Document Properties

Document Number:	D 4.1
Document Title:	**Initial specification of Connectivity Management concepts and architecture**
Document Editor:	Pierrick Seite (FT)
Authors:	Antonio de la Oliva (UC3M) Hassan Ali-Ahmad (FT) Erick Bizouarn (ALBLF) Claudio Cicconetti (INCS) Christian Vitale (IMDEA)
Target Dissemination Level:	PU
Status of the Document:	Submitted to EC
Version:	1.0

Production Properties:

Reviewers:	M. Isabel Sanchez (UC3M) Vincenzo Mancuso (IMDEA)

Document History:

Revision	Date	Issued by	Description
1.0	2013-09-30	FT	Initial release

Disclaimer:

This document has been produced in the context of the CROWD Project. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7) under grant agreement n° 318115.
All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.
For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Abstract:

This document corresponds to the first deliverable tackling exclusively the Connectivity Management (WP4) of the CROWD project. On its pages it can be found a deep analysis of the State of the Art, followed by the first research results achieved during the first 9 months of the project. It also includes the initial architecture and procedures for the SDN-based Distributed Mobility Management Mobility solutions designed, along with initial set of interfaces enabling the communication among modules.

Keywords:

connectivity management, traffic management, mobility, distributed mobility management

Contents

List of Figures iv

List of Tables vi

List of Project Partners viii

List of Acronyms x

Executive summary 1

Key contributions 3

1 Introduction 5

2 State of the art 7

2.1 Connection manager: current practices applicable in CROWD . . . 7
2.1.1 Assistance provided by the network . . . 8
2.1.2 Connection manager architecture and APIs . . . 9
2.1.3 Terminal and network triggers . . . 14
2.2 3GPP mobility management in the EPC . . . 14
2.2.1 Intra-3GPP mobility . . . 14
2.2.2 Connection of IEEE technologies to the EPC: Trusted and un-trusted non-3GPP networks . . . 19
2.2.3 Un-trusted non-3GPP networks . . . 21
2.2.4 Trusted non-3GPP networks . . . 22
2.3 IP mobility management . . . 24
2.3.1 IPv6 mobility management . . . 24
2.3.2 DMM rationale . . . 24
2.3.3 Mobile IP . . . 25
2.3.4 Proxy Mobile IP . . . 25
2.3.5 Distributed Mobility Management . . . 27
2.3.6 SDN-based Mobility . . . 29
2.4 Access selection and vertical HO . . . 30
2.4.1 Overview over Access Selection/Vertical Handover mechanisms . . . 31
2.4.2 Scope and Optimisation goals addressed through Access selection/VHO . . . 32
2.5 IEEE 802.21 . . . 35
2.5.1 802.21 Objectives . . . 35
2.5.2 IEEE 802.21 Architecture . . . 36
2.5.3 IEEE 802.21 amendments . . . 38
2.6 Information gathering in dense networks . . . 38
2.6.1 ALTO protocol . . . 38
2.6.2 802.11u . . . 41
2.6.3 Protocol overlap and information lifetime . . . 42

3 Research activities **43**
3.1 Decision management . . . 43
3.1.1 Analysis of the Connection Manager of different smartphone architectures . . 43
3.1.2 Information gathering and its use for connectivity management . . . 48
3.1.3 ANDSF . . . 50
3.1.4 Energy efficiency optimisation through access selection/vertical HO in presence of dense/heterogeneous scenarios and relay nodes . . . 51
3.1.5 Assessment of the Energy consumption savings with 3G offload . . . 55
3.2 Mobility management . . . 59
3.2.1 DMM landscape . . . 62
3.2.2 Optimised mobility management in the IEEE 802.11 RAN . . . 65
3.2.3 Optimised mobility management in LTE . . . 66

4 Functional architecture **69**
4.1 Functional requirements . . . 70
4.2 CROWD Mobility Management proposed solution . . . 71
4.2.1 Intra-district Mobility . . . 72
4.2.2 Inter-district Mobility . . . 73
4.2.3 Host-based Mobility . . . 75
4.3 Interface definition . . . 77

5 Conclusion **79**

Bibliography **81**

List of Figures

2.1 Enhanced access selection framework. 8
2.2 Network information collection: crowd sourcing. 9
2.3 Network information collection: network based monitoring. 9
2.4 Multiple Interfaces terminal architecture. 10
2.5 OS API communication model. 11
2.6 High-level framework for OpenCMAPI. 14
2.7 S1-based handover 16
2.8 X2-based handover 18
2.9 Interworking architecture for 3GPP and non-3GPP accesses. 20
2.10 Untrusted non-3GPP procedures and user data protocol stack. 21
2.11 Trusted WLAN procedures and user data protocol stack. 22
2.12 Proxy Mobile IPv6 (PMIPv6) scenario. 26
2.13 Heterogeneous Wireless Networks Environment (from [1]). 31
2.14 MCDM framework for access selection (from [2]). 34
2.15 802.21 General Architecture. 37
2.16 ALTO network abstract view. 39
2.17 ALTO routing cost example. 40
2.18 ALTO services. 41

3.1 Initial attachment to the WLAN. 45
3.2 IFOM, MAPCON and NSWO. 52
3.3 ANDSF network support. 52
3.4 Application scenario of the proposed access selection mechanism. 53
3.5 Battery drainage for the different WiFi states. 58
3.6 Battery drainage for the different 3G states. 58
3.7 Mobility management in MIP-based DMM. 63
3.8 Mobility management in PMIP-based DMM. 64
3.9 LTE system architecture (edge only). 67
3.10 Proposed CROWD system architecture. 67
3.11 Procedure for intra-district LTE handover. 68

