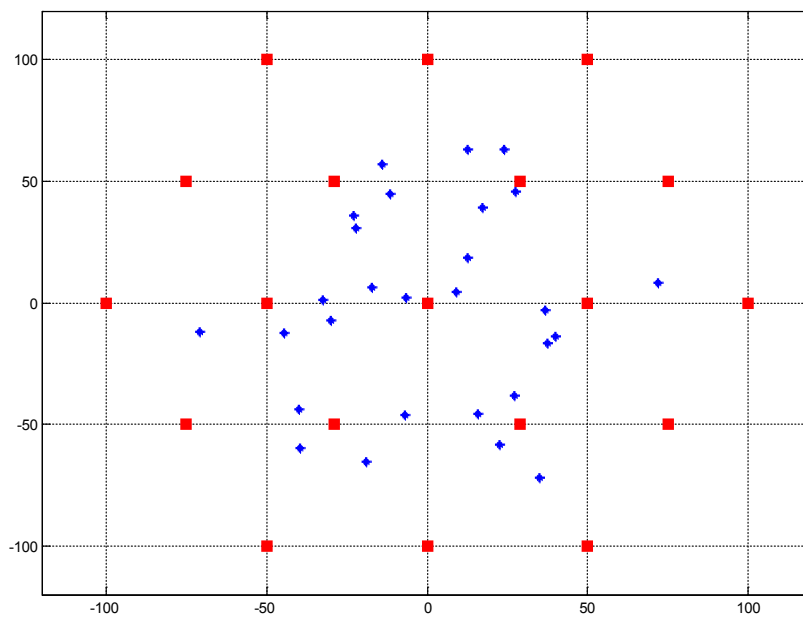


Figure 4-21 Topology – wide Area scenario.

The metrics to be evaluated are: the CDF of Area Throughput (b/s/km^2) which captures the fluctuation of area throughput as observed in different snapshots and the CDF of user Spectral Efficiency (b/s/Hz/user) which shows the per-user achievable rate.

The benchmark in this study is the case where no coordination between iSCs is performed. Hence, each iSC reuses all the available resources using single-cell Proportional Fair (PF) scheduling. Here to mention that a more detailed comparison using numerous benchmarks from the literature was performed for the indoor dense scenario and can be found in [2].

In Figure 4-22, the area throughput is compared against the full interference case with no resource partitioning between the iSCs. Targeting the mean and the median of the CDF curves, we observe a 5% increase at the area throughput using the proposed scheme. This shows that even if the resource utilization per cell is less than reuse-1, the per-cell throughput is still high comparing to the reuse-1 case.

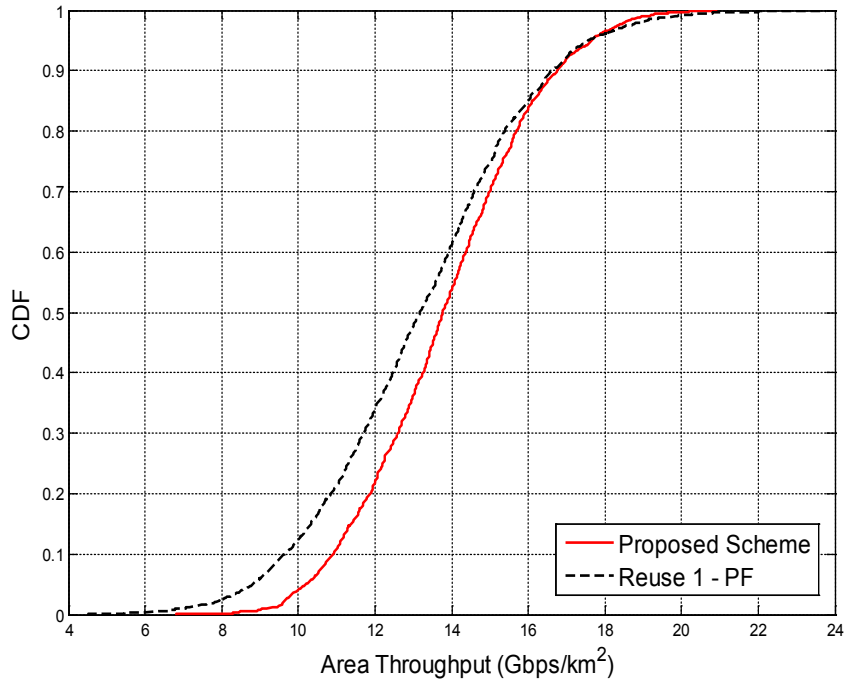


Figure 4-22 Area Throughput comparison

Another metric which captures how the user’s performance is affected by our scheme is the per-user spectral efficiency. In our proposed scheme, the number of users which are allocated resources per snapshot is lower than the Reuse-1 scenario, since our scheme dynamically changes the reuse factor per cell. However, these users as can be seen in Figure 4-23 achieve much higher spectral efficiency due to the fact that the interference from the surrounding cells is lower.



Figure 4-23 User Spectral Efficiency comparison

Targeting the 10th percentile of the CDF, which reflects to the spectral efficiency improvement for the cell-edge users, our proposed scheme shows more than 200% improvement over Reuse-1 with PF scheduling. Additionally, the median of the CDF shows around 150% enhancement and the 90th percentile around 21% gains over Reuse-1.

Therefore, Figure 4-22 and Figure 4-23 show that our proposal outperforms the Reuse-1 (PF) case in both cell/area throughput and user spectral efficiency. For the area throughput, the gain is marginal due to the fact

that the number of resources allocated per cell is lower (on average 70% resource utilization per iSC); hence it might happen that in some snapshots some iSCs utilize a small subset of the shared spectrum. Nevertheless, the throughput per allocated user can be much higher since the interference from other iSCs for each resource block can be lower.

4.6 CT 3.6: Utilization and Energy Efficiency

This section introduces the key performance indicators Utilization Efficiency (UEff) and Energy Efficiency (EEff), which are of particular relevance for the iJOIN system. This section supersedes and complements the corresponding Sections 4.6 in D3.1 [1] and D3.2 [2]. In addition, this section introduces a novel framework which allows for analysing utilisation efficiency in RANaaS instances. Finally, we introduce a joint scheduling method for wireless and data centre resources which allows for utilising RANaaS resource efficiently.

Utilization efficiency is defined as a metric expressing how well the utilized resources are employed to achieve a given performance metric. Therefore, high UEff means the following:

- The system (such as a network) is highly utilized, and therefore not over-provisioned.
- The system is capable to exploit utilized resources efficiently to provide the desired output, such as cell throughput or other targeted metrics.

Measurements in operator networks reveal [20] that 20% of all base stations carry 50% of the overall traffic, meaning that the average utilization ratio is less than 40%. The main reason for this phenomenon is a wide deployment of macro-cells to achieve a high coverage, and the network dimensioning trimmed to peak traffic demands, meaning that a large fraction of deployed resources are underutilized. iJOIN aims at increasing this utilization by means of its two technology pillars, i.e., RANaaS decentralisation and joint RAN/backhaul design.

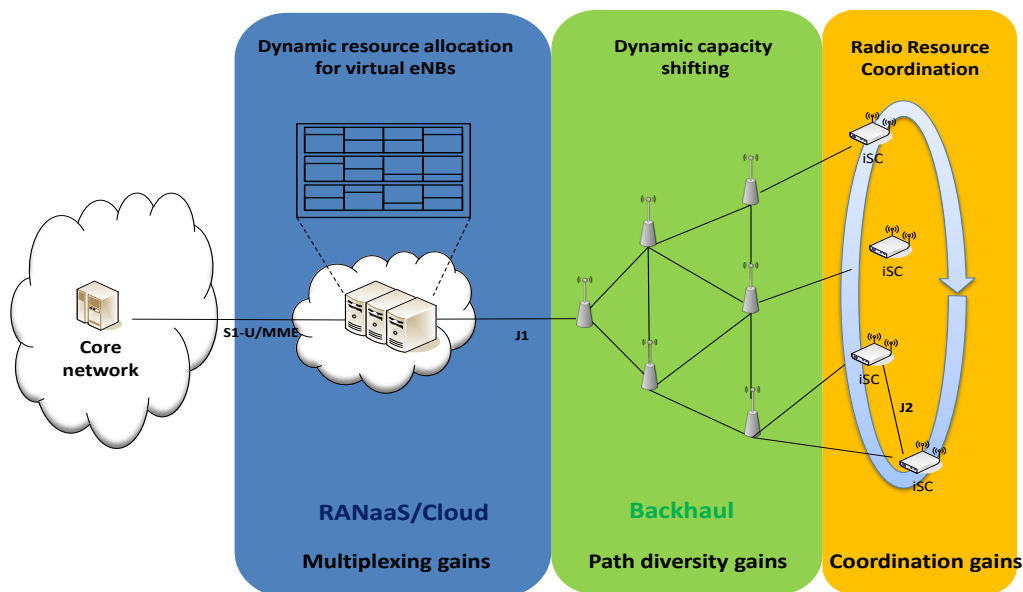


Figure 4-24: Utilization gains in different network domains

Figure 4-24 shows an example of how different resource allocation techniques in different iJOIN network domains can lead to different types of gains (e.g. multiplexing, diversity and coordination gains). It also illustrates a fundamental problem of defining a network-wide metric for utilization efficiency: different network domains (i.e. RANaaS, backhaul, radio access) utilize different types of resources (e.g. CPU cycles, link bandwidth, radio spectrum), such that a simple summation of domain-specific metrics is in general not possible.

At the same time, **Energy efficiency** and sustainability of 5G networks have recently received significant attention from mobile operators, vendors and research projects [21]. Figure 4-25 shows our vision of the EEff evolution in mobile networks toward a sustainable 5G [15], where the exponential mobile traffic growth toward 2020 (blue curve) goes with a stable network energy consumption (red curve), resulting in an increasing EE of the system through the years (green curve).

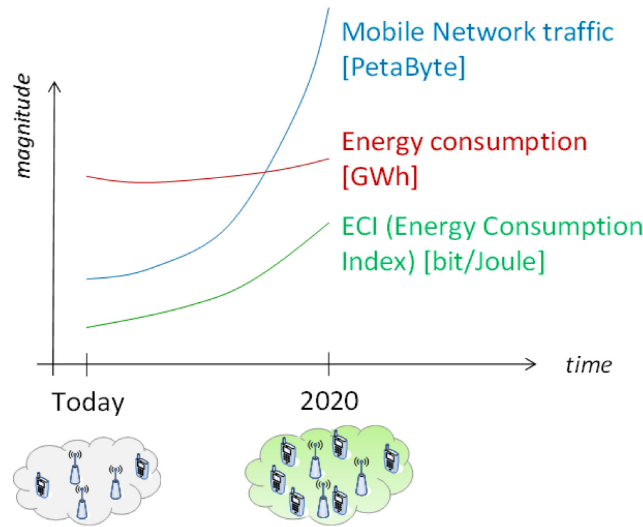


Figure 4-25: Energy efficiency evolution in mobile networks toward a sustainable 5G.

4.6.1 Final definition of utilization efficiency metrics

We define the total UEff of a system as following:

$$\eta_U = \sum_{d \in D} \alpha_d u_d \quad (4.15)$$

where α_d is a scaling factor s.t. $\sum \alpha_d = 1$, and u_d is the *domain utilization* for the considered domain, with D defined as the set of network domains (e.g. RANaaS, backhaul, RAN).

The definition of the domain utilization u_d depends on the resource of interest. As described in (4.15), different network domains have in many cases different resources. However, on a more abstract level, resource normalization can be applied across network domains. We identified the following resource classes which will be investigated in more detail:

- Bandwidth/capacity resources. The domain utilization is defined as

$$u_d^B(X) = \frac{B_{mean,d}(X)}{B_{cap,d}(X)}, \quad (4.16)$$

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 depends on the investigated network scenario and can be the number of cells, user arrival rate, etc.

- Computational resources. Here, the domain utilization is defined by

$$u_d^C(X) = \frac{C_{mean,d}(X)}{C_{outage,d}(X)}, \quad (4.17)$$

where $u_d^C(X)$ is the ratio of expected computational demand and provided computational resources, depending on the number of cells in the scenario, X . The latter is the outage complexity which is defined as the amount of computational resources to make sure that a per-cell computational outage ε is not exceeded. Both are defined through an analytical framework which is described in [22]. This framework resembles the characteristics of computational load of a 3GPP LTE uplink decoder.

4.6.2 Evaluation of RANaaS Utilization Efficiency

4.6.2.1 Analytic evaluation of the computational utilization efficiency

For RANaaS, the computational demand of RAN function execution is of special interest since the question of feasibility of the functional shift towards the centralized network entity needs to be answered for different functional split configurations. One of the main impact factors on the computational demand is the forward error correction (FEC) in the uplink. Consequently, our analysis is based on the formulation of a complexity model for forward error correction which is described in detail in [22]: computational complexity is

measured as the product of information bits and iterations, divided by the number of transmission resources (referred to as channel uses), i. e. bit-iterations per channel use (pcu). This relationship can be formalized as

$$\mathcal{C}(\gamma_k, r_k) = \frac{r_k}{\log_2(\zeta - 1)} \left[\log_2 \left(\frac{\zeta - 2}{K(\hat{\epsilon}_{\text{channel}})\zeta} \right) - 2l_k(\gamma_k, r_k) \right], \quad (4.18)$$

where ζ is a parameter of the model related to the connectivity of the decoder when represented as a graph, and

$$K(\hat{\epsilon}_{\text{channel}}) = -\frac{K'}{\log_{10}(\hat{\epsilon}_{\text{channel}})}, \quad (4.19)$$

$$l(\gamma_k, r_k) = \log_2[\log_2(1 + \gamma_k) - r_k], \quad (4.20)$$

where K' is a parameter of the model and $\hat{\epsilon}_{\text{channel}}$ is a constraint on the channel outage probability. The set of parameters $\{K', \zeta\}$ are selected by calibrating (4.18) with an actual turbo-decoder implementation or message-passing decoder.

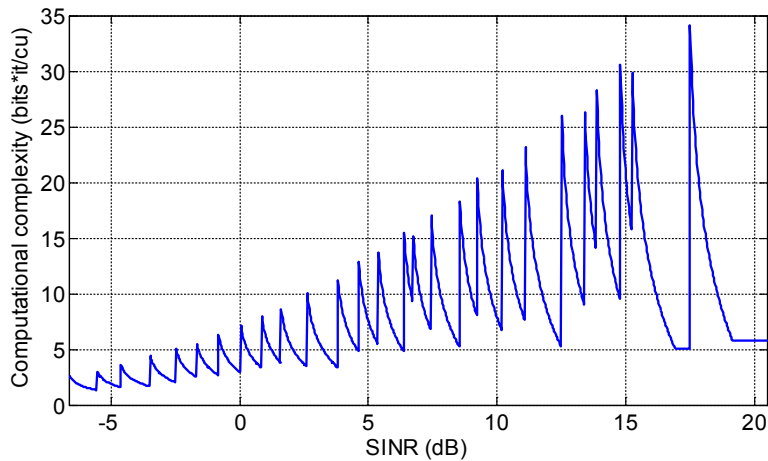


Figure 4-26: Computational effort as a function of the SINR

Figure 4-26 shows an example of the computational effort for FEC decoding versus the instantaneous SINR under the assumption of block fading. The “spiky” behaviour stems from MCS switching, and indicates that significant centralisation gains can be expected. The shown curve was generated under assumption of a target BLER of 10% on the first HARQ transmission.

