Grant Agreement:	**247223**
Project Title:	**A**dvanced **R**adio In**T**erface Technolog**I**es for 4G **S**ys**T**ems ARTIST4G
Document Type:	P (Public)

Document Identifier:	
Document Title:	**Feedback from RAN constraints**
Source Activity:	WP4
Editors:	Guenther Liebl, Avelina Vega
Authors:	Giovanna D'Aria, Mark Doll, Nicolas Gresset, Mourad Khanfouci, Guenther Liebl, Thiago Martins de Moraes, Julien Muller, Danish Nisar, Hanns-Jürgen Schwarzbauer, Eiko Seidel, Avelina Vega
Status / Version:	Version 1.0
Date Last changes:	28.10.11
File Name:	D4.2_v1.0.doc0

Abstract:	The following document contains the current status of the cross-work-package activities between WP1-3 and WP4. This includes an overview of all architecture-relevant innovations in WP1-3, as well as a detailed architectural analysis of the main aspects identified in close cooperation with these work packages.

Keywords:	Cooperation, HetNet, Architecture, S1, X2

Document History:	
04.07.2011	Document created
13.07.2011	Outline of chapters 2-5 available.
24.08.2011	Overview of WP1-3 innovations (chapter 2) available. First draft version of per WP impact study (chapters 3-5) available.
07.09.2011	First draft version of introduction/conclusion available. Harmonized version of per WP impact study (chapters 3-5) available. Ready for first internal WP4 review.
21.09.2011	First review round complete. Discussion of document status (per WP version) in Telco.
05.10.2011	Deadline for addressing the first reviewer comments. Merge of per WP versions into single document. Ready for second internal WP4 review.
17.10.2011	Second review round complete.
27.10.2011	Deadline for addressing the remaining reviewer comments. Ready for final editorial review.
31.10.2011	Ready for delivery to EC.

Table of Contents

Authors

Name	Beneficiary	E-mail address
Mark Doll	Alcatel-Lucent (ALUD)	mark.doll@alcatel-lucent.com
Nicolas Gresset Mourad Khanfouci	Mitsubishi Electric R&D Centre Europe (MERCE)	n.gresset@fr.merce.mee.com m.khanfouci@fr.merce.mee.com
Guenther Liebl Thiago Martins de Moraes Danish Nisar Eiko Seidel	Nomor Research (NR)	liebl@nomor.de moraes@nomor.de nisar@nomor.de seidel@nomor.de
Hanns-Jürgen Schwarzbauer	Nokia Siemens Networks (NSN)	hanns-juergen.schwarzbauer@nsn.com
Giovanna D'Aria	Telecom Italia (TI)	giovanna.daria@telecomitalia.it
Avelina Vega	Telefonica I+D (TID)	avn@tid.es
Julien Muller	Orange-Labs	julien.muller@orange-ftgroup.com

Acronyms

3GPP	Third Generation Partnership Project
AMC	Adaptive Coding and Modulation
AP	Application Protocol
API	Application Protocol Interface
ARQ	Automatic Repeat request
ASN.1	Abstract Syntax Notation 1
BBU	Base-Band Unit
BRAS	Broadband Remote Access Server
BS	Base Station
CN	Core Network
CoMP	Coordinated Multi-Point
CP_COOP(F)	Control Plane Cooperation (Fast)
CP_COOP(S)	Control Plane Cooperation (Slow)
c-Plane	Control Plane
CSI	Channel State Information
CU	Central Unit
D2D	Device to Device
DeNB	Donor evolved Node B
DL	Downlink
DMA	Direct Memory Access
DoS	Deny of Service
D-RoF	Digitized Radio Over Fibre
DS	Differentiated Services
DSLAM	Digital Subscriber Line Access Multiplexer
eNB	evolved Node B
e-RN	End Relay Node
EUTRAN	Evoved Universal Terrestrial Radio Access Network
FDD	Frequency Division Duplex
FIC	Flexible Interference Cancellation
GNSS	Global Navigation Satellite System
GPON	Gigabit Passive Optical Network
GTP	Generic Tunnel Protocol
HARQ	Hybrid Automatic Repeat request
HeNB	Home evolved Node B
HeNB GW	Home evolved Node B Gateway
HO	HandOver
IANA	Internet Assigned Numbers Authority
IE	Information element
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force

if	Interface
I/O	Input/Output
IP	Internet Protocol
IReg	Interrupt Request
ISR	Interrupt Service Routine
LTE	Long Term Evolution
LTE-A	Long Term Evolution Advanced
MAC	Medium Access Control
MDU	Multi Dwelling Unit
MEM	Memory
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MS	Mobile Station
NACK	Negative Acknowledgement
NIC	Network Interface Card
NRT	Neighbour Relation Table
n-RN	Nested Relay Node
OAM	Operations, Administration and Maintenance
ODMA	Opportunity Driven Multiple Access
OFDMA	Orthogonal Frequency Division Multiple Access
ONT	Optical Network Termination
OS	Operating System
PCI	Physical Cell Identifier
PER	Packed Encoding Rules
PHY	PHYsical Layer
PPI	Protocol Payload Identifier
PTP	Precision Timing Protocol
QoS	Quality-of-Service
RAN	Radio Access Network
RFC	Request For Comment
R-FCH	Relay zone FCH
R-MAP	Relay zone MAP
RN	Relay Node
RNL	Radio Network Layer
RRH	Remote Radio Head
RRU	Remote Radio Unit
RS	Relay Station
RX	Receive
S1-AP	S1 Application Protocol
SCTP	Stream Control Transmission Protocol
SeGW	Security Gateway
S-GW	Serving Gateway
SON	Self Optimizing Network
TCP	Transmission Control Protocol

TDD	Time Division Duplex
TEID	Tunnel Endpoint Identifier
TLV	Type-Length-Value
TNL	Transport Network Layer
TX	Transmit
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
IP_COOP	User Plane Cooperation
u-Plane	User Plane
WiMAX	Worldwide Interoperability for Microwave Access (Inc.)
WP	Work Package
X2-AP	X2 Application Protocol

1 Introduction

Within Artist4G WP1 and WP2, a large number of innovations targeting interference avoidance and exploitation have been proposed so far. Many of them are based on the principle of coordination or cooperation among neighbouring evolved Node Bs (eNBs) or Home evolved NodeBs (HeNBs), which implies some form of information exchange between these nodes in the control- and/or user-plane. Since the standard S1 and X2 interfaces in LTE(-A) have not been designed for this additional type of traffic, WP4 has performed an analysis of the principle impact of the WP1 and WP2 innovations on the current system architecture and interfaces and come up with potential solutions for the new requirements.

In addition, Artist4G WP3 considers advanced relay technologies that go beyond what has been specified for LTE-A already by 3GPP. Especially the introduction of multi-hop and moving relays seems challenging from an architectural point of view, since it affects both the layer-3 protocol terminations in the various nodes and has an impact on signalling latency and volume. Hence, WP4 has also performed an analysis of the new requirements and compared potential solutions against each other.

In this document, the outcome of this analysis process is summarized for all three work packages. Hence, it provides valuable feedback for the partners in the other work packages on the implications and possible constraints that should be taken into account when evaluating the efficiency of their proposed innovations.

The remainder of this document is organized as follows: Chapter 2 contains a brief overview of the architecture-relevant innovations proposed in WP1-3. In chapter 3, the architectural impact of the WP1 proposals will be discussed in detail, including the cooperation channel for interference avoidance techniques, CoMP related enhancements, and cooperation aspects in case of heterogeneous networks. Alternatives for time/frequency synchronization and X2 support between femto and macro cells, as well as an X2-AP extension for FIC support, which are relevant for WP2 proposals, are then contained in chapter 4. The architectural impact of multi-hop and moving relay deployments, as proposed in WP3, is evaluated in detail in chapter 5. An overall conclusion on the current status of these cross-work-package activities is drawn in chapter 6.

2 Overview: Architecture-relevant innovations in WP1-3

2.1 Interference avoidance schemes proposed in WP1

2.1.1 Innovation overview and classification

The innovations developed in ARTIST4G WP1 intend to reduce the interference generated in the downlink or uplink of a Radio Access Network (RAN) by the design of new transmission schemes involving Multiple-Input-Multiple-Output (MIMO) techniques and/or multi-user scheduling. A cooperation scheme between base stations allows for reducing or avoiding the generated interference and involves cooperation messages exchange between at least two nodes. The considered nodes are eNBs, Home eNBs (HeNBs), central entities, and User Equipments (UEs). The cooperation message contains Channel State Information (CSI) obtained at the cooperating nodes and/or frequency selective MIMO channel estimates, both in a quantized or statistical form. Also scheduling information might be exchanged on a long or short term basis, in a UE-centric fashion or for a group of UEs. In some innovations, it is also envisioned to exchange user-plane data between several nodes. Thus, the WP1 innovations are sorted into four categories [ARTD11], namely:

- NO_COOP: the category of requirements for innovations that do not require any coordination through the RAN between eNBs.
- CP_COOP: the category of requirements for innovations that require slow (CP_COOP(S)) or fast (CP_COOP(F)) exchange of control-plane (c-plane) traffic through the RAN between eNBs. The slow cooperation does not adapt to the fast fluctuations of the frequency-selective channel. The typical rate of exchange of such slow cooperation messages may vary between 20ms and 200ms. The fast cooperation schemes are constrained by the CSI aging, i.e., the time between the transmission and the CSI measurement. It includes the processing time at each node, the transmission delay on the backhaul, the scheduling delay at the cooperating base station, and is also a function of the CSI feedback periodicity. The typical rate of exchange of such cooperation messages is lower than 20ms.
- UP_COOP: the category of requirements for innovations that require the exchange of user- plane (u-plane) traffic through the RAN between eNBs.
- HETNET_COOP: the category of requirements for innovations developed specifically for heterogeneous networks, and especially for HeNB campus or massive in-band deployments of HeNBs in an eNB network.

The mapping between the classes of WP1 innovations and the categories of requirements is presented in Table 2-1.

Table 2-1: Mapping between the classes of WP1 innovations and the categories of requirements on the RAN architecture.

	NO_COOP	CP_COOP(S)	CP_COOP(F)	UP_COOP	HETNET_COOP
Advanced signal processing innovations					
Advanced 3D-Beamforming	x	x	x		
Single-cell MU-MIMO schemes	x				
Multi-cell MU MIMO schemes			x	x	
Channel estimation					
Feedback design					
Scheduling and cross layer design innovations					
Clustering & user grouping	x	x	x		
Inter cell interference coordination	x	x			x

Coordinated scheduling		x	x		x
Scheduling for joint processing		x	x		
Game theory based scheduling		x	x		

A class of innovations considers advanced 3D beamforming schemes, which involve a new degree of freedom for generating multiple cells for an eNB. This class of innovations can be implemented in ways that do not require any cooperation between nodes when the cooperation is applied among cells of the same eNB. Slow or fast cooperation between nodes is also considered where scheduling information and CSI is exchanged between nodes when cells do not belong to the same eNB.

Inter-cell interference coordination is also addressed for heterogeneous deployments. One target is to limit the impact of the HeNB network on the eNB network, which requires in some modes of realization a central unit that gathers long term CSI and feeds back power setting strategies to the nodes. For HeNB campus deployments, a central unit is also required in order to compute the optimization of the HeNB network, or at least to allow HeNB to HeNB communication if the optimization is distributed.

Most innovations require fast or slow exchange in the c-plane between two eNBs, where the cooperation messages involve CSI and scheduling information for applying a coordinated scheduling or a coordinated choice of the MIMO transmission scheme. Of course, the amount of cooperation data varies according to the innovation, and one goal of WP1 is to select the best strategy for a given overhead and to consider advanced research topics for reducing this overhead, such as game theory.

A class of innovations considers the coordination of UEs inside a cell. Thus, so-called single-cell MU-MIMO schemes do not involve any cooperation between nodes. However, the concept is extended to multi-cell MU-MIMO schemes, which require fast exchange of c- and u-plane data between cooperating nodes. One goal of WP1 is to identify the portion of UEs taking benefit from this class of innovation, in order to evaluate the overhead of such techniques.

For more details on the innovations or the respective requirements on the RAN architecture, the reader is referred to [ARTD11] [ARTD12], and [ARTD13].

2.1.2 Rating of architectural impact of innovations

The WP1 innovations that put the strongest requirements on the architecture belong to the CP_COOP(F), UP_COOP and HETNET_COOP classes, which are the main targets of chapter 3. The main challenges of CP_COOP(F) are to provide new cooperation interfaces between eNBs, that support low latency exchanges of measurements and control and faster protocol to build them. In order to support user plane cooperation between two eNBs(UP_COOP), other solutions than forwarding traffic from one eNB to another must be found. Finally, the number of radio neighbours seen by an eNB when a massive co-channel deployment of small cells is operated under its coverage is the fundamental limit that must be circumvented.

In conclusion, chapter 3 will address the following WP1 innovations listed in Table 2-2.

Table 2-2: Mapping between the classes of WP1 innovations and the categories of requirements on the RAN architecture.

	CP_COOP(F)	UP_COOP	HETNET_COOP
Advanced 3D-Beamforming	x		
Multi-cell MU MIMO schemes	x	x	
Clustering & user grouping	x		
Inter cell interference coordination			x
Coordinated scheduling	x		x

Scheduling for joint processing	x		
Game theory based scheduling	x		

2.2 Interference exploitation schemes proposed in WP2 linked to WP1

The innovations developed in ARTIST4G WP2 intend to take advantage of the interference phenomena taking into account the UE terminal side. They have been clustered into four main groups:

- Channel estimation
- Interference cancellation
- Factor graph based receivers
- Cooperative receivers

The list of all the innovations can be found in section 4.2 of [ARTD24]. Due to their own definition, in the UE side, these innovations have no impact on the architecture studies on WP4.

Nevertheless, even though WP2 innovations alone have no impact on architecture, the mix of both WP1 and WP2 innovations have a clear impact from the architecture point of view. This mix has been named Flexible Interference Control (FIC).

The FIC concepts proposed in WP2 are intended to combine the interference cancellation capabilities supported by advanced receiver architectures defined in WP2 with three different interference avoidance schemes proposed by WP1: joint transmission (i.e., multi-cell MIMO), beamforming and coordinated scheduling.

Four main groups have been identified:

- Joint Transmission Interference Cancellation Receiver (JTICR) (FIC1).

  The concept proposed here is based on the combination of Joint Transmission and successive interference cancellation receivers for a more efficient operation of the downlink. The concept proposal is based on the idea of switching from three different states:

  - Baseline state: single cell SU/MU-MIMO.
  - Inter-cell Joint transmission (JT) state.
  - Inter-cell Interference Cancellation Receiver (ICR) state.

  Transitions between the three states are possible, in both directions. One possible case of use is the UE leaving a CoMP cluster, so Joint Transmission is no longer supported. A second kind of event that can trigger is an UE connected to a JP cluster entering in the area of coverage of CSG femtocell (and consequently not being able to perform handover/reselection to it). A third kind of event that may be associated with switching states is that of UE connected to a femtocell that enters an area where the best server is a macrocell.

- Coordinated beamforming in combination with ABS (FIC2).

  The idea for this concept is based on the fact that almost blank subframes (ABS) are used in heterogeneous deployments with range expansion in downlink direction, to avoid interference to resources used for data channel transmissions. But there is still the need for interference cancellation due to CRS interference. The FIC concept proposes to use ABS only where it is needed and to use non-ABS subframes where data channel interference cancellation allows to do so – even if severe interference is observed.

- Coordinated scheduling for advanced receivers (FIC3).

  The concept looks at the feedback information required to support coordinated scheduling to UEs that incorporate advanced receivers. It is also proposed to use power control to facilitate the interference cancellation.

- Adaptive Resource Allocation Algorithms for Multiuser MIMO Systems with Iterative MMSE-IC Joint Decoding (FIC4).

  The concept proposed looks at resource allocation algorithms that allow to take advantage of Iterative MMSE-IC receivers at the base station to maximize the capacity in Multiuser MIMO systems.

From these four groups only the three first ones (i.e. FIC1, FIC2 and FIC3) have an impact on the architecture and are thus the ones that are going to be studied.

The requirements imposed to backhaul by these innovations are collected in the following table and will be used for their analysis from the architecture point of view.

Table 2-3: Interference avoidance and interference exploitation innovations.

Requirements	FIC 1	FIC 2	FIC 3
X2 interface between macro and femto layers	Yes	Yes	Yes
Cooperation interface between entities: - - Type of entities (femto, macros…) - - Number of entities - - Cooperation schemes - - time distribution - - …			
Enhanced backhaul capacity - bandwidth needed - - ….	No	Not for data, needed for signalling?	No
Frequency synchronization	Femtocells - 50 ppb	Subcarrier level	Yes
Time synchronization	Frame and symbol level	Frame and symbol level	Frame and symbol level
Maximum signalling latency between cooperating cells - 1 way or 2 ways? - different latency requirements (with different performances of the innovation)?	2 ms	X	
Cooperation messages (over X2AP or other): - size of information message - content	Yes	X	Yes

2.3 Advanced relay concepts proposed in WP3

2.3.1 Innovation overview and classification

The innovations developed in ARTIST4G WP3 intend to path the way for advanced relaying concepts to be added to future LTE releases beyond the Rel.10. Especially concepts that enable relays to provide capacity on top of coverage are investigated.

Since deployment of advanced relays is considered to have an impact on several protocol layers, e.g. from physical (PHY) to layer-3, the various innovations proposed in WP3 are rather diverse and cannot be clustered as easily as in the other work packages WP1 and WP2. In the following, the major grouping used so far in WP3 will be shown together with a short description of the primary scope of each element:

- Cooperative PHY Layer Signalling and Protocols: Innovations mainly target different variants of Multiple Input Multiple Output (MIMO) constellations in relay-enhanced systems. This includes aspects like distributed space time codes, which can be used by several relays, or formation of a virtual antenna array by different nodes in order to achieve high throughput.
- Forward Error Coding with Cooperative Hybrid ARQ Relays: Innovations mainly target advanced retransmission strategies with the help of relay nodes (RNs). The latter can, for example, retransmit data in case of a NACK for data previously transmitted by the donor eNB (DeNB).
- Layer 2/3 Protocol Framework for Advanced Relays: Innovations mainly target enhancements to the existing protocol framework, e.g. if carrier aggregation is to be used in conjunction with relay nodes. In addition, the impact on signalling procedures if advanced relays are to be supported is studied in detail.
- Network Coding for Advanced Relays: Innovations mainly target network coding strategies in combination with half duplex or full duplex relays. The goal is to optimize the usage of the available radio resources.
- Interference Management: Innovations mainly target inter-cell interference coordination in relay-enhanced systems. The goal is to cope with the newly generated cell edge users at the boundary between relay and macro coverage areas, as well as between two neighbouring relay coverage areas.
- Cooperation with Relays: Innovations mainly target different levels of cooperation between DeNBs and RNs. This includes aspects like cooperative scheduling and coordinated multipoint transmission in relay-enhanced systems, both on the backhaul and access link.
- QoS & Scheduling: Innovations mainly target quality-of-service (QoS) aware resource allocation and scheduling in relay-enhanced networks. The goal is to apply QoS constraints efficiently across two hops between DeNB and user equipment (UE).
- In-band and Out-band Relay: Innovations mainly target the comparison of different in-band and out-band configuration and resource allocation schemes for relay nodes. Among others, the impact of carrier aggregation is studied in detail.
- Multi-hop Relays: Innovations mainly target extension of relay-enhanced deployments to more than two hops between the DeNB and the UE. The goal is to further increase the cell capacity and extend the coverage by optimal resource allocation for mesh and tree-like structures.
- Moving Relays: Innovations mainly target relays that are no longer statically deployed, but may move between different locations and thus macro cells. The focus here is on dynamic procedures for radio resource management and configuration of the backhaul link. In addition, the impact of moving relays on standard S1 and X2 handover procedures is studied.

Due to the inherent heterogeneity among the above listed groups of innovations, the list of requirement categories introduced for WP1 and WP2 innovations in section 2.1 has been expanded to include further aspects. Thus, the WP1 innovations are sorted into four categories [ARTD11], namely:

- PROT: Category for innovations that require a modification and/or expansion of existing signalling protocols.
- ARCH: Category for innovations that require new network elements and/or protocols in the existing radio access network (RAN) or core network (CN) architecture.

The mapping between the groups of WP3 innovations and the categories of requirements is presented in Table 2-4.

Table 2-4: Mapping between the groups of WP3 innovations and the categories of requirements on the RAN architecture.

	NO_COOP	CP_COOP (S)	CP_COOP (F)	UP_COOP	HETNET_ COOP	PROT	ARCH
Cooperative PHY Layer Signalling and Protocols:		x	x	x			
Forward Error Coding with Cooperative Hybrid ARQ Relays		x		x			
Layer 2/3 Protocol Framework for Advanced Relays						x	x
Network Coding for Advanced Relays		x	x	x			
Interference Management		x	x			x	
Cooperation with Relays		x	x	x			
QoS & Scheduling						x	
In-band and Out-band Relay						x	
Multi-hop Relays						x	x
Moving Relays						x	x

2.3.2 Rating of architectural impact of innovations

A survey on the specific needs for the different requirement categories among WP3 partners has reveiled that the impact of the following categories

- CP_COOP(S)
- CP_COOP(F)
- UE_COOP

is more or less similar to WP1 and WP2 innovations. For this reason, no further studies seem to be required here beyond what is already done for the latter two work packages.

On the contrary, especially the two new requirement categories introduced for WP3 seem to deserve special attention. For the case of multihop relays, this concerns a trade-off analysis of possible deployment architectures, which will be treated in detail in section 5.2.

For the case of moving relays, two aspects seem to be important: The impact on mobility-related procedures, e.g. S1- and X2-based handover, as well as the additional latency that might occur in the control plane signalling. Both of these issues will be treated in detail in section 5.3.

3 Architectural impacts of WP1 proposals

3.1 Cooperation channel for interference avoidance techniques

3.1.1 State of the art

3.1.1.1 Restrictions on setting up an X2 interface

It is assumed that an eNB establishes X2 interfaces using the Neighbour Relation Table (NRT) which is managed by the NRT Management Function based on information provided by UEs when performing measurements (Detected Cell reporting) and also taking into consideration information provided by Operations, Administration and Maintenance (OAM) about further details on the use of a certain X2 interface. Furthermore, it needs to be kept in mind that neighbour cell relations, i.e. the table entries in the NRT, are on a per cell basis, whereas X2 interface is on a per eNB basis ([3GPP36300] clause 22.3.2a).

The number of neighbour cell relations for a given cell of an eNB with cells of a neighbouring eNBs is subject to technical restrictions, and also enforced by ASN.1 definitions. The reason is that the Physical Cell Identifier (PCI) broadcasted by a neighbour cell needs to be unique at least among all the direct neighbours of the source cell and the maximum number of possible PCI values is limited. In the eNB case this restriction is not considered problematic, however, the number of X2 interfaces is likely to be limited in real eNB implementations for practical reasons. It also needs to be taken into consideration that in contrast to the eNB case, the HeNB only supports one single cell. Consequently an eNB with a large number of HeNBs in its coverage area would need to have as many X2 interfaces as there are HeNBs. Note that an eNB can cope with PCI collisions for example by asking a UE to report a neighbor cell's global cell ID.

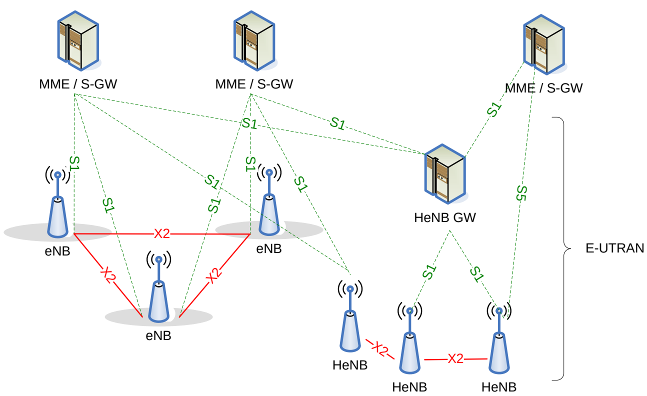


Figure 3-1: Overall E-UTRAN architecture with deployed HeNB-GW [3GPP36300].

In summary the status in 3GPP Release 10 is as follows:

- Generally there are no architectural restrictions regarding establishment of an X2 interface between any two (H)eNBs.
- X2 interface is supported between HeNBs, without any restriction, even in case the neighbouring HeNBs broadcast different CSG-IDs. Restrictions apply on the use of an X2 interface for HO to cases where no access control in the CN is needed. Whether the HeNB is connected to an HeNB-GW is of no importance as the HeNB-GW is an optional element (see [3GPP36300] clause 4.6.1).

- HeNB GW does not support X2 interface, i.e. the HeNB-GW is not involved in any X2 signalling, specifically it is also not acting as X2 proxy where it might act as kind of concentrator to reduce the number of X2 interfaces towards an HeNB. Consequently no X2 interfaces shall be established between the HeNB GW and any other node (see [3GPP36300] clause 4.6.2).
- HeNB can establish a direct X2 interface to a neighbouring HeNB, , i.e. not through a HeNB GW and an X2 interface of it's own per neighbouring HeNB.
- No security solution has been specified by 3GPP SA3 to allow communication between HeNBs inside the private domain only, except via the Security Gateway at operators' premises.

The current restrictive use of X2 interfaces in HetNet scenarios is likely to be removed in Release 11, especially as some operators are pushing for it.

Additionally an X2 proxy function has already been proposed for introduction. This simplifies the establishment of X2 relations from eNBs to HeNBs as an eNB needs to have only one X2 interface towards the HeNB-GW but not individual X2 interfaces per HeNB [3GPP-R3102149] and [3GPP-R3102360].

As there is no X2 between HeNB-GWs, HeNBs served by different HeNB-GWs would still need direct X2 between each other.

3.1.1.2 Traffic types on the X2 interface

On the X2 interface c- and u-plane traffic could be carried. Current motivation of u-plane traffic sent via the X2 interface is along with HO procedure. Consequently, the u-plane tunnel is established in case of HO procedure and lives only for a limited period in time.

Some of the WP1 innovations in contrast need to have the u-plane established all the time the communication relation is in existence. The complete u-plane information might need to be available at both (or even more) (H)eNBs, once received by the anchor (H)eNB from the core network, then multiplied and passed to the partner (H)eNBs according to the information received from WP1.

Current X2 protocol does not foresee any mechanism required for syntonisation/synchronisation of the radio transmitters in the (H)eNBs involved in WP1 innovations for interference avoidance. Current assumption in H(e)NB thinking is that this is achieved via non 3GPP protocols, e.g. Network Time Protocol (NTP), Precision Timing Protocol (PTP) or a Global Navigation Satellite System (GNSS)) (see 4.1).

Considering u-plane traffic on the X2 interface, there are three possible cases according to WP1 innovations:

1. The cooperating eNB receives a complete copy of the u-plane packets forwarded from the anchor eNB.
2. The cooperating eNB receives selected u-plane packets, for example, to achieve additional redundancy.
3. The cooperating eNB receives selected u-plane packets, for example, to achieve load sharing.

In addition to the pure aspect of forwarding u-plane packets between cooperating eNBs, the delay caused by the communication overhead between both nodes is critical and thus limits the use of the X2 interface. The maximum acceptable delay is imposed by the PHY layer based on the following two aspects:

- Channel State Information (CSI) aging: time between the measurement and the actual transmission based on this measurement. It includes the processing time at each node, the transmission delay on the backhaul, the scheduling delay at the cooperating base station, and is also a function of the CSI feedback periodicity (see [ARTD11] for more details).