4.1 WP4 module view and functional internal-interfaces. 69
4.2 Intra-district Mobility: Initial attachment. 72
4.3 Intra-district Mobility: Handover. 74
4.4 Inter-district Mobility: Initial attachment. 75
4.5 Inter-district Mobility: Handover. 76
4.6 Host-based Mobility CROWD solution 76

List of Tables

2.1 Handover triggers 15

3.1 Main characteristics of the analysed smartphones 44
3.2 DNS queried for initial configuration of services on WLAN interface start-up 46
3.3 Information types 48
3.4 Power consumption results 57

List of Project Partners

Name	Acronym	Country
Intecs S.p.A. (*coordinator*)	INCS	Italy
Alcatel-Lucent Bell Labs France	ALBLF	France
France Telecom SA	FT	France
Fundacion IMDEA Networks	IMDEA	Spain
Signalion Gmbh	SIG	Germany
Universidad Carlos III de Madrid	UC3M	Spain
Universitaet Paderborn	UPB	Germany

List of Acronyms

3GPP2 Third Generation Partnership Project 2

3GPP Third Generation Partnership Project

ADN Application Delivery Network

ALTO Application Layer Traffic Optimisation

AN Access Network

ANDSF Access Network Discovery and Selection Function

ANQP Access Network Query Protocol

AP Access Point

API Application Programming Interface

APN Access Point Name

AR Access Router

BA Binding Acknowledgement

BC Binding Cache

BBF Broadband Forum

BER Bit Error Rate

BNG Broadband Network Gateway

BSSID Basic Service Set ID

BU Binding Update

CDN Content Distribution Network

CIR Carrier to Interference Ratio

CLC CROWD Local Controller

CN Core Netweork

CoA Care-of Address

CRC CROWD Regional Controller

CROWD Connectivity management for eneRgy Optimised Wireless Dense networks

CSG Closed Subscriber Group

D2D Device-to-Device

DB Data base

DHT Distributed Hash Table

DM Device Management

DMM Distributed Mobility Management

DMM-GW Distributed Mobility Management Gateway

DNS Domain Name System

DSMIP Dual Stack Mobile IP

DVB Digital video Broadcasting

DVB-T Digital video Broadcasting–Terrestrial

EAP Extensible Authentication Protocol

EAP-AKA Extensible Authentication Protocol - Authentication and Key Agreement

EAP-SIM Extensible Authentication Protocol - Subscriber Identity Module

ECGI Eutran Cell Global ID

eNB Evolved NodeB

EPC Evolved Packet Core

ePDG Evolved Packet Data Gateway

EPS Evolved Packet System

E-RAB E-UTRAN Radio Access Bearer

ESSID Extended Service Set ID

FA Foreign Agent

FMC Follow-Me Cloud

ForCES Forwarding and Control Element Separation

GAS Generic Advertisement Service

GNSS Global Navigation Satellite System

GO Group Owner

GPS Global Positioning System

GRA Grey Relational Analysis

GTP GPRS Tunnel Protocol

HA Home Agent

HAR Home Access Router

HESSID Homogeneous Extended Service Set ID

HFN Hyper Frame Number

HO Handover

HoA Home Address

HSPA High Speed Packet Access

HSS Home Subscriber Server

HTTP Hypertext Transfer Protocol

I2RS Interface to the Routing System

IEEE Institute of Electrical and Electronics Engineers

IETF Internet Engineering Task Force

IFOM IP Flow Mobility

IMEI International Mobile Station Equipment Identity

IMSI International mobile Subscriber Identity

IP Internet Protocol

ISMP Inter-System Mobility Policy

ISP Internet Service Provider

ISR Idle state Signalling Reduction

ISRP Inter-System Routing Policy

JSON JavaScript Object Notation

KPI Key Performance Indicator

L2 Layer 2

LMA Localised Mobility Anchor

LMD Localised Mobility Domain

LTE Long Term Evolution

LIMONET LIPA Mobility and SIPTO at the Local Network

LIPTO Local IP Access

LLC Logical Link Control

MAC Medium Access Control

MAG Mobility Anchor Gateway

MAPCON Multiple-access PDN connectivity

MAR Mobile Anchor Router

MCDM Multi-Criteria Decision-Making

MCS Modulation and Coding Scheme

MEVICO Mobile Networks Evolution for Individual Communications Experience

MFC MobileFlow Controller

MFFE MobileFlow Forwarding Engine

MIF Media Independent Function

MIH Media Independent Handover

MIHF Media Independent Handover Function

MIIS Media Independent Information Service

MIPv6 Mobile Internet Protocol version 6

MME Mobility Management Entity

MN Mobile Node

MO Management Object

MPLS Multiprotocol Label Switching

MWE Multiplicative Weighting Exponent

NB North-Bound

NCSI Network Connectivity Indicator

NFV Network Function Virtualization

NSWO Non-Seamless WLAN Offload

OF OpenFlow

OMA Open Mobile Alliance

OMA-DM Open Mobile Alliance - Device Management

ONF Open Networking Foundation

OpenCMAPI Open Connection Manager Application Programming Interface

OS Operating System

P-GW PDN Gateway

P2P Peer-to-Peer

PBA Proxy Binding Acknowledgement

PBU Proxy Binding Update

PCRF Policy Charging and Rules Function

PDCP Packet Data Convergence Protocol

PDN-GW Packet Data Network GateWay

PDN Packet Data Network

PDP Packet Data Protocol

PIN Personal Identification Number

PLMN Public Land Mobile Network

PMIPv6 Proxy Mobile IPv6

PoA Point of Attachment

POSIX Portable Operating System Interface

PoS Point of Service

PPP Point-to-Point Protocol

PUK PIN Unlocked Key

QoS Quality of Service

RA Router Advertisement

RAN Radio Access Network

RAT Radio Access Technology

REST Representational State Transfer

RF Radio Frequency

RFC Request for Comments

RNC Radio Network Controller

RS Reference Symbols

RSRP Reference Symbols Received Power