For the analysis of the computational utilization efficiency, it is assumed that the RANaaS platform may support a maximum computational complexity \mathcal{C}_{max} such that $\sum_k \mathcal{C}_k \leq \mathcal{C}_{max}$. Furthermore, we define the *outage complexity* \mathcal{C}_{out} as the minimum amount of computational resources required to satisfy a target outage probability ϵ_{out} , where

$$\mathcal{C}_{out} = \arg \min_{\mathcal{C}_{max}} \left\{ P \left[\sum_k \mathcal{C}_k > \mathcal{C}_{max} \right] \leq \epsilon_{out} \right\}. \quad (4.21)$$

The analysis of the outage complexity becomes analytically tractable under the assumption of Rayleigh fading and the central limit theorem which follows from a sufficiently large number of users and centralized iSCs N_c , such that

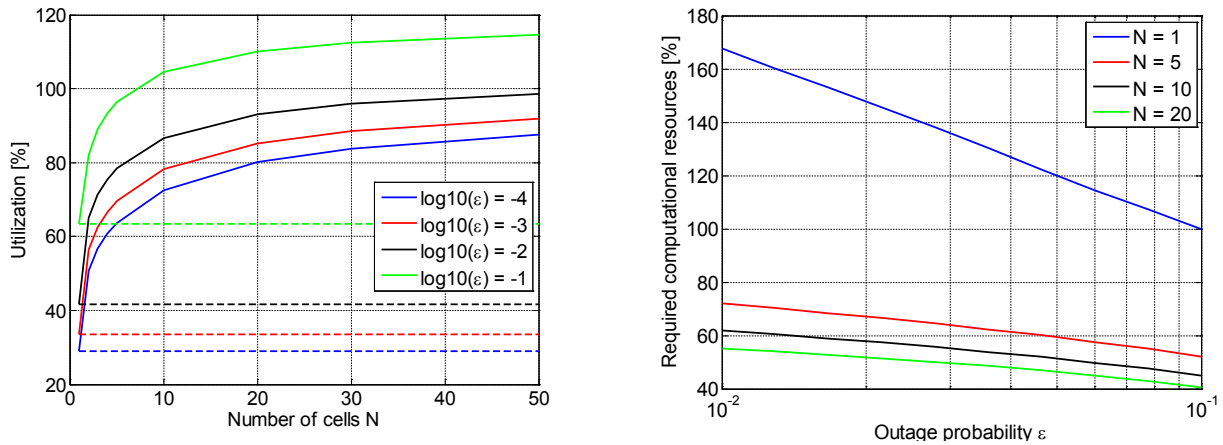
$$\mathcal{C}_{out}(\epsilon, N_c) = \sqrt{\frac{\text{Var}[\mathcal{C}_{tot}]}{N_c}} \mathcal{N}^{-1}(1 - \epsilon_{out}) + E[\mathcal{C}_{tot}], \quad (4.22)$$

where \mathcal{N}^{-1} is the inverse normal CDF, and $E[\mathcal{C}_{tot}]$ and $\text{Var}[\mathcal{C}_{tot}]$ denotes the expectation and the variance of the joint computational complexity process over all cells and UEs, which can be derived from the first moments of the per-user complexities under the assumption of statistical independence of the individual users. A detailed analysis can be found in [23].

Based on this analytical framework, we can calculate the computational utilization efficiency as defined in (4.17) as well as the required computational resources for arbitrary scenarios. For these results typical LTE parameters including actual SNR link-adaptation thresholds have been used. Furthermore, a Rayleigh fading process is assumed with an average SNR of 10dB.

Figure 4-27a) shows numerical results for the utilisation efficiency of the RANaaS instance and, as a comparison in dashed lines, for a local RAN implementation. We can see that depending on the target computational outage, a local implementation runs at 30%-60% utilisation of its processing resources (although a practical system would rather be optimised for very low computational outage resulting in utilisation below 30%). By contrast, a centralised implementation increases the utilisation to at least 90%. In the case of $\epsilon = 10\%$ we can further see that utilisation achieves values above 100%, i.e. the system would run permanently at full load.

Correspondingly, Figure 4-27b) shows the required computational resources as percentage of the distributed case as a function of the per-cell outage probability. It can be seen that for a target outage probability of 10%, only 40% of the computational resources of a fully distributed system are required. Although a certain penalty in form of a slightly reduced system throughput results from this outage probability we will show in the next section how to mitigate these effects.



a) Utilization efficiency as a function of the number of cells b) Required computational resources as a function of the per-cell outage probability

Figure 4-27: Computational utilization efficiency

4.6.2.2 Characteristics of computational complexity in large scale scenarios

In the following, we provide a short overview of the characteristics of the computational complexity by means of a system-level simulator. We employ a calibrated system-level simulator compliant with 3GPP requirements, where channel fading traces are obtained in a 3-tier, wrap-around hexagonal 3-sectored layout with the IMT-Advanced spatial channel model [17], heterogeneous network deployment with clustered small cell and mobility/hand-over modelling. The computational complexity demand is calculated with a link-level implementation of the LTE turbo decoder and rate matching algorithm for an error rate of 0.1 for the first transport block transmission in the HARQ protocol. The evaluation scenario corresponds to the “square” common scenario defined in [9], depending on the parameterization. Note that open loop uplink power control as specified in [24] and Round-Robin scheduling is implemented.

The corresponding histogram and CDF of normalized computational complexities are shown in Figure 4-28. Due to the high SINR advantage of UEs attached to iSCs in downlink, they have a high chance that the highest MCS is selected and decoding is correct without retransmission and deep iteration by the turbo decoder. The strong peak at computational complexity value of 6bit-iter/cu in Figure 4-28a) is caused by this phenomenon, corresponding to an SINR of 18dB or higher in Figure 4-26.

The main direction of interest for computational complexity is the uplink, since otherwise decoding takes place in the UEs which are not subject to centralization. Here it can be observed that micro BSs have a higher demand for computational complexity than macro-cells. The reason is that on average, the computational complexity tends to increase with the SINR until a certain maximum value, as also shown in Figure 4-26.

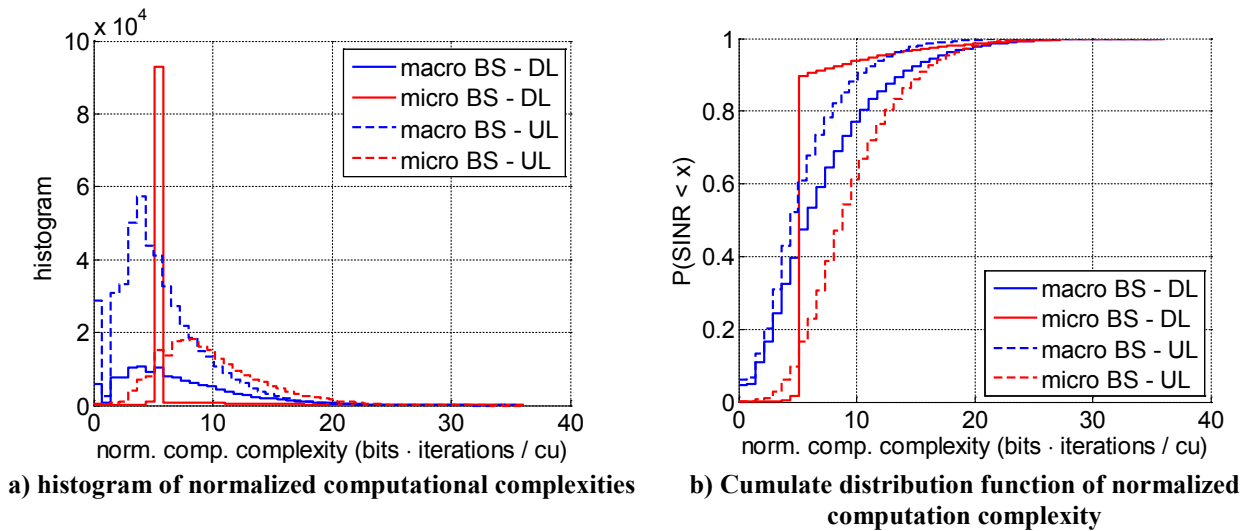


Figure 4-28: Distribution of per-subframe normalized computational complexities

From the perspective of UEff the characteristics of the computational complexity (CC) function is not beneficial, as it indicates that to avoid computational outages, significant over-provisioning of resources is necessary. It is therefore interesting to investigate the impact of resource management, centralization and the corresponding multiplexing gains.

Figure 4-29 shows a trace of the total cell computational complexity, and the corresponding number of UEs in the cell. It can be observed that a dependency exists; trivially in case if there are no UEs in the cell, but also e.g. at time index $3.73 \cdot 10^4$, where three UEs arrive at a cell. However, a strong correlation between the number of UEs and complexity cannot be necessarily concluded. The reason is that the main impact factor is the SINR which leads to potentially very strong changes of the CC function with small changes by its value.

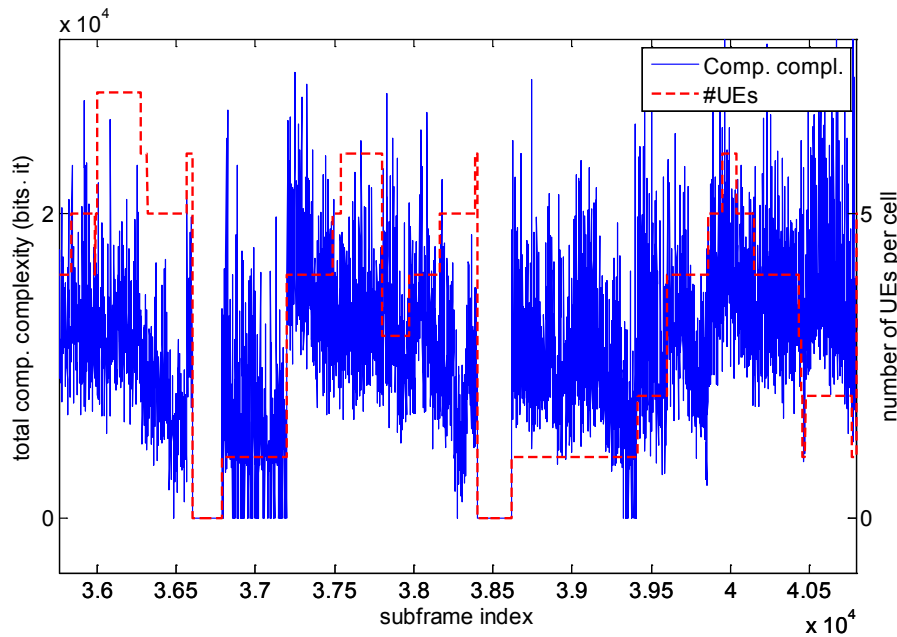


Figure 4-29: Trace of total computational complexity and number of attached UEs per cell (macro cells).

This is further illustrated in Figure 4-30, which shows the mean per-cell CC depending on the number of transmitting UEs. A correlation between the mean CC and the UE/cell density is not observable; the delta between different communication directions (UL/DL) for the same cell type stems from UL and DL having different mean value of SINR.

This result implies that from cell view, multiplexing gains may origin from a reduction of the dispersion of the total CC per-slot.

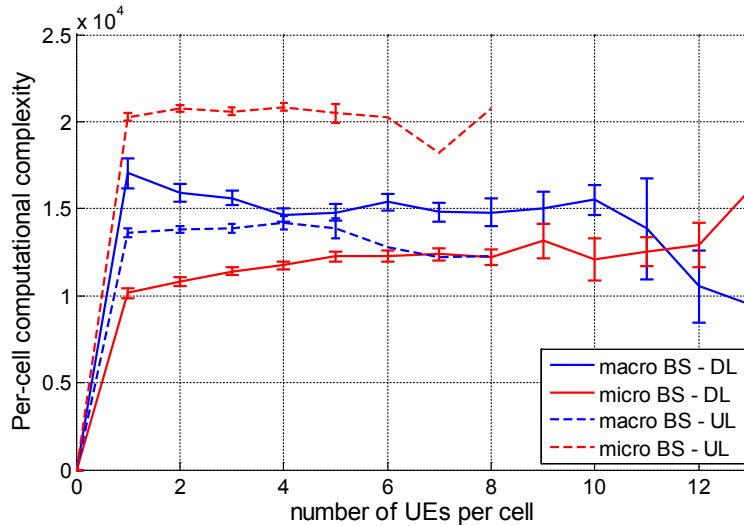


Figure 4-30: Per-cell computational complexity vs. number of attached UEs.

4.6.2.3 Resource allocation with computational complexity constraints

Using the framework presented above, it is possible to determine the required processing resources that ensure the computational outage probability does not exceed ϵ . Instead of accepting these computational outages, we can also apply joint scheduling of RAN and data processing resources. Hence, if the required processing resources exceed the available processing resources, the resource scheduler will adjust the radio resource assignment in order to match the processing resources which are required and available.

We divide the derivation of the resource allocation strategy into two steps. First, we assume that each iSC serves exactly one UE within its cell. Based on this assumption, we derive a strategy to determine the rate r_k and the associated computational complexity \mathcal{C}_k . Then, we show how this strategy is applied in a more practical multi-user scenario.

The objective is to find a scheduling metric which allows for maximizing the sum-rate of the system while avoiding any computational outage. This objective can be formalized by the following optimization problem:

$$\begin{aligned}
 r^* &= \max_{\forall \mathcal{R}: \mathcal{R} \in \mathcal{R}^*} \sum_{r_k \in \mathcal{R}} r_k, \\
 \text{s. t. } &\sum_{r_k \in \mathcal{R}} \mathcal{C}_k \leq \mathcal{C}_{\max} \wedge \forall k: r_k \geq 0.
 \end{aligned} \tag{4.23}$$

This scheduling problem implies a trade-off between the invested computational complexity and the gained achievable rates. Hence, for solving the problem it is important to evaluate whether it is “worth” to invest additional complexity into the processing of a code-word. This requires the derivative $\partial \mathcal{C}_k / \partial r$ which determines the increase in computational complexity for a given increase of data rate. However, the derivative of \mathcal{C}_k is rather complex and difficult to apply in the optimization problem. Therefore, we use the following linearization of $l(\gamma_k, r_k)$:

$$l(\gamma_k, r_k) \approx a_k r_k + b_k, \tag{4.24}$$

using parameters

$$\begin{aligned}
 a_k &= \left. \frac{\partial l(\gamma_k, r_k)}{\partial r} \right|_{r_k} = \frac{-1}{\log(2)(\log_2(1 + \gamma_k) - r_k)}, \\
 b_k &= l(\gamma_k, r_k) - a_k r_k.
 \end{aligned} \tag{4.25}$$

Using this linearization, we can rephrase (4.18) as

$$\begin{aligned}
 \mathcal{C}_k &\approx r_k \left(\underbrace{-\frac{2a_k}{\log_2(\zeta - 1)}}_{=\alpha_k} r_k + \underbrace{\left(\frac{1}{\log_2(\zeta - 1)} \left[\log_2 \left(\frac{\zeta - 2}{K(\hat{\epsilon}_{\text{channel}})\zeta} \right) - 2b_k \right] \right)}_{=\beta_k} \right) \\
 &= \alpha_k r_k^2 + \beta_k r_k
 \end{aligned} \tag{4.26}$$

The solution to the rate allocation problem in (4.23) can then be well approximated by the following theorem:

$$r_k = \frac{1}{2\alpha_k} \left(\frac{1}{\eta} - \beta_k \right)^+ , \quad (4.27)$$

where $0 \leq \eta \leq 1/\beta_k$ and $\sum_{r_k \in \mathcal{R}} C_k \leq C_{\max}$. The proof follows similar arguments and methodology as for the well-known water-filling solution of the power allocation problem. More details are given in Appendix 6 of [23].

The structure of the solution (4.27) is similar to the water-filling solution that is known for the power-allocation problem. The parameter $1/\eta$ determines the water-level which decides whether a UE is served or not. With β_k , the actual cost (in terms of complexity) of transmitting to a UE are determined. If the difference between the selected rate r_k and the capacity $\log(1 + \gamma_k)$ is small, then the term β_k becomes very large (due to the large slope of α_k). Hence, the UE is unlikely to be served. The parameter α_k scales the rate, i. e. if again the rate r_k operates close to capacity, also α_k will be very large and therefore will scale down the assigned data rate in order to reduce the necessary computational complexity. Based on these findings, two computational complexity aware resource allocation schemes have been developed: scheduling with water-filling (SWF), and scheduling with complexity-cut-off (SCC).

4.6.2.4 Scheduling with water-filling (SWF)

Assume that the link-adaptation processes selects an MCS with associated rate $r_k \leq \log(1 + \gamma_k)$. This implies that

$$\frac{1}{2\alpha} \left(\frac{1}{\eta} - \beta_k \right) \geq \sqrt{\frac{C_k}{\alpha} + \left(\frac{\beta_k}{2\alpha} \right)^2} - \frac{\beta_k}{2\alpha}, \quad (4.28)$$

which follows from (4.27) (left hand side) and (4.26) (right hand side solved for r_k), assuming that there is no constraint on computational complexity. This yields

$$\frac{1}{\eta} \geq \sqrt{4\alpha C_k + \beta_k^2} \quad (4.29)$$

which gives us the required 'water-level' for each user. The basic algorithm is outlined below:

Algorithm 4-1: Scheduling with water-filling

- 1.) Initialization: Set \mathcal{R} such that each user k receives the maximum possible rate r_k .
- 2.) While $\sum_{r_k \in \mathcal{R}} C_k > C_{\max}$ do
 - a) Compute $k^* = \arg \max_k \sqrt{4\alpha C_k + \beta_k^2}$
 - b) Reduce the MCS of k^* by one such that $r'_{k^*} < r_{k^*}$
- 3.) End While

4.6.2.5 Scheduling with complexity cut-off (SCC)

The previously described metric still requires the linearization described in (4.26) in order to determine parameters α_k and β_k . In this section, we introduce a slightly simpler and intuitive metric. Instead of applying (4.29), we select the users with the highest complexity and reduce their MCS (and associated rate r_k) until the sum-complexity constraint C_{\max} is fulfilled. In this case, an iterative method can be used as follows:

Algorithm 4-2: Scheduling with complexity cut-off

- 1.) Initialization: Set \mathcal{R} such that each user k receives the maximum possible rate r_k .
- 2.) While $\sum_{r_k \in \mathcal{R}} \mathcal{C}_k > \mathcal{C}_{max}$ do
 - a) Select the user k^* with the highest complexity:
 $k^* = \arg \max_k \mathcal{C}_k$
 - b) Reduce the MCS of k^* by one such that $r'_{k^*} < r_{k^*}$
- 3.) End While

As the result of this procedure, we obtain a rate allocation which satisfies the complexity constraint and always reduces the rate for users which incur the highest complexity. The main difference to the previously discussed metric is that this metric is not necessarily sum-rate optimal. However, it has lower complexity and is very efficient as we will discuss in the next section. In the following, we refer to this empirical metric as SCC.