- Delay constraint from the scheduling depth: when two eNBs cooperate with a coordinated scheduling approach and a decision is made at one of them it must be transferred and processed by the other node in a time allowing for the cooperation strategy to be taken into account by schedulers at both nodes. This also depends on the cell load.

In summary the following limitations apply to the X2 interface in 3GPP Release 10:

- Only c- and u-plane traffic could be sent via the X2 interface.
- The X2 interface is used to carry u-plane traffic only in case of HO.
- Traffic on the X2 interface suffers from additional transmission delay, depending on the physical path chosen for the X2 interface.
- X2 interface is not foreseen to support syntonisation/synchronisation of (H)eNBs, this has to be provided via other communication channels and using non 3GPP protocols or other means.

Possibilities to overcome these limitations in inter-eNB communication, which hinder implementation of WP1 innovations, will be discussed in section 3.1.4.

3.1.1.3 Transport layer protocols considered on the RAN

The relevant[1] Internet transport protocols standardized by the IETF are UDP (User Datagram Protocol, [RFC768]), TCP (Transmission Control Protocol, [RFC793]) and the relatively new SCTP (Stream Control Transmission Protocol, [RFC4960]):

- UDP: unreliable connectionless datagram transport
    - connectionless: no connection setup/termination, just exchange of individual datagrams.
    - unreliable: no protection against reordering, duplication or loss.
    - message-oriented: framing of data into packets by sender is retained and visible to receiver.
- TCP: reliable connection- and stream-oriented transport
    - connection-oriented: connection setup, data exchange with flow and congestion control, connection termination.
    - reliable: in order delivery without duplicates or losses.
    - stream-oriented: byte stream, i.e. no framing of user data.
- SCTP: reliable, message-oriented, multi-streaming and multi-homing capable transport; from the abstract of RFC 4960:
    - acknowledged error-free non-duplicated transfer of user data,
    - data fragmentation to conform to discovered path Maximum Transmission Unit (MTU) size,
    - sequenced delivery of user messages within multiple streams, with an option for order-of-arrival delivery of individual user messages,
    - optional bundling of multiple user messages into a single SCTP packet, and
    - network-level fault tolerance through supporting of multi-homing at either or both ends of an association.

UDP is the simplest transport protocol. It transparently passes through the characteristics of the underlying Internet Protocol (IP), just adding multiplexing capability through ports and a checksum that is mandatory for IPv6 and optional for IPv4 (all zeros in checksum field indicates

[1] Other transport protocols are: RDP (Reliable Data Protocol, RFC 1151, experimental, reliable datagram transport), UDP-Lite (RFC 3828, partial checksum feature) and DCCP (Datagram Congestion Control Protocol (RFC 4332, unreliable datagram transport like UDP with TCP-like congestion control).

no checksum). UDP is used by GTPv1-U [3GPP29281], which in turn provides transport for the u-plane interfaces of E-UTRAN and EPC, e.g. S1-U between eNB and Serving Gateway (S-GW) and X2-U between eNBs. Furthermore, UDP is used for end-to-end communications by the UE, e.g. for Domain Name System (DNS) name resolution.

In contrast, TCP heavily adds functionality, especially reliability and congestion control. No interface of E-UTRAN and EPC uses TCP, but most of the end-to-end communications by the UE do, e.g. with web or mail servers in the Internet.

SCTP was specifically designed for signalling transport, where it is superior to TCP due to its features:

a) message-oriented like UDP (while reliable as TCP), thereby providing framing to higher layers,
b) multi-streaming prevents head-of-line blocking, as opposed to the single stream transport in TCP, i.e. retransmission of a lost message in one stream does not block otherwise independent messages in other streams,
c) an SCTP association endpoint can be multi-homed, i.e. can have multiple transport addresses each of which can possibly be reached over a different path, thereby adding additional robustness against network path failures.

SCTP is used for the signalling interfaces of E-UTRAN and EPC, e.g. S1-MME between eNB and Mobility Management Entity (MME) and X2-C between eNBs. Since SCTP is relatively new, it is typically not used for end-to-end communications by Internet hosts like UEs, although in principle it is well suited for many of them, e.g. for web traffic.

Possible transport options for a new X2 interface are discussed in section 3.1.4.3.

3.1.1.4 Services and applications considered on the RAN

3GPP standardized 9 different QoS Class Identifier (QCI) for IP-Connectivity Access Networks (IP-CANs) like LTE-A, as shown in Table 3-1. The associated performance metrics Packet Delay Budget (PDB) and Packet Error Loss Rate (PELR) are measured between UE and Policy and Charging Enforcement Function (PCEF), a logical part of the PDN Gateway (P-GW). The given performance targets are therefore not valid end-to-end between Internet hosts, i.e. between UE and a server or another UE, but only for the network path between UE and the boundary of the Evolved Packet System (EPS) respectively Public Land Mobile Network (PLMN), to the Internet, cf. Figure 3-2.

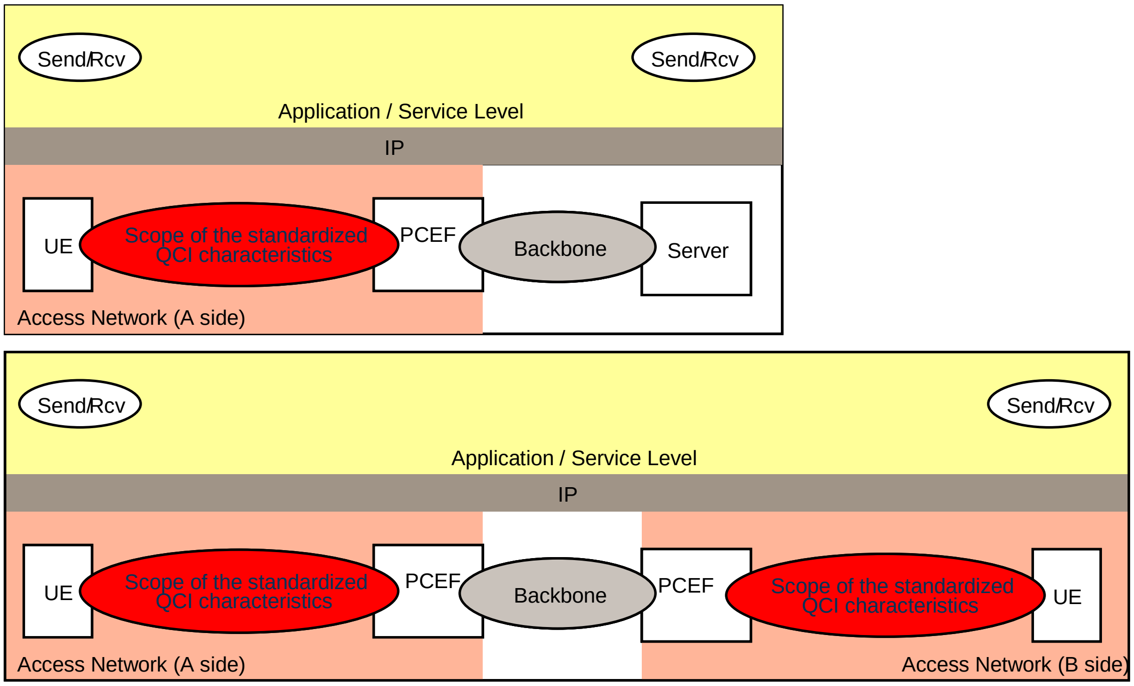


Figure 3-2: Scope of the standardized QCI characteristics for client/server (upper figure) and peer/peer (lower figure) communication (Figure 6.1.7-1 from [3GPP23203], section 6.1.7.2 Standardized QCI characteristics).

Table 3-1: Standardized QCI characteristics (Table 6.1.7 from [3GPP23203], section 6.1.7.2 Standardized QCI characteristics).

QCI	Resource Type	Priority	Packet Delay Budget (NOTE 1)	Packet Error Loss Rate (NOTE 2)	Example Services
1 (NOTE 3)	GBR	2	100 ms	10^{-2}	Conversational Voice
2 (NOTE 3)		4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3 (NOTE 3)		3	50 ms	10^{-3}	Real Time Gaming
4 (NOTE 3)		5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
5 (NOTE 3)	Non-GBR	1	100 ms	10^{-6}	IMS Signalling
6 (NOTE 4)		6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7 (NOTE 3)		7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8 (NOTE 5)		8	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
9 (NOTE 6)		9			

NOTE 1: A delay of 20 ms for the delay between a PCEF and a radio base station should be subtracted from a given PDB to derive the packet delay budget that applies to the radio interface. This delay is the average between the case where the PCEF is located "close" to the radio base station (roughly 10 ms) and the case where the PCEF is located "far" from the radio base station, e.g. in case of roaming with home routed traffic (the one-way packet delay between Europe and the US west coast is roughly 50 ms). The average takes into account that roaming is a less typical scenario. It is expected that subtracting this average delay of 20 ms from a given PDB will lead to desired end-to-end performance in most typical cases. Also, note that the PDB defines an upper bound. Actual packet delays - in particular for GBR traffic - should typically be lower than the PDB specified for a QCI as long as the UE has sufficient radio channel quality.

NOTE 2: The rate of non congestion related packet losses that may occur between a radio base station and a PCEF should be regarded to be negligible. A PELR value specified for a standardized QCI therefore applies completely to the radio interface between a UE and radio base station.

NOTE 3: This QCI is typically associated with an operator controlled service, i.e., a service where the SDF aggregate's uplink / downlink packet filters are known at the point in time when the SDF aggregate is authorized. In case of E-UTRAN this is the point in time when a corresponding dedicated EPS bearer is established / modified.

NOTE 4: If the network supports Multimedia Priority Services (MPS) then this QCI could be used for the prioritization of non real-time data (i.e. most typically TCP-based services/applications) of MPS subscribers.

NOTE 5: This QCI could be used for a dedicated "premium bearer" (e.g. associated with premium content) for any subscriber / subscriber group. Also in this case, the SDF aggregate's uplink / downlink packet filters are known at the point in time when the SDF aggregate is authorized. Alternatively, this QCI could be used for the default bearer of a UE/PDN for "premium subscribers".

NOTE 6: This QCI is typically used for the default bearer of a UE/PDN for non privileged subscribers. Note that AMBR can be used as a "tool" to provide subscriber differentiation between subscriber groups connected to the same PDN with the same QCI on the default bearer.

3.1.2 Physical layer constraints

In the mobile backhaul network, three main categories can be distinguished, regarding the medium support:

- Copper cables
- Optical fibre
- Microwave

The capabilities (delay, bandwidth) of each medium will be analyzed in the following parts, for several technologies, already available or expected.

3.1.2.1 Copper

Digital Subscriber Line (DSL) technologies are extensively deployed in today's mobile backhaul (mostly Symmetrical High Bitrate Digital Subscriber Line (SHDSL)), but will reach their limits for LTE-A deployment. By bonding pairs, capacity can be increased roughly linearly, but the number of pairs is not unlimited on the existing sites, and the distance reached is still quite limited today. Another limiting factor is the non symmetrical bandwidth. Also, the delay values are very dependent on the configured profile, which means that strong optimization and field tests need to be performed in order to comply with LTE-A delay requirements. Table 3-2 describes the main features of the key technologies.

Table 3-2: Copper technologies capabilities.

<table>
<tr><th>Technology</th><th>Capacity</th><th>Latency</th><th># of pairs</th><th>Distance</th></tr>
<tr><td rowspan="2">Bonded G.GHDSL.Bis</td><td>20 Mb/s</td><td rowspan="2">~1 ms (without protection against impulsive noise)</td><td>4</td><td rowspan="2">~ 1 km</td></tr>
<tr><td>40 Mb/s</td><td>8</td></tr>
<tr><td rowspan="3">Bonded ADSL2+</td><td>UL: 40 Mb/s
DL: 2 Mb/s</td><td rowspan="3">min: 4 ms (without protection)
max: 60 ms (with protection)</td><td>2</td><td rowspan="2">max 1 km</td></tr>
<tr><td>UL: 160 Mb/s
DL: 8 Mb/s</td><td>8</td></tr>
<tr><td>UL: 80 Mb/s
DL: 8 Mb/s</td><td>8</td><td>2-3 km</td></tr>
<tr><td rowspan="2">VDSL2</td><td>Profile 17a
UL: 50 Mb/s
DL: 20 Mb/s</td><td rowspan="2">idem ADSL2+</td><td rowspan="2">1</td><td>max 750 m</td></tr>
<tr><td>Profile 30a
UL: 100 Mb/s
DL: 40 Mb/s</td><td>max 350 m</td></tr>
<tr><td rowspan="2">Vectored VDSL2</td><td>Profile 17a
UL: 50 Mb/s
DL: 20 Mb/s</td><td rowspan="2">idem ADSL2+</td><td rowspan="2">1</td><td>max 800 m</td></tr>
<tr><td>Profile 30a
UL: 100 Mb/s
DL: 40 Mb/s</td><td>max 500 m</td></tr>
<tr><td>Bonded VDSL2 + Vectoring</td><td>Profile 17a
UL: 50 Mb/s
DL: 20 Mb/s</td><td>idem ADSL2+</td><td>2</td><td>max 900 m</td></tr>
</table>

	Profile 30a UL: 100 Mb/s DL: 40 Mb/s			max 800 m
	Profile 17a UL: 100 Mb/s DL: 40 Mb/s		4	max 900 m
	Profile 30a UL: 200 Mb/s DL: 80 Mb/s			max 800 m

3.1.2.2 Microwave

Microwave is also commonly deployed in today backhaul. The capabilities depend on technology of course, but also on modulation (e.g. 256QAM, QPSK...), channel frequency (6 GHz, 36 GHz, 80 GHz…), and channel bandwidth (56 MHz, 250 MHz, 1 GHz…). There are too many microwave solutions to detail all associated capabilities in a table. But standard technologies commonly used can reach a capacity up to 400 Mb/s, on quite a long distance. And solutions up to 1 Gb/s are already in use, thanks to new sophisticated formats (e.g. 256/512 QAM) and waves (e.g. E-Band @80 GHz), but on shorter distance. Microwave solutions have also a low latency (~100 µs).

3.1.2.3 Optical Fibre

Optical fibre is the preferred medium for high bandwidth, low delay and long distance communications. The capabilities of the two available topologies (point-to-point and point-to-multipoint) are described in Table 3-3.

Table 3-3: Fibre technologies capabilities.

Technology	Capacity	Latency	Distance
Optical Ethernet	100 Mb/s	5 µs/km	80 km
	1 Gb/s		100 km
	10 Gb/s		80 km
GPON	**Shared** DL : 2.48 GB/s UL : 1.24 Gb/s	~100 µs	20 km
XG-PON1	**Shared** DL : 10 GB/s UL : 2.5 Gb/s	~100 µs	20 km

3.1.2.4 Suitability to LTE-A

The physical layer in mobile backhaul poses strong constraints concerning delay and bandwidth. To comply with LTE-A capacity, the best candidates are optical fiber (Gigabit Passive Optical Network (G-PON) or direct link) and microwave. The delay generated by these physical layers is small enough (tens of microseconds) and thus there is no need to optimize it.

However, copper may also be used in some cases (bonded copper pairs …). The delay and bandwidth here strongly depend on the Digital Subscriber Line (DSL) technology (Asymmetric DSL (ADSL2+), Very high bit-rate DSL (VDSL2), Vectored VDSL2 …), the bonding (number of copper pairs) and the configuration (protection, …). But as listed in Table 3-4, for micro eNBs, working with 10 MHz bandwidth and supporting cat3 UEs (see Figure 3-3 for UE categories description), bonded VDSL2 technology can support LTE traffic via short distances. For other configurations, the limits of copper technology are reached and other technologies are needed: For example, with cat4 UEs in 20 MHz, microwave and fibre will be the only solutions to support 150 Mbit/s peak rate.

<table>
<tr><th colspan="2">Category</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th></tr>
<tr><td rowspan="2">Peak rate Mbps</td><td>DL</td><td>10</td><td>50</td><td>100</td><td>150</td><td>300</td></tr>
<tr><td>UL</td><td>5</td><td>25</td><td>50</td><td>50</td><td>75</td></tr>
<tr><td colspan="7">Capability for physical functionalities</td></tr>
<tr><td colspan="2">RF bandwidth</td><td colspan="5">20MHz</td></tr>
<tr><td rowspan="2">Modulation</td><td>DL</td><td colspan="5">QPSK, 16QAM, 64QAM</td></tr>
<tr><td>UL</td><td colspan="4">QPSK, 16QAM</td><td>QPSK, 16QAM, 64QAM</td></tr>
<tr><td colspan="7">Multi-antenna</td></tr>
<tr><td colspan="2">2 Rx diversity</td><td colspan="5">Assumed in performance requirements.</td></tr>
<tr><td colspan="2">2x2 MIMO</td><td>Not supported</td><td colspan="4">Mandatory</td></tr>
<tr><td colspan="2">4x4 MIMO</td><td colspan="4">Not supported</td><td>Mandatory</td></tr>
</table>

Figure 3-3: LTE UE categories [3GPP36306].

Table 3-4: Technology capacity summary.

Admissible technologies	DL limit (Mbit/s)	UL limit (Mbit/s)
All	100	40
Bonded Vectored VDSL2 + Microwave + Optical Fibre	200	80
Microwave + Optical Fibre	400	400
E-band Microwave + Optical Fibre	10 000	10 000

3.1.3 Security constraints

The LTE / EPC security architecture is defined in [3GPP33401]. From the UE point of view, the Radio Link Control (RLC) and the Packet Data Convergence Protocol (PDCP) layer dealing with radio signalling and u-plane data transport are terminated in the eNB, where u-plane data is transmitted to or received from the Core Network. The radio protection (encryption of the u-plane; encryption and integrity protection of the RRC radio signalling) therefore only applies between UE and eNB.

In contrast, the NAS (Non-Access Stratum) layer dealing with mobility signalling is terminated in the MME. Hence, NAS messages are protected (encryption and integrity protection) end-to-end between UE and MME, which means that the eNB cannot access the content of these messages. Instead, it forwards them transparently to the MME.

From the eNB point of view, communication with S-GW, MME, and neighbour eNBs is not protected by default: On the S1-MME interface, only NAS messages are end-to-end protected between UE and MME, as already mentioned above. Otherwise all signalling messages transported directly over S1-AP between eNB and MME are not protected. The same is true for signalling via X2-AP on the X2-C interface, as well as for u-plane data on the S1-U and X2-U interface. Hence, 3GPP recommends to apply the Network Domain Security specifications (i.e.

Internet Protocol Security (IPsec)) on the backhaul links carrying the above interfaces, if the backhaul network is considered as untrusted.

3GPP has chosen to use SCTP [RFC4960] as transport protocol for S1-AP and X2-AP applications. This protocol has been designed by the Internet Engineering Task Force (IETF) to efficiently manage multiplexing of streams, multi-homing, and packet reordering. An SCTP association is initiated with a mechanism that theoretically prevents spoofing of source IP addresses and injection of traffic into an existing SCTP association by an external node. However, security threats like SCTP association hijacking still exist, as explained in [RFC5062]. Furthermore, by design, SCTP does not provide any confidentiality and also no integrity protection of the data exchanged.

Traffic can be protected at different levels: at application level, at network level, or at transport level. Multiplication of protection (e.g. encapsulation of 2 IPsec tunnels) is useless and to be avoided. Hereafter, the interest is to see whether it is useful to protect the traffic between E-UTRAN and EPC at transport level. First of all, two kinds of u-plane traffic can be identified, depending on the service architecture that is going to be deployed with the EPC network:

1. Case of a normalized packet service core (IP Multimedia Subsystem (IMS)): The u-plane carries Session Initiation Protocol (SIP) signalling and mainly Real-Time Transport Protocol (RTP) and RTP Control Protocol (RTCP) media flows. The IMS standard implies the mandatory protection of SIP signalling via IPsec between UE and Proxy Call Session Control Function (P-CSCF). For the protection of RTP/RTCP streaming traffic, a specific 3GPP standard is on going on IMS media security and will define 2 solutions:
    - One based on a protection handled in IMS equipments (under operator control).
    - One which is independent of IMS equipments (under third party control such as the police or other National Security and Public Safety services). This protection of RTP/RTCP traffic is optional in IMS network.
2. Case of a non standardized packet service core (case of access to the Internet): Here, u-plane confidentiality protection will depend on each UE and each service platform characteristics: Whether they support traffic protection, via Transport Layer Security (TLS) or Hypertext Transfer Protocol Secure (HTTPS) at application level, or IPsec at network level. If this type of protection at application or network level is considered sufficient, further protection of u-plane traffic at transport level across the S1-U interface is not useful.

Note that for any of the above cases, the signalling traffic in the c-plane across S1-MME and X2-C will have to be protected at transport level.

From a security perspective, three different types of eNBs can be distinguished:

- Trusted eNB: These are eNBs deployed by the Mobile Network Operator (MNO) in his trusted network, with physical security (e.g. lockers with intrusion alarms).
- Non-trusted eNB: These are eNBs deployed by the MNO, but without efficient physical security (e.g. pico cells on walls).
- HeNB: These are located and operated inside customer premises.

Correspondingly, an operator has different options regarding the security of mobile traffic over the transport network:

- Option A: He could decide not to protect the mobile traffic at all. With this option, the operator considers that the eNB is located in a trusted area (e.g. thanks to a physical locking of the eNB) and leave the traffic going transparently. In this case, all the security risks have to be accepted by the MNO and mitigated by physical means. This option can be used for Trusted eNBs.
- Option B: He could decide to protect only the c-plane on the S1-MME and X2-C interfaces. With this option, the operator will protect the eNB and the MME from any

malicious attacks on the signalling. This is not necessary for Trusted eNBs, but strongly recommended for Non-Trusted eNBs and/or HeNBs.

- Option C: He could decide to protect both c- and u-plane on the S1 and X2 interfaces. With this option, the operator provides additional confidentiality on the u-plane to all mobile traffic. However, this option presents many constraints to be applied due to the large volume of u-plane traffic, the low-latency needed for it, the overhead inherent to IPsec encryption, and the costs related to the Security Gateway (SEG) dedicated to u-plane traffic protection. This is not necessary for Trusted eNBs.
- Mix of options: Multi cell (macro + pico/micro) deployment may imply a mix of options. This will be the case if the macro eNBs are in a trusted area (Trusted eNBs), whereas pico/micro eNBs are located in insufficiently trusted areas (Non-Trusted eNBs). This case may necessitate 2 different options: option A for macro, option B/C for pico/micro eNBs. The consequence is additional complexity in the transport design due to separation of the flows.

NB: The IPsec Encapsulating Security Protocol (ESP) can be either used in transport mode or tunnel mode. In the first case, the IPsec security applies end-to-end between two nodes. In the 2nd case, an internal IP header is added in the ESP layer which allows a node to be routed over a trusted network after passing an IPsec security gateway. IPsec ESP transport mode is more suitable for X2-C interface protection while IPsec ESP tunnel mode is more suitable for S1-MME interface protection. However, IPsec ESP tunnel mode can be used for X2-C interface security as well, if the X2-C traffic transits through an IPsec security gateway; it can also be used for end-to-end security, with specific IP configuration on the two nodes. The tunnel mode has the benefit of hiding the network topology and IP routing existing on the transport and core network, but has also the disadvantage of an additional IP header in the ESP layer and therefore increased overhead. 3GPP mandates the support of IPsec ESP in tunnel mode in eNBs.

3.1.4 Towards a new cooperation interface

The WP1/WP2 innovations require the cooperating eNBs to have a permanent communication channel available between them, adding an as low delay as possible on top of the pure end-to-end delay due to transport layer. Current 3GPP specifications only define the X2 interface between eNBs as shown in the figure below. Only the c-Plane is permanently available, once established. Unidirectional u-Planes are established per UE whenever required for Hand-over. Hence considerations for possible re-use need to start here. Additionally the evolution of core networks towards use of carrier grade Ethernet needs to be taken into account.

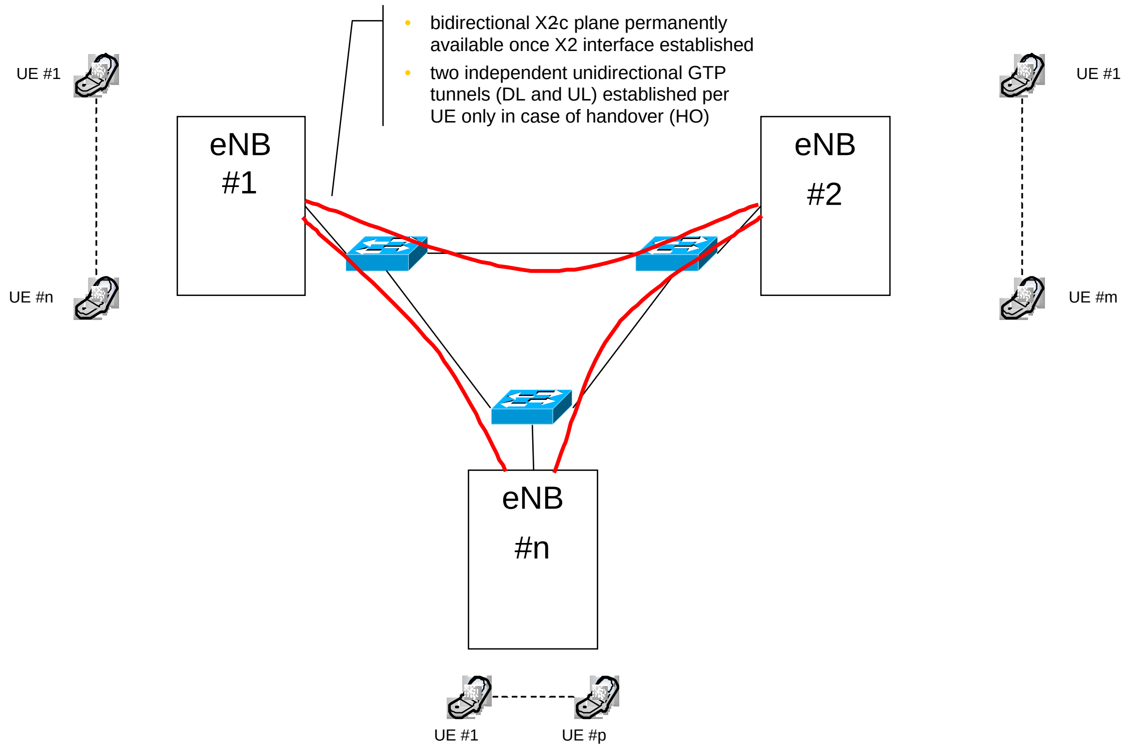


Figure 3-4 Meshed network to support Artist4G innovations.

As for the S1 interface, the X2 interface provides for a control plane that is permanently available, once established. Additionally unidirectional GTP tunnels might be established per UE in case of handover.

In contrast to S1/X2 practise, u-plane data exchange needs to be permanently available but no requirement for having an UE individual communication channel was found. Note, even a common channel seems to be sufficient, means to allow identification of UE individual u-plane is considered necessary.

3.1.4.1 Considerations based on existing protocols and architecture

Innovations regarding advanced signal processing algorithms for interference avoidance have been described in detail in [ARTD11]. A common characteristic of the “cooperative” algorithms is the requirement to have a communication channel between those nodes allowing them to exchange necessary information amongst them.

The term “necessary information” is used as a generic term, as the type of information exchanged varies depending on the algorithm applied. It might comprise both u-plane and also c-plane information. C-plane information might also well include new kind of interference control information exchanged between cooperating nodes.

In short the requirements on this inter eNB communication channel could be summarized as:

- low delay and jitter;
- provide a bidirectional communication between eNBs in the cooperation area according to the needs of the innovation:
    a. star type communication for innovations using a centralised approach;
    b. mesh type communication for innovations using a decentralised approach;
- permanently available channel, independent whether the eNB serves UEs or not;
- communication channel between cooperating eNBs is used for all UEs served by the two peer eNBs. There is no dedicated channel per UE.