RSRQ Reference Symbols Received Quality

RSS Received Signal Strength

S-GW Serving Gateway

SAP Service Access Point

SAW Simple Additive Weighting

SDN Software Defined Network

S-GW Serving GateWay

SIM Subscriber Identity Module

SINR Signal to Interference plus Noise Ratio

SIPTO Selected IP Traffic Offload

SIR Signal to Interference Ratio

SLAC Stateless auto-configuration

SMS Short Message Service

SNR Signal to Noise Ratio

SN Sequence Number

SSID Service Set ID

SSP Subscription Service Provider

TAI Tracking Area Identity

TA Technology-Agnostic

TA Tracking Area

TEID Tunnel Endpoint Identifier

TOPSIS Technique for Order Preference by Similarity to Ideal Solution

TWAG Trusted WLAN Access Gateway

TWAN Trusted WLAN Access Network

TWAP Trusted WLAN AAA Point

UE User Equipment

UICC Universal Integrated Circuit Card

UI User Interface

UMTS Universal Mobile Telecommunication System

USSD Unstructured Supplementary Service Data

VHO Vertical HandOver

VLAN Virtual LAN

VoIP Voice over IP

VPN Virtual Private Network

WG Working Group

WiFi Wireless Fidelity

WiMAX Worldwide Interoperability for Microwave Access

WLAN Wireless Local Area Network

WP Work Package

Executive summary

The purpose of this document is to present the first WP4 achievements regarding connectivity management in CROWD environment. The document presents the different operations and current solutions proposed for the connectivity management in the context of extremely dense networks. Then, the document presents first thoughts regarding the application of these concepts to the CROWD context. Basically, the document proposes an architecture for connection manager and mobility management framework based on DMM and SDN directions given by deliverable D1.1. DMM and SDN mobility are jointly used to manage respectively the inter-district and the intra-district mobility. This document presents the first results and does not pretend to address all connectivity management issues from CROWD. Next WP4 deliverables (D4.2 and D4.3) will cover missing requirements and refine the concepts presented in this document.

Key contributions

The main technical contributions of this deliverable can be summarised as follows:

- State of the Art regarding connectivity management, including access selection, IP mobility management, triggers and assistance mechanisms provided by the network (e.g. ANDSF).
- First analysis of the behaviour of different connection managers on the three major mobile Operating Systems (Android, iOS, windows).
- Connectivity management architecture, clarifying the role of the different modules and its logical placement.
- Definition of the intra-domain and inter-domain mobility solution.
- Connection Manager: specification of the architectures, functionalities and identification of relevant APIs.
- Definition of mechanisms to provide feedback information between ALTO servers and IEEE 802.21MIIS.
- Application of the DMM concepts to CROWD environment, including motivations for DMM and design of different DMM approaches based on MIPv6 and PMIPv6.
- Global mobility support (between different operational domains) beside local mobility support.
- Application of SDN-based approach defined in CROWD to intra-domain and inter-domain mobility management in LTE.
- Preliminary definition of optimised handover procedures for LTE in CROWD.

The key contributions to the research (contribution type P) and standardisation (contribution type S) communities resulting from the work performed within the aforementioned deliverables are shown in the table below.

Summary of Research (P) and Standardisation (S) contributions

Contributions	Type	Outline	Organisation
Energy consumption savings with 3G offload	P	This paper discusses energy consumption savings brought by the offloading of 3G traffic to WLAN	VTC2013
CROWD mobility architecture	P	Presents the CROWD architecture as defined in this document	SDN4FNS
Single Radio handover	S	UC3M form part of the Ballot resolution committee	IEEE 802.21c
Technical edition of the draft standard with respect to group management	S	Group management aims to optimise the handover of multiple terminals at the same time	IEEE 802.21d
Requirements for Distributed Mobility Management	S	This document defines the requirements for Distributed Mobility Management (DMM) in IPv6 deployments	IETF/DMM WG
Distributed Mobility Management: Current practices and gap analysis	S	The document discusses possible deployment of existing mobility protocols in a distributed mobility management environment	IETF/DMM WG
Mobility Anchor Selection in DMM: Use-case Scenarios	S	This document presents and discusses different use-case scenarios of mobility anchor selection in Distributed Mobility Management (DMM)	IETF/DMM WG

1 Introduction

This deliverable focuses on connectivity management in the CROWD environment (i.e. dense access network context). Connectivity management covers different operations such as access selection, initial attachment, access monitoring, access re-selection or handover management among others.

These are regular connectivity management operations already supported in many mobile systems. However, current mechanism and protocols have been designed for centralised systems and may not meet CROWD requirements. This is what the WP4 will have to show and address.