4.6.2.6 Numerical results

The evaluation was conducted with a 3GPP compliant system-level simulator in a wide-area scenario. The details of the evaluation parameters are shown in Table 4-6.

Table 4-6: CT 3.6 Simulation parameters

Parameter	Value
Spatial distribution of users	PPP
Density of UE per unit area	$\lambda = 1$ UEs/Km ²
Path loss exponent	$\alpha = 3.7$
Number of centrally processed iSCs	$N_c = 10$
Computational outage	$\epsilon_{out} = [10, 1, 0.1]\%$
Channel outage constraint	$\hat{\epsilon}_{channel} = [10]\%$
Fading	Rayleigh
Fractional Power-Control Factor	$s=0.1$
Transmit power	$P_0 = 10$ W
Noise power	$W = 100$ mW
Simulation trials	$N_{trials} = 10^7$

Figure 4-31 shows the CDFs of the computational complexity, and Figure 4-32 of the corresponding achieved sum-rate over $N_c = 10$ iSCs. The figures show the curves for both SWF and SCC, when the system is designed such that $\epsilon_{out} = [10\%, 1\%, 0.1\%]$ computational outage holds (the notches in the magnification show the corresponding value of \mathcal{C}_{max}). The figures also show as a benchmark the curve for the unconstrained case, which selects the maximum possible rate that can be used since there is no computational constraint.

It can be observed that a stronger constraint on the available computational resources implies higher computational outage. More importantly, the figures highlight that while the required computational complexity is significantly reduced by dimensioning the system to an higher computational outage, the sum-rate only decreases slightly for both SWF and SCC, i. e. the average sum-rate only decreases by $\approx 0.28\%$ for $\epsilon_{out} = 10\%$ and by $\approx 0.07\%$ for $\epsilon_{out} = 0.1\%$.

This shows the efficiency of the proposed schedulers, which impact the achievable sum-rate only marginally, while they reduce the required computational resources significantly. Furthermore, Figure 4-31 shows that the introduced schedulers are able to completely avoid any computational outage, which would lead in the worst case to drop the connection to the UE. Even if one solution could be to dimension the system for a very low computational outage, i. e. $\epsilon_{out} = 10^{-6}$, the drawback is that the system will be significantly over-provisioned and most of the time the allocated resources are under-utilized. By contrast, the complexity-aware schedulers allow to avoid computational outage, while maintaining a high utilization efficiency.

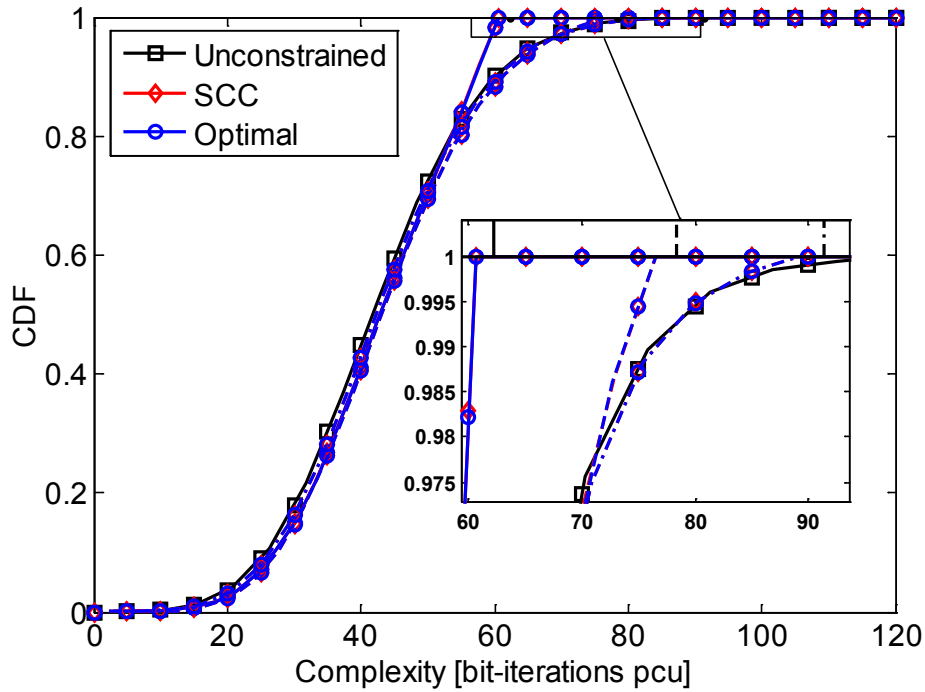


Figure 4-31: CDF of computational complexity.

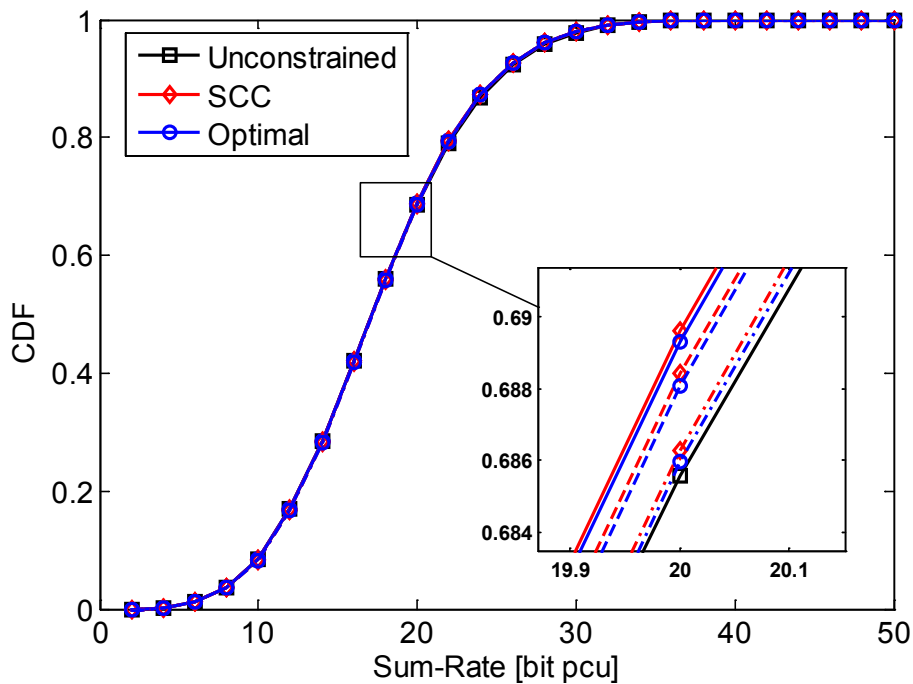


Figure 4-32: CDF of sum-rate.

Figure 4-33 shows the mean sum-rate as a function of the number of iSCs, which are centrally processed at the RANaaS platform. The performance are shown for SWF and SCC as well as for Max-Rate Scheduling (MRS), which does not account for the computational constraint. The figure shows the impact of the computational outage targets on the different schedulers. As it can be noticed, for all values of N_c , the impact of the constraint on the computational resources is marginal for the computational aware schedulers. By contrast, as ϵ_{out} increases, the impact on a system which uses a MRS increases linearly with ϵ_{out} due to the increasing computational outage. Furthermore, the magnification shows that there is only a marginal difference between SWF and the SCC, emphasizing the fact that the less complex SCC algorithm achieves almost the same performance as SWF.

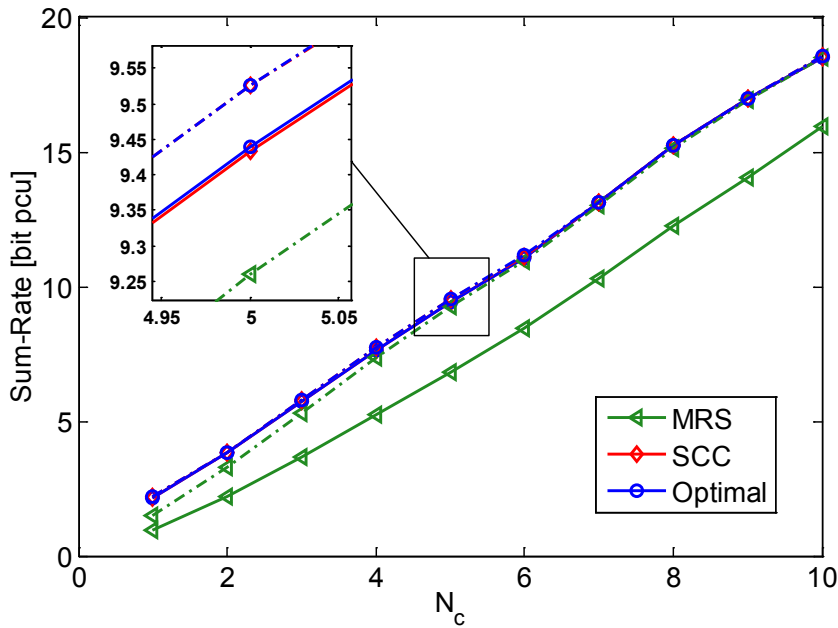


Figure 4-33: Average sum-rate vs. number of centralized iSCs.

Finally, Figure 4-34 shows the sum-rate of SWF, SCC and MRS as function of the density of UE per km² when the maximum computational resources available lead to a computational outage of 10% and 0.1% in the case of $\lambda = \text{UEs}/\text{km}^2$. It shows the ability of the schedulers to provide services to all the users while slightly reducing the system-throughput as the computational outage increases. Furthermore, it is shown that the proposed scheduling algorithms are able to accommodate the increasing traffic demand, while the MRS suffers gradually from the higher computational outage.

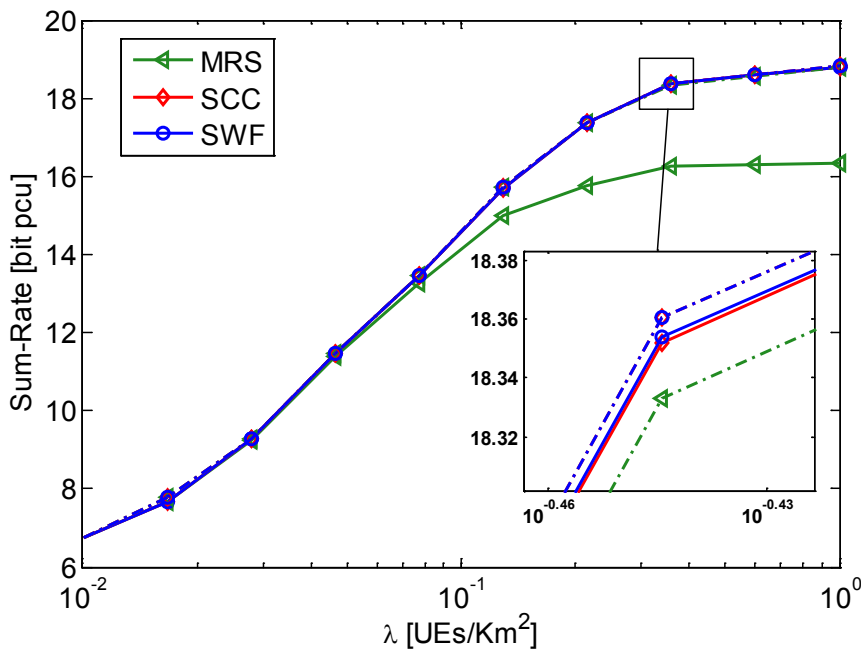


Figure 4-34: Mean sum-rate vs. user density.

In conclusion, the computational complexity of RAN functions is one of the main obstacles for the introduction of cloud computing principles into the mobile network radio access. In this CT, we have developed a framework which solves the user resource allocation problem under the assumption of limited computational resources in a centralized cloud platform, maximizing the computational utilization efficiency of the system. We showed that the underlying optimization problem can be solved with an adapted water-filling approach, making it feasible to fulfill the strict timing requirements of the wireless access (e.g., several milliseconds in LTE). Furthermore, we have shown that an intuitive complexity-cut-off approach delivers near-to-optimal results as well, even further decreasing the complexity of the resource allocation

problem. Finally, the numerical evaluation confirms that meeting computational complexity constraints does not lead to significant penalties in terms of throughput, a fact which underlines the applicability of the approach in practical systems.

4.6.3 Evaluation of RANaaS Energy efficiency

As highlighted several times in this and in past deliverables, CT 3.6 is different from the other candidate technologies. It is not a “functional” technology, i.e. a functional component implementing a step of the LTE-A processing stack. It is rather a metric-based CT, whose purpose is to help assessing the actual impacts of the iJOIN key pillars (functional split, and joint RAN-BH optimization) on the overall usage efficiency and energy efficiency system metrics. Besides, Energy Efficiency is a metric which cannot be derived from other iJOIN KPIs, whereas User Efficiency might have a relation to Area Throughput. As a matter of fact, in order to perform an evaluation of EEff in iJOIN, it must be bound to one of the “functional” candidate technologies, to allow defining an actual test scenario and to execute real measurements.

Such a real measurement campaign was not part of the iJOIN project plan, since the amount of effort and resources it would request is not in the actual scope of iJOIN, and would not provide an added value to iJOIN’s concept proofing in line with the cost of performing it. Thus, we turned to a more theoretical model for deriving an indicative figure, letting up to future research an extensive measurement and evaluation campaign to assess energy efficiency. We report anyhow in this section an outline of the steps to perform, if a real measurement procedure were to be executed.

4.6.3.1 Proposed Measurement Methodology

The measurement campaign should start from selecting one of the available CT’s, looking for the ones offering the most potentially interesting cases for the evaluation of Energy Efficiency. Then, an appropriate scenario should be selected, and a related functional split type should be applied. The comparison must be done between the same data transmission in a standard LTE mode and a RANaaS-optimized mode. The number of virtual machines instantiated in the RANaaS should duly be set, and a test configuration, defined by a certain set of parameters should be defined. At this point, the evaluation can be performed through the following steps:

- Set a baseline measurement, measuring (or estimating) the energy consumption E_1 of all the iSCs under control of the RANaaS, and of the backhaul, for a given number of packets of defined size sent from the core network over a given time interval;
- Turn the iSCs off and the RANaaS instance on;
- Activate the functional split configuration, with the RANaaS managing optimal iSC activation;
- Repeat the packet transmission in the same way, measuring the total energy consumption E_2 at the RANaaS, backhaul and iSCs.
- Compare E_2 vs E_1 ;
- Repeat the measurement with changes of some configuration parameters.

The transmission time interval should have a minimal length allowing a reliable power measurement at the RANaaS instance. Ideally, it should allow more than one single measurement, to improve the energy consumption accuracy since the actual measured value is power, not energy.

The parameters characterizing a given test configuration are:

- Power consumption profile of RANaaS, backhaul and iSCs;
- Number of small cells under the RANaaS control;
- Transmission time interval;
- Packet transmission mode (regular vs bursts);
- Packet size;
- Packet transmission frequency (number of transmitted packets per time unit).

The same test can be replicated by changing one of these parameters at a time, and finding significant parameter combinations. For instance:

- Changing the number of available iSCs, to assess the improvement of efficiency when less iSCs are utilized inside the requested QoS boundaries;
- The transmission time interval length should not be relevant, except for improving the energy consumption computation accuracy;
- Packet transmission mode: burst transmission could cause more iSCs switch on/off than regular transmissions, hence decrease the efficiency gain;
- Packet transmission frequency: more packets should cause more iSCs switch on/off than regular transmissions, hence decrease the efficiency gain;
- Packet size could have a similar effect, and in the end not cause different effects from the previous ones;
- Cross-variation of number of iSCs and packet transmission frequency or mode could give interesting figures.

An evaluation of RANaaS power profile impact can be done theoretically, if no different server classes are available in the testbeds. Also, backhaul changes, wherever compliant with the chosen physical scenario, can be evaluated, however their impact should be limited to the energy consumption of the backhaul itself, and not have an effect on the combined energy efficiency of the whole system.