As cooperating nodes need to also be neighbours of each other, they typically have an X2 interface established between them.

The current primary usage of the X2 interface between neighbouring nodes is to support the handover (HO) procedure for UEs when moving from the currently serving eNB to one of its neighbours and SON procedures between eNBs. Triggered by the X2 HO process a temporary forwarding of UE user plane information in downlink (DL) direction is enabled. The temporary forwarding is stopped as soon as the DL path has been switched by the S-GW towards the new serving eNB. It needs to be recognised that DL/UL tunnels as currently defined for X2 HO are dynamically setup per UE.

In contrast to this temporary establishment of an unidirectional DL path from source to the target, the algorithms for cooperative interference avoidance and Flexible Interference Cancellation (FIC) require a permanent bidirectional communication possibility between all eNBs in the cooperation area. Furthermore, the information exchange has stringent near real-time requirements depending on the applied innovation.

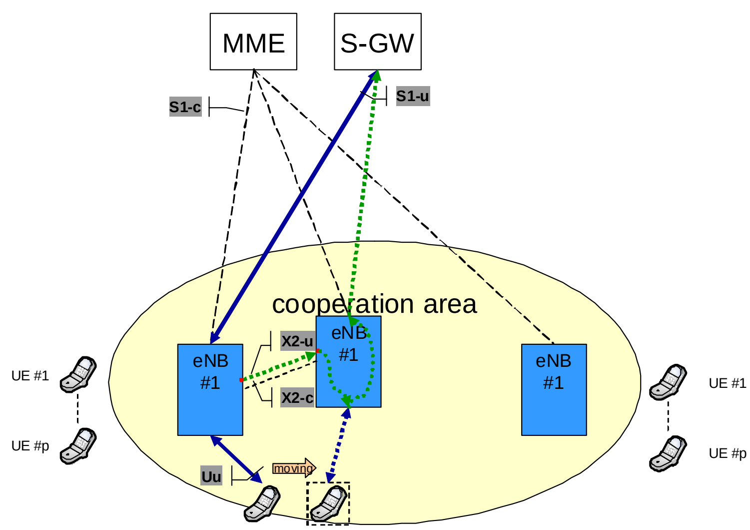


Figure 3-5: Downlink Forwarding GTP tunnel for X2 hand-over.

X2 DL forwarding allows for a UE specific unidirectional GTP tunnel from source to target eNB. Optionally also UL forwarding could be established. The TNL addresses and TEIDs of the GTP tunnel endpoints are exchanged during the X2 Handover message exchange.

It needs to be recognised that using a tunnelling protocol for UE originated/destined IP traffic is due to the fact that UEs are allowed to move, but IP routing in the operators' network must be kept untouched to remain static and stable. Taking this into account there is no justification to use the GTP tunnelling mechanism when transferring messages between stationary eNBs in the cooperation area.

- Finding:
  Do not use the GTP encapsulation for exchanging messages between eNBs in the cooperation area.

X2 HO procedure is heading to establish a user plane path from source eNB to a single target eNB for a specific UE. As this procedure is only initiated when required, adapting it to establish a UE specific cooperation (“CoMP”) channel from the serving eNB to all neighbours would result in a significant increase in signalling traffic and, even worse, endup in a completely new X2 procedure.

- Finding:
  Adapting the mechanisms for UE HO is not adequate for establishing a permanent communication channel to exchange UE specific “CoMP” information.

Taking into account that the communication channel carrying the “CoMP” information is not required to be UE specific but could be modelled as a permanently available common channel between neighbouring (H)eNBs having multiplexing capabilities. Therefore it seems preferential to consider the use of the available procedures for exchanging X2 related Transport Network Layer (TNL) information toadditionally convey the TNL addresses required for establishment of an optional common the “CoMP” channel between (H)eNBs.

- Proposal:
  Adapt the S1-AP messages for eNB and MME Configuration Transfer to also convey “CoMP” channel related TNL configuration information.

Considering the fact that UE specific information needs to be passed via the “CoMP” channel, a mechanism is needed to identify, if necessary, the UE the “CoMP” information is related to. Furthermore as the length of required “CoMP” information is flexible and could also be Bit oriented, it is necessary to identify the length of the information in bits, especially in case padding to byte boundaries has to be done. As communication requires low delay and latency the protocol overhead needs to be kept at a minimum.

- Proposal:
  Use a “Type/Length/Value” (TLV) structure to identify the information passed and length of the information field.

There has been no work done to elaborate the messages that need to be exchanged between nodes in the cooperation area, consequently also the requirements on e.g. transport aspects like reliable versus not reliable transport or use of IP addresses versus Ethernet MAC addresses are still at its infancy.

- Proposal:
  Allow for ultimate flexibility regarding TNL addresses and refrain from defining a “CoMP” AP along with related messaging for the time being.

Last but not least, as there is no logical relation between the “CoMP” interface and the X2 interface, except for the fact that “CoMP” related TNL addresses are communicated using the same procedures as for communicating X2 TNL addresses.

- Proposal:
  Do not refer to the "CoMP" interface as being integral part of X2 interface.

3.1.4.2 Requirements on the “CoMP” channel

To get an improved understanding and acquire a more complete view of the requirements raised by WP1 and WP2 innovations feedback was requested, based on a questionnaire covering the aspec

- data volume;
- message transport;
- architectural aspects

Table 3-5: Questionnaire to derive requirements on the "CoMP" channel.

Data volume
• dependency on number of UEs
• identification of source UE
• average message length per direction
• average number of messages/second/direction
Message transport
• requirement for reliable transfer
• requirement for in-sequence delivery
• requirement to support discard of "aged messages"
• multiple priorities for messages
• average/maximum transmission delay
• requirements on maximum jitter
Architecture
• "CoMP" messages carry control-plane information
• "CoMP" messages carry user-plane information
• prioritisation between "client/server" and "peer-to-peer"
• "CoMP" messages sent to all "clients"/"peers" of cooperation area
• future support for UE mobility

The responses received from WP1 so far are given below. They also show where requirements differ between the various WP1 innovations. Feedback from WP2 is currently communicated only verbally. The essential message, however, is that the responses would not be too different from those already available from WP1.

With respect to the "architecture" topic the preliminary responses received provided guidance for further analysis in WP4 regarding possible reuse of already existing SON procedures.

Data volume

- Is the data volume depending on the number of UEs served by the sending eNB?

  CP_COOP(S): No, for non-UE specific cooperation.
  Yes, when applying clustering of UEs, the cluster creation involves a traffic which depends on the number of UEs in the cluster. But, when applying the cooperation, the traffic is not dependent anymore on the number of UEs.

  CP_COOP(F): Yes, only UE-specific cooperation.

  UP_COOP: Yes

- Is it required to identify the source, i.e. the UE in the data sent?

  CP_COOP(S): For most innovations, no

  CP_COOP(F): It depends on the implementation, but in general Yes

  UP_COOP: Yes, as soon as the cooperating node needs to know the UE context for sending UP data. (X2 or Central unit cooperation?).
  No if it is allowed to send data+ressource allocation+MCS only.

- Is there a timing restriction regarding data sent on the u-Plane?

UP_COOP: Yes, the same constraints as for no cooperation (transport/application constraints). But less limited than the CP_COOP(F).

- What is the average length of a message in UL and DL? (bytes/msg)?

  CP_COOP(S): These are typically small packets sent periodically.

  CP_COOP(F): These are typically small packets sent very often. Then the amount of data really depends on the implementation (e.g., do we take the precoders from a codebook, or do we forward a quantized version…)

  UP_COOP: worst case for overhead: 1.5 to 3 x the average data rate without CoMP

- Is there a dependency on the number of UEs (bytes/UE/msg)?

  See answer above. It still needs to be clarified if grouping of several feedbacks in a single "packet" is possible.

- Average sending interval of these messages in UL and DL (msec)?

  CP_COOP(S): more than 20ms, but 100ms can also be supported.

  CP_COOP(F): around 5ms-10ms.

  UP_COOP: This is usually non periodical for cooperating UEs, and depends on the User Plane traffic. Worst case is like CP_COOP(F) (5-10 ms).

- What is the channel information? Which type of measurements does the UE perform?

  CSI (frequency selective feedback) already in Rel. 10. It needs to be clarified whether MU-MIMO measurements can be used for CoMP.

- What is the aging problem?

  The channel might have changed between the time the measurement is performed and the time the cooperation is applied, the cooperation scheme is not adapted anymore to the actual channel. This might result in worse performance than with no cooperation. This depends on the propagation channel/environment.

Message Transport

- Is reliable transfer and in-sequence delivery of the messages required?

  CP_COOP(S): Yes, because the latency is not a problem, but the cooperation is maintained for a "long time".

  CP_COOP(F): We can live without acknowledged data: the final performance is a tradeoff between lattency with acknowledgement, CSI aging and probability of error on the cooperation channel

  UP_COOP: Yes

- What is the average/maximum transmission delay required in [ms]?

  CP_COOP(S): no delay constraint

  CP_COOP(F): 1ms on the cooperation interface (ideally)

  UP_COOP: worst case 1ms

- Should there be a possibility to discard "out-dated" messages in the transport layer at the earlierst point in time, e.g. already on Tx side?

  Yes

- Is there a requirement to have different priority levels for the messages exchanged?

  CP_COOP(S): no

  CP_COOP(F): Yes cooperation message must be of high priority

  UP_COOP: yes

3.1.4.3 Considerations on implementation options

3.1.4.3.1 General architecture

As one of the fundamental requirements on the "CoMP" channel is to allow for an as low delay as possible, this section provides some initial thoughts on a possible implementation in the eNB.

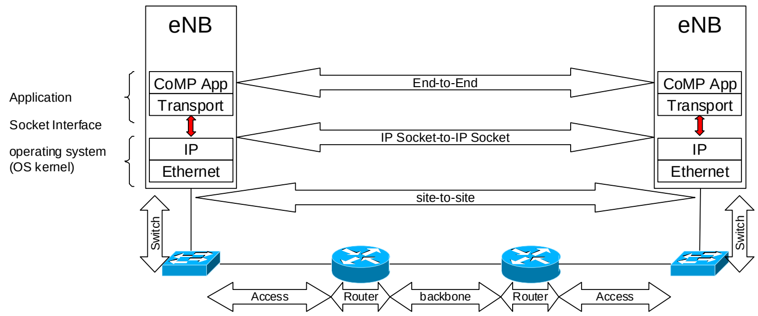


Figure 3-6 Deriving the contributors to the E2E delay.

The "site-to-site" delay could be determined using appropriate traffic analysing devices. The requirement to have a site-to-site delay of about 0.1 ms is expected to be achievable in the next years using the latest available technology in the transport backbone.

Measurements have been done by various operators, but no detailed figures are publically available.

The predominant contribution towards end-to-end delay is expected to be caused by processing delays in the involved eNB's. The eNB internal decomposition is a rough approximation only to show the various sources and to assist in understanding the suggested implementation.

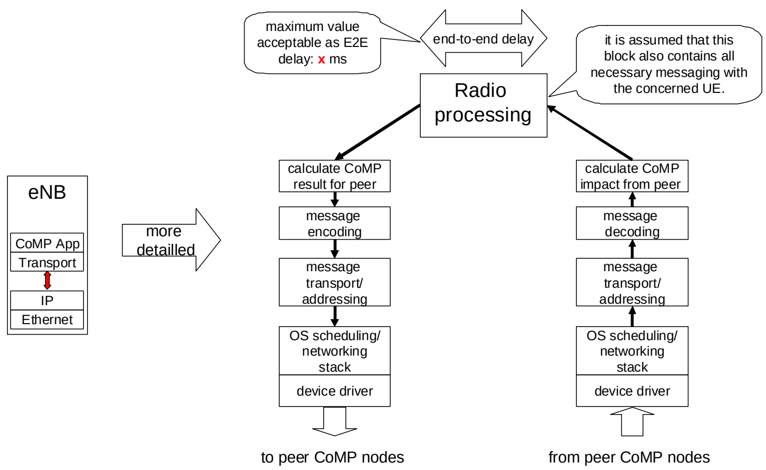


Figure 3-7 eNB internal contributors to the E2E delay.

Assuming that the delay contributions originated by "Radio processing" and "calculate CoMP impact for/from peer" are unavoidable, the subsequent parts need to be examined as to whether their presence is ultimatively required for proper operation.

These delay contributions can be observed in more detail.

- Message encoding/decoding
  - Encoding messages using ASN.1 and PER is certainly fully aligned with 3GPP RAN practise, but not necessary. E.g. Networking protocols as defined by IETF are using simple message structures to achieve high performance. TCP or SCTP are not using ASN.1 at all.
  - It is suggested to not use ASN.1

- Message transport/addressing
    - Assuming that the application doesn't require reliable and in sequence delivery of messages, there is no need to use TCP or SCTP as transport protocols, but use a datagram protocol only, e.g. UDP.
    - In case the transport between nodes forming a cooperation area is based on Ethernet technology, there is not even a need to use IP addressing scheme, but messages could be sent/received using Ethernet MAC addresses only. It should be noted that this is not something exotic but e.g. machine automation protocols or other applications with strict real-time requirements make frequent use of Ethernet MAC addresses only. One example is e,g, the IEEE 1588 Precision Timing Protocol (PTP).
- OS scheduling/network stack
    - Avoiding the use of IP addressing allows to bypass the IP stack on TX and RX side.
    - OS scheduling could be also well bypassed by implementing the functionalities supporting WP1 and WP2 innovations on a dedicated hardware, using a dedicated real-time capable OS kernel.
- Device driver/NIC
    - The remaining critical parts are the Device driver and the Network Interface Card (NIC). Two essentially different implementations could be envisaged.

    1. Implementation suggestion (1)

        NIC receives an Ethernet frame from the interface

            Frame is stored in the NIC's I/O buffer

            Appropriate Interrupt is raised

        Interrupt Service Route (ISR) is started

            "blind" transfer from NIC I/O buffer to common memory is initiated (DMA setup)

        DMA completes, indicating it to the OS

        NIC driver code started in eNB and frame is accessed in common memory for further analysis and processing (possibly transferred to CoMP function by another MEM transfer causing additional Interrupt)

    2. Implementation suggestion (2)

        NIC receives an Ethernet frame from the interface

            the frame is stored in the NIC's I/O buffer

            NIC needs to be able to differentiate on e.g. Ethertype field to raise associated Interrupt

        Interrupt Service Route (ISR) is started

            based on IntReq subtype, transfer from NIC I/O buffer to I/O buffer of target (e.g. CoMP or common memory) is initiated (DMA setup) DMA completes

        Further processing takes place in the receiving unit.

- Note:
  "Real-time" means to have a deterministic, guaranteed maximum service time upon a request. This applies to data transfer from NIC to MEM and also for a subsequent data processing by the NIC driver. Therefore, the process of transferring the frame to the target should not be interrupted once again (non-interruptable IReq) as nesting of interrupts results in unpredictable service time.

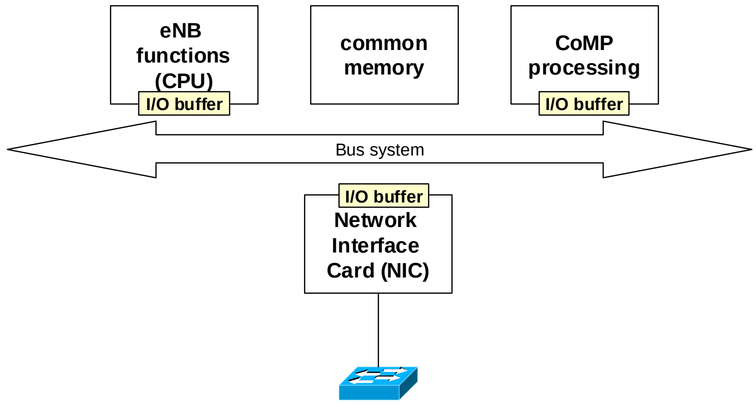


Figure 3-8 Implementation suggestion for eNB supporting WP1/2 innovations.

The implementation suggestion shown above is assuming that the X2 interface and the interface used to exchange “CoMP” information (c-/c-plane) could be differentiated already on the Network Interface Card level.

Background and more details about the approach shown above is given in [LHa04] and LHb04]. A discussion on predictability for a communications interface supporting real-time requirements is given in [MCU08]. This document also provides the motivation for pushing time critical applications “down the stack”.

3.1.4.4 Summary

Three basically different architectural solutions applicable to WP1 innovations are shown in the figures below.

- Fully distributed solution
- Centralised solution “1”
- Centralised solution “2”

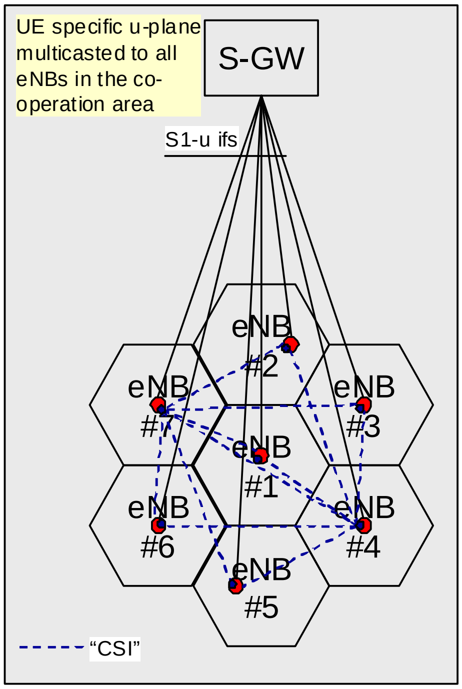


Figure 3-9: Multicasting u-plane to eNBs in cooperation area.

Every eNB in the cooperation area receives additionally the u-plane information (shown as black solid lines) sent to its neighbouring nodes. The total traffic per cooperation area is multiplied by the number of cooperating nodes. Control information is exchanged between the eNBs in the cooperation area (shown as dashed blue lines denoted "CSI").

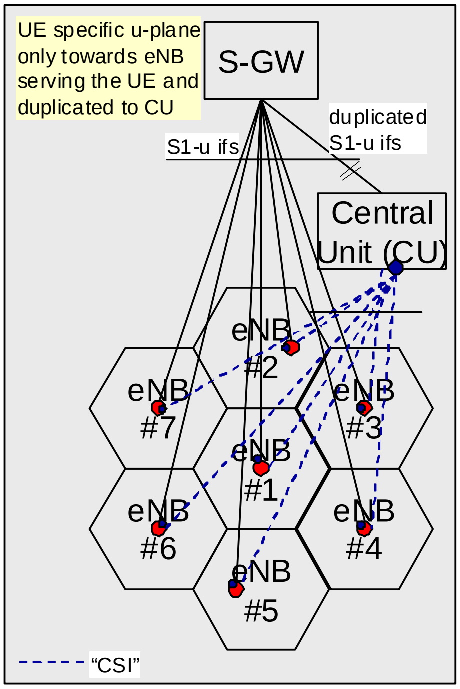


Figure 3-10: CN bi-casting u-plane to eNBs and Central Unit in cooperation area.

Every eNB in the cooperation area receives the u-plane information destined for the UEs it serves. Additionally this u-plane information is bi-casted to the Central Unit where the “CoMP” processing is performed and the derived control information distributed to eNBs in the cooperation area.

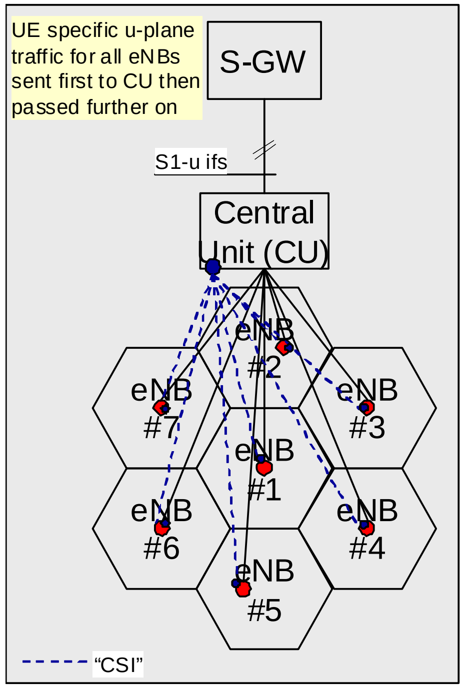


Figure 3-11: U-plane only sent to CU of cooperation area.

The u-plane information for UEs served by eNBs in the cooperation area is only sent to the Central Unit. Here “CoMP” processing is performed and the derived and the modified u-plane along with appropriate control information is further distributed to eNBs in the cooperation area.

Regarding the “CoMP” Interface protocol the findings/proposals from this section are repeated below:

- Finding:
  Do not use the GTP encapsulation for exchanging messages between eNBs in the cooperation area.
- Finding:
  Adapting the mechanisms for UE HO is not adequate for establishing a permanent communication channel to exchange UE specific “CoMP” information.
- Proposal:
  Adapt the S1-AP messages for eNB and MME Configuration Transfer to also convey “CoMP” channel related TNL configuration information.
- Proposal:
  Use a “Type/Length/Value” (TLV) structure to identify the information passed and length of the information field.
- Proposal:
  Allow for ultimate flexibility regarding TNL addresses and refrain from defining a “CoMP” AP along with related messaging for the time being.
- Proposal:
  Do not refer to the "CoMP" interface as being integral part of X2 interface.

3.2 Architectural enhancements for CoMP

3.2.1 Clustering for cooperative schemes

Most CoMP innovations are based on the definition of suitable cooperation areas, which result in clusters of eNBs and, in turn, in architecture requirements.

In particular, some examples of WP1 innovations which involve clustering are:

- joint processing, performed among base stations grouped into a cooperating set (cluster)
- HeNB campus, where radio resource management and/or allocation is optimised based on clusters,
- eNB/HeNB ICIC, in which HeNBs and eNBs are clustered and send measurement reports to a central unit, which gather all information and multicast commands to the nodes.

In general, the definition of the cooperation areas can be based on either a network centric, or a user centric, or a combined network-user centric approach.

In the network centric clustering approach, the network is subdivided into non-overlapping (i.e., non-intersecting) coordinated clusters. The definition of the cooperation areas is static, typically based on the radio cell neighbourhood and on the cells that appear as strong interferers as they result from network planning considerations. The main advantages are simplicity and fairly low scheduling complexity: for example, the set of eNBs included in the cluster is broadcasted to the UEs once, without any additional measurement, and scheduling is only required among the eNBs in the cluster for serving any user located in the same cluster. Moreover, in order to include as many strong interferers as possible, the clusters of eNBs may turn out to be very large, increasing the amount of feedback signaling and related measurements in the uplink. This and the limited throughput gains are the main disadvantages of the network centric approach.

In the user centric approach, the cooperation areas are defined per each single UE and they typically result in dynamic and overlapping clusters of eNBs. In fact, each UE measures and reports to the serving cell its set of strongest interferers and includes the cells that result to be the strongest interferers in the cooperation area, without any particular restriction. This means that an eNB can belong to different clusters, in order to optimally serve different UEs, which on the other hand results in higher CoMP algorithms complexity (e.g., the eNB may face situations where different CoMP strategies need to be applied with respect to different UEs, depending on which cluster the UE belongs to). Another issue to be taken into account when considering the user centric clustering approach is that the X2 signalling load increases, as the eNB needs to exchange as many control data flows as the number of clusters the eNB belongs to. Given a fixed cooperation area (cluster) size, the user centric clustering approach provides higher throughput gains, but requires higher scheduling complexity than the network centric approach, i.e., more complex resulting coordination between cooperation areas selection and multi-cell packet scheduling.

In the combined network-user centric approach, in an effort to get a trade-off between complexity and gain, the cooperation area is selected at each UE from a set of cooperation areas which have been pre-defined by the network. Such an approach results in a semi-static definition of the cooperation areas with reduction in signalling and CoMP algorithm and scheduling complexity.

Several issues should be examined when defining clusters for cooperative schemes. For example, it is possible to consider the class of non-overlapping network-centric clusters where eNBs are connected through the existing physical communications links. In this case the eNBs in the cooperation cluster, when not co-located, share the same Passive Optical Network (PON) for their backhauling. This on the one hand reduces the number of hops (switches/routers) on the physical backhaul path between cooperating eNBs and thereby the latency on the X2 interface; on the other hand, clusters are limited to the physical RAN topology, i.e., cells of two different clusters cannot cooperate with each other, even if such a cooperation would be beneficial.

For overlapping clusters, where the backhaul topology may more easily not be aligned with the physical RAN topology, a more optimised cooperation is implemented at the possible cost of higher delays between cooperation entities.

In conclusion, the main characteristics of the network centric (generally non overlapping) and the user centric (generally overlapping) clustering approaches can be be summarised as follows:

- **Network centric clustering approach:**
  - static cooperation areas' definition (e.g., the network is subdivided into non-overlapping coordinated clusters, the set of the eNBs included in the cluster is broadcasted to the UEs once and there's no need for specific additional measurement at UEs to define the cluster);
  - limited impact on UE signalling load (see above);
  - fairly low scheduling complexity (e.g., scheduling is required only among the eNBs in the cluster to serve UEs in that cluster);
  - clusters may turn out to comprise very large numbers of eNBs, in order to include as many strong interferers as possible, so increasing the amount of feedback signaling and related measurements in the uplink;
  - limited throughput gains.
- **User centric clustering approach:**
  - cooperation areas' definition per each specific UE;
  - higher throughput gains than in the network centric clustering approach;
  - more complex cluster topology than in the network centric clustering approach (e.g., cooperation areas typically result in dynamic and overlapping clusters of eNBs);
  - impact on UE signaling load;
  - greater backhaul load;
  - fairly high scheduling complexity;

When dealing with CoMP with eNB clustering, two main architectures can be considered, namely, the centralized and the distributed Control Unit architecture.

The different CoMP transmission schemes proposed in ARTIST4G can be combined with either one of the two architecture variants, considering that the implementation complexity degree may vary from one scheme to the other.

In the centralized Control Unit architecture depicted in Figure 3-12, a central entity is needed which gathers the channel information coming from all the UEs in the area covered by the coordinating eNBs. The central entity then performs user scheduling and some signal processing operations (e.g., precoding).

For each UE, there is one single cell in the set of cooperating eNBs that acts as the serving cell of that UE (namely, the anchor cell): based on proper measurements, each UE estimates the channel information related to the whole set of cooperating eNBs and feeds back such information to its anchor cell. Once the information is gathered, each eNB forwards it to the central entity which performs scheduling and transmission parameters setting and sends new resource allocation information back to the eNBs.

In order to successfully carry out such operations, eNBs need to be tightly time synchronised among them. Moreover, user data need to be available at all collaborating nodes.

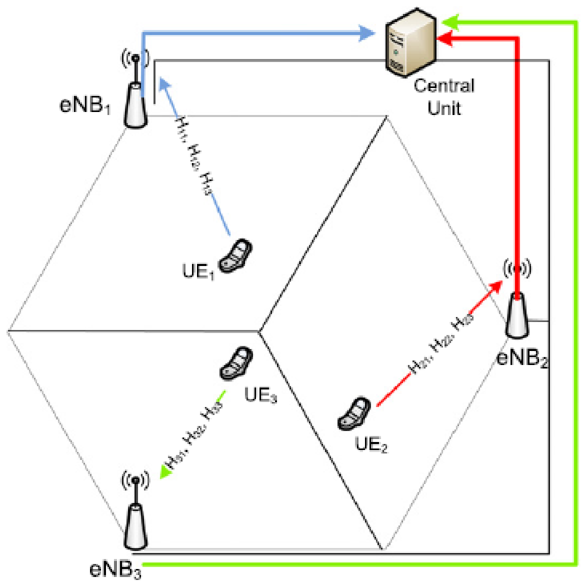


Figure 3-12: Centralized CoMP approach.