Connection management includes the access selection problem, which is usually tricky since it may depend on plethora of criteria, sometimes inconsistent. The access selection usually depends, on the terminal capabilities, the user preferences, the Quality of Service (QoS) of the access network, the quality of the radio link, the application preferences or the level of battery of the terminal. The problem is that these criteria are not always consistent, for example, the best access for an application may not correspond to the user preferences. The access selection thus requires sophisticated mechanisms where different players come in (users, connection manager, network operator, network monitoring entities, etc.). Once the selection is made, the connectivity management system of the terminal may be required to move some IP flows from one access to another, it includes reattachment operation (e.g. Layer 2 (L2) connection and authentication to the new access point) and, sometimes, the preservation of the IP address. Chapter 2 of this document explores current solutions and practices for connectivity management. Typically, the document describes current behaviour of connection manager, and solutions for mobility management (i.e. the way to handle the handover) in IP networks as well as in cellular networks. The solution space identified in Chapter 2 is then applied to the specifics of the CROWD scenario in Chapter 3. Finally, before going into the conclusions and proposal for next steps, Chapter 4 defines the mobility architecture based on Distributed Mobility Management (DMM) and Software Defined Network (SDN) directions given by deliverable D1.1. DMM and SDN mobility are jointly used to manage respectively the inter-district and the intra-district mobility. Ideas presented here are initial proposals and will be refined in next WP4 deliverables.

2 State of the art

This section is devoted to provide a deep analysis of the state of the art on technologies used for Connectivity Management in CROWD. Herein we provide an overview of the different technologies that can be used for providing to the terminal information regarding surrounding points of attachment in order to perform an educated network selection. We also provide an overview of the current trends on Connection Manager architectures currently deployed in terminals and the mobility protocols that can be used to support IP mobility management. Finally we also provide an overview of current IP mobility management approaches followed by the 3GPP.

2.1 Connection manager: current practices applicable in CROWD

It is a common practice for mobile handsets to use a connection manager that performs network interface selection based on local information (e.g. available accesses, radio signal strength, ongoing applications, battery level, etc.) and user preferences (e.g. preferred networks). The connection manager is a dedicated application providing interaction with the user and the operator and implementing interfaces with functional modules of the operating system (i.e., access monitoring, routing management, authentication function, etc.).

A connection manager is organised over three main functions that should come into play concerning connection management:

1. *Initiation function*: this function monitors events which possibly require connection management operations. Events may be related to access link characteristics, application hints, user triggers, etc. When detecting an event which may require multiple interfaces management operations (e.g. selection of a new provisioning domain), the initiation function triggers the decision function.

2. *Decision function*: upon triggers and information fired by the initiation function, the decision function makes a decision about multiple interfaces management (e.g. select the best appropriate provisioning domain to be used, source address selection, etc.). The decision is also made using local information (e.g. policies, user preferences) fetched from a local or remote repository.

3. *Execution function*: once the decision has been made, the decision function triggers the execution function if required, e.g. to perform the attachment to the targeted interface configuration, control of associated mechanisms for communication continuity if needed (e.g. configuration of a virtual interface, mobile IP procedures, and so on,....).

Connection manager behaviour can be enriched by network guidance provided to the terminal, as illustrated in Fig. 2.1. For example, the network can recommend the access network taking into account the available quality of communications. There are currently two ways to provide such guidance:

1. Relying on a **policy server** (e.g. Third Generation Partnership Project (3GPP) Access Network Discovery and Selection Function (ANDSF)): provides access selection rules and

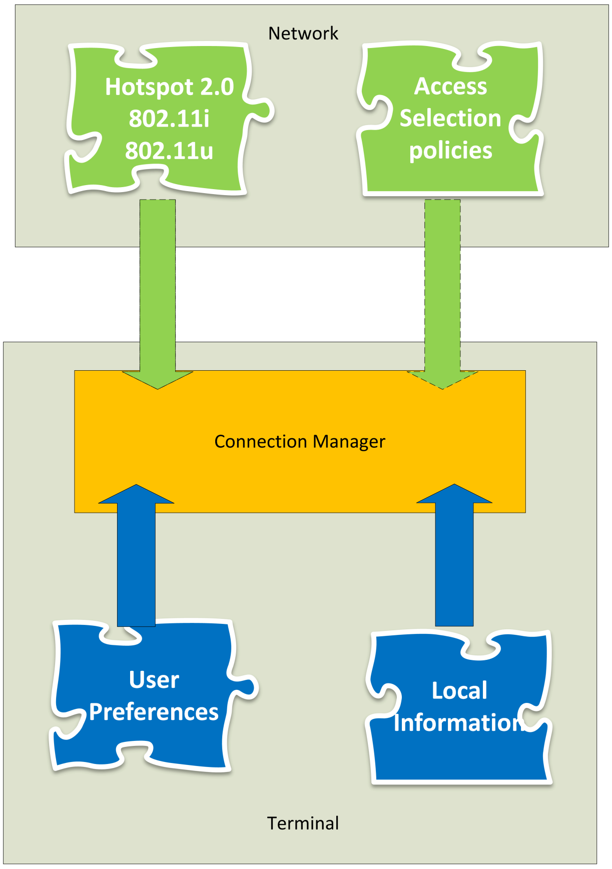


Figure 2.1: Enhanced access selection framework.

discovery information[1] to the terminal.