4.6.3.2 Theoretical Results

Exemplary results have been derived to have a first general appraisal of what can be expected by applying iJOIN principles in terms of energy efficiency. The basis for this theoretical evaluation has been described in detail in [15], to which the reader is referred for more details about the foundation of made assumptions and all the details of derived formulas.

The steps for this evaluation are essentially the same described in the former subsection, with the only difference that numerical values are not measured on a real physical demonstrator; they are instead computed by theoretical formulas, whose derivation and explanation is available in the referenced paper.

The selected candidate technology for this exemplary evaluation was CT 2.2 (Multi-Processing Turbo Decoder). As physical scenario, the Stadium case was picked, with the application of a type-A functional split. The benchmark configuration includes a number of 10 iSCs, whose BB processing functions, in case of type-A functional split, are all fully centralized and shifted into the RANaaS layer.

We estimate the power consumption of this configuration in case of traditional (iJOIN-less) processing. In this case, the consumption of each small cell is computed through the formula (4) provided in [15]:

$$P_{iSC-n} = \frac{N_{ant} \frac{W}{10[MHz]} (P'_{BB} + P'_{RF}) + y_n P_{PA-max}}{(1 - \sigma_{DC}) (1 - \sigma_{MC})}$$

As configuration parameters, in line with the referenced paper, we make the following assumptions:

Table 4-7: Parameters for the energy efficiency evaluation of the baseline system.

Parameter	Value
Type of cell	Baseline small cell (all BB processing in the iSC)
Cell load (y_n)	10%
Maximum TX power per antenna	30 dBm
Number of antennas in the iSC (N_{ant})	2
System bandwidth (W)	10 MHz
BB engine base consumption (P'_{BB})	6.8 W for baseline small cells
RF transceiver base consumption (P'_{RF})	1.5 W for TX power = 30 dBm
Power amplifier maximum consumption (P_{PA-max})	3.2 W for TX power = 30 dBm
DC-DC conversion losses (σ_{DC})	6.4%
Main supply losses (σ_{MC})	7.7%

With these numerical assumptions, the overall estimated power consumption of iSCs is:

$$P_{iSC} = 10 * \{[2 * (10/10) * (6.8 + 1.5)] + 0.1 * 3.2\} / (1 - 0.064) * (1 - 0.077) = 195.85 W$$

In case of iJOIN type-A split, all the processing, and accordingly the related power consumption, is moved from the small cells into the RANaaS layer. Still in [15], formula (9) expresses the power consumed by the RANaaS layer to perform the same amount of BB processing which was previously done by the small cells, assuming all the servers in the RANaaS have equal processing power, and the load is fully balanced among them:

$$P_{RANaaS} = \left[\frac{X(y)}{X_{Cap}} \right] \cdot \left\{ P_0^{srv} + \sum_{n=1}^{N_{iSC}} y_n \cdot \frac{W}{10 \text{ MHz}} \cdot \left(30N_{TX} + 10N_{TX}^2 + 20 \frac{e_{MSC}}{6} e_{MIMO} \right) \Delta_p^{srv} P_{max}^{srv} \right\} X_{Cap} \left[\frac{X(y)}{X_{Cap}} \right],$$

with the parameters described in Table 4-8.

Table 4-8: Parameters for exemplary energy efficiency evaluation of the iJOIN system

Parameter	Value
Number of iSCs (N^{iSC})	10
Processing demand to servers in the RANaaS - $X(y)$	1.056 GFLOPS
Processing capacity of RANaaS servers - X_{Cap}	384 GFLOPS (4 cores * 96 GFLOPS/core)
Idle power consumption of servers (P_0^{srv})	120 W
Maximum power consumption of servers - P_{max}^{srv}	215 W
Number of antennas serving a UE per iSC - N_{TX}	2
Average number of data bits per symbol per user - e_{MSC}	4/3
Average number of MIMO layers per user - e_{MIMO}	1.1
Load of cells (y_n)	Assumed 10%
System bandwidth (W)	10 MHz
Slope of the equivalent server linear power model - Δ_p^{srv}	0.44

In order to obtain the value of $X(y)$, we leverage the cost study in [19]. The server configuration analysed in this study is the same considered in our exemplary case¹. From the analytical elaboration presented in Section II.B, the conclusion is that, to support such configuration, considering the system at capacity (i.e., with the highest performance level, and highest number of decoding iterations), we obtain $X(y) = 1.056$ GFLOPS. With this last value, the obtained value for the RANaaS power consumption is:

$$P_{RANaaS} = 55.4 W$$

To make the figure realistic, it should include the actual power consumption in the cloud datacentre not directly tied to the usage of computational power (server power), but induced and requested by the actual running of computational workload on the datacentre IT equipment (servers, storage, network). This power consumption is typically given by HVAC (heating, ventilation and air conditioning) and datacentre facility (lighting, connectivity, etc.) consumption factors. A good approximation of this figure can be obtained by multiplying the IT power consumption for the PUE parameter². In a modern datacentre, PUE's typically assume values among 1.5 and 2. Taking PUE = 2, we obtain a final value of

$$P_{RANaaS} = 110.8 W$$

This figure shows a significant reduction of power consumption (more than 40%) with respect to the power consumed by the iSCs without any functional split. It is a theoretical computation, but makes an interesting statement to start reasoning about energy efficiency. To be complete, the estimation should include a measurement of backhaul power consumption difference in iJOIN and non-iJOIN case, considering that

¹ Intel Xeon 4870 10-core processors achieving 96 GFLOP/s, packaged on quad-socket boards

² For an explanation of the PUE metric, see for instance http://www.thegreengrid.org/~media/WhitePapers/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf?lang=en

functional split is theoretically increasing the bandwidth usage on control plane and accordingly the power consumption of the backhaul. However, with reasonable approximation, we can consider this potential power variation not significant in the present case.

The theoretical models above allow evaluating also the impact of changing any of the used parameters. For instance, as highlighted by [19], if we decide to minimize the number of decoding iterations, we reduce the required processing power, and the consequent RANaaS consumption can be reduced by about one third.

In conclusion, from this specific but realistically built up exemplary case, we can say that the iJOIN flexible functional split paradigm has a theoretical potential of improving the power efficiency for a RAN system, speculating on the better incremental power optimization capabilities of an industry standard datacentre with respects to a number of base stations distributed in the geographical area served by the RAN itself.

4.7 CT 3.7: Radio Resource Management for Scalable Multi-Point Turbo Detection

To increase the performance of the (edge) users in dense small cell deployment in the uplink direction, we proposed to rely on a multi-point version of the turbo detection principle [1], [2]. Indeed, the turbo detection in a multi-user context allows significant performance improvement by exchanging information (extrinsic log-likelihood ratios) between the detection stage and the decoding stage in an iterative way among the users which are transmitting on the same resources.

To deal with the created interference among the “aligned” users in the uplink, we introduce a scalable form of the turbo detection principle, where the processing is either performed centrally at the RANaaS data centre (multi-point turbo detection - MPTD) or locally within each involved iSC (single-point turbo detection – SPTD). In the first case, a high quality backhaul with low latency is required to handle the LTE HARQ strict timing in the uplink. Signal associated to each user is sent from the iSC to the RANaaS, which will perform the turbo- detection of all “aligned” users. Performance is improve due to the increase of the receive diversity (all iSCs receive antennas signals are exploited). In the second case, the processing is done locally in each iSC. With less receive diversity available at each iSC (coming only from its own receive antennas), this solution cannot achieve the same gain as the first case when no cooperation among the iSCs is considered as shown in D3.2 [2].

While the link-level evaluation of the MPTD/SPTD is performed in WP2 (CT2.2, [5]), the work carried here in WP3 aims at evaluating this approach at the system level, by also focusing on the RRM part enabling such process to take place.

4.7.1 Final Implementation of CT

In D3.2 [2], we introduced an RRM algorithm to determine which iSCs and which UEs will benefit from this advanced multi-user detection. In both SPTD/MPTD cases, this centralised RRM algorithm is always running in the RANaaS data centre, on a per-iSC demand basis. It provides the set of iSCs and UEs to be involved in the turbo processing process, thus giving a “long-term” scheduling framework for each involved iSC. The “short-term” scheduling takes place at the iSC level under this framework. To limit the complexity, we consider only two UEs to be part of a single turbo-detection process (MPTD or SPTD).

Compared to the implementation of the CT presented in D3.2 [2], the RRM algorithm has been internally a bit modified in the way users are “paired” together, but its general implementation within the final iJOIN architecture still holds, with the same Message Sequence Charts (MSCs) being used between all the logical entities. In D3.2, one user only reported the most dominant interfering iSC it could sense through RSRP. We allow the user to report all iSCs being within the threshold detection. During the search of a potential UE to pair, we can therefore request more iSC.

4.7.2 Evaluation of the CT

4.7.2.1 Scenario

We investigate a deployment scenario targeting the Airport/Shopping mall common scenario (CS4) described in D5.2 [9], with indoor dense hotspots operating on the same channel. Figure 4-35 shows the topology envisaged, where solid lines represent the minimum requirements assumed on the interface (High Quality, Medium Quality, Low Quality related to the bandwidth/latency capability of the link). High quality

links clearly refer to optical BH link, while medium ones are millimetre waves. Low quality links can be any other BH link such as xDSL or sub-6GHz technologies.

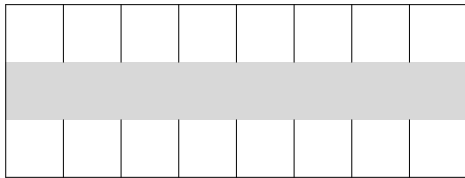


Figure 4-35: Scalable multi-point turbo detection scenario (solid lines are minimal requirements).

4.7.2.2 System Model

We use the same layout as in D3.2 [2] associated with the airport/shopping mall scenario, but with more UEs and a higher transmission power (controlled through the open loop power control parameters). Following a Monte-Carlo approach, UEs will be uniformly dropped in the area and attached to an iSC based on the best received power criteria. The UEs have full-buffer traffic to send in the uplink, with the same QoS priority. To simulate the load balancing operation, the same number of UEs will be attached to each iSC during their drop. Table 4-9 gives the iSC and UE parameters. Table 4-10 shows the main static assumptions valid for both baseline and MPTD/SPTD investigations, while Table 4-11 gives the parameters used for the dynamic system-level simulations related to the airport/shopping mall scenario (CS4).

Table 4-9: Node parameters.

iSC Parameters		
Antenna	Number	2 (Uniform Linear Array)
	Spacing	0.5λ (Wavelength associated to f_c)
	Polarization	Vertical
	Pattern	Omnidirectional
	Gain	0dBi
	Height	6m
Transmit Power		24dBm
Noise Figure		5dB
UE Parameters		
Antenna	Number	1
	Polarization	Vertical
	Pattern	Omnidirectional
	Gain	0dBi
	Height	1.5m

Transmit Power	Maximum	23dBm
	Minimum	-40dBm
Noise Figure		7dB

Table 4-10: Indoor Hotspot system level simulation static parameters.

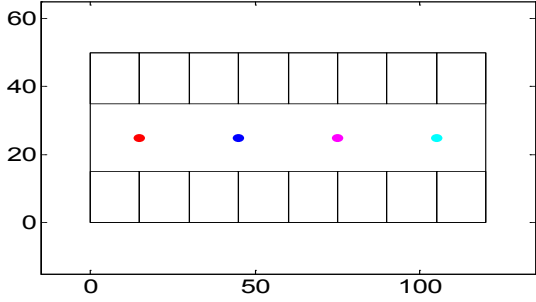
LTE Parameters		
Bandwidth		10MHz
Frequency		$f_c = 2.6\text{GHz}$
Layout Parameters		
		
Block	Number	2 Rows of 8 Blocks
	Size	15m x 15m
Hall Size		20m x 120m
Deployment Parameters		
Number of iSCs		4 (fixed position)
Number of UEs		40 (random drop)
Propagation Parameters		
Thermal Noise Density		$N_0 = -174\text{dBm/Hz}$
Channel Model		ITU-R InH [17]
Line of Sight Probability (d is the iSC-UE 2D-distance in meters)		$P_{LoS} = \begin{cases} 1, & d \leq 18 \\ e^{-\frac{d-18}{27}}, & 18 < d < 37 \\ 0.5 & 37 \geq d \end{cases}$
Pathloss	LoS	$PL = 16.9 \log_{10}(d) + 32.8 + 20 \log_{10}(f_c)$
	NLoS	$PL = 43.3 \log_{10}(d) + 11.5 + 20 \log_{10}(f_c)$
Shadowing Std Dev	LoS	$\sigma = 3\text{dB}$
	NLoS	$\sigma = 4\text{dB}$

Table 4-11: Indoor Hotspot system level simulation dynamic parameters.

Dynamic Parameters	
Number of RBs for PUSCH	50
HARQ	8 synchronous HARQ processes. Chase combining
Overhead	DMRS assumed (12 SC-FDMA symbols available per frame) SRS not simulated
Power Control (outer loop)	$P_0 = -96\text{dBm}$, $\alpha = 1.0$ [18]
Fast Fading	ITU-R InH [17]
Traffic	Full Buffer
Scheduler	Equal resource repartition / CT 3.7 scheduler
Physical layer abstraction	MIESM compression LTE-compliant AWGN look-up tables per MCS and RB

Equal resource allocation is assumed, meaning that each UE will be scheduled on 5RBs since 10 active UEs are attached per iSC on a 10MHz bandwidth (50RBs).

4.7.2.3 Approach

For the turbo detection to be possible (either locally or centrally), some users served by different iSCs need to be “paired”, i.e., aligned on the same resources. Indeed, the main idea behind the turbo principle is to use the interference coming from a different user as a source of information in an iterative process to decode the user of interest’s signal. Once one user’s signal has been detected, its (estimated) contribution in the received signal (which is the sum of all users’ signals) is subtracted to improve the detection of another users’ signal. This detection and interference cancellation process is repeated iteratively.

In D3.2 [2], we proposed a centralised pairing algorithm where one UE reports to its serving iSC its main dominant interferer (i.e., another iSC) in the downlink, if the RSRP difference between the two iSCs is within a threshold $\Delta_{\text{threshold}}$. The UEs reporting such a situation are included in a candidate list, and pairs are then formed based on the input from this list. This pairing based on downlink measurement is used to detect quickly users which may be in a cell edge position between two or more iSCs.

To represent the turbo detection receiver strategy at system-level, perfect user signal cancellation is assumed when computing the SINR, i.e., when computing the SINR of user #1, the contribution of user #2 is not taken into account and vice versa. Such approach is quite optimistic, but the turbo-detection tends to usually reach this bound for the considered channel models (see D2.2 [4]).

We compare the SPTD performed locally at each iSC (without iSC-iSC cooperation during the physical processing) and the MPTD performed centrally at RANaaS data centre to the baseline scenario, which assumes an Enhanced Linear Minimum Mean Square Error receiver strategy with Interference Rejection Combining (E-LMMSE-IRC) at each iSC (see D2.1 [3]).

4.7.2.4 Final Results

Figure 4-36 gives the CDF of the user uplink throughput for all three receiver strategies. For the SPTD/MPTD evaluation, we use a threshold value of $\Delta_{\text{threshold}} = 9\text{dB}$, leading to an average pairing percentage equal to 52% (almost half of the UEs are paired together). Compared to the baseline (MMSE), MPTD provides significant gain in the average uplink throughput (3.35 Mbps vs 2.59 Mbps). More importantly, the edge users (5-percentile) do really benefit from our approach as targeted by our pairing algorithm. Indeed, MPTD (1.79 Mbps) and SPTD (1.57 Mbps) give a 118% and a 92% boost to the uplink throughput of the edge users (0.82 Mbps), respectively.

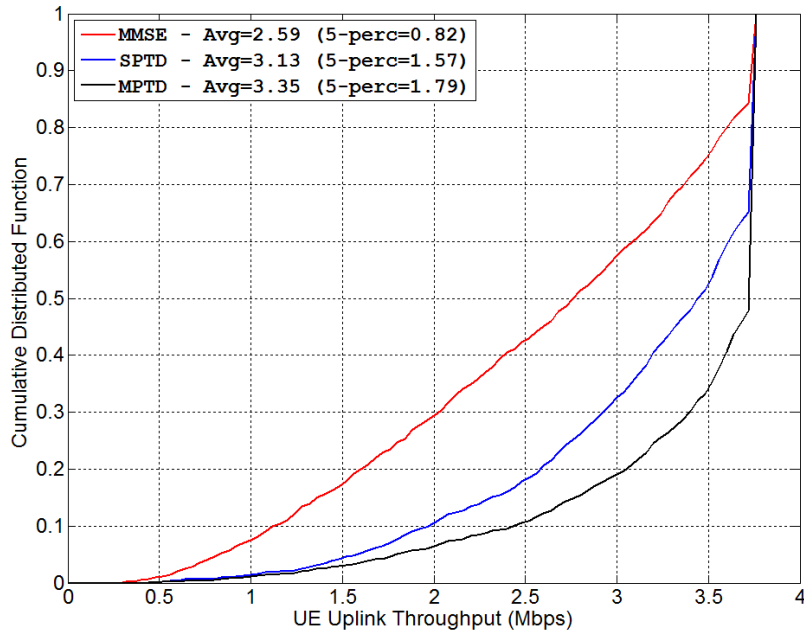


Figure 4-36: Comparison of the user uplink throughput CDF.