The main advantages of the centralized Control Unit architecture are that the eNB functions remain limited to their typical and ordinary ones and the backhaul topology can be easily optimised. However, the main drawbacks are that an additional centralised and stand-alone node (the central entity) needs to be deployed and an additional logical interface and relevant communication protocols need to be identified. Finally, the trade-off between the number of deployed central entities on the one side and the delay and backhaul load on the other side has to be investigated. In fact, in order to limit delay and backhaul load, as many deployed central entities as possible would be preferred, with the disadvantage of an increased network cost.

In the distributed Control Unit architecture depicted in Figure 3-13, cooperation is supported without the requirement of a central entity, but maintaining the communication links among the different eNBs and the decisions are jointly made by the connected eNBs.

Dealing with the distributed Control Unit architecture, the main advantages are that the network deployment optimisation is simpler than in the centralized architecture. The delay and the backhaul load may be kept low, especially as clusters are kept limited. Moreover, the X2 interface can be possibly re-used. On the other hand, eNB complexity increases, proprietary Control Unit implementations may have a negative impact on interoperability, and the backhaul topology is not easily optimised.

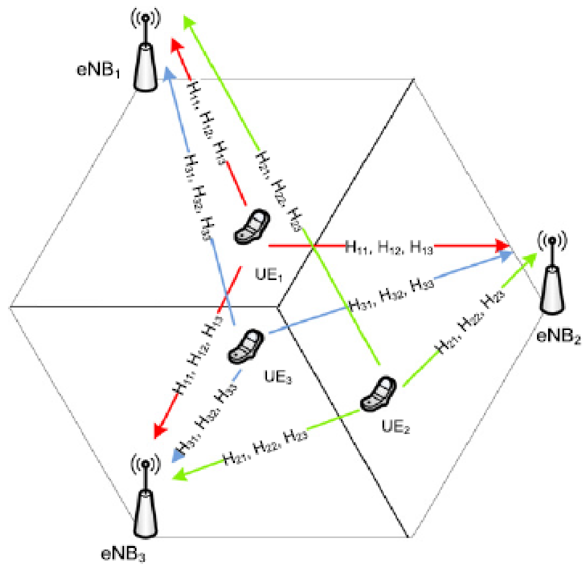


Figure 3-13: Distributed CoMP approach.

In order to minimize the infrastructure and signalling protocol cost associated with inter-eNB communication links, distributed architectures may completely get rid of inter-eNB communication links to perform cooperation. The underlying assumption is then that all eNBs perform the same scheduling and that channel information regarding the whole coordinating set can be made available to all cooperating nodes. Based on proper measurements, the UE estimates the channel information from all the coordinating eNBs and sends back such information to all cooperating eNBs. Scheduling is then independently performed in each eNB.

The drawback of such an architecture solution lies in suboptimal CoMP performance gains. This might be due to the fact that the UE channel-condition information is reported to the different cooperating eNBs via different wireless feedback links, which may suffer from different propagation conditions and impairments.

3.2.2 X2 dynamic set-up using ANR

Currently, the only defined interface between eNBs is the X2 interface, cf. Figure 3-14. X2 logical channel is a direct connection between eNBs. This means that there will be one X2 connection per cooperating eNB. But due to deployment constraints (no direct physical link between eNBs, need to secure u-plane with Security GW…), X2 physical link can pass through a certain number of routing equipment, and therefore add a significant delay. The X2 interface is split into the X2-C part for the c-plane and the X2-U part for the u-plane:

- X2-C relies on one SCTP set up between two neighbouring eNBs. This SCTP association is carried over the IP connectivity link established between these two RAN nodes.
- X2-U interface relies on GTP tunnels, set up dynamically "on demand" (e.g. HO with data forwarding, which means one tunnel per UE to forward the packets during HO), between neighbouring eNBs. Multiple GTP tunnels are carried over the IP connectivity link established between these two RAN nodes.

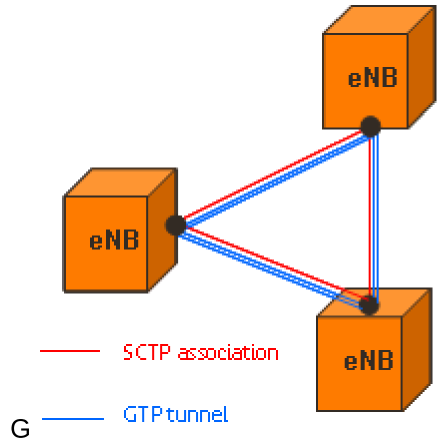


Figure 3-14: X2 interface requiring IP connectivity between neighboring eNBs.

3.2.2.1 X2-C self configuration using Automatic Neighbour Relation function

Automatic Neighbour Relation (ANR) is part of the Self Organizing Networks (SON) topic defined in [3GPP36300]. It relies on the measurement reports done by the UEs to allow the serving eNB to build and maintain a neighbouring relations list. This list can be used to dynamically set up X2 interfaces.

Figure 3-15 contains a simple example: eNB1 discovers eNB2 as neighbour thanks to the ANR procedure. Note here that an eNB is uniquely identified by its eNB Identity (ID).

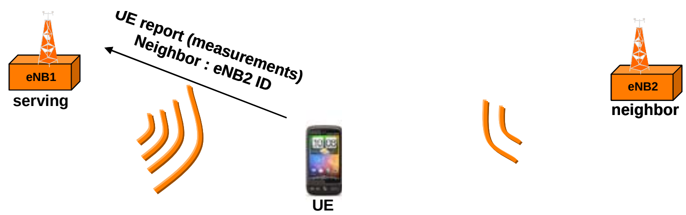


Figure 3-15: ANR procedure overview.

Once the target eNB ID has been discovered with ANR, the source eNB queries the MME for the respective IP address which must be used as the remote endpoint for the IP connectivity supporting the X2-C interface (see Figure 3-16). After successful connection establishment, X2-C allows the eNBs to communicate with each other in the c-plane and to agree on a handover decision.

1. eNB1 Neighbor discovery (ANR procedure)
   UE measurements reported to eNB1
   eNB1 neighbor: eNB2 ID

2. Neighbor CP IP@ discovery
   "*Configuration Transfer*" procedure with the MME in order to discover the CP IP@ of eNB2

3. X2-C setup
   SCTP association with eNB2 establishing
   X2-C interface used to support X2 AP messages

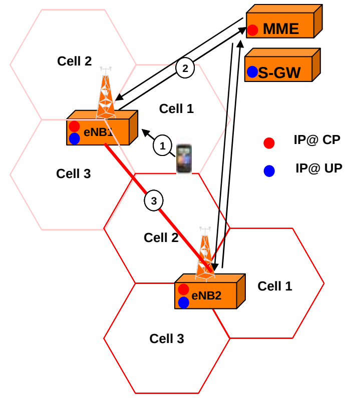


Figure 3-16: X2-C setup based on ANR information.

3.2.2.2 X2-U self configuration

Once the X2-C connection has been set up, the source eNB will use it to inform the target eNB about an ongoing handover. Note that the handover decision at the source eNB also relies on the measurement reports done by the User Equipment (UE). In that case a GTP tunnel must be set up on the fly when the handover occurs. For that purpose the source eNB queries the target eNB via X2-C for the IP address which must be used as the remote endpoint for the IP connectivity supporting the X2-U interface (see Figure 3-17). After successful connection establishment, X2-U can then be used to transport u-plane from the source to the target eNB during the handover procedure.

1. Neighbor UP IP@ discovery
   use of X2-C to transport the *Handover Request* message to learn the UP IP@ of eNB2

2. X2-U setup
   one or multiple GTP tunnels with the eNB2 to support UP traffic over X2-U. User Plane IP packets can be directly transferred from eNB1 to eNB2 ("Data Forwarding") during the Handover.

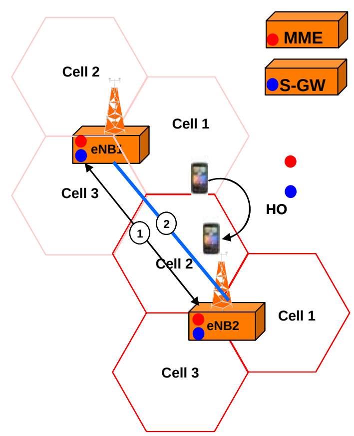


Figure 3-17: X2-U setup for Handover.

X2-U interface is established on a per UE basis, which means that there is one X2-U interface per UE. Therefore this innovation could be used in the particular case, when few nodes are concerned, to accelerate CoMP communication channel establishment.

3.2.3 Distributed base station with D-ROF, BBU Hostelling – Security aspects

The distributed base station using Digitized Radio Over Fibre (D-RoF), as shown in Figure 3-18, proposes a clean separation between the Remote Radio Head (RRH), which is installed at antenna sites, and the Base-Band Unit (BBU), which can then be installed in more restricted areas (operator access-controlled local or cabinet). Thus, from a security point-of-view, this separation of the eNB into two different units makes it similar to the UMTS architecture:

- The RRH is responsible for the analog processing of the LTE-Uu interface layer 1 (Radio Frequency). This is similar to the Remote Radio Unit (RRU) for 3G NodeB managing lower radio layers.
- The BBU is responsible for the digital processing of the Uu interface (upper part of the LTE-Uu layer1, MAC, RLC, PDCP, and RRC layers), and hence of all cryptographic and security operations of u-plane traffic and RRC signalling handled by the PDCP layer. This is similar to the BBU for 3G RNC managing higher radio layers including the security processing.

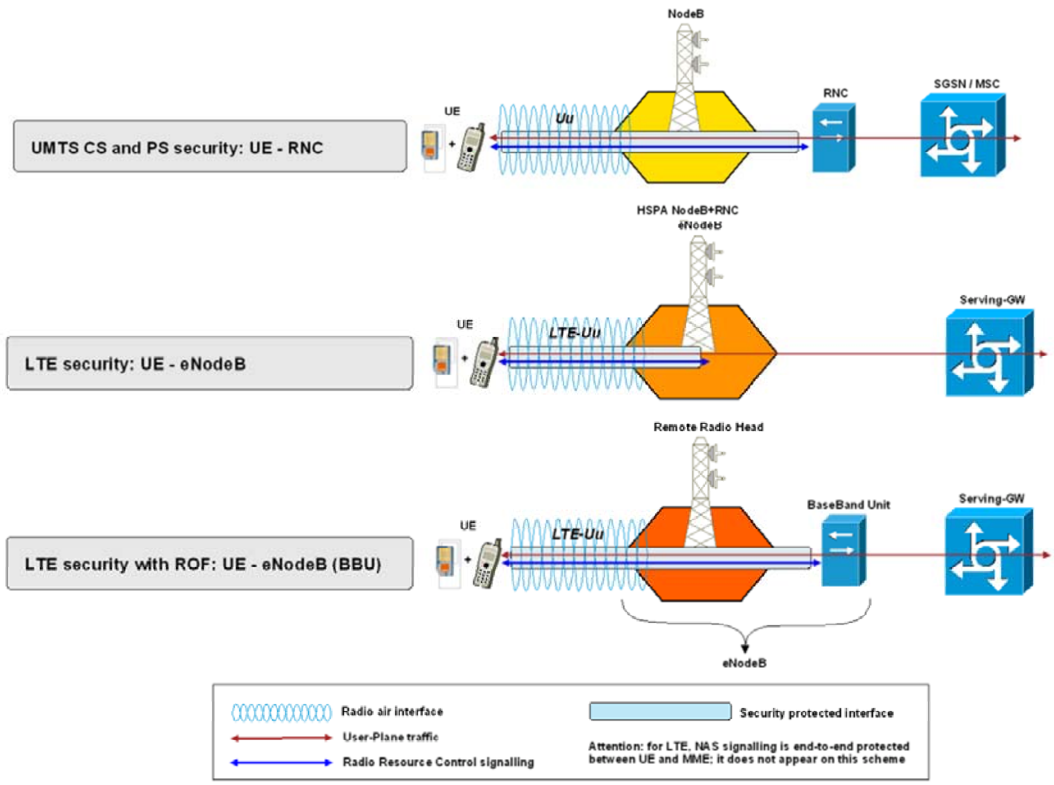


Figure 3-18: Comparison of 3G security architecture and LTE with/without D-ROF.

Consequently the issue of terminating the radio security protection in the eNB seems to be solved: This approach allows extending the LTE-Uu interface security up to the BBU which can be more easily installed in a safer area with much less common interfaces than Ethernet (optical). However, some care must be taken with respect to the following aspects:

- Cells and eNB that use D-RoF and thus do not implement IPsec protection on their S1 and X2 interfaces must not be interfaced with any "unsecured" equipment (e.g. having an X2 interface with a non-trusted eNB). If the latter is unavoidable, IPsec protection must be used on the X2 interface.
- O&M specific traffic used for configuring the D-ROF interface between the RRH and the BBU has to be secured. In particular, it should not allow any reconfiguration of the BBU from unsecured equipment.

More details on D-ROF and BBU Hostelling can be found in [ARTD41].

3.3 Cooperation in heterogeneous networks

In case of heterogeneous networks (HetNet) not only the introduction of HeNBs in a co-channel deployment needs to be taken into account, but also the fact that access to HeNBs is subject to additional restrictions as HeNBs can operate in one of the following three modes: open, closed, hybrid.

The definitions of the different modes of operation are taken from the stage 2 description given in [3GPP36300].

The concept of “Closed Subscriber Group” (CSG) is commonly used since the advent of computerised communication systems to either

- keep communication inside a defined group of subscribers,
- restrict the usage of certain infrastructure elements to a defined group of subscribers.

Of course, this concept is at odds with the principle of providing “public” communication services, hence, it could be only applied by selected network elements that are not essential in providing public services in the regulatory sense. One such case is the deployment of HeNBs, where the “owner/subscriber” of the HeNB has a vital interest to restrict the use of the HeNB to certain other subscribers only, e.g. members of his family, or employees of his enterprise.

Cell selection procedures in the UE ensure that only in case of manual CSG selection the UE could attempt to attach to a CSG cell in case the Universal Subscriber Identity Module (USIM) does not have an appropriate entry in the Access Control List (ACL). Consequently only (H)eNBs with identical CSG-IDs could form cooperating clusters.

[3GPP36300] specifies some of the most important terms for this type of access control:

- Closed Subscriber Group (CSG): A Closed Subscriber Group identifies subscribers of an operator who are permitted to access one or more cells of the PLMN but which have restricted access (CSG cells).
- CSG cell: A cell, part of the PLMN, broadcasting a CSG indication that is set to TRUE and a specific CSG identity. A CSG cell is accessible by the members of the closed subscriber group for that CSG identity.
- CSG identity: An identifier broadcasted by a CSG or hybrid cell/cells and used by the UE to facilitate access for authorised members of the associated Closed Subscriber Group.
- closed or CSG cell:
    - [3GPP36300]: A cell broadcasting a CSG indicator set to true and a specific CSG identity.
    - [3GPP36304]: A cell, part of the PLMN, broadcasting a CSG indication that is set to TRUE and a specific CSG identity. A CSG cell is accessible by the members of the closed subscriber group for that CSG identity.
- hybrid cell:
    - [3GPP36300]: A cell broadcasting a CSG indicator set to false and a specific CSG identity. This cell is accessible as a CSG cell by UEs which are members of the CSG and as a normal cell by all other UEs.
    - [3GPP36304]: A cell, part of the PLMN, broadcasting a CSG Indicator that is set to FALSE and a specific CSG identity.
- open cell: Even [3GPP36300] doesn’t explicitly specify “open cell”, it is obvious that a cell in this mode of operation neither broadcasts a CSG identity nor the CSG indicator.

In WP1 [ARTD13], several innovations are addressing femto base stations, i.e. Home eNBs (HeNBs), either in an eNB/HeNB cooperation or in a cooperating HeNB-only network (campus scenario). In this section, an architecture and control solution in order to support eNB/HeNB ICIC when the number of femto base stations deployed in the coverage of a macro network is high is provided. The HeNB campus scenario and present architecture solutions for enabling inter-HeNB cooperation are also addressed.

3.3.1 Enablers for eNB/HeNB ICIC in heterogeneous networks

During the standardization phase of recent systems, like 3GPP-LTE, Inter-Cell Interference Coordination (ICIC) techniques have been extensively discussed. The UE reports measurements to its serving eNB and eNB-to-eNB messages have been standardized in order to allow for efficient ICIC. The eNB-to-eNB messages are, for instance, conveyed through an eNB-to-eNB link, the so-called X2 interface in LTE terminology. However, this link might not exist between an eNB and a HeNB since a massive deployment of HeNBs will prevent having an X2 interface between an eNB and all HeNBs within its coverage area. Furthermore, when access control restricts the connection to a set of UEs (via a closed subscriber group CSG), the closed HeNBs may strongly interfere with eNBs and even create coverage holes.

In order to mitigate the interference between several base stations, more sophisticated techniques can be employed, such as soft-reuse ICIC, where a part of the resource is reserved to cell-edge UEs. To this end, a power pattern is defined for each base station or home base station, where the maximum transmit power to be used (at the UE for uplink and at the eNB or HeNB for downlink) in a given frequency resource varies. This pattern changes from one eNB/HeNB to another. Especially, two neighbouring base stations or home base stations should have different patterns. The UE allocation is then optimized depending on the level of interference it suffers from or creates. Part of the coordination requires the allocation (dynamically or statically) of one pattern to each base station. This scheme is particularly well adapted to macro-cell ICIC between eNBs. Figure 3-19 illustrates the way three patterns can be allocated to cells (sectors) in a macro-cell deployment. Each of the three patterns is characterised by a high-power part of the frequency resources which is different from the other patterns. For instance, cell-edge UEs can be allocated in this high-power part.

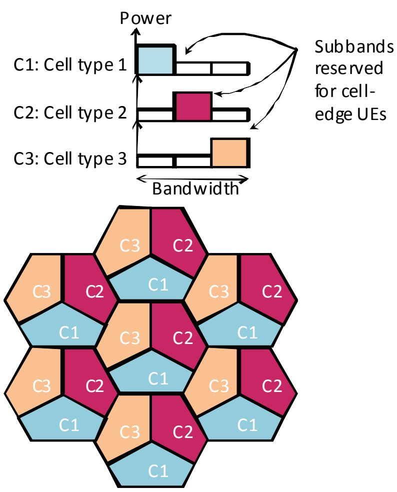


Figure 3-19: Example of downlink ICIC for a hexagonal deployment of tri-sectorised eNBs.

Soft-reuse ICIC cannot be applied to eNB/HeNBs ICIC (between two HeNBs or between one eNB and one HeNB) for the following reasons: In a heterogeneous network, the number of neighbors is very high: Thousands of femto HeNBs might lie under the coverage area of a macro eNB. The number of patterns dedicated to each transmitter cannot be multiplied. Typically, the algorithms work with 3 to 7 patterns. Indeed, multiplying too much the number of patterns increases the signaling overhead in the system. Having a more flexible soft reuse scheme like in LTE has also a cost in signaling, as it requires bi-directional communication between neighboring cells. Furthermore, if the patterns are defined as in Figure 3-19, increasing the number of patterns is equivalent to decreasing the amount of resources corresponding to the high-power part. Thus, it reduces the number of UEs that might take significant benefit from ICIC. From a network point of view (signaling overhead on the backhaul, i.e., on the core network), it is not realistic for a base station to cooperate with thousands of neighbors. Only cooperation with some tens of neighbors is targeted.

In 3GPP LTE Rel.10, it is not possible to establish X2 interfaces between eNBs and HeNBs or between HeNBs considered in residential deployment scenario. Thus, femto/macro ICIC as it is foreseen for macro eNBs (i.e., through cooperation between individual eNBs/HeNBs) is not possible.

In WP1 [ARTD13], one innovation presents solutions for eNB/HeNB ICIC in downlink and in uplink. The conclusions are that the ICIC can efficiently be achieved by choosing a HeNB power setting in a blind fashion in downlink, i.e., without any control exchange between nodes. The performance can be further improved by allowing some cooperation between the eNB and the HeNB, which can be addressed in a clustered fashion. For the uplink, it is stated that coordination through a central entity is needed. A strategy is presented that can be implemented in three ways: In the first case, a *central entity* collects long term measurements from the HeNBs and the eNB, computes a maximum tolerated transmit power for the UEs of each node, and provides the result of the optimization to each node. In the second case, the central entity broadcasts a power setting function. Each HeNB chooses its parameter according to its previously transferred measurements. In the last case, the HeNBs are clustered, and the central entity provides the power setting strategy to each cluster of HeNBs.

3.3.1.1 Enabler for HeNB clustering

In order to cope with the large number of neighbouring eNBs and HeNBs induced by an heterogeneous deployment of macro and femto base stations on the same frequency resources, the grouping of femto HeNBs for uplink ICIC and downlink ICIC is proposed. For this reason, an additional *Physical Cell Identity* (PCI) is associated to each group, which is named *Group Physical Cell Identity* (GPCI). Each HeNB transmits a signal or signals linked to the GPCI in addition to the signals related to its own PCI. The macro UEs feed back some measurements to their serving eNB, including measurements related to GPCIs. The eNB communicates with a central entity, which manages coordination between eNBs and HeNBs and between HeNBs and gives instructions to the different groups. Each neighbouring group of HeNBs is seen by an eNB as a single neighbouring eNB for ICIC management.

The grouping is illustrated in Figure 3-20, where blue groups are neighbouring groups of eNB1 and red groups are neighbouring groups of eNB2. Note that grouping for uplink and downlink ICIC may be different. For instance, downlink groups are composed of neighbouring HeNBs and uplink groups in slices of HeNBs located at roughly the same distance from an eNB. Each HeNB may belong to several groups for UL ICIC and several groups for DL ICIC. A HeNB can easily belong to several groups by sending signals linked to several GPCIs in addition to its own PCI. A group may not be geographically localized. The groups may be distributed in space, paving the space for virtual eNBs. In order to move a HeNB from one group to another, its GPCI just has to be changed.

Figure 3-21 illustrates the architecture including eNBs, HeNBs, and central entities. One HeNB may be connected to several central entities. The signals linked to the GPCI may be of the same nature as signals linked to a classic PCI as defined for a regular eNB. In other words, the GPCI appears like a PCI and the signals associated to the GPCI are part of the signals which would be associated to this GPCI if it were a regular PCI. For instance, in LTE, the signals linked to the GPCI might be the synchronisation signal (PSC and SSC), the reference signals for channel estimation, and part of the system information (for instance the broadcast channel BCH and possibly some SIBs (System Information Blocks)) associated to a PCI equal to the GPCI. In this case, the measurement corresponding to a PCI or a GPCI would be the same and would be transparent for the UE, reusing the PHY signalling (measurement reporting) as already defined in the specification. Likewise, the report from the UE would be the same for the GPCI and the PCI. For ICIC, the eNB could ignore which part of the PCIs corresponds to GPCIs and communicate with the central entity, as if it were communicating with a regular eNB, reusing the signalling (X2 + architecture) as already defined in the specifications. However, the mobility from an eNB to a

GPCI must be precluded, for example, by allowing a GPCI to represent a closed cell. The signals linked to the GPCI may also be a new signal dedicated for ICIC.

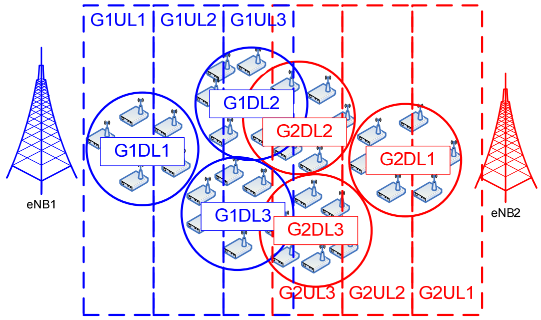


Figure 3-20: Example of grouping.

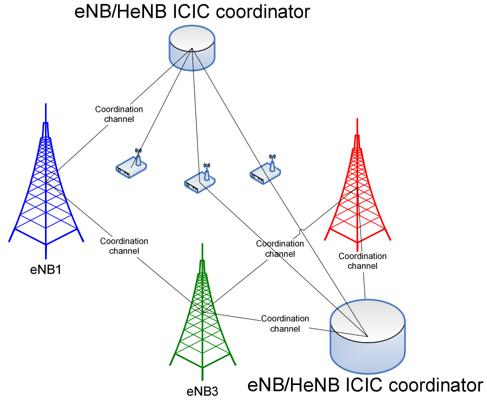


Figure 3-21: Example of architecture.

A HeNB transmits the different signals linked to its GPCIs superimposed with all signals (including control + data) it has to transmit for its normal operation, i.e., for communicating with femto UEs it serves. The HeNB behaves as if it were serving several cells (i.e., with distinct PCIs) with same coverage area and on same carrier frequency and system bandwidth. In order to cope with HeNB power limitation, the signals linked to a GPCI can be deboosted. This deboosting may take into account the number/density of HeNBs of the group. Indeed, the higher the density, the higher the cumulated signal power at a UE receiving the signal from several neighbouring HeNBs.

All HeNBs in a group share the same ICIC parameter, for instance, the same pattern for soft reuse. There are different ways/means to control the parameters of a HeNB, either by changing the group the HeNB belongs to or by changing the ICIC parameters of the group itself.

3.3.1.2 Central entity location

A central entity may be an independent device, communicating with several eNBs and the HeNBs under its coverage area. The communication with different eNBs is useful in order to manage a HeNB located close to the border between two eNBs. Indeed, such a HeNB may be part of a group for one eNB and another group for the other eNB (see for instance the intersection of G1DL2 and G2DL2 in Figure 3-20). It would be beneficial to choose the parameters of the two groups such that they are compatible. This can be achieved easily if there is a single central entity for the two eNBs.

Note that a central entity may also be located in an eNB.

3.3.1.3 Control exchange

In a first solution, there is a transmission of interference information from eNBs to the central entity as if it were a regular eNB, i.e., neighbouring eNBs might cooperate via the S1 or X2 interfaces with the central unit, which is seen as a radio neighbor, the radio interface of the central unit being a cluster of HeNBs sharing a GPCI. Measurements from HeNB are also gathered at the central entity. Decisions and updated parameters are transmitted from the central entity to the HeNBs. This solution is transparent for eNBs.

In a second solution, the eNB is informed by the central entity about ICIC parameters and GPCI-to-ICIC parameter vector association, and can take this information into account when scheduling UEs. Indeed, when a UE reports a measurement for a given GPCI, the eNB knows that there is at least one HeNB in the UE neighbourhood using ICIC parameters associated with the GPCI. Compared to the first solution, only the behaviour of the central entity and the eNB changes, not the ICIC architecture or the information exchanged.

3.3.2 HeNB/HeNB ICIC

As discussed in the previous sections, HeNB grouping/clustering is necessary for enabling cooperation with neighbouring macro base stations. For intra-group HeNB/HeNB ICIC, two solutions can be considered:

- Direct X2 signalling based ICIC between HeNB within each group.
- Central entity based ICIC for HeNBs in the group.

These options are depicted in Figure 3-22 where (a) describes the direct signalling option and (b) the central entity based coordination.