2. Expose additional information on Wireless Fidelity (WiFi) networks thanks to **Hotspot 2.0** amendment: enhanced WiFi discovery services with advertisement of network characteristics (e.g. security method), prior to the association. Hotspot 2.0 also delivers to the terminal information about the network.

2.1.1 Assistance provided by the network

Although not specified by the 3GPP ANDSF, one of the main interest of the policy server based solution is to take into account the access network conditions in the policy sent to the terminal. Currently, two methods can be used to reflect the network condition into the selection policies:

- *Terminal based monitoring, a.k.a crowd sourcing:* (see Fig. 2.2): a dedicated server establishes a picture of the networks quality using QoS report sent by all the terminals. This picture is then used by the policy server.

- *Network based monitoring* (see Fig. 2.3): Terminals are not requested to send QoS measurements; the policy server learns the network condition (i.e. load, QoS) directly from appropriate network entities (e.g. 3GPP Radio Network Controller (RNC), WiFi controller, Broadband Forum (BBF) Broadband Network Gateway (BNG), etc.).

[1] Discovery information: list of networks that may be available in the vicinity of the User Equipment (UE) and information assisting the UE to expedite the connection to these networks.

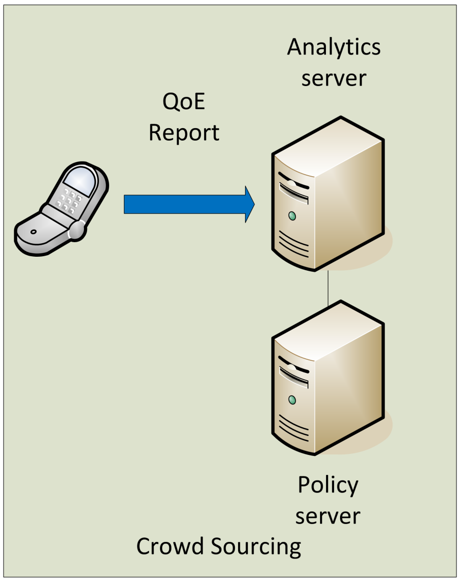


Figure 2.2: Network information collection: crowd sourcing.

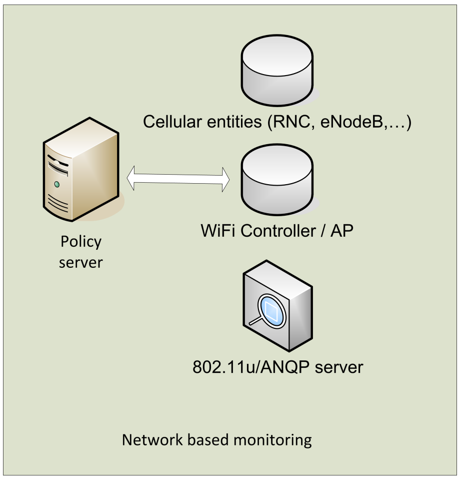


Figure 2.3: Network information collection: network based monitoring.

The crowd sourcing solution is easily deployable and has no direct impact on the network. However, it induces additional signalling load and accuracy of QoS map is questionable. The network based solution has minimal impact on terminal and is supposed to give more accurate information, taking benefit of the knowledge an operator has from its network. However, interfaces from policy server to network entities and available Key Performance Indicator (KPI) are not standardised and are still to be defined.

2.1.2 Connection manager architecture and APIs

The connection manager is a dedicated application providing interaction with the user and the operator and implementing interfaces with functional modules of the operating system (i.e., access monitoring, routing management, authentication function, etc.). This section provides updated information on current connection manager architectures being discussed on the Multiple Interfaces (MIF) Internet Engineering Task Force (IETF) Working Group (WG), in which CROWD members participate.