We also evaluate the Area Throughput (AT) gain achieved by the MPTD/SPTD. Figure 4-37 gives the normalised AT gain when the threshold parameter $\Delta_{\text{threshold}}$ is changing from 6 dB to 18 dB for both approaches, while Figure 4-38 gives the corresponding average pairing ratios. As expected, the higher the threshold is, the more candidates are paired together as a UE can report more potential iSCs to be paired with its serving one. Regarding the AT, gains of 32% and 23% are achieved by the MPTD and SPTD for a threshold of 12 dB, respectively. This corresponds to an average pairing ratio of users of 65%. Increasing the threshold rises up these gains, but not by a high margin. For instance, an 18dB threshold leads to 35% and 24% increase, while the average paired ratio goes up to 77%, putting more pressure on the backhaul part for the MPTD and on the processing part of an iSC in the SPTD.

Indoor Hotspot

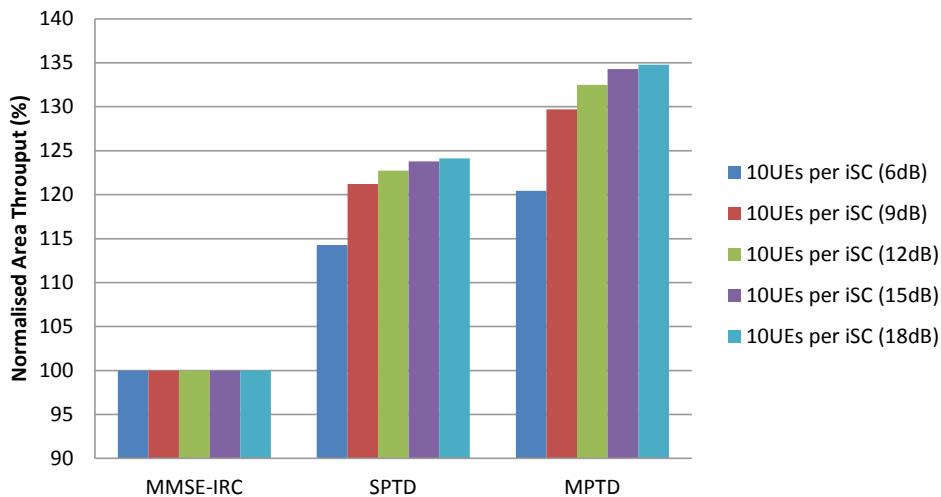


Figure 4-37: Normalised Area Throughput gain for different thresholds.

Indoor Hotspot

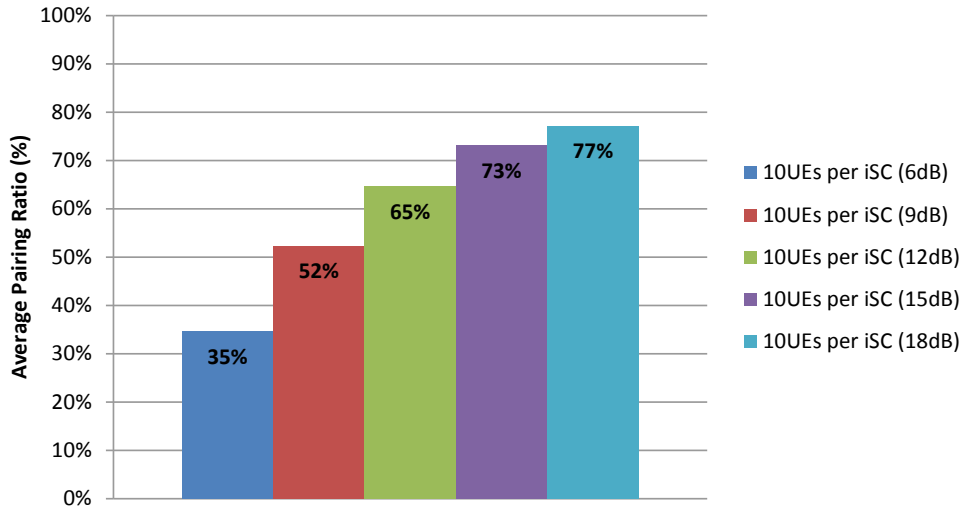


Figure 4-38: Average pairing ratios (%) for different thresholds.

The gain in AT seems less pronounced in the indoor scenario, because the cell throughput for the baseline is already quite high. With the power control parameters defined in Table 4-11, the average iSC throughput with the baseline receiver is around 27.5 Mbit/s per iSC on average. Considering only one antenna and 5 RBs per UE, the peak uplink throughput on 10MHz is then equal to 37.5 Mbit/s if the 10 transmitting UEs use the best supported MCS. Therefore, we only have “room” for a 10 Mbit/s improvement, if we do not change the UEs and the number of deployed iSCs. With MPTD, we have around 34.25 Mbit/s on average per iSC with threshold of 12dB, which means that we are almost maximising the air interface based on the UE capabilities. This can be seen with the throughput CDF of Figure 4-36 where more than 50% of the UEs have the maximum MCS.

Equivalent evaluations on the Stadium (CS1) and the Wide-Area Continuous Coverage (CS3) scenarios have also been performed for the project-wide analysis reported in D5.3 [10]. Using the layout assumptions described in D5.2 (Section 4.1.4 for CS1 and Section 4.3.4 for CS3), the AT gains compared to the baseline are given in Table 4-12, with an aggressive pairing threshold $\Delta_{\text{threshold}} = 15$ dB and the following power control parameters $P_0 = -86$ dBm, $\alpha = 1.0$ [18].

Table 4-12: Area Throughput Gain for MPTD/SPTD.

Common Scenario	#iSCs	#UEs	Pairing Ratio	MPTD AT Gain	SPTD AT Gain
Stadium (CS1)	15	15	90%	56%	Not Simulated
	15	75	85%	60%	Not Simulated
Wide Area (CS3)	19	38	56%	51%	20%
	19	190	72%	65%	26%

All in all, MPTD and SPTD approaches manage to provide significant gains which can be controlled by the threshold parameter of our centralised RRM pairing algorithm. This allows for backhaul bandwidth adaptation in case of congestion.

4.8 CT 3.8: Radio Resource Management for In-Network Processing

The scope of this CT is the investigation of the RRM aspects of the distributed Multi-User Detection (MUD) technique developed in WP2 CT 2.1. A detailed explanation of this technique and the different variants of the actual distributed algorithm such as, e.g., the Distributed Consensus-based Estimation (DiCE) algorithm, can be found in the deliverables D3.2 [2] and D2.2 [4]. The MUD is performed in a distributed fashion across a set of cooperating iSCs and is able to jointly detect all users within the coverage area of these iSCs, i.e., is able to cancel all interference of these UEs among each other. This enables the reuse of physical resources, e.g., in neighbouring iSCs without the need for blanking of subframes, thus allowing a high spectral efficiency, and consequently, an increased area throughput.

4.8.1 Final Implementation of CT

The cooperation among iSCs requires communication via the J2 links. Since several iterations with information exchange are performed for each subframe, a low latency/high capacity backhaul of category 1c or better is required regardless of the functional split and the employed backhaul technology for the J1 link. The decision which iSCs are communicating directly with each other is up to the resource management, since it will affect the total amount of backhaul traffic. However, we assume that every cooperating iSC can reach its peer iSC with at maximum 2 hops, i.e., allowing for a star or a mesh as physical backhaul topology. Possible configurations were shown in D3.2 [2]. Since the detection is performed in an iterative fashion, the processing time grows with the number of iterations. In the light of PHY layer constraints on the total processing time, this implies that the allowed number of iterations depends on the backhaul properties. Details on this issue can be found in the discussion of CT 2.1 in D2.3 [5].

While in CT 2.1 the single link performance for a given MCS is investigated, this CT will use these results in order to suggest guidelines for a resource allocation algorithm for the proper choice of MCS and the decision which UEs to detect jointly and to which allocate orthogonal resources.

4.8.2 Evaluation of the CT

In order to investigate this CT, we will build on the link level simulations performed in WP2. Therefore, details on the implementation of the link level simulations can be found in [5]. Shortly described, we replace the local equalization in each iSC by a cooperative one among several ones using the DiCE algorithm. Please note that the results and gains presented here cannot be directly compared with the ones presented in D2.3 [5]: In D2.3, the number of INP iterations is varied between 2 and 5 for the different INP variants, respective of the practical backhaul constraints. Here, a fixed number of 5 iterations is assumed. Furthermore, here MCS selection is performed for every UE, which is not the case in D2.3, where all UEs are assumed to use the same MCS. Combining these 2 aspects, the achievable gains presented here are larger than the ones for the CT 2.1 evaluation. For the project wide performance evaluation, the results from CT 2.1 were used, even though they are denoted as CT 2.1+3.8, in order to point out that the corresponding candidate technology spans two work packages.

The primary scenario investigated is the CS3 “Wide Area Continuous Coverage”, defined in the D5.2 document [9] with two UEs per iSC resulting in a total number of $N^{UE} = 38$ UEs. This deployment is depicted in the following Figure 4-39:

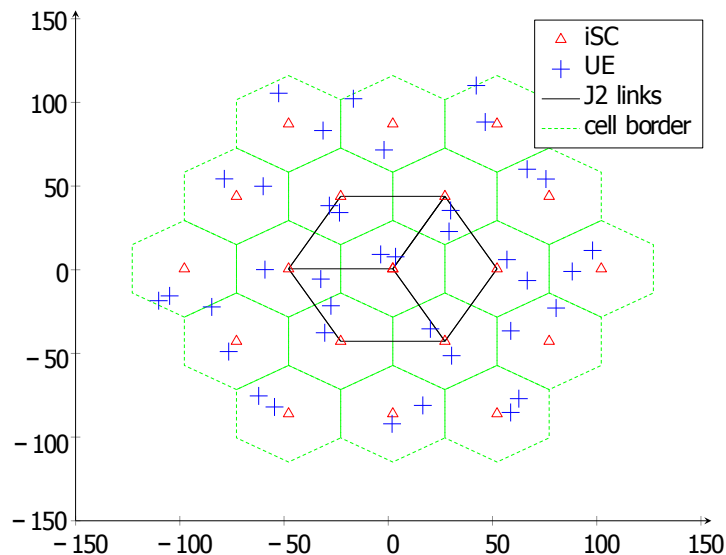


Figure 4-39: Deployment model for CT3.8 in the Wide Area Coverage scenario.

The access channel between UEs and iSCs is modelled via the UMi channel model with the “Pedestrian A” power profile defined in [18]. The pathloss is modelled via the urban micro NLoS pathloss formula. Each iSC is equipped with $N_R^{iSC} = 2$ receive antennas while each UE has $N_T^{UE} = 1$ transmit antenna. For the sake of simplicity, all 38 UEs are assumed to use the same MCS configuration. We will compare 2 cases. In the first case, the whole available bandwidth of $B = 10\text{MHz}$ is allocated to both UEs in the cell, causing a

complete overlap of allocated resources. In the second case, the bandwidth is split equally among both UEs per cell.

For evaluation of the area throughput, only the 14 UEs placed in the seven inner hexagonal cells are considered, while the remaining 24 UEs of the outer cells produce a constant interference level at the detecting iSCs. The logical interconnection topology of the iSCs is depicted in Figure 4-39 using black lines, i.e., an iSC is communicating with either 2 or 3 adjacent iSCs. Due to the restrictions of the simulator, the UE dropping was static throughout all simulations.

For one run of a simulation, the link level simulator assumes identical MCSs for all of the UEs, with a given power allocation per UE. Based on this, the per-UE throughput over the ratio of UE transmit power P_{UE} to per-subcarrier noise level σ_n^2 (which shall be denoted as signal-to-noise ratio in the following) will be obtained, which is then summed up and normalised to area.

The results for a power allocation to all UEs with equal powers P_{UE} are shown in Figure 4-40. It shows the resulting area throughput for MCS 1-9 for the first case of complete spectral overlap, while Figure 4-41 shows the case of split bandwidth. 5 iterations of the DiCE algorithm were performed. It can be seen that if the increased interference level is taken care of by choosing a lower MCS, a higher area throughput can be achieved if the entire bandwidth is allocated to all UEs. Allocating dedicated bandwidth to the UEs of one cell allows for a higher MCS choice, but cannot compensate the reduced spectral bandwidth.

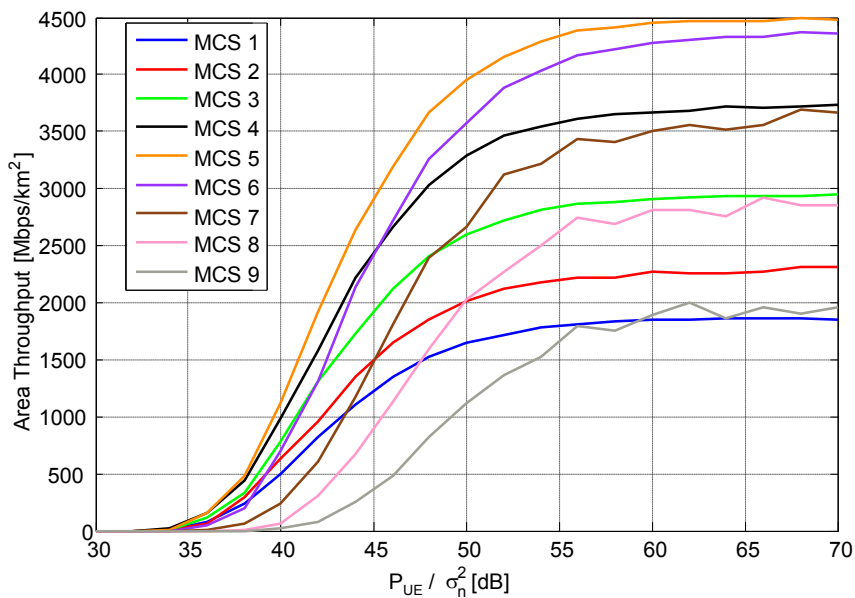


Figure 4-40: Area throughput for CT3.8 with whole bandwidth allocated.

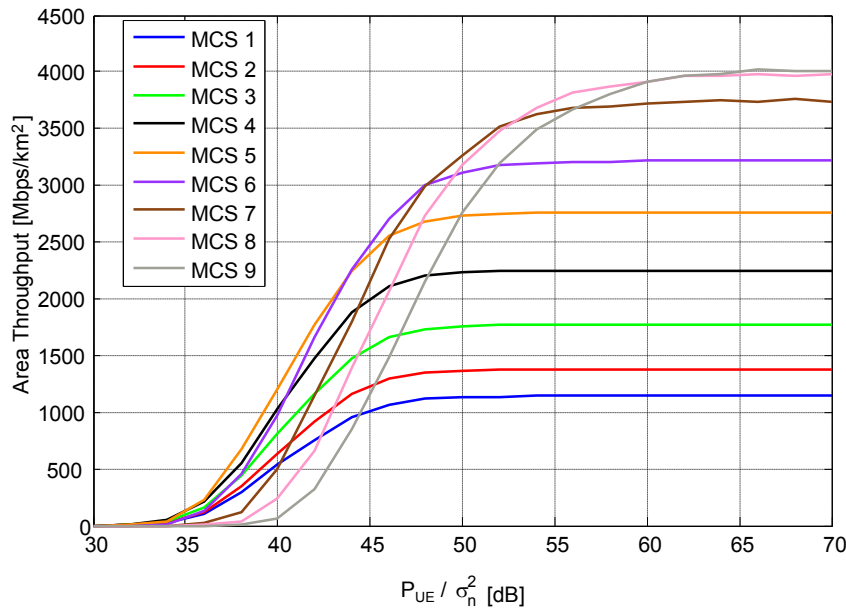


Figure 4-41: Area throughput for CT3.8 with equal bandwidth split among UEs of one cell.