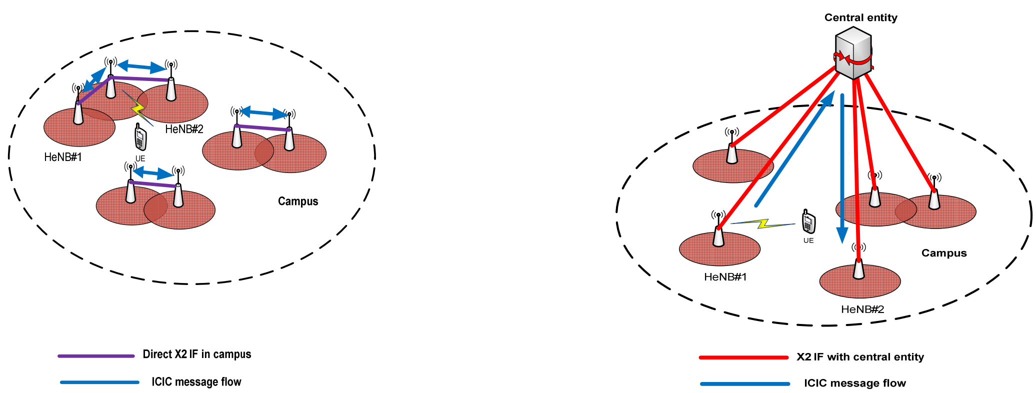


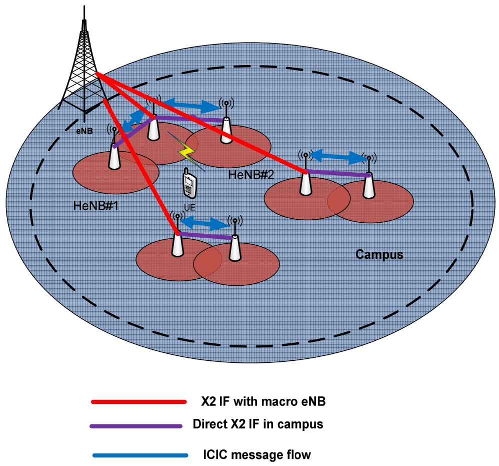
Figure 3-22: ICIC options for campus HeNB deployment.

In the former case direct signalling X2 interfaces are set up between neighbouring HeNBs and ICIC is performed as for macro eNBs by exchanging load indication information over X2 as described in [ARTD41]. Hence, each HeNB can autonomously perform ICIC based on the received X2-AP messages (i.e. RNTP, UL OI, UL HII) from the neighbouring HeNBs.

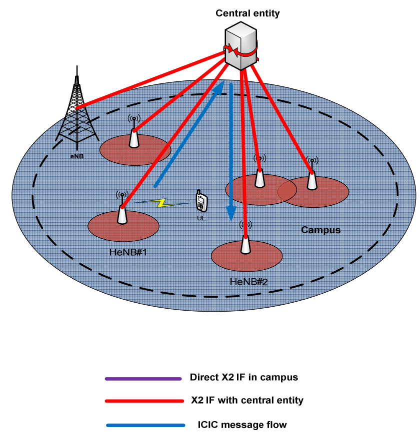
In case of central entity based coordination, all HeNBs deployed on the campus are connected to the central entity. ICIC decision can either be performed in the central entity or coordinated between the central entity and the HeNBs on the campus.

However, the campus HeNBs are not isolated from macro cells, and functionalities such as ICIC or CoMP for the HeNBs on the campus can be highly improved by coordination between the neighboring macro eNB network and HeNBs within the campus. For example, ICIC decision between HeNB#1 and HeNB #2 in scenario (a) above may introduce extra interference for macro

UEs in the neighbourhood of HeNB#1, if the ICIC is made locally without coordination with the macro network. The two solutions depicted above in Figure 3-22 can be extended to introduce coordination between campus and overlapping macro eNBs. The result is shown in Figure 3-23. In option (b), the central entity is in charge of the coordination, although in option (a), the eNB is connected to a significant subset of the campus HeNBs. Indeed, there is no need for neighbouring eNBs to be connected to all the HeNBs on the campus.



(a)



(b)

Figure 3-23: Coordinated ICIC options for campus HeNB deployment.

In solution (a), the HeNBs to be connected through X2 with the neighbouring macro network need to be identified during the campus deployment so this solution is less scalable than solution (b).

Moreover, in case of a large/dense campus deployment, the planning of the nodes to be connected with neighbouring macro networks may be a difficult and time consuming task. This planning task can be avoided with option (b).

3.3.3 Cooperation central entity in the HeNB-GW

In this section, we further extend the role of the central unit as a prerequisite for a higher cooperation between HeNBs or between HeNBs and eNBs, for example with Joint Processing (JP) schemes.

The current 3GPP architecture, however, does not foresee a “central unit” and in turn there are also no interface specifications available.

In the pure HeNB/HeNB case the HeNB-GW could additionally serve as “central unit” for the cooperating HeNBs deployed in a defined campus network. In case of eNB/HeNB cooperation, the central unit might be located in the eNB, for example, but a HeNB-GW seems to be preferable. The reason is that the introduction of an HeNB-GW allows a set of HeNBs to be advertised like a single node supporting a number of cells. This would greatly simplify coordination. The HeNB-GW needs to convert the group-cooperation information to HeNB-centric info and route it to concerned HeNBs.

In Figure 3-24 below a central HeNB-GW serving many HeNBs is located inside the operator domain behind the Security Gateway. As the HeNB-GW acting as “central unit” for WP1 innovations is not located in the private domain, central coordination is suffering from twice the transmission delay between the HeNBs and the HeNB-GW.

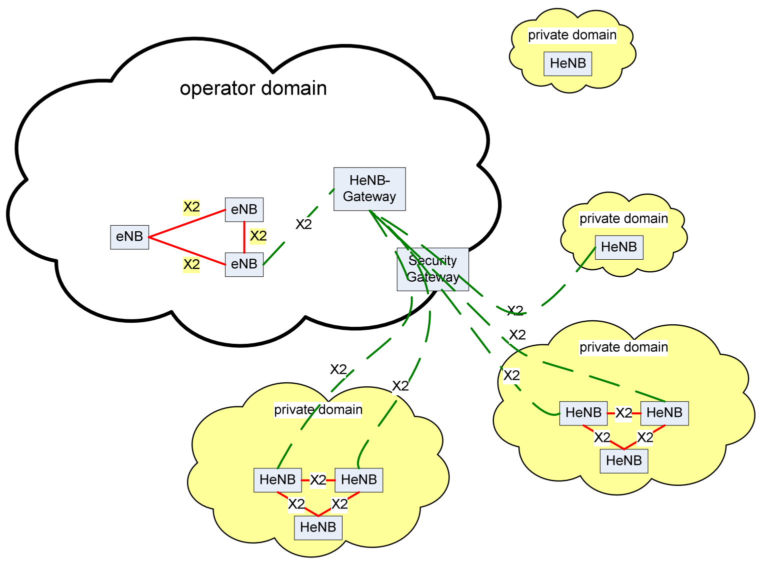


Figure 3-24: X2 interfaces in a HetNet scenario with central HeNB-GW.

The HeNB-GW located in the operator domain is assumed to act as X2 proxy on behalf of the HeNBs served. The HeNB's located in the same private domain might have "direct" X2 interfaces between each other. Details of the standardised solution are subject to ongoing work in Rel-11.

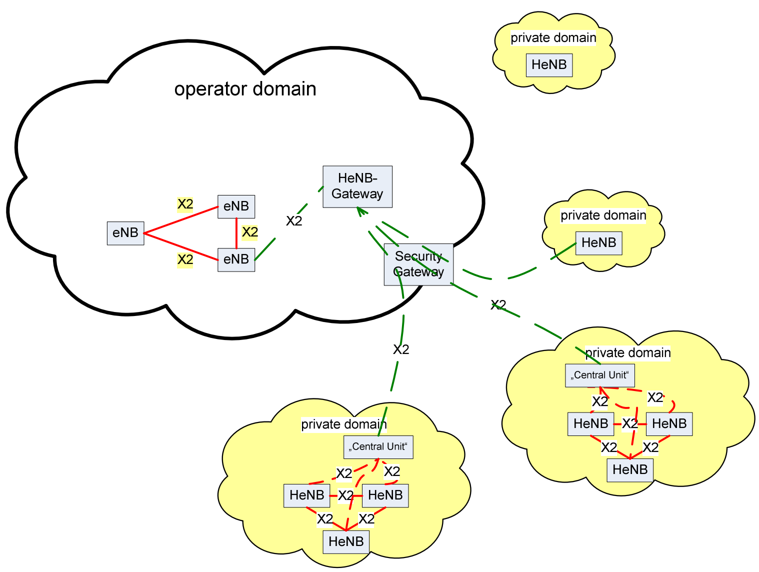


Figure 3-25: X2 interfaces in a HetNet scenario with "Central Unit" per private domain.

In case the envisaged interference mitigation scheme requires extremely short transmission delays, a "Central Unit" needs to be deployed inside the private domain, as shown in Figure 3-25. This "Central Unit" might be a local HeNB-GW.

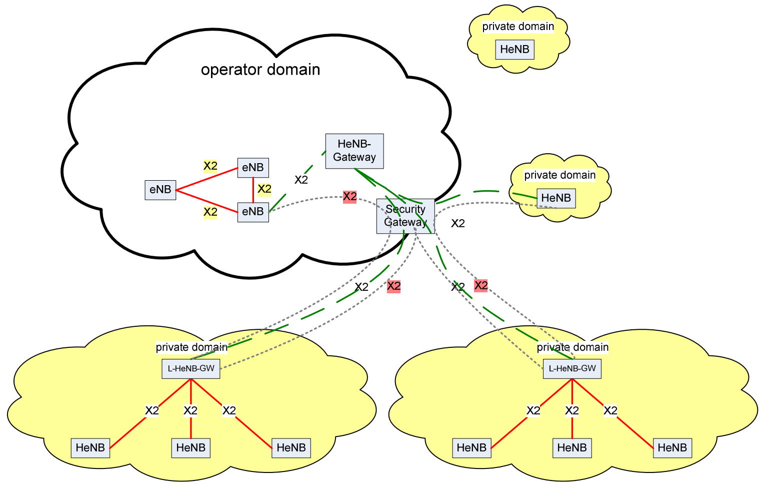


Figure 3-26: Cascading X2 interfaces to allow for a local HeNB-GW per private domain.

If such a solution needs to be further taken into account, the list below is a starting point showing potential areas that require further investigation:

- "Central Unit" uses X2 interface and acts as a local HeNB-GW (L-HeNB-GW):
  - Any new requirements on X2 protocol in case of cascading multiple HeNB-GWs at different hierarchical levels (see Figure 3-26).
  - Direct X2 interfaces between L-HeNB-GWs at same hierarchical level.
  - Direct X2 interfaces from eNBs to HeNB-GWs at different hierarchical levels.
- "Central Unit" uses a newly defined interface with the HeNBs for interference avoidance:
  - Is it sufficient to include only HeNBs in the private domain?

In summary the status in 3GPP Release 10 is as follows:

- HeNB-GW acting as X2 proxy is not supported.
- Optional HeNB-GW is considered as part of the operator domain, located behind the Security Gateway.
- There are currently no requirements to allow cascading of HeNB-GWs.

4 Architectural impacts of joint WP1 and WP2 proposals

4.1 Alternatives for time and frequency synchronization between macrocells and femtocells

One of the requirements associated with the support of both interference avoidance and interference exploitation innovations in heterogeneous environments is the need for tight frequency and time synchronization between cells of different hierarchical layers. These are new requirements that have not been considered in the first releases of the standard, as stated in [ARTD41]. In LTE Release 8, frequency stability requirements for macro, micro and femtocells are assigned different values, in order to recognize the difficulty of providing accurate frequency reference over the backhaul infrastructure used (xDSL, cable or fiber), as well as the use of low cost and low accuracy oscillators such as voltage controlled temperature compensated crystal oscillator (VCTCXO).

The solutions proposed may be classified into two main groups: over the air (OTA) solutions and backhaul based solutions. The option of considering that the femtocell may have direct access to an accurate clock reference (e.g., GPS) is not considered feasible due to installation requirements and cost reasons.

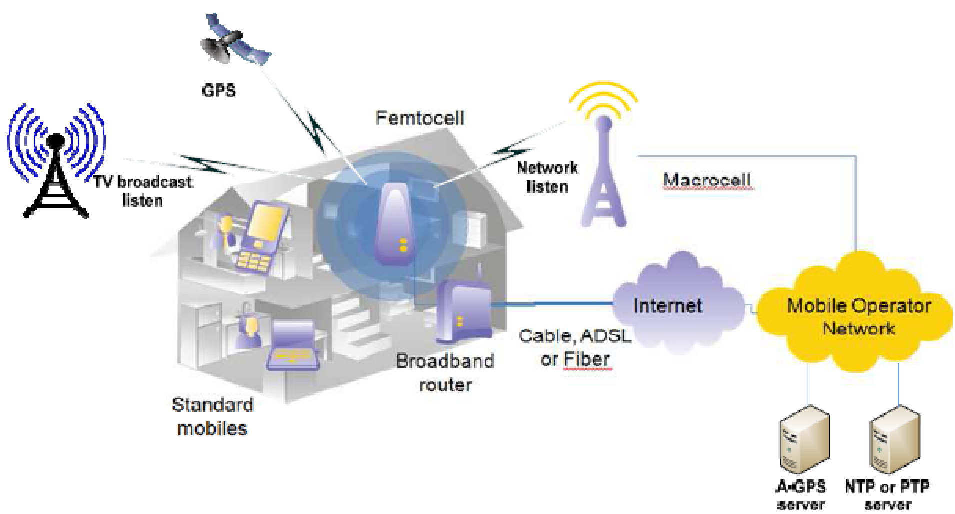


Figure 4-1: Femtocell synchronization solutions [FF1].

4.1.1 Over The Air based solutions

4.1.1.1 OTA time and frequency synchronization

The simplest option to achieve the required synchronization between FDD macro and pico/femto cells would be that the latter obtain the required clock reference from the former through the air interface. For this option to be supported, several requirements must be fulfilled:

- The femtocell needs to incorporate a receiver in the downlink frequency band. This may increase the cost and complexity of the femtocell.
- The femtocell needs to be able of listening to the synchronization and reference signals transmitted by the macrocell. If the signal from macrocells is significantly attenuated, the

use of macrocell signals for the purpose of synchronizing femtocells can result in delay and/or severe timing offset estimation errors. It is a cost-effective technique which does not imply extra infrastructure needed, the femtocell only needs a receiver that can perform a correlation on the synchronization signals. This receiver may be simply the same receiver that the femtocell uses to receive uplink traffic. In this case, the femtocell may be configured with longer guard period or it may suspend regular traffic from time to time in order to tune to the desired macro downlink carrier.

For synchronization purposes, the LTE standard defines two synchronization signals, namely the primary (PSS) and the secondary (SSS). In FDD mode, these signals are located on the 62 subcarriers symmetrically arranged around the DC-carrier in the first slot in the sixth and seventh OFDM symbols of the first and the sixth subframes. Also, cell-specific reference signals utilize 4-QAM modulated symbol alphabets and are mapped to the resource elements.

This solution is not considered appropriate for the requirements on FIC because the reception of the signal from the macro cannot be guaranteed so time and frequency synchronization cannot be also guaranteed.

4.1.1.2 TDD based OTA time and frequency synchronization

The solution proposed makes use of the TDD mode of the LTE standard for the synchronization and coordination of femtonodes serving final users in FDD or TDD mode.

Thanks to the establishment of a TDD signalling and communication interface supported by licensed TDD spectrum in a frequency band different from the one used to provide the final users with the required service, self-organizing femtonode networks are implemented, providing time synchronization for femtonodes as well as a signalling interface to exchange information about radio resource usage and inter-cell coordination. In this way, the interference among femtonodes will be minimised or even eliminated, providing coverage optimisation. Hereafter, in order to give the reader a clearer vision of the procedure developed in this solution, only the FDD serving mode will be considered in the text and the figures.

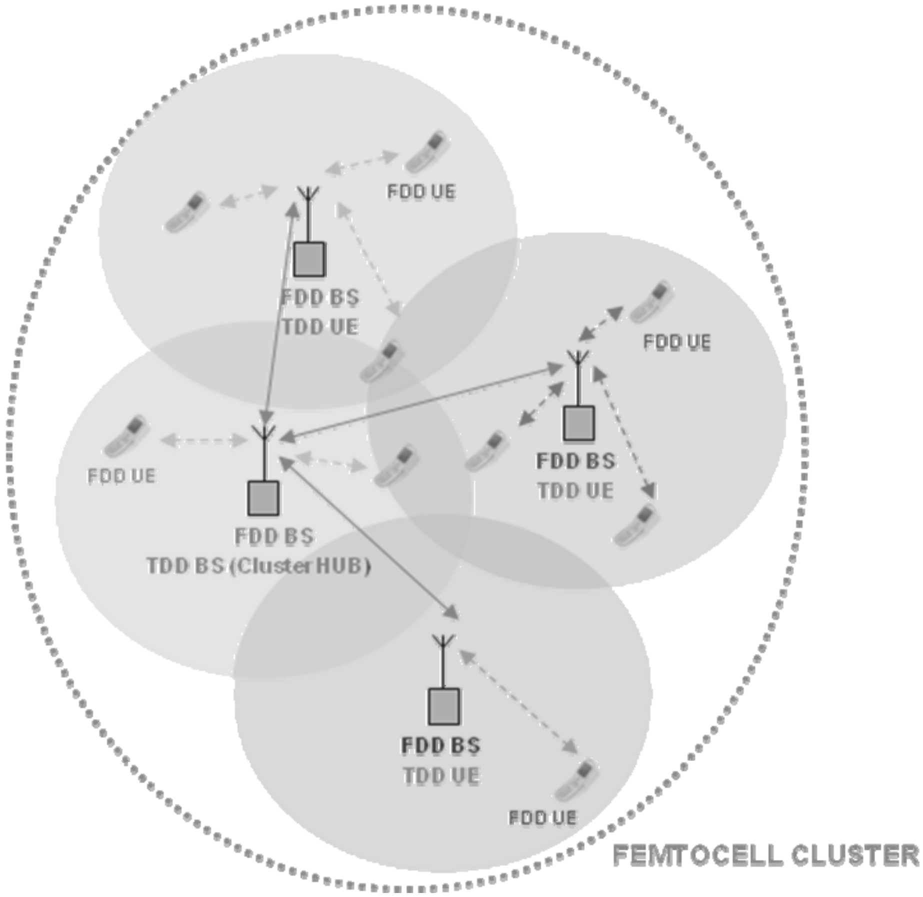


Figure 4-2: Femtonode cluster.

Neighbouring femtonodes are grouped into **clusters**. A cluster is similar to a TDD cell, with a femtonode playing the role of a TDD base station that is called the **cluster hub**, and several femtonodes playing the role of TDD user terminals (see Figure 4-2). Therefore, communication between any two femtonodes in the cluster is guaranteed, and can be used to exchange coordination messages or any other information required, in a similar way to the inter-eNB signalling X2 interface. The TDD base station, or cluster hub, transmits subframes in the DL and radiates reference and synchronization signals, as a standard TDD base station. The femtonode acting as TDD UEs transmit subframes in the UL and join the cluster using the UE registration method.

Time and frequency synchronization for all the femtonodes within the cluster is achieved using the standard TDD synchronization method. The cluster hub femtonode is the clock reference for the whole cluster and the rest of the cluster members synchronize their reference clocks in frequency and time with this master clock. As FDD and TDD modes present the same frame length (10 ms) and slot partitioning (20s in a frame), the TDD frame timing obtained by the use of this time and frequency synchronization method will be directly applied to the FDD mode.

Self-configuring is one of the main features of the cluster. Any femtonode can be the hub “a priori”. Its role is determined at power-on depending on the environment (initial scanning in TDD band). At any time, a femtonode may appear or disappear (the owner may switch it on/off), therefore network roles are dynamically reassigned in a decentralized way with the same procedure as the self-configuring at power on.

In order to avoid any potential interference among clusters, different carriers are used by each cluster. This is achieved by means of a random allocation of carrier frequencies during the self-configuration process of the cluster hub femtonode.

As in previous solution, as the hub gets the synchronization from the macro over the air the solution is not considered appropriate for the requirements on FIC because the reception of the signal from the macro cannot be guaranteed so time and frequency synchronization cannot be also guaranteed.

4.1.1.3 TV signal based time and frequency synchronization

A different option for OTA synchronization is the use of the TV signal for providing both frequency and time synchronization as all global standard TV signals include information that can be used for precise time and frequency. TV signal over the UHF band would have a number of significant advantages over satellite based options:

- High transmission power (more than 1 MW for some transmitters).
- Use of lower frequencies with better indoor coverage characteristics.
- TV signals are wideband, enabling efficient mitigation of multipath effects.
- TV signals have stable timing and are highly reliable, even during disasters.
- TV signals are broadcast across every metro area on Earth.
- Less attenuation compared with satellite options (GPS, Galileo) due to horizontal signals.
- Frequency diversity due to the transmission of several channels per transmitter.

The technology has been patented by an American company named Rosum [ROS1] (although its patent portfolio has been recently sold to TruePosition). In terms of the frequency accuracy that can be achieved, Rosum indicates it is possible to get 10 parts per billion, whilst in terms of timing accuracy it is possible to get to 1 microsecond accuracy. The TV based solution can also be combined with GPS for enhanced accuracy in those areas where TV signal is not available. Timing hybrid TV-GPS indoor is better than 1 microsecond and is typically better than 300 ns. Frequency accuracy is generally better than 10ppb and is typically better than 5 ppb.

The main barrier to the adoption of this technology (which can be supported over different digital TV standards, like ATSC, DVB-T or ISDB-T) seems to be the large size of the chips, which would be difficult to integrate into mobile phones. This, however, may be a not so significant problem for its use in femtocells.

This option, however, would require a more widespread support of the technology to make sense from an economic viewpoint.

This option meets the requirements on Table 2-2 but a detailed studied on the economic impact of inserting the chips in all femtos should be performed. It is a possible solution but needs further analysis.

4.1.2 Clock recovery from the backhaul

A second set of solutions is based on the transmission of the clock reference from the fixed network. The most straightforward solution would be the use of IEEE 1588 v2 or NTP synchronization protocol, that supports both time and frequency synchronization. These options have been analyzed in [ARTD41].

The main problem here is associated with the synchronization requirements of femtocells. Possible congestion issues in the backhaul may preclude achieving the accuracy requirements associated to some of the innovation proposed by ARTIST4G. It is thus interesting to locate time servers in points where the risk of congestion is as low as possible.

A possible example of a backhaul network for a femtocell deployment with several aggregation points is represented in the following Figure 4-3.

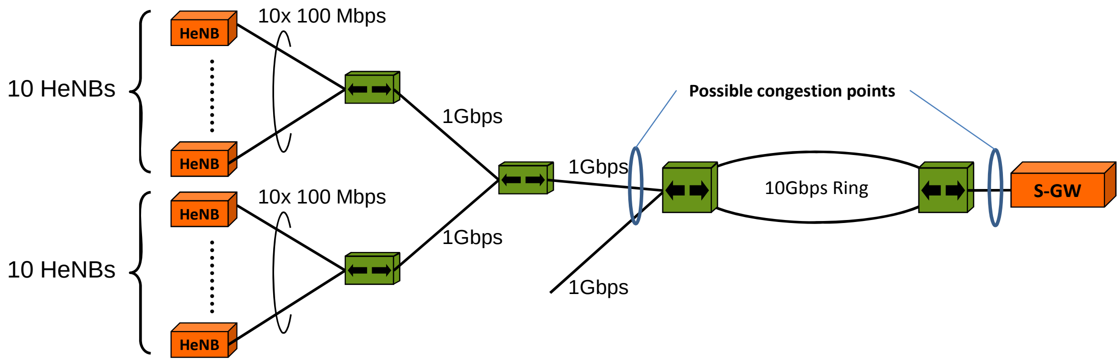


Figure 4-3: Example of backhaul network for a femto deployment.

In the case represented above it would be interesting to locate the clock server as close to the eNBs as possible and any case 'below' the possible congestion points.

In this sense two alternatives are considered for clock distribution from the backhaul that are:

- Clock distribution from the macrocell.
- Clock insertion in the backhaul network.

4.1.2.1 Clock distribution from the macrocell

For those innovations that require a high level of cooperation, there may be direct links between macro cell and the micro/pico/femto cell. These links should use an enhanced X2 interface or be based on a new interface, as explained in section 3.1.4. So it could be possible to enhance the X2 interface in order to support the distribution of the clock signal from the macrocell to the neighboring small cells, as stated in next section 4.2. For these purposes, the use of IEEE 1588 is a natural option, as it provides the most accuracy in both time and frequency synchronization.

According to the recommendations provided in [ARTD41] and the requirements of the innovations in Table 2-2 the natural recommendation to provide time and frequency synchronization to these innovations is the use of PTP IEEE 1588 v2 or NTP. All of them will fall in Groups C or D of the classification in section 5.1.2 in [ARTD41]. The only issue not considered is that FIC1 needs 50 ppb in frequency synchronization instead of 250 ppb which will be the baseline. With both of these solutions, NTP or PTP, this requirement will be met. A specific study where each innovation is going to be deployed should be done in order to find out which will be the most interesting of the two solutions for each specific situation.

4.1.2.2 Clock insertion in the backhaul network

One possible solution to reduce the number of hops between the clock source and the femtocell is to insert the clock signal in an intermediate node. It seems likely that aggregation nodes will be deployed in order to reduce backhaul costs for femtocells. It is a solution to the possible lost in accuracy for IEEE 1588 or NTP where the network is congested, but any of the two protocols should be present to transmit the synchronization.

One option is to incorporate the timing server in the BRAS (Broadband Remote Access Server) or the ONT (Optical Network Termination), depending on whether the femtocell backhaul is based on xDSL or fiber. However, it would be preferable to insert the clock signal in a node closer to the femtocell, as to avoid congestion issues that would compromise the synchronization accuracy. For example, in FTTN or FTTC deployments, the clock signal may be incorporated in the MDU (Multi Dwelling Unit).

4.2 Alternatives for X2 support between femtocells and macrocells. Estimation of the associated latency

3GPP release 10 [3GPP36300] permits X2 establishment between femtocells but no X2 establishment between femtocells and macrocells. Artist4G innovations, which need HetNet cooperation, can not currently use X2 to exchange information with Release 10 specifications. In order to solve this problem, this section aims to study alternatives for X2. As delay is a strong constraint for many cooperation schemes, this section will do a first estimation of the associated delay.

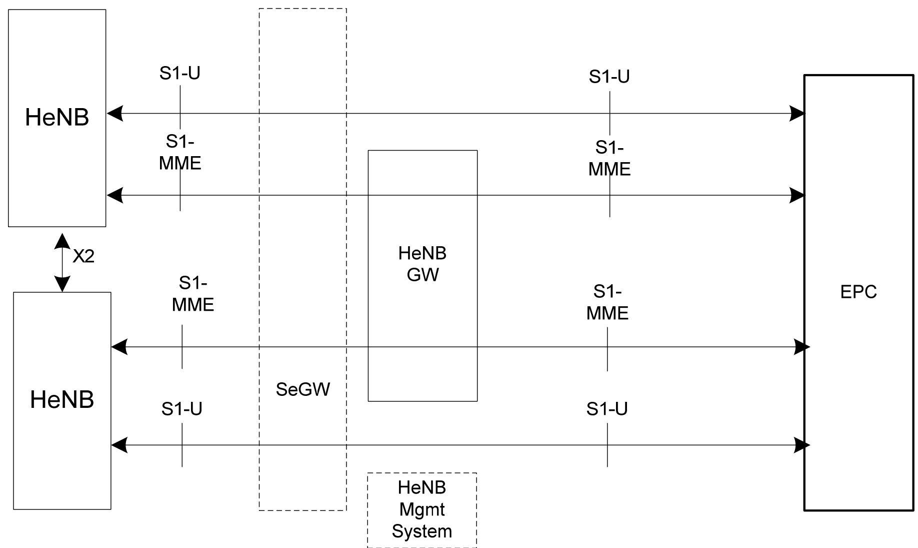


Figure 4-4: E-UTRAN HeNB logical architecture [3GPP36300].