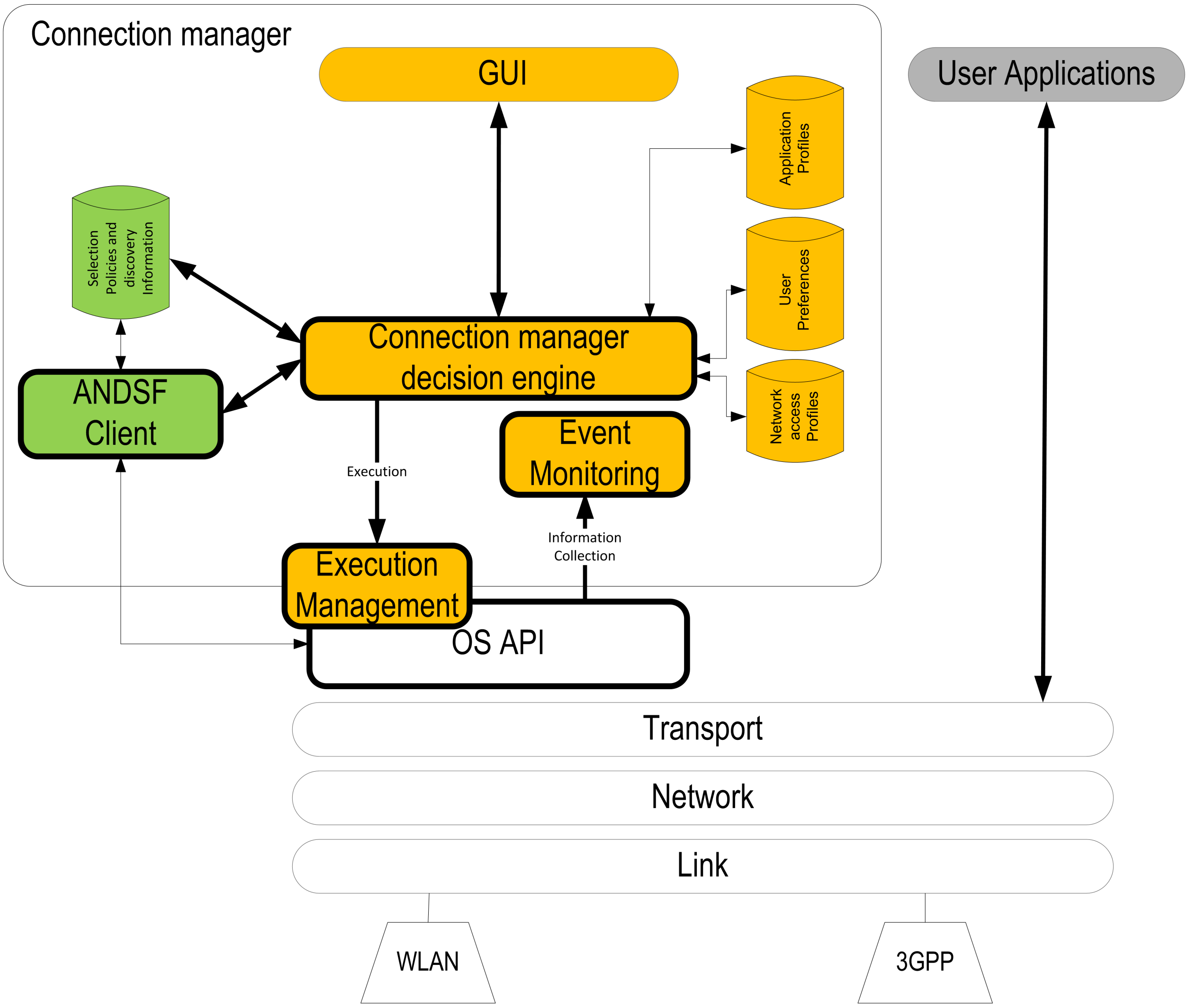


Figure 2.4: Multiple Interfaces terminal architecture.

The suggested connection manager architecture is given on Fig. 2.4. The connection manager is a module interacting with the different layers of the terminal. The connection manager implements a set of dedicated managers (authentication manager, Domain Name System (DNS) manager, etc.) in charge of making decisions with respect to specific issues. Each manager implements an initiation function which is expected to compute information issued from lower layers (e.g. system drivers) or provided by the operating system (e.g. level of battery). After making decision, the connection manager applies it via specific interfaces allowing the manipulation of configuration objects (e.g. routing table). The connection manager architecture, presented in this document, may be based on the standardised Application Programming Interface (API) concept. For instance, Fig. 2.4 shows the connection manager interfaced with the IETF MIF API, expected to be issued from standardisation effort.

The connection manager can be handled at the application layer; however, performance issues (e.g. scan duration, application transfer duration, etc) may raise up. The other option is the native selection logics in the device developed as part of e.g. 3GPP ANDSF or HotSpot2.0 specifications which would be included in the Operating System (OS) stack or even reported in the chipset with tight integration with other necessary modules (e.g. baseband). This may enhance selection performances in comparison to the solution based on APIs.

An API may expose objects useful to deal with multiple interfaces to user applications, or connection managers. APIs are typically the interfaces that the operator wishes to see standardised (in order to retrieve consistency in connection management behaviour). However, it is reminded that the APIs are not supposed to support sophisticated selection mechanisms, which must remain under the control of the operator.

2.1.2.1 Network Link API

The network link API is responsible for managing whatever network links are present on a node, whether these are physical links or tunnels. What precisely this functional box contains may vary greatly from device to device, but usually this API allows to activate/de-activate an interface, monitor interfaces, request scanning, get information on the interface status (e.g. signal strength), etc. However, many devices, especially handsets, do not provide a direct access to the network API but the Operating System can include similar services in the OS API (e.g., Android). In the following, we will assume an access to the interface services via the OS API.

2.1.2.2 OS API

The OS API provides a set of mechanisms allowing the applications to interact with the operating system. For instance, it allows applications to retrieve location information, the battery status or manage the connectivity.

The OS API usually follows a communication model as the one depicted in Fig. 2.5. Applications can subscribe to notifications (see list below), get on-demand information (e.g. status of an interface) or send commands (e.g. request specific configuration operation). In the following, we will assume that the OS API follows that communication model.