Since the interference power among UEs does not depend on the chosen MCS, but only on the transmit power, it is possible to select the optimum MCS per UE after the simulation. This was performed for the following results, however equal power allocation was maintained. The accumulated UE throughput was normalized to the area in order to obtain the average area throughput, which is depicted in Figure 4-42. As a baseline, local detection of the 2 UEs of each iSC at the corresponding iSC was referred to, i.e., one iSC will only detect the UEs located within its cell area and will treat the other iSCs' signals as noise. Here, orthogonal, equal resource allocation of the 2 UEs was assumed, so no intra-cell, but only inter-cell interference is encountered. Due to the severe interference from neighbouring iSCs, an area throughput of only 4.5 Gbps/km² can be achieved. The joint decoding of the inner 7 iSCs achieves an area throughput of roughly 3 times that value.

In order to qualitatively assess the impact of the power allocation to the achievable rate, simulations were performed in which the power was allocated exactly such that the path loss was compensated and the signals were received with equal power. For the given frame conditions, these simulations showed no improvement over the depicted results. Of course, significantly differing results can be expected from a power allocation mechanism that considers the selected MCS and occasionally mutes certain UEs. However, such a mechanism cannot be implemented in the employed multi-link-level simulator due to its practical restrictions.

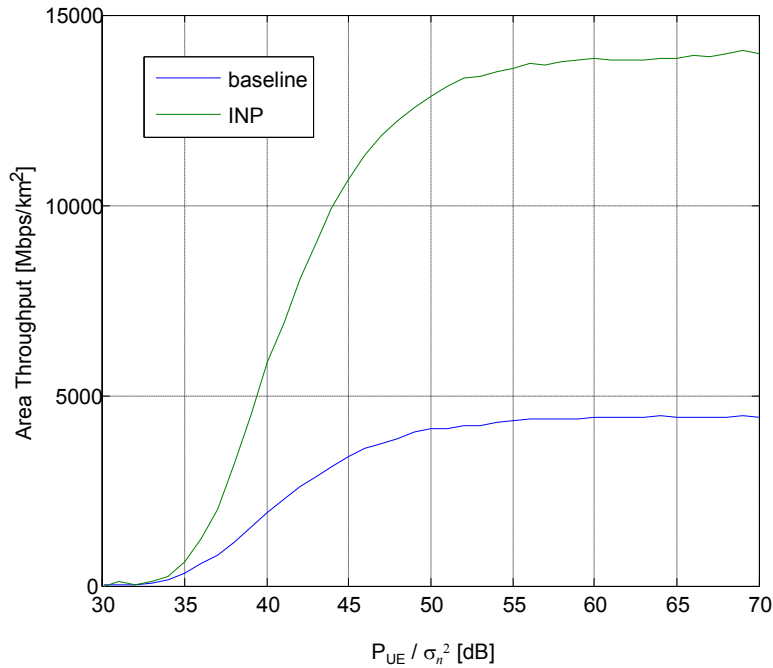


Figure 4-42: Area throughput gains of INP compared to baseline.

The second scenario under investigation is the indoor scenario. For this scenario we assume 4 iSCs, one per room, with 5 UEs each. This scenario is depicted in Figure 4-43. Like in the wide area case, the UMi channel model with the “Pedestrian A” power profile is used. For UEs located within the hall area, LoS path loss is assumed, while for the channels to UEs located in one of the rooms, NLoS path loss is assumed. Each UE is equipped with $N_T^{UE} = 1$ transmit antenna and each iSC has $N_R^{iSC} = 2$ receive antennas. Also here, equal power allocation with P_{UE} was performed. For the baseline, the bandwidth was split equally among the 5 UEs of one iSC.

Compared to the wide area case, the joint detection does not encounter any external interference and thus is able to achieve a better performance. Even for the baseline case of local detection of all UEs of one cell, the interference levels are lower, since only 3 interfering UEs (from the other iSCs) are operating within the same bandwidth.

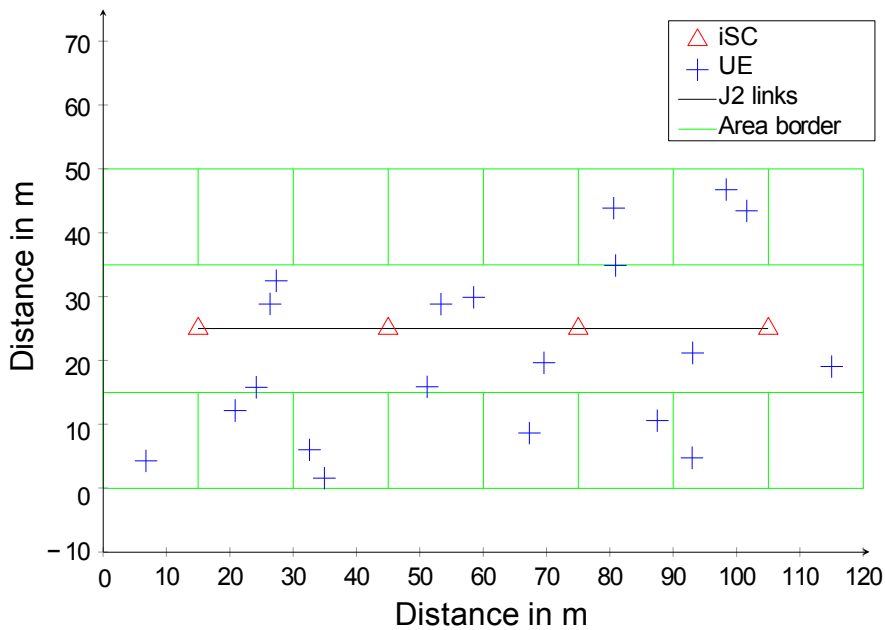


Figure 4-43: Deployment model for CT3.8 in the indoor scenario.

The achieved area throughput is depicted in Figure 4-44. Due to the reduced interference compared to the wide area case, the gain compared to the baseline is only approx. 8% for high UE transmit powers, but significantly larger for lower signal-to-noise ratios: E.g. for 25 dB, 650% gain is observed, 114% for 30 dB, and still 28% gain for 35 dB. The shift between the two curves is caused by the fact that due to the joint detection, higher-rate MCS can be selected for mid-range signal-to-noise ratios between 20 and 40 dB.

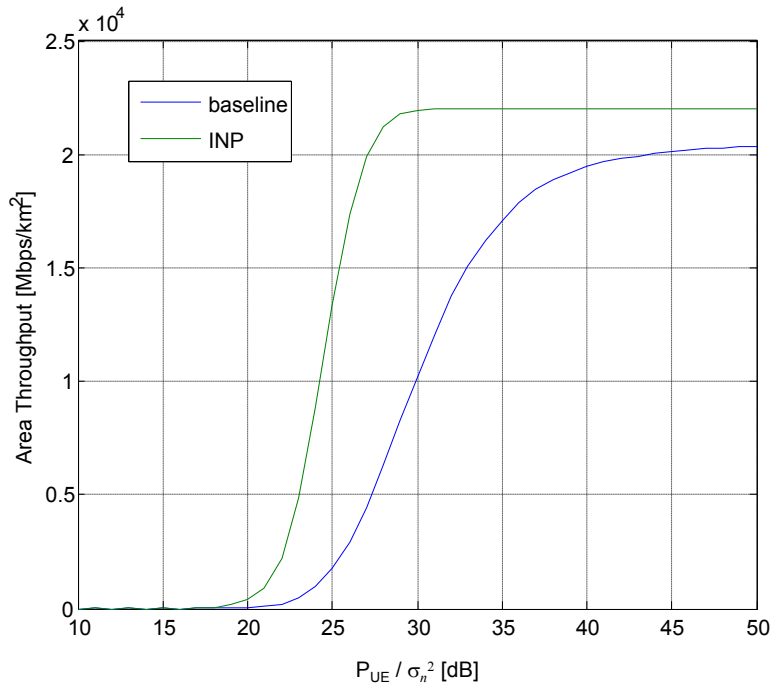


Figure 4-44: Area throughput gains of INP compared to baseline in the indoor case

4.9 CT 3.9: Hybrid local-cloud-based user scheduling for interference control

The goal of this CT is to develop new cooperative scheduling algorithms which efficiently exploit any backhaul topology available and deal with strong requirements in terms of exchange of information that characterize classic cooperative RRM schemes. A joint centralized scheduler across multiple iSCs allows to fully exploit the multi-user diversity of the network in order to increase the strength of the desired signal while at the same time reducing the interference between the adjacent iSCs (see Figure 4-45). This centralized scheduling comes however at a high cost in terms of backhaul resources since all the channel coefficients (including the cross channels) of each user have to be provided to the centralized scheduler. It is often not possible to acquire all this information in a timely manner, or at least too expensive to be scaled with the size of the network. Thus, distributed scheduling or partially centralized scheduling appears as promising approaches. This problem is a very challenging one and the generality of the scenario leaves room for many different approaches. In this context, we present here one possible approach to adapt to the backhaul conditions. It aims at exploiting a backhaul link to the RANaaS which introduces large delay and exploit the locally available interactions and CSI.

In that configuration, only the statistics of the channel (i.e., the topology) are available in the RANaaS while each iSC can exploit local CSI and has the possibility to locally interact with its neighbouring iSCs. The local interactions are defined through *scheduling functions* and *cooperation protocols* between the iSCs. On the basis of the long term information available and the knowledge of the scheduling functions and of the cooperation protocols, a centralized optimization is carried out in the RANaaS in a centralized manner, possibly through a learning process as it needs only to be slowly updated. More details can be found in [2] and [16].

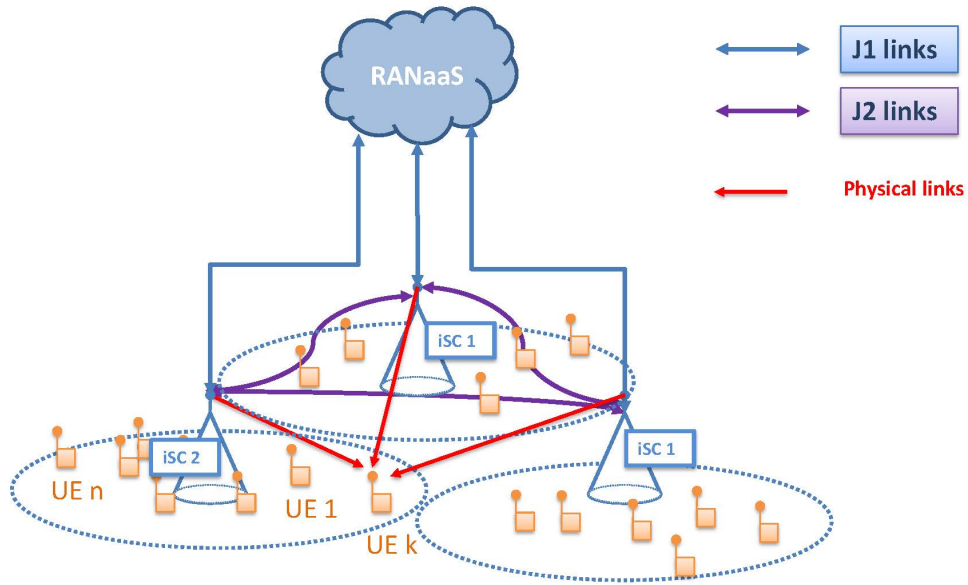


Figure 4-45 Schematic presentation of the architecture as studied in CT3.9

4.9.1 Final Implementation of CT

As a first step, we have considered a simple, yet practical configuration, where each iSC only knows its own direct channel. The optimal distributed scheduling policy at each iSC consists then in applying a *threshold function* to decide whether or not to transmit: iSC i transmits if the instantaneous realization of the channel gain is larger than a given threshold value [16]. The value of the different thresholds can then be optimized during the centralized optimization to manage interference. This partially centralized scheduling algorithm is described. This example corresponds to an extreme case where the iSCs do not interact locally. Depending on the backhaul topologies, CSI or interference management information (e.g., inhibiting bits to avoid collisions) could be studied in future works. For the sake of clarity, the main steps of the partially centralized scheduling studied here are summarized in the message sequence chart in Figure 4-46.

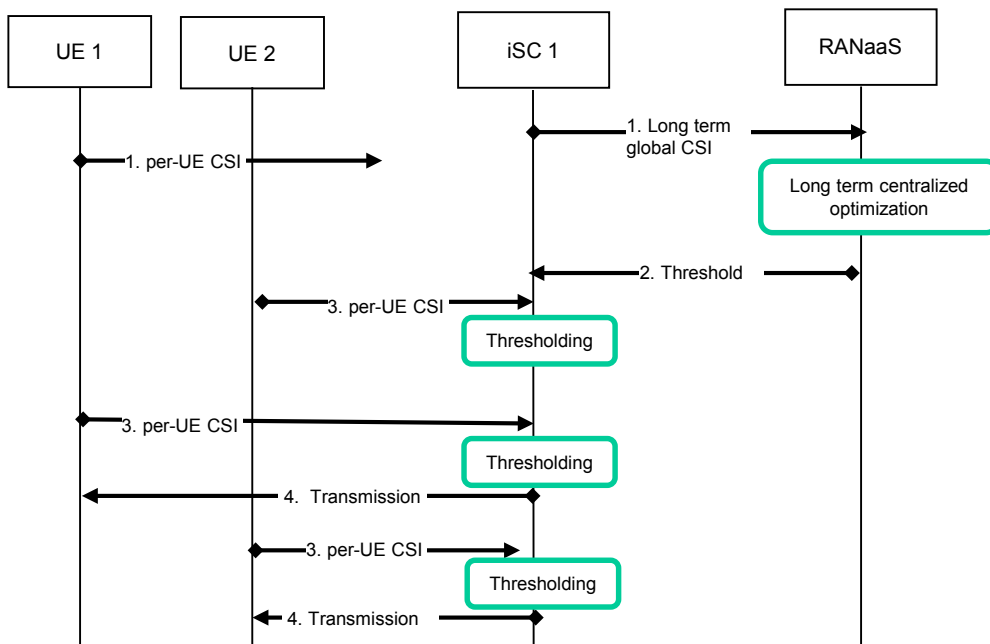


Figure 4-46: Message sequence chart for CT3.9.

4.9.2 Evaluation of the CT

We consider the Wide Area Coverage scenario as described in [9]. For the sake of completeness, we recall the simulation parameters in Table 4-13.

Table 4-13: Simulations parameters for the performance evaluation of CT3.9

Parameters	Values
Number of small cells	19
Number of antennas at each iSC	2
Number of antennas at each UE	1
Power constraint at each iSC	30dBm
Number of UEs	1 UE per small cell (avg.)
Small cell dropping	Regular on Hexagonal Grid
ISD	50 m
UE dropping	Random dropping on the edge of the cells
Pathloss model	ITU-R NLoS UMi
Backhaul Capacity / Latency	~50 - 100 Mbps, 1-10ms 10 Gbps, 5 μ s

We further show in Figure 4-47 one realization of the random dropping of the UEs, where the squares denote iSCs while the red crosses are the UEs. The area throughput is evaluated in the 7 green cells in the middle while the outside iSCs simply generate uncoordinated interference. The UEs are randomly dropped on the edge of the cells. This corresponds to the fact that this CT is used to coordinate heavily interfering iSCs and ensure a sufficient quality of service to these weak users.

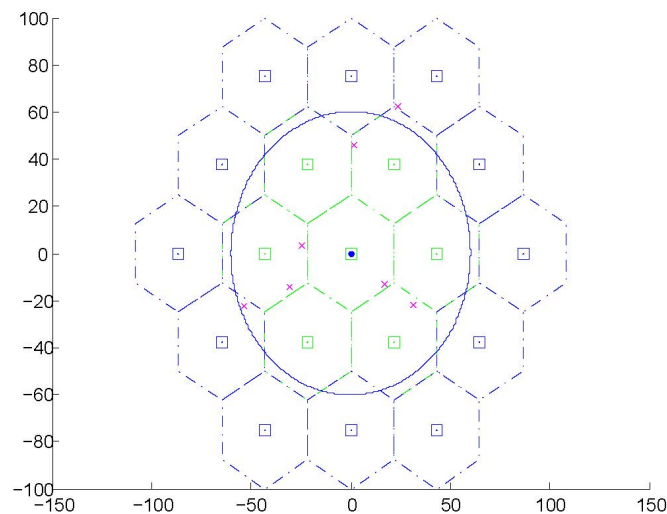


Figure 4-47: Topology of the simulated network for one random dropping of the UEs.

The distribution of the area throughput is shown in Figure 4-48. We compare the partially centralized approach to the two conventional scheduling approaches, being Round Robin and non-cooperative transmission. The first one achieves perfect coordination but does not exploit the multi-user diversity while the second one leads to losses due to the interference. It can be seen that the partially centralized approach outperforms both approaches. The average gain in area throughput from the partially centralized approach is equal to 20%.