Due to huge improvement in the number of X2 interfaces manageable by a single eNB, depending on manufacturer's implementations, 3GPP may permit and define X2 interface between femtocells and macrocells. This section will do some assumptions about this normalization possibility and estimate associated latency.

4.2.1 Alternatives for X2 support between femtocells and macrocells

4.2.1.1 Direct cooperation interface

As discussed in chapter 3.1.4 of this document, a direct cooperation interface could be a good candidate to meet CoMP requirements regarding INCF exchange between cooperating nodes. As this chapter will not define again the requirements of this interface, it will focus on the main differences and limitations between femto-macro and macro-macro cooperation.

- **Security concerns/SeGW**

  SeGW and HeNB GW are optional in current specifications. But it should be considered, at least for integrity control. As HeNB are not controlled by the operator, DoS attacks on macro network, like a huge number of link establishment demands from many non-controlled HeNBs, can also be a risk. SeGW and/or HeNB GW can cancel this risk by acting as a proxy for direct connection establishment.

- **Non permanent availability of femtocell and/or cooperation interface**

  HeNB are not controlled by the operator which means that they can be switched off/on without operator control. Connection interface (i.e. ADSL, …) can be erratic and force HeNB to connect and reconnect again and again with the network and the neighbouring

HeNBs. This can lead to a non negligible increase of signalling between femtocells and macrocells. The direct connection process shall therefore include a functionality to avoid the multiple connection retries and its signalisation messages. Or again, a proxy function, in HeNB GW for example, can cancel this risk.

The questionnaire in section 3.1.4.2 can be used for innovations involving femto-macro cooperation.

Site-to-site delay as defined in Figure 3-6 is quite different in femtocell case. Having a look on Table 3-2 indicates that ADSL2+ will add few ms to the E2E delay.

4.2.1.2 BBU Hostelling for business use case

As seen in chapter 7.1.1 in [ARTD41], BBU Hostelling will separate eNBs into two entities (RRH and BBU), and gather BBUs of several eNBs in a same place (called BBU Hostel). Communication between eNBs of the same BBU Hostel becomes local.

The same principle can be use with HeNBs, separate in a RRH and a BBU connected by D-ROF. It can be imagined a building with x HeNBs at the same stage, with each BBU in a single physical room, directly connected to the operator network. Cooperation between these femtocells and macrocells covering the area will be improved, due to the quality of the link between the BBU Hostel and the operator network, and due to less network equipments used. Cooperation between HeNBs will also be improved.

BBU Hostelling using D-ROF can only be used with optical fiber, which means the use only in a business use case scenario (building with optical cable).

4.2.2 X2 between femtocells and macrocells

4.2.2.1 3GPP Release 11 assumptions

Study Item *"Further enhancements for HNB and HeNB (LTE/HSPA*)" has been approved in 3GPP RAN WG3. There is a good chance that X2 between femtocells and macrocells will be discussed. In this case the best assumption is that restriction on X2 establishment between macrocells and femtocells will simply be removed. And principles from macro to macro X2 connection will be reused without any change.

HeNB GW proxy function for X2 also appeared in the specifications. The use of this GW can help to manage scalability issues on HeNBs and eNBs number of X2 interfaces, due to increasing number of HeNBs in the same geographical area. In this case only one X2 will be established between a specific HeNB and its HeNB GW.

4.2.2.2 Latency estimation

In HeNB deployment, backhaul constraints are more important, due to the access technology available to deploy a large number of femtocells. Copper access with xDSL technologies will be the majority. But some deployment could use optical fiber like GPON or P2P fiber (business use case). This means that the CoMP information has to go through equipment like DSLAM or OLT. Regarding Figure 3-6, site-to-site delay will change a lot compare to CoMP with only macrocells.

Copper access

See Table 3-2 delay added by the technology. In this case X2 physical link will pass through DSLAM.

Estimated delay (one way) is in the order of 10ms.

GPON

See Table 3-3 delay added by the technology.In this case X2 physical link will pass through OLT. Estimated delay (one way) is 5ms.

P2P Fiber

See Table 3-3 delay added by the technology.

Estimated delay (one-way) can be approximately equal as baseline X2 delay (few ms) between macrocells.

4.3 X2AP extension to support the signalling messages associated with the FIC procedure

4.3.1 Overview of architectural alternatives

When considering architectural alternatives a basic LTE philosophy needs to be taken into account: Core network (CN) is not aware about Radio Access network (RAN) aspects. A consequence of this principle is that in case a cooperation between (H)eNBs is required CN nodes will not be aware of this. The cooperation has to be dealt with inside the RAN.

Generally, Artist4G innovations from WP1 and WP2 require the support of communication among eNBs in the cooperation area. From a protocol point of view, there is no significant difference regarding necessary extensions to S1/X2 messages. Basically, this mesh of communication paths resembles the mesh of X2 interfaces between neighbours. However, considering the near real-time requirements for exchanging "CoMP" related information, it needs to be carefully considered whether a simple "re-use" of the X2 u-plane communication is adequate or not. Whether or not the communication for "CoMP" needs to be further subdivided into separate control and user channels is not ultimately clear for the time being. However, based on the responses received from WP1 it seems pretty much clear that such a separation is indeed not needed. Main reason is that there seems to be no requirement for an UE specific "CoMP" channel. In the following only one single communication channel is considered that is always available, once established.

Starting with the currently defined X2 protocol stack between neighbouring eNBs, the figure below introduces step by step various alternatives how to realise the "CoMP" interface. Furthermore, evolution of carrier backbones towards carrier-grade Ethernet is taken into account.

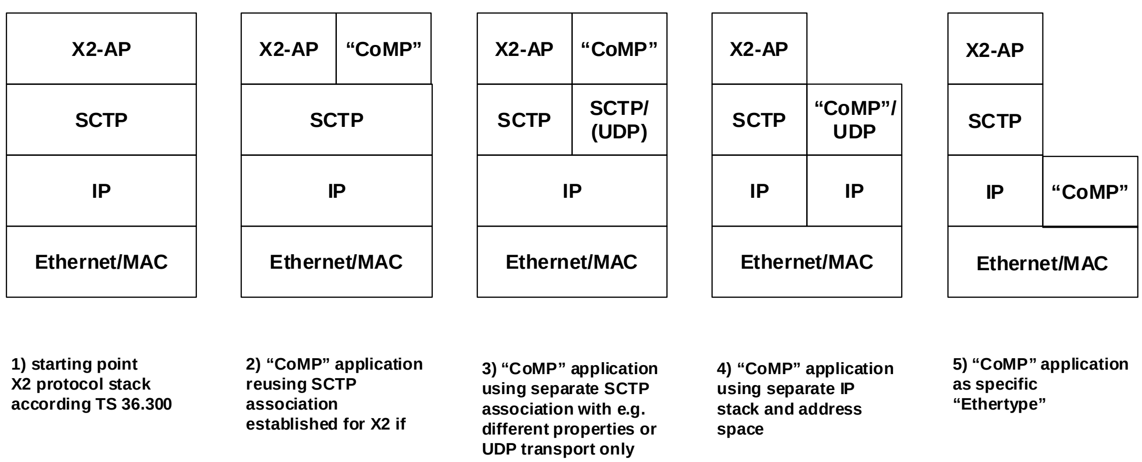


Figure 4-5: Implementation alternatives for the "CoMP" interface.

1. Current protocol stack of the X2 interface. The information about the TNL addresses to be used for SCTP association setup are communicated between the neighbouring nodes using S1 messages defined for eNB and MME configuration transfer. X2 uses IANA assigned SCTP port number (value = 36422) and Payload Protocol Identifier (PPI) (value= 27).

2. This alternative puts the "CoMP" application side-to-side with X2-AP and requires, from standardisation point of view, to specify "CoMP" application itself within 3GPP and to request from IANA the assignment of a PPI value for "CoMP" to identify and separate messages already on the transport protocol layer. Using the Partial Reliablity extensions for SCTP [RCF3758] the "CoMP" application might request SCTP to provide "partial reliability", whereas X2-AP still insists on reliable transfer of its messages."CoMP" messages need to be sent to the SCTP API and compete with X2-AP messages for transport resources. Standard SCTP does not provide its applications with the possibility to impact the SCTP internal scheduling process. Additionally, as SCTP provides for bundling of user data from several sources, it is pointless from an application's point of view to request a specific setting of the DS bits [RFC2474] in the IP header, as the result could be undesired usage of preferential packet handling by other traffic sources. Albeit not explicitly ruled out, the common practise is to assign identical DS values to all SCTP packets of a given association. "CoMP" messages need to be passed to SCTP layer using OS provided inter-process communication methods.
3. This alternative provides "CoMP" with a transport connection of its own. Depending on requirements this might even be pure UDP based datagram service in case no reliable and in-sequence delivery of messages is finally required. This alternative solves race conditions at the transport layer, would allow consistent setting of the DS field but still requires the "CoMP" messages to be passed to the common IP stack using inter-process communication.
4. This alternative allows "CoMP" to use an IP stack of its own, at the expense of requiring an additional IP address. The OS, however, needs to be able to distribute received messages based on destination IP address. From Ethernet/MAC view, frames to both applications have the same Ethertype value (e.g. 0x0800 for IPv4 protocol) assigned. Thus differentiation could take place only in the IP layer.
5. This alternative bypasses the complete IP stack, but requires the definition of a new "Ethertype" and supporting changes in the Ethernet driver. Nevertheless, this usage is commonly used for all applications requiring "real-time" behaviour. The majority of these applications originate from factory automation (e.g. EtherCAT, Ethernet Powerlink, …) Also the Precision Time Protocol (PTP) [IEEE1588] has been assigned a specific Ethertype (0x88F7) to allow for the envisaged accuracy. It would be only this alternative that allows the "CoMP" messages to be passed to a dedicated platform at the earliest point in time, not further impacting the already existing eNB software.
    - Ethertype: This is a two-octet field in an Ethernet frame. It is used to indicate which protocol is encapsulated in the payload of an Ethernet Frame. Effectively, it allows to use multiple protocols simultaneously on an Ethernet interface.

To summarize, alternative 5 guarantees for shortest possible delay and latency as it allows to bypass the IP protocol stack and further protocols on top of it. Furthermore, an implementation independent of existing eNB Software based on e.g. a separate "CoMP" platform is possible. Alternatives 1, 2, and 3 require optimisations to existing protocol implementations to ensure the required low delay. Furthermore, additional internal communication within the (H)eNB is required, the complexity, however, depends on the actual implementation.

A characteristic common to all alternatives is that peer nodes need to be informed about TNL addresses used to exchange cooperation information. Even in case this "CoMP" mesh is considered as fully independent of X2 interfaces, the mechanism about how to inform participating eNBs about "CoMP" capabilities and "CoMP" TNL addresses is suggested to be merged with the current methods for the eNB to get information about X2 TNL addresses of X2 peers. The message flow below shows the relevant details of the current way of exchanging information about X2 TNL addresses.

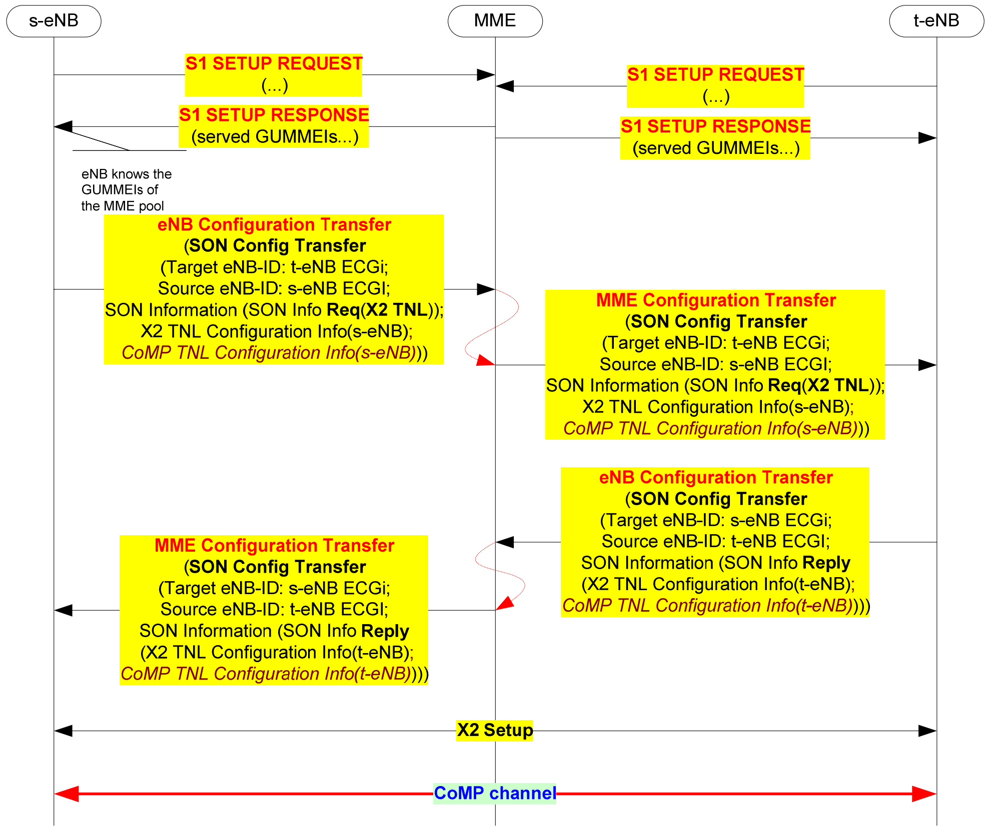


Figure 4-6: S1-AP message flow to exchange X2/Comp TNL addresses.

The message exchange already includes the exchange of CoMP TNL addresses. The benefit of introducing a separate IE to carry the CoMP TNL address information is that it allows for maximum flexibility. If this new IE reproduces the TNL address also used for X2 interface, CoMP message exchange could simply re-use the SCTP association established for X2 message exchange. Consequently, the “CoMP” messages would need to be defined as additional X2 messages, e.g. as transparent “CoMP” container.

4.3.2 Extensions to S1-AP messages

As shown in Figure 4-6 in a first step the S1 messages eNB and MME configuration transfer need to be modified to allow for sufficient flexibility in setting up the “CoMP” interface. Furthermore, the presence of the TNL address for the “CoMP channel” is used as an indication that the partner node supports CoMP. Therefore, no modifications to other S1 messages/IEs than eNB and MME configuration transfer are seen necessary for the time being.

As both messages make use of the SON Configuration Transfer IE only, necessary modifications would be required in few places only.

Current 3GPP specifications [3GPP36414] and [3GPP36424] state that:

“The Transport Layer Address signalled in S1AP/X2AP messages is a bit string of

a) 32 bits in case of IPv4 address according to IETF RFC 791; and

b) 128 bits in case of IPv6 address according to IETF RFC 2460.”

The definitions in both specifications would need to be extended to also allow IEEE 802.3 MAC addresses to be transported.

“The Transport Layer Address signalled in S1AP/X2AP messages is a bit string of

a) 32 bits in case of IPv4 address according to IETF RFC 791 and

b) 128 bits in case of IPv6 address according to IETF RFC 2460 .

c) 48 bits in case of IEEE 802.3 address is used for the "CoMP channel"

There are two possible ways to implement the necessary ASN.1 modifications in [3GPP36413]:

- add a new optional IE "CoMP TNL Configuration Info" to the SON Configuration Transfer IE and introduce the definitions for this IE;

or

- modify the X2 TNL Configuration Info IE to also contain the "eNB CoMP Transport Layer Address".

Both implementation alternatives require extensions to existing S1-AP messages and are described using ASN.1 in the following chapters. The two alternatives convey identical information, but use modifications in different Information Elements. As the newly added information is optional, backwards compatibility is ensured.

4.3.2.1 Enhancement of SON Configuration Transfer IE

This implementation alternative modifies the SON Configuration Transfer message by explicitly adding the optional CoMP TNL addresses, thus leaving the X2 TNL Configuration Information unmodified. Formally speaking, CoMP could be considered here as independent from X2.

This IE contains the SON Information and additionally includes the eNB identifier of the destination of this SON information and the eNB identifier of the source of this information.

Table 4-1: Enhanced SON Configuration Transfer.

IE/Group Name	Presence	Range	IE type and reference	Semantics description
SON Configuration Transfer				
>Target eNB-ID	M			
>>Global eNB ID	M		9.2.1.37	
>>Selected TAI	M		TAI 9.2.3.16	
>Source eNB-ID	M			
>>Global eNB ID	M		9.2.1.37	
>>Selected TAI	M		TAI 9.2.3.16	
>SON Information	M		9.2.3.27	
>X2 TNL Configuration Info	O		9.2.3.29	Source eNB X2 TNL Configuration Info
>CoMP TNL Configuration Info	*O*		*9.2.3.xx*	*Source eNB ComP TNL Configuration Info*

Necessary definition for the CoMP TNL Configuration Info (referred to as 9.2.3.xx)

The CoMP TNL Configuration Info IE is used for signalling CoMP TNL Configuration information for exchanging CoMP related messages between eNBs in a cooperation area.

Table 4-2: New CoMP TNL Configuration Info.

IE/Group Name	Presence	Range	IE type and reference	Semantics description
IP-Sec Transport Layer Address	O		9.2.2.1	Transport Layer Address for IP-Sec end-point
Transport Layer Address	M		9.2.2.1	Transport Layer Addresses for CoMP end-point.

4.3.2.2 Enhancement of X2 TNL Configuration Information IE

This implementation alternative modifies the X2 TNL configuration IE to allow a third set of TNL addresses to be specified. Formally speaking, CoMP could be considered here as part of X2.

The X2 TNL Configuration Info IE is used for signalling X2 TNL Configuration information for automatic X2 SCTP association establishment.

Table 4-3: Enhanced X2 TNL Configuration Info.

IE/Group Name	Presence	Range	IE type and reference	Semantics description
eNB X2 Transport Layer Addresses		*1 to <maxnoofeNBX2 TLAs>*		
>Transport Layer Address	M		9.2.2.1	Transport Layer Addresses for X2 SCTP end-point.
eNB CoMP Transport Layer Addresses				
>IP-Sec Transport Layer Address	*O*		*9.2.2.1*	*Transport Layer Address for IP-Sec end-point*
>Transport Layer Address	*O*		*9.2.2.1*	*Transport Layer Addresses for CoMP end-point.*
eNB X2 Extended Transport Layer Addresses		*0 to <maxnoofeNBX2 ExtTLAs>*		
>IP-Sec Transport Layer Address	O		9.2.2.1	Transport Layer Addresses for IP-Sec end-point
>**eNB GTP Transport Layer Addresses**		*0 to <maxnoofeNBX2 GTPTLAs>*		
>>GTP Transport Layer Address	M		9.2.2.1	GTP Transport Layer Addresses for GTP end-points (used for data forwarding over X2).

4.3.3 Extensions to X2-AP messages

With the exception of “Alternative 1” as described in the previous chapter, “CoMP” could be considered as a separate Application Protocol. None of the currently defined X2-AP procedures is considered useable for “CoMP” procedures. This view eases the definition of “CoMP” messages once they could be specified. Even in this case if it is later on felt that “CoMP” messages need to be described as part of X2-AP, they would form a completely different set of messages as compared to currently defined X2-AP messages and the independently defined “CoMP messages could be just incorporated into X2-AP.

As a consequence, no extensions for X2-AP messages between eNBs in the cooperation area are seen as being requiring Exchange of “CoMP” related TNL addresses is achieved by S1-AP.

As both proposed methods how to convey the additional “CoMP” TNL information between concerned nodes are considered equivalent, the decision which solution to take should be left for 3GPP discussions.

5 Architectural impacts of WP3 proposals

5.1 State-of-the-art in advanced Type-1 relaying

LTE-A relays as specified in the upcoming 3GPP Rel-10 [3GPP36814] will only be of Type-1, i.e. non-transparent: From the perspective of the user equipment (UE), a relay node (RN) appears in the same way as a regular evolved NodeB (eNB), i.e., it terminates all layer-2 and layer-3 protocols at the air interface. Furthermore, relay deployment is subject to the following two restrictions:

- There are at maximum two hops on the wireless transmission path between the donor eNB (DeNB) and the UE. Hence, data is forwarded via one single RN only.
- The deployed relay nodes are stationary (e.g. located on lamp-posts, walls, etc.) and thus always remain in the coverage of the same DeNB.

These restrictions prevent advanced setups, e.g. for fine-granular capacity enhancement in the macro cell via multiple layers of very-low-power RNs or relay-enhanced coverage on moving vehicles, like trains or buses. While the consequences of waiving these restrictions on the existing LTE-architecture and protocols will be discussed in detail in section 5.2 and 5.3, a brief overview of how these issues are already addressed in competing wireless systems is given in section 5.4.

5.2 Multi-hop relay deployments

The consequences of waiving the main restrictions on the existing LTE-architecture and protocols as described in the previous section are discussed in detail in the following.

5.2.1 Overview

Despite the terminology “multi-hop”, relay networks are considered to be limited in the number of allowed “hops”. Hence, their nature is not as general as in the area of e.g. “sensor networks”, allowing in general for an unlimited number of hops between source and sink.

For obvious reasons, e.g. delay, management overhead, and complexity, studies performed in ARTIST4G consider only two intermediate/transit relay nodes, before reaching the “root” or ”anchor” node, i.e. the donor evolved NodeB (DeNB), which has a direct physical connection to the mobile Core Network. The limitation allows a clear nomenclature to be used in designating the nodes, but already shows in an exemplary way the problems encountered with this technology.

Deliverable D3.4 [ARTD34] introduced two different structures for multi-hop relays.

- Tree structure
- Mesh structure.

The main characteristic of the tree-structured architecture is to allow communication paths only along the hierarchical structure depicted in Figure 5-1. Even more important, the tree structure easily ensures that there is only one single communication path available for c/u-plane communication (S1/X2 interface) between any two nodes. This is also reflected in the specifications allowing only one single S1 or X2 protocol instance between any two nodes.

Before going into more details, some definitions need to be introduced and explained.

- n-RN: A nested Relay Node, is a relay node connected via the Un interface with the DeNB and additionally serving other end relay nodes and UEs.
- e-RN: A end Relay Node, is a relay node connected via the Un interface with a nested relay node and serving only UEs.
- intra DeNB: Communication supported only between e-RNs and n-RNs served by the same DeNB.

- inter DeNB: Communication also supported between e-RNs and n-RNs served by the different DeNB.
- intra n-RN: Communication supported only between e-RNs served by the same n-RN.
- inter n-RN: Communication also supported between e-RNs served by the different n-RNs.
- same level: Communication supported between nodes of the same type (e-RN or n-RN).
- different level: Communication supported only between nodes of different types (relevant only in case of inter-nRN and inter-DeNB).
- physical connection: This term is used in case the two communicating entities are exchanging messages over a communication channel where no routing based on address information on layers above the Media Access Control (MAC) needs to be done. In case of two neighbouring e-RNs this could be e.g. a radio channel established between them. The source needs to have as many physical connections as there are communication partners.
- logical connection: This term is used in case the two communicating entities are exchanging messages over a communication channel where routing is based on address information being part of the transported payload. In case of two neighbouring e-RNs this could be a communication first sent to the n-RN and then the address information of the IP layer is used to pass the packet to the radio link towards the target e-RN. The source needs to have only one physical connection to the intermediate node performing the address evaluation.

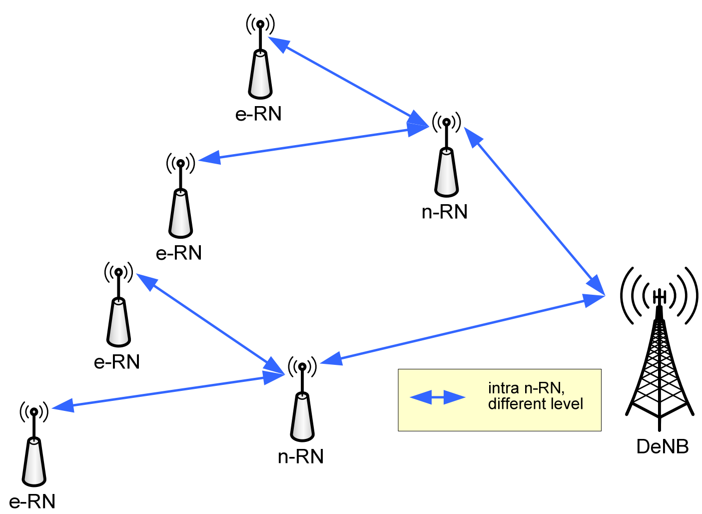


Figure 5-1: Tree-structured multi-hop RN architecture.

Using the definitions as introduced before the tree structured multi-hop RN architecture as shown in Figure 5-1 above could be described as “intra DeNB/intra-RN” having communication between “different levels” only. It has to be remembered that the communication links are carried via the radio interface between the nodes. Therefore, the tree structure naturally follows the physical layout of the communication network. Consequently, the X2 interface between e-RNs connected to the same n-RN needs to be carried via that n-RN or via the DeNB. The situation becomes even worse, if two adjacent e-RNs want to establish an X2 interface, but are served by different n-RNs: Their X2 interface would need to be carried via the DeNB!

Based on these facts, the addressing schemes used are the key to allow for ultimate flexibility and low delay. Current 3GPP solutions mainly rely on “application level routing” by introducing “proxy” functionalities in the node having multiple communication interfaces. The number of parallel active protocol instances is thus kept to a minimum in the node having only a single communication interface. This is the current solution when introducing relay nodes in Rel 10.

Starting with the pure tree structure as shown in the Figure 5-1 and introducing additional communication possibilities, different kinds of communication relations could be recognised in a mesh structure as depicted in Figure 5-2:

1. e-RN to e-RN, both attached to same n-RN: “Intra n-RN” (yellow line);
2. e-RN to e-RN attached to different n-RNs or n-RN to n-RN attached to same DeNB: “Inter n-RN” at “same level” (brown line)
3. e-RN to a second neighbouring n-RN but both n-RNs attached to same DeNB: “Inter n-RN” at “different level” (red line).

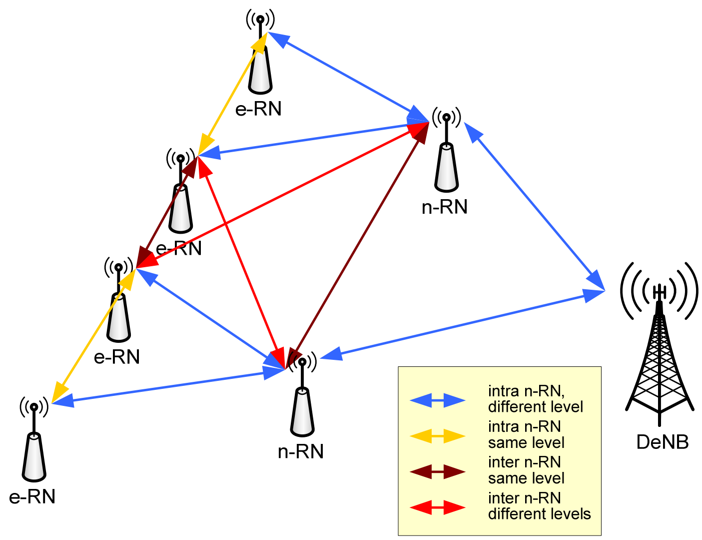


Figure 5-2: Mesh-structured multi-hop RN architecture.

Yet another generalisation is to allow for having direct communication channels between (e/n)-RNs that are served by different DeNBs. For the sake of completeness this even more advanced mesh-structured architecture is introduced below. This architecture is characterised by the fact that now n-RNs are allowed to have Un interfaces to more than one single DeNB.

The following additional communication relations are thus introduced in Figure 5-3 in addition to the previous figure:

1. communication between nodes (e-RN to e-RN or n-RN to n-RN) each served by a different DeNB within same hierarchical level ("Inter DeNB 1", purple line);
2. communication between nodes (e-RN to second n-RN or n-RN to second DeNB) each served by different DeNB crossing hierarchical levels ("Inter DeNB 2", magenta line).