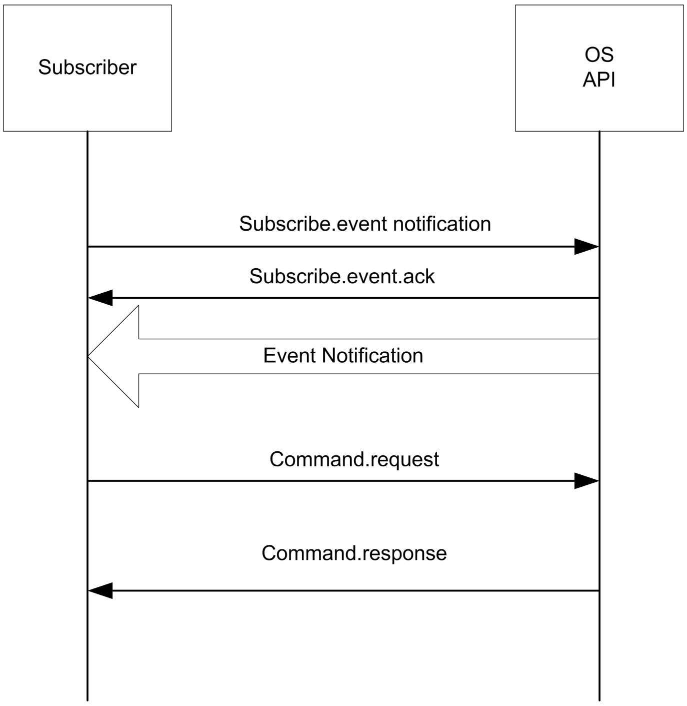


Figure 2.5: OS API communication model.

2.1.2.3 Communication API

When an application wants to communicate with a remote node and similarly, when an application receives a connection from a remote node, it must communicate with that remote node. The communications API is used for this communication. Popular examples of such APIs include the Portable Operating System Interface (POSIX) socket API.

2.1.2.4 MIF API

The IETF/Media Independent Function (MIF) working is specifying a dedicated API for multiple interfaces terminals. A document [3], led by China Mobile, focusing on an abstracted and generic API concepts, will be issued by mid-2014.

The MIF API is intended to describe the minimal complete set of API calls required to implement higher level APIs, or applications (like connection manager), that solve connection management issues on multiple interfaces terminals. The API specified in [3] is an abstract API, meaning that only functionalities are specified, but it is not provided specific bindings for any programming language.

The MIF API provides notification and command services. Basically, it allows API users to be notified about the interface availability, get information on configuration and open TCP connection with specific source and destination IP addresses. The communication model is similar to the one in the OS API (Fig. 2.4); the MIF API provides both push and pull services to higher layer functions. In push mode, an application using the MIF API has first to subscribe to announcements it want to receive. Then notifications are sent by the MIF API when subscribed event occurs. In pull mode, the application sends on demand requests to get information from the MIF API.

2.1.2.5 OMA API

This API is not represented in the terminal architecture but could be used for management of WiFi and 3GPP interfaces. The Open Mobile Alliance (OMA) has specified the requirements for an Open Connection Manager Application Programming Interface (OpenCMAPI). The requirement focuses on 3GPP and Wireless Local Area Network (WLAN) interfaces but put aside virtual interfaces, like Virtual Private Network (VPN). Multihoming is also not in the scope of the OpenCM API, meaning that the 3GPP and the WiFi interface cannot be used at the same time. The use-case where different flows are mapped simultaneously to different interfaces is thus not supported. In this context, the aim of the OMA OpenCMAPI is to address requirements for all connectivity and relevant connection management aspects such as:

- Connection/Disconnection.
- Interface selection.
- All relevant elements related to the connection or the device and more specifically all elements necessary and useful to be provided to any User Interface (UI) and user experience or to any application needing information status on the connection.
- Additional services such as Short Message Service (SMS), Unstructured Supplementary Service Data (USSD), Global Positioning System (GPS), etc., when associated to connection management and relevant for the device considered.
- Data services, especially push service configuration on the device side.

This API shall have the following features:

- It is OS-independent.
- It supports Multi-Instance (several applications/services can use it in parallel if necessary).
- It fits different types of devices requiring access to mobile Internet such as mobile broadband devices, wireless routers, M2M, smartphones, tablets, cloud devices, etc.
- It is UI-independent.

The high-level framework is depicted on Fig. 2.6. The OpenCMAPI is composed of the following functions and services:

1. Device discovery: search for new devices, open/close devices, get information on devices,
2. Cellular network management: get information about Radio Frequency (RF), home network (radio system (3GPP/Third Generation Partnership Project 2 (3GPP2)), Public Land Mobile Network (PLMN) name, ID, and icon) and serving network of the subscriber.
3. Connection management: add/delete/update a profile information, get information of a profile, select current PLMN, connect/disconnect to/from a cellular network, etc.
4. Network management: get connection status, set up automatic connection mode, set default cellular profile, restrict bearer, etc.
5. Device service: handling of the different information related to the device. (e.g. manufacturer information, device name, hardware version, International mobile Subscriber Identity (IMSI), International Mobile Station Equipment Identity (IMEI)).
6. Personal Identification Number (PIN)/PIN Unlocked Key (PUK) management: verify, change, unblock, enable/disable PIN.
7. Universal Integrated Circuit Card (UICC) management: get the last terminal profile sent by the device to the smart card, indicate to the smart card the terminal supported by the connection manager.
8. WiFi: function used to connect/disconnect to WiFi network, scan, set connection parameters, add/remove known WiFi network, etc.
9. Statistics: get connection statistics information (e.g. average signal strength, etc.).
10. Information status: get information on the current connection (e.g. IP address, PIN status, Radio Access Technology (RAT) type, radio state, Access Point Name (APN) information, etc.).
11. SMS management: e.g. send, get, delete, update the status, of the SMS, etc.
12. USSD management: build an USSD request and release the USSD session.
13. Global Navigation Satellite System (GNSS) API: used to manage the GPS (get position, set tracking mode, set/get GPS config, etc.).
14. Data push service management: enable/disable data push service.