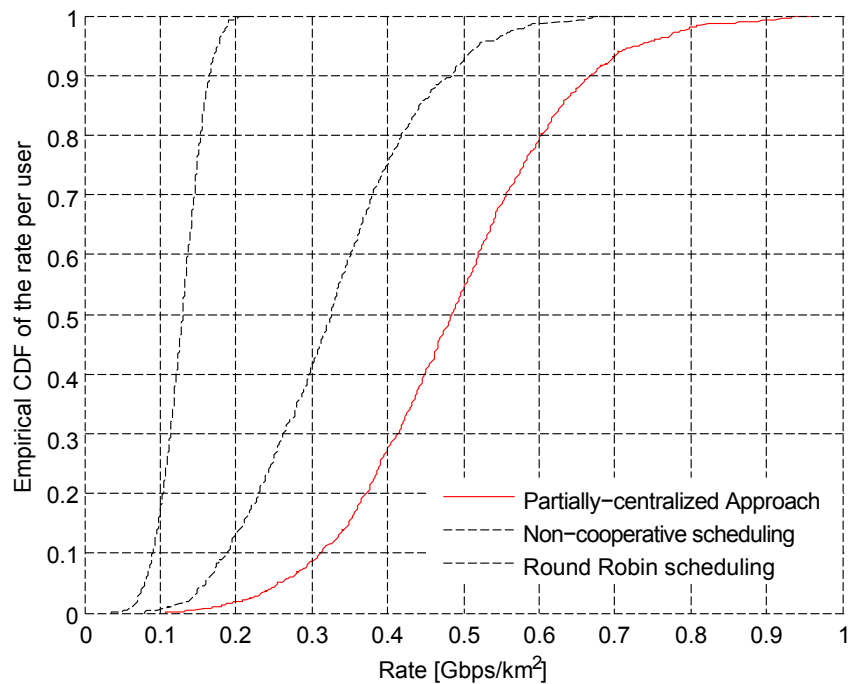


Figure 4-48: Distribution of the area throughput.

We have shown how considering a distributed binary power control, i.e., the possibility for the iSC to serve a user or not could be considered together with a central scheduling on the basis of the long term statistics. This approach is new and has led to the publication [16]. A major interest of this approach is the fact that coordination gains can be achieved with low CSI requirements and low signal processing capabilities at the iSCs since the difficult optimization is carried out in the RANaaS while only the simple thresholding step is done at the iSCs. This coordination scheme is very robust since no sharing of the instantaneous CSI is needed.

Extending this approach to consider more local interactions is an interesting direction for further research. Indeed, the exploitation of more local interactions could help bridge the gap between the proposed robust coordination approach and fully “centralized” scheduling with full CSI exchange.

5 iJOIN MAC/RRM Overall Evaluation

In this section we detail the overall performance gains achieved by the WP3 CTs in relevant iJOIN scenarios. WP3 mainly focuses on the wide area coverage use case where MAC solutions have been developed to improve the network area throughput. However, this section also provides results for the stadium, the square, and the indoor scenarios. Moreover, quantitative results are also provided for some CTs in terms of EEff and UEff.

5.1 CT Interoperability

Table 5-1 presents the relationships between CTs in terms of compatibility. Compatible CTs can be successfully integrated to jointly enhance the performance of a reference system. This integration may require some kind of coordination between CTs that operate on the same resources but on a different time scale, e.g. inter-cell interference coordination and scheduling algorithm. On the contrary, CTs that conflict with each other cannot be used jointly since they simultaneously operate e.g. on the same resource or functionality.

In this table, we define two CTs as interoperable when CTs operate on different resources or in different operational domains. However, algorithmic coordination and/or information exchange with iJOIN network entities (e.g. iNC or iveC) may be necessary. Different domains/resources include the backhaul domain (such as channel resources, links, routing constraints), and RAN (e.g., transmit power, downlink/uplink radio resources, connection control). Note that we assume that uplink and downlink can be handled separately with some coordination efforts.

It can be observed that WP3 CTs are characterized by a limited number of possible conflicts:

- CT 3.4 “Computational Complexity and Semi-Deterministic Scheduling”, CT 3.5 “Computation complexity and semi-deterministic scheduling”, and CT 3.9 “Hybrid local-cloud-based user scheduling for interference control” cannot be implemented jointly since they all focus on downlink radio resource management.
- CT 3.7 “Radio resource management for scalable multi-point turbo detection/In-network Processing” may not be compatible with CT3.8 “Radio Resource Management for In-Network-Processing”, which is a concurrent uplink RRM method.

Table 5-1: Compatibility of WP3 CTs.

	CT 3.2	CT 3.3	CT 3.4	CT 3.5	CT 3.7	CT 3.8	CT 3.9
CT 3.1	Y	Y	Y	Y	Y	Y	Y
CT 3.2		Y	Y	Y	Y	Y	Y
CT 3.3			Y	Y	Y	Y	Y
CT 3.4				N	Y	Y	N
CT 3.5					Y	Y	N
CT 3.7						N	Y
CT 3.8							Y

5.2 Evaluation methodology

Evaluation metrics and reference scenarios are detailed in D5.2 [9]. Additionally, for sake of completeness, Section 5.2.1 briefly recapitulates the definition of relevant metrics for WP3, specifically, the area throughput, the energy efficiency, and the utilization efficiency. Moreover, Section 5.2.2, will detail relevant evaluation parameters used to capture the iJOIN scenario in our simulations.

5.2.1 Metrics

The studies conducted in this WP focus mainly on improving the system area throughput. Additionally, CT 3.3 and CT 3.6 investigate methodologies to evaluate and improve the energy and utilization efficiency of future cellular networks. In the following we briefly present these evaluation metrics for sake of clarity, while further details on their definition can be found in D5.2 [9].

Area Throughput is expressed in terms of *bits/second/area*. It measures the utilization of the radio spectrum over a given geographic area (A) and also represents the capacity that a mobile operator offers to its subscribers. Denoting by $r_i(t)$ the bits correctly delivered at (from) the UE i at the time slot t , the area throughput in a network comprising N UEs is calculated as

$$AT = \frac{R}{A} = \frac{1}{A \cdot T} \sum_{i=1}^N \int_{t=0}^T r_i(t) dt \quad [\text{bps}/\text{km}^2], \quad (5.1)$$

where R is the average network throughput measured in the time interval T .

In iJOIN, the metric for evaluating the system energy efficiency is the consumed *energy per information bit*:

$$EEff = \frac{E_{Total}}{R \cdot T} = \frac{\int_{t=0}^T P_{Total}(t) dt}{\sum_{i=1}^N \int_{t=0}^T r_i(t) dt} = \frac{\int_{t=0}^T (P_{RANaaS}(t) + P_{BH}(t) + \sum_{i=1}^N P_i(t)) dt}{\sum_{i=1}^N \int_{t=0}^T r_i(t) dt} \left[\frac{J}{\text{bit}} \right], \quad (5.2)$$

where P_{RANaaS} , P_{BH} , and P_i are the RANaaS platform, the BH network and the small cell network power consumption, respectively.

We define the total *utilization efficiency* of a system as following:

$$UEff = \sum_{d \in D} \alpha_d u_d, \quad (5.3)$$

where α_d is a scaling factor s.t. $\sum \alpha_d = 1$, and u_d is the *domain utilization* for the considered domain, with D is the set of network domains, e.g. RANaaS instance, backhaul network, and RAN.

Finally, results discussed in the following are presented as performance gains G with respect to baseline solutions assessed in the evaluation scenarios presented in details in D3.2 [2]:

$$G = \frac{P_{Prop} - P_{Base}}{P_{Base}}, \quad (5.4)$$

where P_{Prop} and P_{Base} indicate the performance achieved by the proposed solution and the baseline technology in terms of energy efficiency, utilization efficiency, or area throughput.

5.2.2 Evaluation parameters

Table 5-2 shows the evaluation parameters organized according to the common scenarios, the number of UEs, the functional split option and the backhaul assumptions. If not stated otherwise, the evaluations are conducted by means of simulations with implement the common evaluation scenarios (including detailed RAN assumptions, e.g. for pathloss models, device configuration, simulation layout, etc.) as defined in [2].

Table 5-2: CT evaluation scenario assumptions

Common Scenario	#UEs	Split option	Backhaul assumptions					CTs
			Reference	Technology	#Hops	Latency	Bandwidth	
Wide area	2 – 10 per iSC	B	1c	Millimeter wave	J1: 3 hops J2: 2 hops	J1: 600µs J2: 400 µs	2.5Gbps	3.1, 3.8, 3.3, 3.6 ¹
		C.1	3a	Sub-6 GHz Ptp NLoS TDD	2 hops	10ms	500 Mbps	3.4, 3.5, 3.6, 3.8, 3.9
		D.1	5	xDSL	end-to-end	35ms	100 Mbps	3.2
Square	15 – 30 per iSC	C.1	3a	Sub-6 GHz Ptp NLoS TDD	2 hops	10ms	500 Mbps	3.4
		D.1	5	xDSL	end-to-end	35ms	100 Mbps	3.2, 3.3
Stadium	320 in area	B	1c	Millimeter wave	J1: 3 hops J2: 2 hops	J1: 600µs J2: 400 µs	2.5Gbps	3.7
Indoor	5 – 10 per iSC	A	4a	Dark fibre	1 hop	5µs	10 Gbps	3.7
		B	1c	Millimeter wave	J1: 3 hops J2: 2 hops	J1: 600µs J2: 400 µs	2.5Gbps	3.8
		C.2 (only scheduling)	3a	Sub-6 GHz Ptp NLoS TDD	2 hops	10ms	500 Mbps	3.7 (distr.)

Notes: ¹CT 3.6 employs a different layout for evaluation.

5.3 Evaluation Results

The different CTs presented in Section 3 each achieve improvement on the KPIs defined in D5.2 [9][8]. For every CT, the relevant common scenarios and the primary KPI has been identified. This information was initially presented in the previous deliverable D3.2 [2]. The scope of this section is the presentation of achieved KPI improvements per common scenario in a concentrated, consolidated fashion. For this purpose, for every scenario, the achieved relative gains of every relevant CT are juxtaposed in a tabular fashion.

Table 5-3 provides an overview of the KPIs, the scenarios and the main input evaluation parameters of all CTs in WP3. It can be stated that the majority of the CTs address area throughput. However, most CTs also lead to improvements of other KPIs, from which we selected the most relevant one as “secondary KPI”. The valuation of the secondary KPI was mainly done in a qualitative way in D3.2 [2] and in D5.2 [9], and is summarized in Table 5-4 in this section.

Table 5-3: Overview of KPIs and input evaluation parameters per CT.

CT	Main KPI	Sec. KPI	Secondary KPI evaluation	Scenarios	Main input evaluation parameters
CT 3.1	AT	UEff	qualitative	Wide area	BH latency, #UEs.
CT 3.2	AT	UEff	quantitative	Square, wide area	BH capacity, #UEs, #iSCs
CT 3.3	EE	CEff	qualitative	Square, wide area	BH latency, #UEs
CT 3.4	AT	CEff	qualitative	Square, wide area	BH latency, #UEs
CT 3.5	AT	UEff	qualitative	Wide area	#iSCs, #UEs
CT 3.6	Ueff,	none	N.A.	Wide area	#UEs, #iSCs
	EEff	none	N.A.	Wide area	functional split
CT 3.7	AT	UEff	qualitative	Indoor, wide area, stadium	#UEs
CT 3.8	AT	UEff	qualitative	Indoor, wide area	#UEs
CT 3.9	AT	EE	qualitative	Wide Area	#UEs, #iSCs

It needs to be pointed out that every CT might make certain assumptions (e.g., the used functional split) which are described in the corresponding CT subsections, and that assumptions of different CTs might even contradict each other. Therefore, as described in Section 5.1 on CT interoperability, not all CTs might be combined with each other, or, if they can, their gains might not necessarily linearly add up.

The qualitative impact of the CTs towards the iJOIN objectives are summarized in Table 5-4. The “++” symbol indicates that a given CT mainly affects positively a specific objective, “+” accounts for beneficial side effect and the “0” represents a negligible impact. Most of the WP3 CTs target to improve the AT; however, a number of CTs also ameliorate the system UE by enhancing the effectiveness in the resource usage.

Table 5-4: Qualitative impacts of WP3 CTs.

	AT	Eeff	CEff	UEff
CT 3.1	++	0	0	+
CT 3.2	++	0	0	++
CT 3.3	0	++	+	0
CT 3.4	++	0	+	0
CT 3.5	++	0	0	+
CT 3.6	0	+	+	++
CT 3.7	++	0	0	+
CT 3.8	++	0	0	+
CT 3.9	++	0	0	+

5.3.1 Area Throughput

5.3.1.1 Wide Area Continuous Coverage Scenario

The wide area coverage scenario assumes a uniform spacing of iSCs as an “infinite” tile of hexagonal cells with an iSC-iSC distance of 50 metres. For simulations, 3 tiers, corresponding to 19 cells are considered. The density of the randomly distributed UEs is rather low and lies at 1-3 UEs per iSC. The used backhaul technology is not strictly specified, therefore, the investigated splits range from A to C. Details on this scenario as well as on all other scenarios can be found in D5.2 [9] and in Table 5-2.

Table 5-5 shows the numerical evaluation results for the area throughput KPI.

Table 5-5: Results for area throughput in the wide area coverage scenario.

CT	3.1	3.2	3.4	3.5 ³	3.7 ⁴	3.8	3.9
Area throughput gain	140% ¹	49% ²	57%	5% to 18%	Split A: 51% Split C: 20%	200% ⁵	5% to 20%

Notes: ¹The results refer to the backhaul area throughput per cell; ²The results correspond to the average area throughput when each small cells serves 2 UEs; ³5% gain is the median, for cell edge users up to 18% gain can be achieved; ⁴2 UEs assumed (10UEs results available in CT3.7 specific subsection); ⁵Equal power allocation to all UEs is assumed, and all UEs are scheduled to be active all the time.

In this table, the gains of CT 3.1 demonstrate what can be achieved by tackling the problem of joint path selection and backhaul link scheduling in a dense iSC deployment considering 60GHz multi-hop backhaul where RANaaS operates as a coordinator and traffic aggregator for the group of iSCs. More specifically the proposed solution of dynamically adjusting the number of hops from RANaaS to the destination iSCs leads into efficient utilisation of short-distanced LoS backhaul links and increases throughput performance. When a non-adaptive single-hop approach is adopted, performance will be deteriorated due to the fact that NLoS links may be chosen for transmission in backhaul. The 140% gain refers to the median perceived backhaul throughput per cell by using 2 to 5 paths (in high delay bound, i.e. by considering a maximum of 20 timeslots

delay from the CU to reach each destination iSC) when compared to the case where a single-hop approach (with no delay bound) is considered.

CT 3.2 aims at improving the system area throughput by optimizing the cell association mechanism through a BH-aware load balancing. Specifically, the proposed CT results in notable gains since it jointly takes into account the radio link quality (as the baseline 3GPP solution) as well as the BH capacity and the cell load. Accordingly, this mechanism avoids bottleneck and also improves the RAN & BH utilization efficiency, leading to a maximum gain of 49% in the evaluation.

CT 3.4 demonstrates the gains that can be achieved by applying multi-stage coordinated scheduling instead of local round robin scheduling. For the latter base line system, frequency reuse 1 is assumed, while both, user scheduling and link adaptation is performed locally at the iSCs. The gains of coordinated scheduling come from the flexibility to select iSCs, which are not transmitting at particular time slots. Consequently, strong interference especially to users, which are mostly located at cell edge regions, is omitted. However, the theoretical gains of coordinated scheduling are typically reduced due to backhaul delays, which cause inaccurate CSI. The CT allows the iSCs to update the global scheduling decision based on newer CSI.

CT 3.5 considers limited capacity and non-negligible latency backhaul and demonstrates the gains that can be achieved by employing ICIC mechanisms for dynamic resource allocation instead of the Reuse-1 with PF scheduling approach. The proposed solution of dynamic resource partitioning between the iSCs improves the average per-cell throughput despite the fact that resource utilization is less than in reuse-1 case; this is due to the fact that the inter-cell interference is significantly mitigated. Moreover, the interference mitigation leads to a much higher improvement in terms of cell edge users' throughput; therefore, our solution provides better rate fairness within the system.

CT 3.7 demonstrates the gains that can be achieved when iSCs can rely on the RANaaS to perform RRM and user association for (local or central) turbo detection pairing. Results are obtained through (calibrated) system-level simulations. With split C, no information other than RRM message coordination is required on the J1 interface, while split A will require PHY received signals (after first FFT block) to be sent back for a joint (multi-point) turbo-detection. The later approach needs high backhaul bandwidth and low latency for the J1, but obviously provides the higher gains. The J2 interface is not used.

CT 3.8 demonstrates the gains that can be achieved by full cooperation of iSCs for joint detection compared to purely local detection. Since this scenario is interference limited, joint detection can bring enormous gains, however, at the expense of high inter-iSC J2 communication effort, and thus backhaul load. The backhaul load can be reduced by limiting the number of INP iterations at the expense of AT. Using a full system level simulator correctly implementing the scheduling and power control algorithms, the achievable gain might be smaller due to an increased throughput in the baseline.