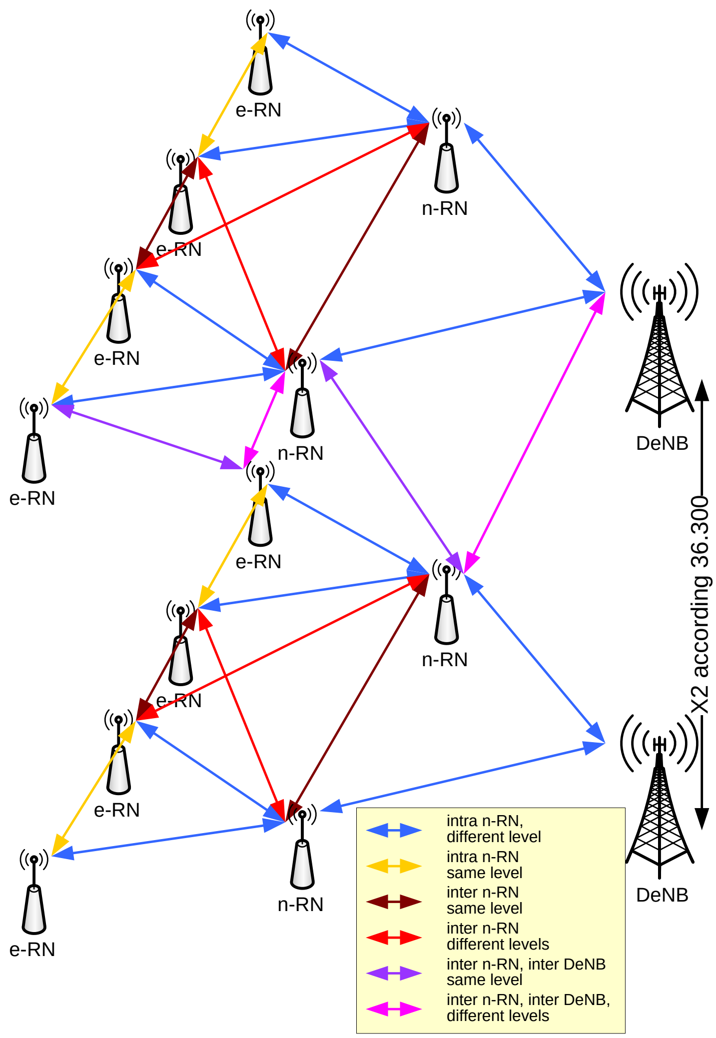


Figure 5-3: Generalised mesh-structured multi-hop RN architecture.

Considering these additional communication interfaces to span the mesh, it needs to be recognised that e-RN and n-RN are only able to use a radio based interface to communicate with neighbours. The only physical interface defined between relay nodes according to current specifications [3GPP36300] is the Un interface. This would, however, require an e-RN to also provide n-RN functionality. From an e-RN view, there would be then two different flavours of n-RNs. First a single n-RN the e-RN has the S1 and an X2 interface with. And second, a multitude of neighbouring "pseudo" n-RNs the e-RN has X2 interfaces with. However, requiring the e-RN to also provide n-RN functionality in order to allow a direct X2 interface with neighbours is not considered a viable way towards a multi-hop RN architecture, especially as setting up the Un interface is a rather complex procedure.

In general as long as these additional interfaces carry only X2 communication, it is only necessary to ensure that a certain neighbour is reachable via one X2 interface only. Direct X2 interface and proxying the X2 at higher hierarchical level in parallel is not supported by current protocol specifications.

Another approach would be to exchange messages with neighbouring nodes by using the Un interface between e-RN and n-RN and let the n-RN to route the message using Transport Network Layer (TNL) address information. Note, in this approach the n-RN does not need to act as proxy. The communicating nodes, however, need to have TNL address information instead, whereas in the proxy case Radio Network Layer (RNL) address information is used to determine the destination node of the message and to select the appropriate point-to-point connection. The main benefit in using RNL address information is the fact that this information could be easily derived from information broadcasted on the radio interface by the neighbouring node. Only the proxy needs to have the ability to map RNL to TNL address information.

5.2.2 Architectural analysis for tree-structured multi-hop RN

5.2.2.1 Introduction

Three different protocol stack approaches have been presented in [ARTD34]:

A) n-RN is fully transparent for DeNB – e-RN communication (blind packet forwarding, see Figure 5-4).

B) n-RN is acting as IP router for DeNB – e-RN communication (see Figure 5-5).

C) n-RN provides proxy functionality for the attached e-RNs (cascaded proxy, see Figure 5-6).

There is yet one more approach possible not involving any changes to the protocol stacks at all:

n-RN acting as an Amplify and Forward repeater (AF-relay) which would not be visible from a protocol stack view. This solution does not have any architecture impact at all. Having an AF relay can be suitable in case there are no UEs that need to be served at this intermediate location. However, it is known from the literature that performance of AF relays is poor compared to relays acting on Layer 3 [BRR+09].

Depending on the chosen protocol stack variant used in the tree structure, c-plane communication between (e/n)-RNs might be urged to be passed via the DeNB forming the “root”.

For ease of reference the S1 stacks as shown in the figures in section 7.1 of [ART34] are reproduced in the subsequent chapters. The diagrams for X2 would be similar, except not having the NAS layer extending to the CN.

5.2.2.2 Analysis Architecture A “blind forwarding”

All communication needs to go via the DeNB, designated “Proxy” in Figure 5-4. This adds at minimum (20 + 8 + 12) = 40 bytes to the Un interface between DeNB and n-RN for extra IP, UDP, and GTP-u headers. Due to header compression in PDCP, this number will be reduced significantly, but currently there is no profile defined for further compression of the embedded next IP header [3GPP36323]. In section 5.1 of 3GPP TR 36.806 [3GPP36806] this issue is already investigated and shows the benefits of either introducing an additional profile or applying a two step compression scheme.

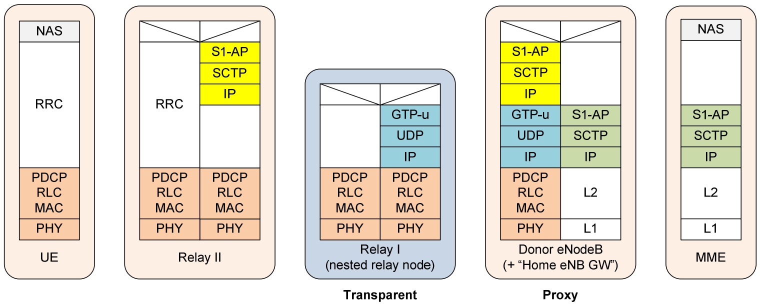


Figure 5-4: Architecture A: "blind packet forwarding".

Apart from the additional GTP tunnel header on the interface between DeNB and n-RN for traffic to/from e-RNs, no additional functionality on c-/u-plane is required. Conceptually, this architecture needs to have an additional definition for the Un interface between n-RN and DeNB. This finally impacts the DeNB as it needs to have different Un variants towards Rel-10 RNs and n-RNs (later releases). The main functional difference is that in order to support e-RNs, DeNB and n-RN have to agree on TEIDs for the additional tunnelled Un interface.

For establishment of the SCTP association carrying the S1 interface, the IP address to be used at the DeNB side is made available to the e-RN already via the Rel 10 procedures.

The transparent n-RN (referred to as Relay I) is assumed to have as many e-RN (referred to as Relay II) specific GTP-u tunnel towards the DeNB as there are e-RNs served by the n-RN.

Functionally, the n-RN, when receiving a packet from a certain e-RN determines the GTP tunnel the packet has to be sent to. The DeNB needs to run individual SCTP instances per e-RN and n-RN. The SCTP association originating at the e-RN is transparent for the n-RN. Therefore, no S1/X2 proxy functionality on behalf of the e-RN is present at the n-RN.Logically, there is no necessity for the DeNB to differentiate between e-RN and n-RN as there are individual S1-AP instances per node. At the DeNB there are individual SCTP associations per e-RNs and per n-RN.

5.2.2.3 Analysis Architecture B "IP routing"

The protocol stack depicted in Figure 5-5 shows that no additional bytes need to be transported on the Un interface between the DeNB (referred to as proxy") and the n-RN (referred to as router), routing on IP layer using TNL addresses in n-RN possible. To keep the picture simple, there is no SCTP/S1-AP shown on the n-RN as this is considered to be naturally present at the n-RN and only the additional IP routing is shown here. The local SCTP/S1-AP is addressed via the n-RNs TNL (IP) address, whereas the SCTP/S1-AP of the e-RN is addressed via the e-RNs TNL (IP) address.

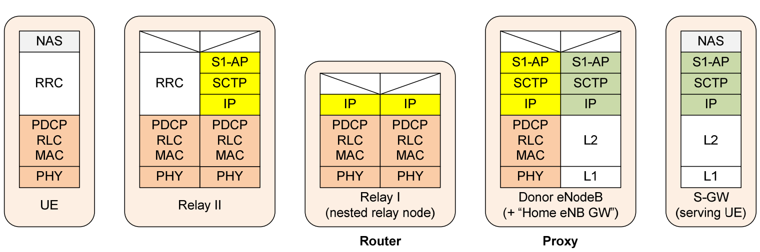


Figure 5-5: Architecture B: "IP routing".

Taking into account the fact, that the e-RN is informed about the IP address to be used for setting up the S1 interface with its DeNB, a simplified approach could be used at the n-RN. The e-RN considers the n-RN as its "default gateway" from IP perspective. This requires the n-RN, after having received any IP packet via the radio link from one of its e-RNs to pass the packet to IP routing process. In a first step, the routing table is defined in a way to forward all messages to the DeNB. This is due to the fact that procedures allow only communication towards the DeNB.If in a second step, direct communication between e-RNs would be allowed, e.g. for X2 interfaces, then n-RN could keep a more sophisticated IP routing table to route messages also between e-RNs if the proper IP addresses are used. Introducing this direct e-RN to e-RN communication, however, requires the e-RN to support multiple SCTP associations, one per partner e-RN. The SCTP association originating at the e-RN is transparent for the n-RN. Therefore no S1/X2 proxy functionality on behalf of the e-RN is present at the n-RN. At the DeNB there are individual SCTP associations per e-RN and per n-RN.

5.2.2.4 Analysis Architecture C "cascaded proxy"

No additional bytes to be transported on the Un interface between the DeNB and the n-RN. X2 proxy functionality in n-RN could route to target.

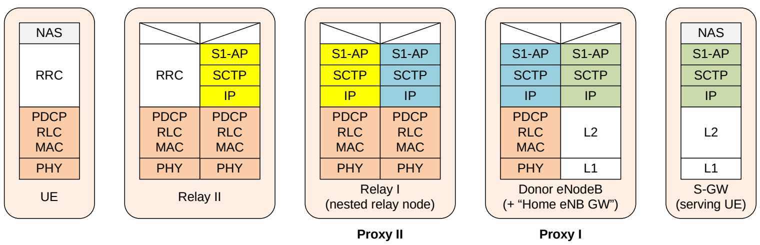


Figure 5-6: Architecture C: "cascaded proxy".

The DeNB needs to run one single SCTP instances per n-RN but the additional SCTP instances per e-RN are no longer required as e-RNs appear as cells served by the corresponding n-RN.Logically, the DeNB would consider the cells spanned by the e-RNs connected to a given n-RN as cells spanned by this n-RN.The SCTP association originating at the e-RN is terminated at the n-RN and S1/X2 proxy functionality needs to be provided.

5.2.3 Architectural analysis for meshed multi-hop RN

With respect to S1 interface, a fully meshed architecture potentially collides with basic requirements of the LTE-Advanced architecture.

- Only one single S1 interface is allowed between an eNB and it's serving MME.
- In case of LTE relay this applies to the interface RN to DeNB.
- In case of multi-hop relay this applies to the interface e-RN to n-RN or n-RN to DeNB.
- There is only one X2 interface allowed between any two neighbouring eNBs not having the same eNB ID.
- Cells of (e/n)-RN are required to have an E-UTRAN Cell Global Identity (ECGI) sharing the leftmost 20 bits with the ECGI of the serving DeNB.
- Consequently not more than 4096 (2**12) nodes, n-RNs and e-RNs could be served by one D-NB.

Consequently the additional interfaces shown in Figure 5-2 and Figure 5-3 could only carry an X2 interface to the adjacent node.

The remaining advantage regarding the S1 interface, when introducing physically fully meshed (e/n)-RNs would be path-diversity. The only way to benefit from path diversity is to exploit the capabilities of SCTP in using multiple paths when establishing an SCTP association (multi-homing). However, this comes only at the additional burden to properly manage the IP addresses assigned to the (e/n)-RNs to let the path diversity be reflected in the IP layer.

Another requirement originating from support of “multi-homed” SCTP associations is that nodes in the path need to provide IP routing functionality, but not “route” on application layer, i.e. no Proxy functionality.

5.2.4 Summary and conclusions

An important aspect to be considered along with the introduction of a tree structured multi-hop architecture is the backward compatibility issue. The table below is comparing different protocol stack proposals made for the tree structured multi-hop RN.

	“blind forwarding”	“IP routing”	“cascaded proxy”	“AF relay”
e-RN	Rel 10	Rel 11 [1)]	Rel 10	Rel 10
n-RN	Rel 11 or later [2)]	Rel 11 [3)]	Rel 11 or later [4)]	-
DeNB	Rel 11 or later [5)]	Rel 11 [6)]	Rel 10	Rel 10

Table 5-1: Compatibility of tree structured alternatives.

Notes:

1. e-RN is required to know the IP address of the other e-RNs belonging to the same DeNB. This is to avoid that X2 messages go through the DeNB. The additional complexity on top of Rel 10 is probably not too high.
2. n-RN is required to remove the additional GTP tunnel headers.
3. n-RN functionality is Rel 10 enhanced with the IP routing functionality. The additional complexity on top of Rel 10 is probably not too high, especially as IP routing has no impact on 3GPP specifications.
4. The Rel 10 RN and Rel 10 DeNB are combined to provide the n-RN functionality. It is assumed to become a highly complicated and complex device.
5. DeNB has to add one more GTP tunnel if message needs to be sent to e-RN.
6. DeNB has to distribute the IP addresses of all RNs belonging to it in order to avoid that X2 messages continue going via the DeNB. The additional complexity on top of Rel 10 is probably not too high.

According to the table above, it would be possible to introduce the tree structured multi-hop based on existing infrastructure elements already starting using Rel 10 compliant equipment, depending on the solution selected.

Meshing of multiple nodes could be done by using either a

- logical mesh, where communication between nodes is “routed” via a limited number of intermediate nodes (see e.g. the Internet), or a
- physical mesh, where communicating nodes have a direct physical connection.

Main benefit of using a physical mesh instead a logical mesh is that this allows to have “path diversity”. This, however impacts the characteristics of the relay node: Whether it is acting as e-RN or n-RN is no longer preserved when applying the mesh structure. The node characteristic even changes depending on the current communication relation, as two nodes might act as e-RN and n-RN for one communication, but switch the role for another communication relation.

As current LTE architecture is not suited to support physically fully meshed multi-hop architectures and due to the additional burden that in physically meshed multi-hop architecture every e-RN needs to also have n-RN functionality, it is suggested to not further consider physically meshed multi-hop architectures and continue with tree structured multi-hop architecture only.

Regarding the different architectural options from protocol view, architecture B “IP routing” is the preferred solution as it has the least impact on additional functionality in both the DeNB and the n-RN. Furthermore, not having an additional S1 proxy at the n-RN is also beneficial from a delay perspective and having end-to-end SCTP transport, failure detection is also simple and does not require additional messaging on S1-AP level.

5.3 Mobility and latency aspects of moving relays

5.3.1 Overview

Relay nodes seamless mobility is a key enabler for the deployment of relays in public transportation systems (PTS). In such deployment scenario, type 1 relay nodes can provide coverage and capacity enhancement and improve the QoS experienced by the users of the PTS.

In previous study [ARTD34] the feasibility of X2 handover of type 1 moving relay in the framework of the current 3GPP architecture alternative A 2 has been investigated. More precisely, it has been compared the X2 handover latency of the alternative A 2 moving relay to the X2 handover latency of the moving relay considered within architecture alternative A 1.The conclusion of this study is that architecture alternative A 2 needs to be upgraded to support relay nodes mobility. In this section one possible upgrade is presented and a evaluation of the corresponding X2 handover latency is performed.

5.3.2 Optimized architecture for moving relays

5.3.2.1 Moving relay handover procedure

The basic idea of alternative A 2 is to move the S/P-GW of the relay node into the DeNB in order to improve the latency of the inbound handover of the UEs attached to the relay node. The DeNB terminates the S1 and X2 interfaces and provide connectivity to the MME and neighbouring DeNBs. This principle is maintained in the case of moving relays so the S1 and X2 interfaces needs to be re-established during the moving relay handover. Moreover, the natural architecture alternative A 2 handover procedure described in [ARTD34] shows that connectivity with OAM will be lost and, so, the moving relay node may fail to get configuration information of the target DeNB that will lead to radio link failure of UEs attached to it.

The optimized architecture proposed in this section intends to solve both problems through:

1. Fast S1/X2 proxy relocation at the target DeNB during handover execution phase so the latency of the setup of S1/X2 interfaces to the target is minimized.
2. Maintain the connectivity with source DeNB during the handover execution phase such that the moving relay node can get the configuration information of the target DeNB from relay node OAM through OAM signalling carried to the relay node by the source DeNB.

The overall handover in the optimized architecture includes a phase of soft handover-like procedure that is executed during the handover execution phase. In this soft handover phase, the moving relay is having connectivity to both source and target DeNB while the source DeNB sends *proxy relocation* message to the target DeNB that is similar to UMTS fast serving RNC relocation. Group path switch of the users attached to the moving relay is performed based on this proxy relocation message.

Figure 5-7 illustrates the architecture elements involved in the optimized moving relay architecture.

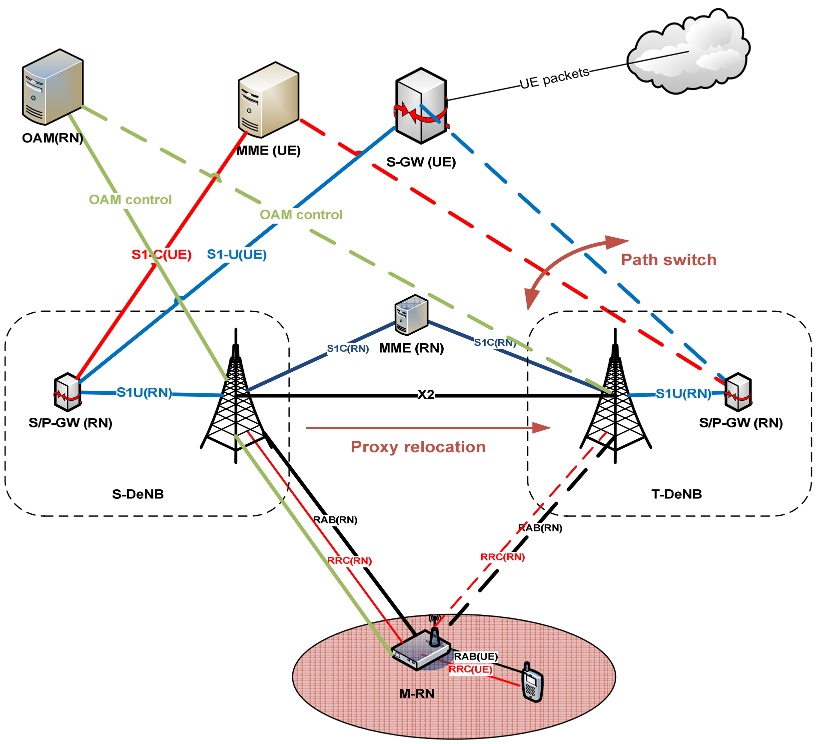


Figure 5-7: Architecture elements involved in moving relay handover (optimized architecture).

Detailed X2 handover process for the optimized architecture will be described next. The X2 handover can be cast as for the standard alternative A 2 into handover preparation, handover execution, and handover completion phase:

- Moving relay handover preparation phase

  In this phase, the following steps are performed between the S-DeNB and T-DeNB:

  - The moving relay sends a measurement report to the source DeNB in order to trigger handover to the target DeNB. The measurement report includes the eNB ID of the target DeNB.
  - A handover decision is made at S-DeNB based on measurement reports of m-RN and S-DeNB issues a message to T-DeNB including the necessary information to prepare the moving relay handover at T-DeNB. This message includes E-RAB and QoS parameters of the moving relay.
  - Admission control is performed at T-DeNB dependent on the received E-RAB/ QoS parameters of the m-RN. Two possible admission control solutions can be envisioned :
    - o Admission control based on the E-RABs and QoS parameters of the Un interface of the moving relay.
    - o Admission control based on the E-RABs and QoS parameters of the Uu interfaces of the UEs attached to moving relay.

  In the first alternative, the moving relay is admitted to T- DeNB if it has enough resources to support the Un bearer of the moving relay. The second alternative admits the moving relay on UE/group of UEs basis and has the advantage of minimizing UE call blocking in case of highly loaded environment. Since in this case S-DeNB knows the mapping between Un RABs and the UE RABs, it can transmit the UEs RABs attributes to T-DeNB in order to perform alternative 2 admission control for the UEs. It is important to note that this admission control scheme needs additional information elements (IEs) to be transmitted in the X2 handover message since S-eNB indicates in the HO preparation message the type of E-RAB (i.e. E-RAB of moving relay or UE).

- T- DeNB sends a Handover request ACK message to S-DeNB that includes RNL/TNL information for data forwarding for m-RN and necessary security information of T-DeNB to be forwarded to m-RN.
- S- DeNB sends HO command to the moving relay in order to execute the handover.

* Moving relay handover execution phase

  In this phase the following steps are performed, involving S-DeNB, T-DeNB and m-RN.

  - T-DeNB generates an RRC message to perform handover. This message may include (as for standard UE handover) some modified *RRC Connection Reconfiguration* message and the *mobilityControlInformation* message to be sent by S-DeNB to m-RN. These RRC messages contain the necessary information for backhaul reconfiguration of m-RN towards T-DeNB. Note that the necessary RRC information used in this step is new and specific to the Un interface (i.e. different from the RRC configuration information exchanged during the handover of a regular UE). The moving relay gets control and configuration information of T-DeNB from its OAM through the OAM connectivity with S-DeNB.
  - S-DeNB sends an SN STATUS TRANSFER message to T-DeNB conveying PDCP SNR receiver status for UL and DL of m-RN in order to ensure in-sequence data forwarding of the m-RN SDUs from S-DeNB to T-DeNB.
  - S-DeNB starts forwarding for the user plane of the moving relay to T-DeNB. In the case selective admission control is performed, more sophisticated methods regarding data forwarding could be envisaged like forwarding to multiple T-DeNBs in parallel. However, this needs more investigation, especially as it drastically increases the amount of traffic sent back and forth between DeNBs.
  - Moving relay node performs backhaul reconfiguration to T-DeNB and accesses the resources signalled previously in the Un specific RRC message and OAM message from S-DeNB. When compared to a regular UE handover, no RACH procedure is needed in this case. On the uplink the data of the UEs are buffered in the m-RN, on the downlink the incoming data for m-RN are buffered in the source DeNB. This backhaul reconfiguration is performed while maintaining the connectivity with S-DeNB. The moving relay waits until OAM configuration information is received before initiating the backhaul reconfiguration to T-DeNB. The OAM connection is maintained with OAM via S-DeNB in order to transmit the configuration from m-RN to OAM. This step corresponds to the soft handover like phase introduced previously.
  - Moving relay node sets up S1 control and user plane with T-DeNB.
  - S-DeNB sends to T-DeNB *proxy relocation message.* This message is including TNL information used by S-DeNB for the user plane of the UEs attached to the moving relay. This proxy relocation message will be used by the T-DeNB to prepare or to group path switches of the UEs attached to m-RN.

* Moving relay completion phase:

  In the last phase of the moving relay node handover, the following steps are performed:

- Based on the received *proxy relocation* message T-DeNB initiates the PATH SWITCH request for MMEs of the different UEs attached to the m-RN. The MME of the UEs sends a update user plane message to the corresponding S/P-GW.
- The target DeNB uses the proxy relocation message to group the user plane PATH SWITCH request messages to the MMEs of the UEs which send an Update user plane message to the S/P-GW(UE).of the different UEs.
- The S/P-GWs of the UEs switch the downlink data path to the target DeNB.
- The target DeNB sends one or more end marker packets and then S-DeNB releases any u-plane/TNL resources of the m-RN.
- The S/P-GWs sends an UPDATE USER PLANE RESPONSE to the MMEs of the UEs attached to the m-RN.
- T- DeNB confirms the target DeNB PATH SWITCH message with the PATH SWITCH ACKNOWLEDGE message and sends path switch response to the m-RN.
- T- DeNB sends m-RN context release message to S-DeNB and the connection between m-RN and S-DeNB is released.

The advantages of the proposed moving relay handover compared to the handover procedure described in previous studies [ARTD34] can be summarized as:

- The connectivity between m-RN and source DeNB is maintained during the m-RN handover so the m-RN can obtain the necessary configuration information from the relay node OAM.
- T- DeNB initiates user plane path switch for the m-RN through the proxy relocation message. Improvement is expected for handover latency since the procedure saves at least one m-RN path switch message over Un interface compared to the case where the m-RN initiates the user plane path switch as seen in Figure 5-9. The impact on the overall delay of this user plane handling is analysed further in section 5.3.2.2. .

Figure 5-8 summarizes the different steps of the m-RN handover procedures described previously.

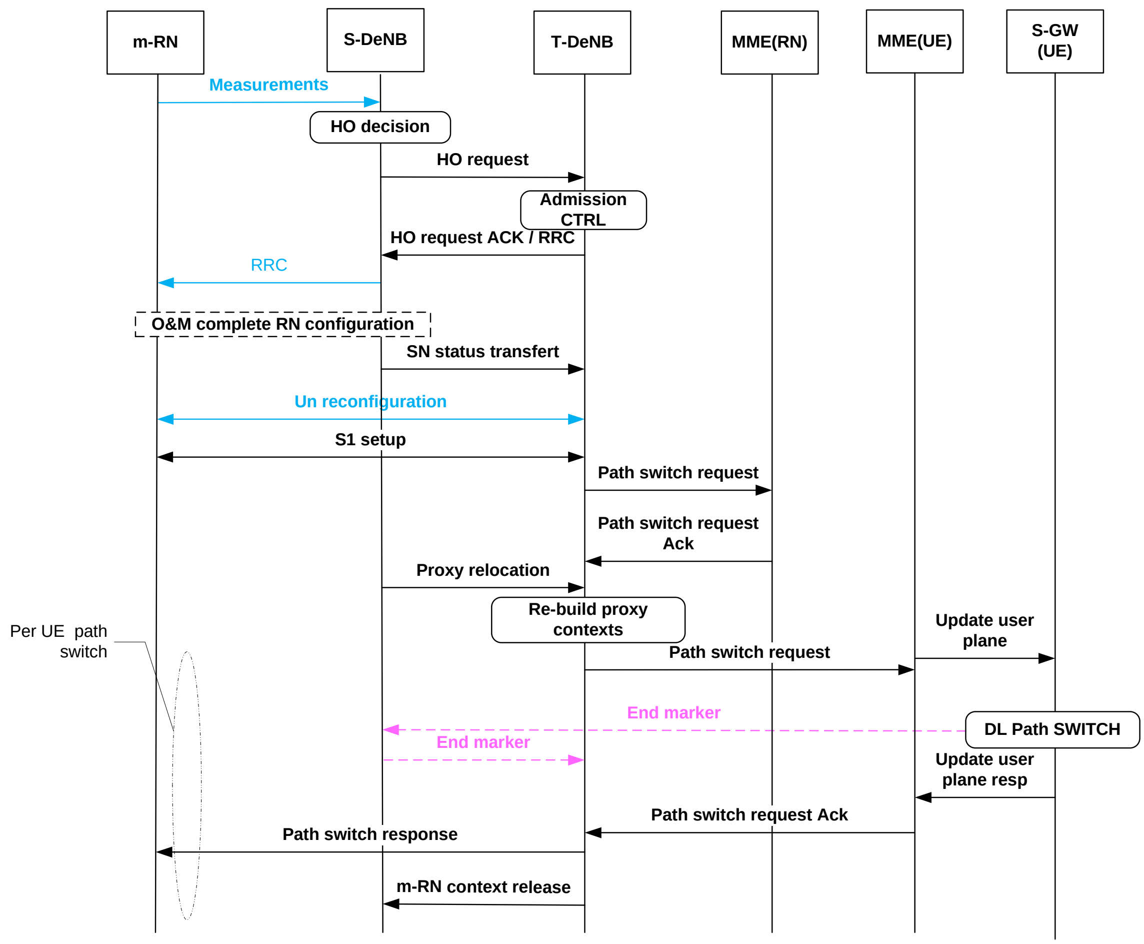


Figure 5-8: Flow chart for m-RN handover in the optimized architecture.