CT 3.9 shows that exchanging solely long term statistics can already bring some scheduling gains both from the opportunistic transmission and the interference management. The gains can become as high as 20% when considering edge users. As we allow for the exchange of some limited amount of instantaneous CSI, the gains will become larger. The proposed approach is interesting as it allows improving the performance in the most challenging settings where either the backhaul links are very weak or high mobility users are being served.

5.3.1.2 Stadium

The stadium scenario features a very high iSC and UE density ($\sim 1 \text{ UE/m}^2$), resulting in an enormous target area throughput of 100-400 Gbps/km². Therefore, a fibre ring backhaul is assumed, suggesting functional split A.

Table 5-6: AT results for the stadium scenario.

CT	3.7
Area throughput gain	Split A: 56%

CT3.7 demonstrates the AT gain when multi-point turbo detection is used with functional split A, where the physical processing is performed at the RANaaS. Up to 56% of gain is measured through system-level simulations when one UE is served by one iSC on the whole 10MHz bandwidth (equivalent to 5% of the

covered stadium seats being in transmission). As seen in the section dedicated to 3.7, this gain can go up to 60% if more UEs are served by an iSC.

5.3.1.3 Square

The square scenario is similar to the wide area case, since in total only 15-30 UEs with 4-10 iSCs are considered. However, these iSCs are randomly and not uniformly placed. For backhaul, like in the wide area case, different backhaul techniques can be applied.

Table 5-7: AT results for the square scenario

CT	3.2	3.4
Area throughput gain	52%	61%

In the square scenario, CT 3.2 achieves slightly better performance than in the wide area case due to the higher user density, which increases the multi-user diversity. The presented results are obtained by Monte Carlo simulations, where 50 UEs are present in each macro cell sector and the BH link capacity is set equal to 120 Mbps.

Simulating CT 3.4 in the square scenario results in up to 61% gain compared to Round Robin scheduling. This is about 4% more compared to CS3, the wide area continuous coverage scenario. The simulations are performed with 30 users, which is about twice as much as for CS3. Therefore, the square benefits from a larger scheduling gain.

5.3.1.4 Indoor

The indoor scenario targets an airport or shopping mall deployment. Here, only 2 or 4 iSCs are investigated, with 5-10 UEs per iSC. For backhaul, wireline technology is assumed, which does also include low-capacity DSL.

Table 5-8: AT results for the indoor scenario

CT	3.7	3.8
Area throughput gain	Split A: 32% Split C: 23%	28%

CT 3.7 assumes 10 UEs per iSC. The gain is less pronounced than in other scenarios as the baseline already exhibits good results. Indeed, the baseline average uplink throughput per iSC has been measured to 27.5 Mbit/s. The peak uplink throughput achievable when UEs attached to a single iSC are not scheduled on the same resources is equal to 37.5 Mbit/s with equal resource partitioning. The maximum AT gain with the same number of iSCs would be around 36%. With 32% achieved considering the functional split A, we are quite close to maximizing the air interface.

CT 3.8 here shows a smaller gain compared to the wide area case, since much less interference is encountered in the indoor scenario. The restrictions of the link level simulations detailed above also apply here. 8% gain is the minimum obtained for very high UE transmit powers. For lower transmit powers, the gains are significantly higher, as detailed in Section 4.8.2. E.g., for an SNR of 35 dB, 28% gain is observed.

5.3.2 Energy Efficiency

Table 5-9: Energy efficiency results.

CT	3.3	3.6
Energy Efficiency	70% (WA) 60% (Square)	43% ¹

Notes: ¹For theoretical evaluation.

The functional split model transfers functions, and accordingly the energy consumption requested to perform such functions, from the BB processing layer in base stations to the datacentre (PoP) where the RANaaS instances are executed. This can potentially increase the energy efficiency, if certain conditions are satisfied.

For instance, the functional split can allow to fully switch off base stations (apart from the RRHs), completely removing their power consumption whereas the corresponding consumption in a datacentre is incremental, only given by the software resources allocated and activated to execute the RANaaS functions. Also, energy efficiency strategies can be (in general) more easily enacted on a datacentre, given the large number of computational nodes and resources, allowing a better optimization than in single dedicated base stations. Of course, this must be checked against a number of factors, which can influence the amount of energy saving, and in some case even revert the balance.

In iJOIN, two kinds of research have been performed about energy efficiency:

- CT 3.3 studied the case of energy efficiency obtained through optimized Radio Resource Management in the RAN and in the backhaul. The CT was evaluated in two different scenarios (Square and Wide Area), where the solution was compared to a classic RAN DTX scheme and to a joint RAN/backhaul scheme. The best results (shown in the table above) come in respect to the RAN DTX, with savings of 60% and 70%. Wide Area continuous coverage scenario leads to improved gains due to the reduced number of active UEs considered in this scenario.
- CT 3.6 made a theoretical general assessment of energy efficiency in an iJOIN system, starting from analytical models of power consumption respectively in case of pure-LTE systems, and of iJOIN based functional split RAN. The formulas in this case are quite complex and depending upon a number of distinct factors. To appraise even one single exemplary figure, following the indications of two scientific papers produced by iJOIN, a given configuration scenario was defined, setting realistic values of the involved parameters. This exemplary figure came out as 43% saving. It's important to restate that this is just one very specific case, nonetheless it shows that, from a general theoretical standpoint, we can expect the application of iJOIN to improve the energy consumption of a RAN system.

5.3.3 Utilization Efficiency

Table 5-10: Utilization efficiency results.

CT	3.3	3.6
Utilization Efficiency	49% (WA) 52% (Square)	80% (WA)

CT 3.2 achieves notable improvements in terms of RAN and BH utilization efficiency in both the wide area coverage and square scenario. The results presented here are computed by averaging the gain on both the RAN and BH side i.e., $u^{RAN/BH} = \frac{C}{2 \cdot C_{BH}^T} + \frac{C}{2 \cdot C_{RAN}^T}$, Where C is the average area throughput and N_{iSC} is the number of deployed small cells. Moreover, by defining C_{BH} as the BH capacity at each small cell, and C_{RAN} as the maximum cell capacity, we have $C_{RAN}^T = N_{iSC} \cdot C_{RAN}$, $C_{BH}^T = N_{iSC} \cdot C_{BH}$.

The results obtained by CT 3.6 for utilization efficiency cover two scenarios. In the first, the required computational resources are obtained in case of centralization of forward error correction functions (i.e., for functional split B). In this case, due to the large multiplexing gains the utilization efficiency of computation resources can be increased by 80%, as shown in Section 4.6.

In the second scenario it is assumed that the computational resources are provided in such a way that computational outage may occur – which corresponds to a utilization of the system of 100%. In such cases, computational-aware scheduling mitigates the effects of outage such that only a minimal impact on the throughput is observable.

It should be noted that the utilization efficiency gains may vary slightly according to the evaluated scenario, although the main impact factor is the number of centralized iSCs and correspondingly, processed UEs.

5.4 Summary

The overall evaluation in the scope of WP3 has shown that significant gains can be achieved by exploiting centralization and coordination on MAC and RRM layer of RAN and backhaul. This holds true for all investigated common scenarios, as well as for all investigated KPIs.

This insight is especially relevant for deployment scenarios where the backhaul capacity is limited to such a degree that a RAN functional split below MAC layer is not feasible. The results in this section shows that even in such cases, centralized algorithmic approaches for coordinated connection control and radio resource management are capable of improving the overall performance of dense small cell networks significantly. For more details on the achievable gains per functional split see Section 3.2.

6 Summary and Conclusions

In this report, we provided the final definitions and evaluations of WP3 concepts and candidate technologies. In particular, the key paradigm of flexible functional split between local and centralized processing and joint RAN and backhaul operation design has been deeply investigated. The most relevant functional split options have been identified and assessed with respect to their advantages and shortcomings in terms of promising centralization gains, requirements, protocol and architectural aspects as well as backhaul latency and bandwidth. According to this analysis, key decision factors were identified in order to match the best functional split configurations with specific transport network deployments. Additionally, joint RAN and backhaul operation design is discussed in details by presenting the impact of backhaul characteristics on the RAN performance, proposing relevant technological approaches, and defining protocol aspects.

Moreover, the set of candidate technologies introduced in deliverable D3.1 [1] and developed in D3.2 [2] are finalized and potential gains identified. The implementation of the CTs in the iJOIN architecture has been discussed in terms of practical constraints and requirements. Finally, the overall evaluation of the iJOIN MAC/RRM solutions has been provided in the context of the iJOIN evaluation framework described in D5.2 [9], and by taking into account the possible interactions amongst different technical solutions. This evaluation describes the contribution of the WP3 CTs to the iJOIN key objectives, i.e., Area Throughput, Energy Efficiency, and Utilization Efficiency.

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).

Appendix I Backhaul classification reference

In this appendix we summarize the parameters of the considered set of backhaul technologies as discussed in D4.2 [6].

Table A-6-1: Backhaul classification reference

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 μ sec	≤ 1 Gbps	PtP (LoS)	FDD	--
1c		70-80GHz Light licensed	≤ 200 μ sec	≤ 2.5 Gbit/s	PtP (LoS)	FDD	--
2a	Microwave (28-42 GHz) Licensed		≤ 200 μ sec	≤ 1 Gbps	PtP (LoS)	FDD	--
2b			≤ 10 ms	≤ 1 Gbps	PmP (LoS)	TDD	TDMA
3a	Sub-6 GHz Unlicensed or licensed		≤ 5 ms	≤ 500 Mbps	PtP (NLoS)	TDD	--
3b			≤ 10 ms	≤ 500 Mbps (shared among clients)	PmP (NLoS)	TDD	TDMA
3c			≤ 5 ms	≤ 1 Gbit/s (per client)	PmP (NLoS)	TDD	SDMA
4a	Dark Fibre		$5 \mu\text{s}/\text{km} \times 2$	≤ 10 Gbps	PtP		--
4b	CWDM		$5 \mu\text{s}/\text{km} \times 2$	$\leq 10 \cdot N$ Gbps (with $N \leq 8$)	Ring		WDM
4c	Metro Optical Network		$250 \mu\text{s}$	≤ 1 Gbps	Mesh/Ring		Statistical Packet Multiplexing
4d	PON (Passive Optical Networks)		≤ 1 ms	100M – 2.5Gbps	PmP		TDM (DL)/ TDMA (UL)
5	xDSL		5-35 ms	10M – 100Mbps	PtP		--

References

- [1] INFSO-ICT-317941 iJOIN, “D3.1 - Final report on MAC/RRM state-of-the-art, Requirements, scenarios and interfaces in the iJOIN architecture”, Nov. 2013.
- [2] INFSO-ICT-317941 iJOIN, “D3.2 - Definition of MAC and RRM approaches for RANaaS and a joint backhaul/access design,” Nov. 2014.
- [3] INFSO-ICT-317941 iJOIN, “D2.1 - State-of-the-art of and promising candidates for PHY layer approaches on access and backhaul network,” Nov. 2013.
- [4] INFSO-ICT-317941 iJOIN, “D2.2- Definition of PHY layer approaches that are applicable to RANaaS and a holistic design of backhaul and access network,” Nov. 2014.
- [5] INFSO-ICT-317941 iJOIN, “D2.3- Final definition and evaluation of PHY layer approaches for RANaaS and joint backhaul-access layer,” Apr. 2015.
- [6] INFSO-ICT-317941 iJOIN, “D4.2- Network-layer algorithms and network operation and management: candidate technologies specification,” Nov. 2014.
- [7] INFSO-ICT-317941 iJOIN, “D4.3 - Network-layer algorithms and network operation and management: candidate technologies specification”, Apr. 2015.
- [8] INFSO-ICT-317941 iJOIN Project, “D5.1- Revised definition of requirements and preliminary definition of the iJOIN architecture”, Nov. 2013.
- [9] INFSO-ICT-317941 iJOIN, “D5.2 - Final Definition of iJOIN Requirements and Scenarios”, Nov. 2014
- [10] INFSO-ICT-317941 iJOIN, “D5.3 - Final definition of iJOIN architecture,” Apr. 2015.
- [11] INFSO-ICT-317941 iJOIN Project, “D6.2- Final proof-of-concept results for selected candidate algorithms,” May 2015.
- [12] 3GPP, “TR 36.932 V12.0.0; Scenarios and Requirements for Small Cell Enhancements for E-UTRA and E-UTRAN”, Dec. 2012
- [13] C. J. Bernardos, A. de la Oliva, J. C. Zuniga, and T. Melia. *Applicability Statement on Link Layer implementation/Logical Interface over Multiple Physical Interfaces. Internet Engineering Task Force, draft-bernardos-netext-ll-statement-01.txt (work-in-progress)*, Mar. 2010.
- [14] P. Rost, et al., *Cloud Technologies for Flexible 5G Radio Access Networks*, IEEE Communications Magazine, May 2014.
- [15] D. Sabella, A. De Domenico, E. Katranaras, M.A. Imran, M. Di Girolamo, U. Salim, M. Lalam, K. Samdanis, A. Maeder, *Energy Efficiency Benefits of RAN-as-a-Service Concept for a Cloud-Based 5G Mobile Network Infrastructure*, IEEE Access, vol.2, no., pp.1586,1597, 2014
- [16] P. de Kerret, S. Lasaulce, D. Gesbert, and U. Salim, “Best-response team power control for the interference channel with local CSI”, accepted for publication in IEEE International Conference on Communications (ICC), 2015.
- [17] ITU-R, “REPORT M.2135, Guidelines for evaluation of radio interface technologies for IMT-Advanced”, 2008.
- [18] 3GPP, “TR 36.814 V9.0.0; Further advancements for E-UTRA physical layer aspects (Release9)”, Mar. 2010.
- [19] V. Suryaprakash, P. Rost, G. Fettweis, *Are Heterogeneous Cloud-Based Radio Access Networks Cost Effective?*, IEEE Journal on Selected Areas in Communication, 2015.
- [20] H. Guan, T. Kolding, and P. Merz, “Discovery of Cloud-RAN,” in Cloud-RAN Workshop, Apr. 2010.
- [21] A. De Domenico, E. Calvanese Strinati, A. Capone, “Enabling Green cellular networks: A survey and outlook”, Computer Communications, vol. 37, pp. 5-24, Jan. 2014.
- [22] P. Rost, S. Talarico, and M. Valenti, *The complexity-rate tradeoff of centralized radio access networks*, IEEE Transactions on Wireless Communications, 2015, *submitted as minor revision*

- [23] P. Rost, A. Maeder, S. Talarico, M. Valenti, "Centralized Wireless Resource Scheduling under Computational Complexity Constraints", submitted to IEEE GLOBECOM 2015
- [24] 3GPP, "TS 36.213 V10.4.0; Physical layer procedures", Dec. 2011
- [25] R. Fritzsche, P. Rost, G. Fettweis, "Robust Proportional Fair Scheduling with Imperfect CSI and Fixed Outage Probability," in Proc. IEEE PIMRC 2014
- [26] R. Fritzsche, P. Rost, G. Fettweis, "Robust Proportional Fair Scheduling with Imperfect Channel State Information," accepted for publication in IEEE Transactions on Wireless Communications 2015
- [27] E. Pateromichelakis, A. Maeder, A. De Domenico, R. Fritzsche, and P. de Kerret, "Joint RAN/Backhaul Optimization in virtualized 5G RAN," accepted to EuCNC 2015.
- [28] E. Pateromichelakis, M. Shariat, A. Quddus, and R. Tafazolli, "Joint Routing and Scheduling in Dense Small Cell Networks using 60GHz backhaul," accepted to IEEE ICC - Workshop on Cloud-Processing in Heterogeneous Mobile Communication Networks (IWCPM), London, 2015.
- [29] A. De Domenico, V. Savin, D. Kténas, and A. Maeder, "Backhaul-Aware Small Cell DTX based on Fuzzy Q-Learning in Heterogeneous Cellular Networks," submitted to IEEE Globecom 2015.
- [30] P. Rost and A. Prasad, "Opportunistic Hybrid ARQ—Enabler of Centralized-RAN Over Nonideal Backhaul," IEEE Wireless Communications Letters, vol.3, no.5, pp.481,484, Oct. 2014.
- [31] E. Pateromichelakis, M. Shariat, A. Quddus, and R. Tafazolli, "Graph-Based Multicell Scheduling in OFDMA-Based Small Cell Networks," Access, IEEE , vol.2, pp.897,908, 2014
- [32] Z. Mheich, A. De Domenico, and V. Savin, "Uplink Capacity and User Association for Cooperative Heterogeneous Cellular Networks" IEEE SPAWC 2015.