5.3.2.2 Moving relay handover latency analysis

In this section an initial analysis of the moving relay handover latency for the optimized moving relay architecture is presented. Two assumptions are made for the analysis:

- It is assumed that the latency of OAM reconfiguration of m-RN is around 12 ms which is the same as the latency of the communication between eNB and MME and 2 ms processing in the m-RN.
- It is assumed that the latency of S1 setup of m-RN to the target eNB is calculated as the exchange of 8 SCTP control messages over Un and S1AP control messages (S1 setup request, S1 setup response). This leads to an estimate of 24 ms delay for the S1 re-establishment with T-DeNB.
- It is assumed that the delay of rebuilding proxy contexts is 2 ms

The handover latency for the moving relay in the optimized architecture is given in the following table:

Table 5-2: Handover latency for the optimized architecture.

Message	Latency
m-RN transmit measurements to S-DeNB and processing in DeNB	2 ms + 2 ms
HO request S-DeNB → T-DeNB and processing at T-DeNB	5 ms + 2 ms
Admission control at T-DeNB	5 ms
HO-request Ack T-DeNB→ S-DeNB + RRC to m-RN	5 ms + 2 ms
SN status transfer S-DeNB → T-DeNB	5 ms
OAM reconfiguration of m-RN	12 ms
Path switch request T-DeNB → MME(RN)+processing in MME (RN)	5 ms + 2 ms
Path switch request ACK MME(RN) → T-DeNB + processing in T-DeNb	5 ms + 2 ms
Backhaul reconfiguration	20 ms
S1 setup	24 ms
Proxy relocation S-DeNB → T-DeNB	5ms
Rebuild the proxy context in T-DeNB	2ms
Path switch request (m-RN) → T-DeNB + processing in T-DeNB	5 ms + 2 ms
Path switch request T-DeNB → MME(UE)+processing in MME(UE)	5 ms + 2 ms
Update user plane MME(UE) → S-GW (UE) and DL path switch in S-GW (UE)	5 ms + 5 ms
Update user plane resp S-GW(UE) → MME (UE) and processing in MME (UE)	5 ms + 2 ms
Path switch request ack MME (UE) → T-DeNB and processing in T-DeNB	5 ms + 2 ms
Path switch response T-DeNB → m-RN and processing in m-RN	2 ms + 2 ms
m-RN context release (T-DeNB → S-DeNB)	5 ms
Total latency	**110+Nx42(ms)**

For the purpose of comparison with the results already provided in previous studies, it is important to recall the moving relay handover procedure that was detailed in [ARTD34]. The different steps of handover procedure for m-RN in alternative A 2 based architecture are shown in Figure 5-9.

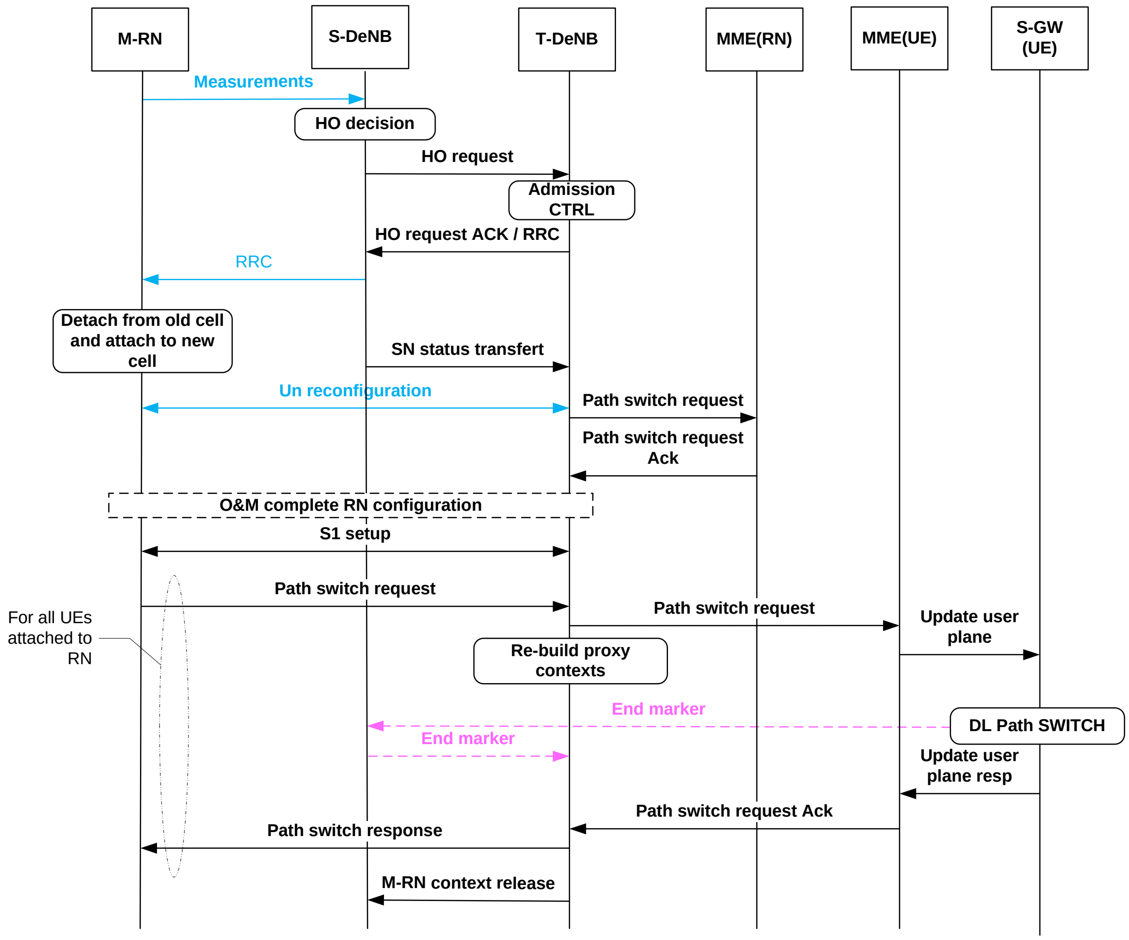


Figure 5-9: Flow chart for m-RN handover in the non optimized alternative A 2 based architecture.

As seen from the flow chart, the basic features of the non optimized architecture are:

- OAM connectivity to the m-RN is obtained after the Un reconfiguration to the target DeNB.
- m-RN initiates user plane path switch for the UEs attached to it. It can group these path switches based on the MMEs of the UEs. Target DeNB uses these path switch messages to rebuild the S1/X2 proxy contexts at the target DeNB.

The evaluation of the handover latency for the moving relay in this case relies on the same assumptions described previously. The handover latency budget is shown in the Table 5-3.

Table 5-3: Handover latency for the non optimized alternative A 2 architecture.

Message	Latency
m-RN transmit measurements to S-DeNB and processing in DeNB	2 ms + 2 ms
HO request S-DeNB → T-DeNB and processing at T-DeNB	5 ms + 2 ms
Admission control at T-DeNB	5 ms
HO-request ACK T-DeNB→ S-DeNB + RRC to m-RN	5 ms + 2 ms
SN status transfer S-DeNB → T-DeNB	5 ms
Backhaul reconfiguration	20 ms
Path switch request T-DeNB → MME(RN)+processing in MME (RN)	5 ms + 2 ms
Path switch request ACK MME(RN) → T-DeNB + processing in T-DeNb	5 ms + 2 ms
OAM reconfiguration of m-RN	12 ms
S1 setup	24 ms
Path switch request (m-RN) → T-DeNB + processing in T-DeNB	5 ms + 2 ms
Path switch request T-DeNB → MME(UE)+processing in MME(UE)	5 ms + 2 ms
Proxy context rebuilding in T-DeNB	2 ms
Update user plane MME(UE) → S-GW (UE) and DL path switch in S-GW (UE)	5 ms + 5 ms
Update user plane resp S-GW(UE) → MME (UE) and processing in MME (UE)	5 ms + 2 ms
Path switch request ack MME (UE) → T-DeNB and processing in T-DeNB	5 ms + 2 ms
Path switch response T-DeNB → m-RN and processing in m-RN	2 ms + 2 ms
m-RN context release (T-DeNB → S-DeNB)	5 ms
Total latency	**105 + N x 42 (ms)**

As seen from the table above the overall latency is composed of m-RN and UE path switch handover latencies. The m-RN handover latency corresponds to messages in plain text in the table. UE path switch handover latency corresponds to messages in italic in the table. If it is assumed *N* UEs attached to the moving relay, the overall worst case latency depends on *N* as shown in the table.

5.3.3 Summary and conclusions

In this section an optimization for the architecture and procedures has been presented and a preliminary X2 handover latency analysis for the handover of moving relays (m-RN) has been performed. The handover latency of the optimized architecture has been compared with the handover latency of the 3GPP alternative A 2 architecture described in [ARTD34]. The basic result of the study is that handover latency of the optimized architecture is similar to the handover latency of the alternative A 2 architecture. However, the optimized architecture shows the advantage of providing continuity of the connectivity between the relay OAM and the moving relay so the moving relay gets necessary configuration information of the target DeNB during the handover and radio link failure of the moving relay is minimized. The proxy relocation message sent by the source DeNB to the target DeNB will help target DeNB to trigger grouped path switch request for the moving relay node at the target DeNB, saving some backhaul signalling. Another advantage of the proposed architecture is to allow *selective admission control* of the UEs attached to the moving relay node so optimizing further the resource usage at the DeNB.

5.4 Relaying capability in other systems

5.4.1 UEs performing as moving relays

There already exist some attempts to modify the traditional role of UEs, i.e. acting as data sources or sinks for users in the macro cell, such that they can also relay user data to/from another UE in their vicinity, which has a weaker link to the base station.

In 3GPP, a solution called Opportunity Driven Multiple Access (ODMA) has been proposed to extend the network coverage via UEs acting as relays for other UEs closer to the cell edge [3GPP25924]. Furthermore, UE-to-UE (also called Device-to–Device or D2D) communication has been addressed in the WINNER+ project for an LTE network in TDD operation [WIND13]. Assuming that the eNB still controls all resource in the cell (including the D2D links), it could be shown in [WIND13] that this type of communication enables new local services and also achieves a noticeable gain in cell capacity for two different deployment scenarios.

For a more detailed overview of D2D communication, the reader is referred to the WP3 deliverable D3.4 [ARTD34], section 2.2 and the references therein.

5.4.2 Multi-hop relaying in Mobile WiMAX

Mobile WiMAX [IEEE80216] is possibly one of the closest competitors to LTE, since many parts of the air interface are similar, like the usage of Orthogonal Frequency Division Multiple Access (OFDMA) and the availability of Multiple-Input Multiple-Output (MIMO) transmission modes, Adaptive Coding and Modulation (AMC), and Hybrid Automatic Repeat Request (HARQ). The only major difference is the limitation to pure Time Division Duplex (TDD) in Mobile WiMAX compared to the choice between the latter and Frequency Division Duplex (FDD) in LTE.

While the principle organisation of the WiMAX radio protocol stack also shows a physical (PHY), medium access control (MAC), and layer-3 structure, the overall network architecture is simpler than in 3GPP, since there only exist two types of nodes[2]:

- Mobile Stations (MS).
- Base Stations (BS).

Hence, compared to 3GPP, everything apart from the air interface procedures is outside the scope of the Mobile WiMAX specification.

With the completion of the IEEE-802.16j amendment [IEEE80216j], relaying capability has been introduced in Mobile WiMAX. Compared to LTE-A, the standard amendment already defines both transparent and non-transparent Relay Stations (RS). Furthermore, the PHY extension already supports more than one RS on the path between the BS and the MS, but only if all the RSs are operated in non-transparent mode.

Based on the existing TDD structure in WiMAX, separation of the backhaul link (BS – RS) and the access link (RS – MS) is only possible in the time domain by defining different zones in the downlink (DL) and uplink (UL) part of the TDD frame. This corresponds to in-band relaying as defined for Type-1 relays in LTE-A.

Figure 5-10 depicts a possible deployment with 2 RSs on the path between the BS and the MS. As can be seen, each additional relay hierarchy level requires reservation of a separate relay zone in both DL and UL subframe. Due to the unavoidable guard intervals for switching between transmitter and receiver part at each node and the extra allocation info inside each DL relay zone (R-MAP, R-FCH), the overhead increases significantly. Hence, more than 3 hops as shown in this example may not be of practical relevance.

[2] Note: In contrast to all other parts of this document, the specific WiMAX terminology is used throughout this subsection when referring to a mobile terminal or a base station. Hence, the description is in-line with the respective IEEE specifications.

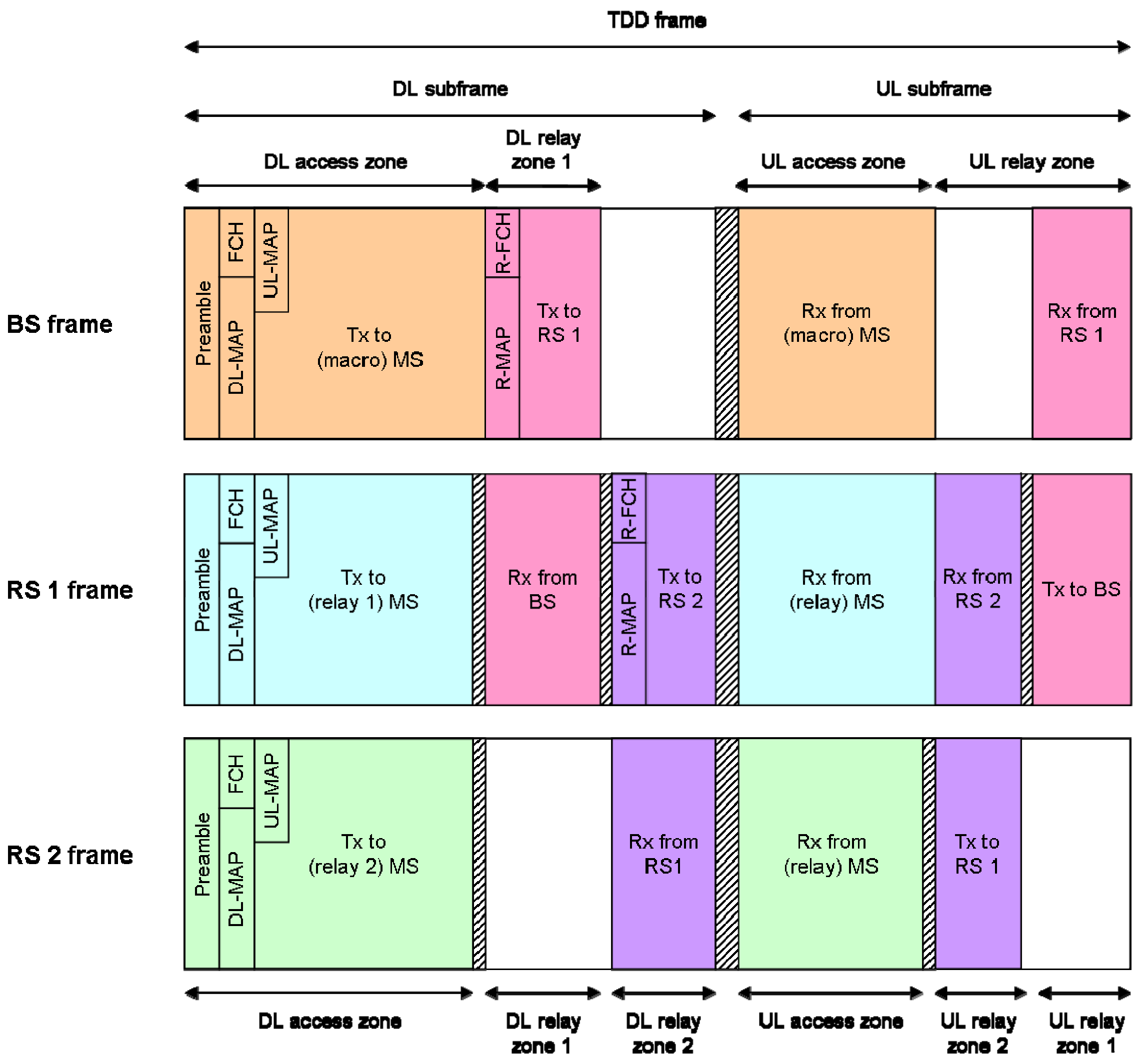


Figure 5-10: Example of a multi-hop deployment with 2 RS on the path between the BS and a MS [PH09].

For a more detailed overview of the relaying capabilities of Mobile WiMAX, the reader is referred to the WP3 deliverable D3.4 [ARTD34] and the references therein.

6 Conclusion

In this document several recommendations on the different architectural requirements placed forward by the innovations proposed in WP1-3 have been provided. Among the largest group of innovations that have an impact on the architecture are those that require some form of cooperation among (H)eNBs, so frequent information exchange among them is needed. An analysis on the current protocols and RAN architecture has been done and has resulted in the recommendation of a new cooperation interface. This interface should not use the GTP encapsulation for exchanging messages between eNBs in the cooperation area, but rather

- adapt the S1-AP messages for eNB and MME Configuration Transfer to also convey "CoMP" channel related TNL configuration information,
- use a "Type/Length/Value" (TLV) structure to identify the information passed and length of the information field when later defining the "CoMP" related messages,
- allow for ultimate flexibility regarding TNL addresses and refrain from defining a "CoMP" AP along with related messaging for the time being,
- do not refer to the "CoMP" interface as being integral part of X2 interface.

Another topic considered in this document is the need for defining suitable cooperation areas for performing CoMP strategies. These clusters can be based on either a network centric, a user centric, or a combined network-user centric approach. The last one aims at a good trade-off between complexity and achievable performance gain. Regarding CoMP and eNB clustering two different architectures can be considered, centralized and distributed Control Unit architecture.

Other innovations require cooperation either among eNB/HeNBs or HeNB-only. An architecture and control solution in order to support eNB/HeNB ICIC when the number of femto base stations (HeNB) deployed in the coverage region of a macro base station (eNB) is high has been provided. HeNBs are clustered and a central unit provides the power setting strategy to each cluster of HeNBs associated with an additional Group Physical Cell Identity (GPCI). Dealing with intra-group HeNB/HeNB ICIC there are two possible solutions:

- Direct signalling based ICI between HeNB within the group.
- Central entity based ICIC for HeNBs in the group.

In case of the latter solution, the location of the central unit is another topic addressed, i.e. whether it can be co-located with the HeNB-GW in the operator domain or requires a localized placement.

When innovations from WP1 and WP2 are mixed they have a clear impact on architecture. This mix has been named Flexible Interference Control (FIC). The different alternatives for providing time and frequency synchronization between macro and femtos have been described and the main requirements for the three types of FIC innovations have been analysed. As in the previous deliverable D4.1 [ARTD41] it has been concluded that the use of packet based solutions is needed to provide synchronization. Either NTP or, most suitable, IEE1588 v2 are the recommended solutions to provide time and frequency synchronization in these heterogeneous scenarios.

Dealing with the support of X2 interface between femto and macro base stations, it is not possible to use X2 to exchange information with Rel-10 specifications, but will most likely be solved in future specifications. Hence, different alternatives to support this X2 interface between femto and macro cells have been provided. Also a latency estimation on the delay when using cooper access, GPON, and P2P Fiber as transport technology for the X2 interface is included.

Finally, for the FIC innovations an analysis of S1/X2-AP extensions needed to support the signalling messages has been done. Different architectural alternatives have been described concluding that the last one, Alternative 5 (which bypasses the complete IP stack), is the one that guarantees for shortest possible delay and latency, but requires the definition of a new Ethertype value. It needs to be noted that Alternative 5 requires that nodes in the cooperation area are connected via Ethernet. No extensions to X2-AP messages are needed and dealing with extensions to S1-AP messages they need to be modified to allow for sufficient flexibility in setting up the CoMP interface. The definition of the TNL address needs to be extended to also allow for a bit string of 48 bits in case of IEEE 802.3 address is used for the CoMP channel.

Moreover, in this document the impact of innovations in the field of multi-hop and moving relays has been treated. Two different structures are possible for multi-hop relays:

- tree structure
- mesh structure.

An architectural analysis for each of them has been done and it has been concluded that it would be possible to introduce the tree structured multi-hop based on existing infrastructure elements already starting using Rel 10 compliant equipment. Meshing of multiple nodes could be done by using either a logical or a physical mesh. However, the latter solution is considered as practically infeasible.

Finally the mobility and latency aspects of moving relays have been discussed. An upgrade to the relay alternative A2 architecture has been proposed and an evaluation of the corresponding X2 handover latency has been performed. The optimized architecture proposed consists of:

- Fast S1/X2 proxy relocation at the target DeNB during handover execution phase so the latency of the setup of S1/X2 interfaces to the target is minimized.
- Maintaining the connectivity with the source DeNB during the handover execution phase such that the moving relay node can get the configuration information of the target DeNB from relay node OAM through OAM signalling carried to the relay node by the source DeNB.

A detailed analysis of the moving relay handover latency has been also done and compared with the handover latency of 3GPP alternative A2 architecture.

References

[3GPP25924] 3GPP TR 25.924, “Opportunity Driven Multiple Access (ODMA)”.

[3GPP29281] 3GPP TS 29.281, “General Packet Radio Service (GPRS) Tunnelling Protocol User Plane (GTPv1-U) (Release 10)”.

[3GPP33401] 3GPP TS 33.401, “3GPP System Architecture Evolution (SAE); Security architecture".

[3GPP36300] 3GPP TS 36.300, “Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10)”.

[3GPP36306] 3GPP TS 36.306, “User Equipment (UE) radio access capacities”.

[3GPP36323] 3GPP TS 36.323, “Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Packet Data Convergence Protocol (PDCP) specification (Release 10)”.

[3GPP36413] 3GPP TS 36.413, “Evolved Universal Terrestrial Radio Access (E-UTRAN); S1 Application Protocol (S1AP)”.

[3GPP36414] 3GPP TS 36.414, “Evolved Universal Terrestrial Radio Access (E-UTRAN); S1 data transport”.

[3GPP36424] 3GPP TS36.424, “Evolved Universal Terrestrial Radio Access (E-UTRAN); X2 data transport”.

[3GPP36806] 3GPP TR 36.806, “Evolved Universal Terrestrial Radio Access (E-UTRA); Relay architectures for E-UTRA (LTE-Advanced) (Release 9)”.

[3GPP36814] 3GPP TR 36.814, “Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)”.

[3GPP-R3102149] 3GPP, Samsung , “X2-GW for supporting HeNB mobility enhancement”, R3-102149.

[3GPP-R3102360] 3GPP, Nokia Siemens Networks, Qualcomm Incorporated, Ericsson, Orange, “X2 interface for HeNB mobility enhancement”, R3-102360.

[ARTD11] Artist4G, “D1.1 – Definitions and architecture requirements for supporting interference avoidance techniques”, 2010

[ARTD12] Artist4G, “D1.2 - Innovative advanced signal processing algorithms for interference avoidance”, 2010.

[ARTD13] Artist4G, “D1.3 - Innovative scheduling and cross layer design techniques for interefence avoidance”, 2011.

[ARTD34] Artist4G, “D3.4 – Relay configurations”, 2011.

[ARTD41] Artist4G, “D4.1 - Impact on RAN architecture and latency”, 2011.

[BRR+09] A. Bou Saleh, S. Redana, B. Raaf, T. Riihonen, J. Hämäläinen, R. Wichman, “Performance of Amplify-and-Forward and Decode-and-Forward Relays in LTE-Advanced,” 70th IEEE Vehicular Technology Conference Fall (VTC 2009-Fall), 2009.

[FF1] Femtoforum, “Femtocell Synchronization and Location”, 2010.

[GMY+08] V. Kenc, S. Murphy, Y. Yu, and J. Murphy, “IEEE 802.16j Relay-based Wireless Access Networks: An overview,” *IEEE Communications Magazine*, Oct. 2008.

[IEEE80216] IEEE 802.16, “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems”, 2009.

[IEEE80216j] IEEE 802.16j, “IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, Amendment 1: Multi-hop Relay Specification”, 2009.

[IEEE1588] IEEE 1588 “IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems”, 2002.

[LHa04] J. Loeser and H. Haertig, “Using Switched Ethernet for Hard Real-Time Communication,” Proceedings of the 16th Euromicro Conference on Real-Time Systems (ECRTS’04), pages 13-22.

[LHb04] J. Loeser and H. Haertig, “Low-latency hard real-time Communication over switched Ethernet,” Proceedings of the International Conference on Parallel Computing in Electrical Engineering (PARELEC 04), pages 349-353.

[MCU+08] M. V. Micea, G. N. Carstoiu, L. Ungurean, D. Chiciudean, V. Cretu, V. Groza, “Predictable Data Communication Interface for Hard Real-Time Systems,” ROSE 2008 – IEEE International Workshop on Robotic and Sensors Environments Ottawa – Canada, 17-18 October 2008, pages 98 – 101.

[PH09] S. W. Peters and R. W. Heath, Jr., “The future of WiMAX: Multi-hop relaying with IEEE 802.16j,” *IEEE Communications Magazine*, Jan. 2009.

[RFC768] IETF RFC 768, “User Datagram Protocol”, Aug 1980.

[RFC793] IETF RFC 793, “Transmission Control Protocol”, Sep 1981.

[RFC2474] IETF RFC 2474, “Differentiated Services Field”, December 1998

[RFC3758] IETF RFC 3758, “SCTP Partial Reliability Extension”, May 2004

[RFC4960] IETF RFC 4960, “Stream Control Transmission Protocol”, Sep 2007.

[RFC5062] IETF RFC 5062, "Security Attacks Found Against the Stream Control Transmission Protocol (SCTP) and Current Countermeasures", Sep 2007.

[ROS1] http://www.gpsworld.com/lbs/indoor-positioning/news/rosum-siano-introduce-alloy-chip-using-tv-signals-take-gps-indoors-9607

[WIND13] WINNER+, "D1.3 - Innovative concepts in Peer-to-Peer and Network Coding", Jan. 2009.

[WIND61313] WINNERII, “D6.13.13 v01 Deployment cost assessment of the WINNER II system concept”, 2007.

[ZSZ+08] M. Zhang, J. Shi, T. Zhang and Y. Hu, “Hard Real-Time Communication over Multi-hop Switched Ethernet,” International Conference on Networking, Architecture, and Storage, (NAS’08)” pages 121 -128