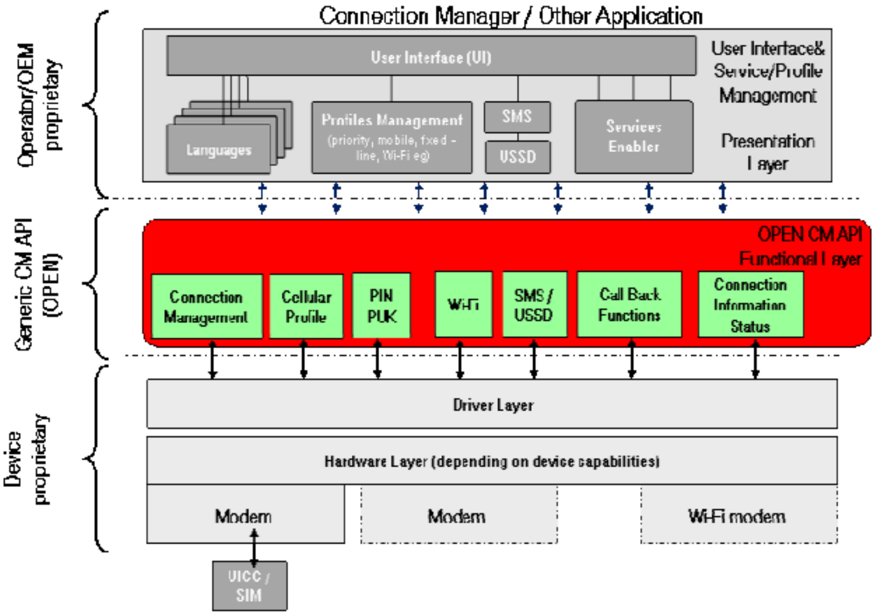


Figure 2.6: High-level framework for OpenCMAPI.

2.1.3 Terminal and network triggers

There are various events which can require the connection manager to make a decision regarding the selection of a new access. For example, the QoS of the current access degrades and the link is going down. Table 2.1 reports several events and decision inputs that have been identified from Connectivity management for eneRgy Optimised Wireless Dense networks (CROWD) use-cases [4]. These parameters come into play in the decision process either as a decision event, as a decision input, or both.

The list of parameters is theoretically established from CROWD use-cases, however some parameters may not be provided by current handsets OS (e.g., Android) and would require specific CROWD enhancement. The column *Origin* indicates if the parameters can be provided by the OS or not.

2.2 3GPP mobility management in the EPC

This section is devoted to the analysis of the different mechanisms used by 3GPP specifications for managing IP mobility within the 3GPP core.

2.2.1 Intra-3GPP mobility

The term Handover (HO) refers to the process of transferring an ongoing call or data session from one channel connected to the core network to another. There are two types of HO procedure in Long Term Evolution (LTE) for UE in active mode:[2] S1-HO procedure and X2-HO procedure. For Intra-LTE mobility, the eNBs use S1 interface only if X2 interface is not available. The 3GPP

[2] When an UE switches to *active mode* the Evolved Packet System (EPS) Bearers (either dedicated or default), S1 connection between the Evolved NodeB (eNB) and the Serving GateWay (S-GW) is active and the Radio Bearers over the air are activated. In *Idle mode* S1 connection between the eNB and the S-GW and the Radio Bearers over the air are torn down.

Table 2.1: Handover triggers

Parameter	Origin	Event	Decision input
Terminals location (i.e. 3GPP cell-ID according to assumption)	OS API	x	x
Users velocity	Dedicated Function	x	x
Battery status	OS API	x	x
Time of the day	OS API	x	x
Interface WiFi on/off	OS API	x	x
Application Start	OS API	x	
Application Stop	OS API	x	
Application requests a specific interface	OS API	x	
Application requirement (e.g. QoS required)	Dedicated Function		x
User add/remove a WiFi profile in the access network profile	Dedicated Function	x	
User Profile modification	Dedicated Function	x	x (user profile taken into account into the decision)
User preferences modification	Dedicated Function	x	x (user preferences taken into account into the decision)
L1/L2 QoS of access link	Few information provided OS API and drivers	x	x
L3 QoS of of access link	Dedicated function	x	x
IP connectivity Check	Dedicated Function		x
Network trigger	Dedicated function	x	

reference standard is [5]. Furthermore, a recent survey on 3GPP data traffic offloading can be found in [6].

The procedure starts with the measurement reporting of a handover event by the UE to the serving eNB. The UE periodically performs downlink radio channel measurements based on the Reference Symbols (RS); namely, the UE can measure the Reference Symbols Received Power (RSRP) and the Reference Symbols Received Quality (RSRQ) . If certain network configured conditions are satisfied, the UE sends the corresponding measurement report indicating the triggered event. In addition, the measurement report indicates the cell to which the UE has to be handed over, which is termed target cell.

We start with analysing the S1-HO procedure.

2.2.1.1 S1-HO procedure

In Fig. 2.7 it is shown the S1-based handover. In the following the procedure's steps are explained:

Step 1: The source eNB decides to initiate a S1-based handover to the target eNB. This can be triggered e.g., by the absence of X2 connectivity to the target eNB, or by an error indication from the target eNB after an unsuccessful X2-based handover, or by dynamic information learnt by the source eNB.

Step 2: The source eNB sends Handover Required (Direct Forwarding Path Availability, Source to Target transparent container, target eNB Identity, Closed Subscriber Group (CSG) ID, CSG access mode, target Tracking Area Identity (TAI) , S1AP Cause) to the source Mobility Management

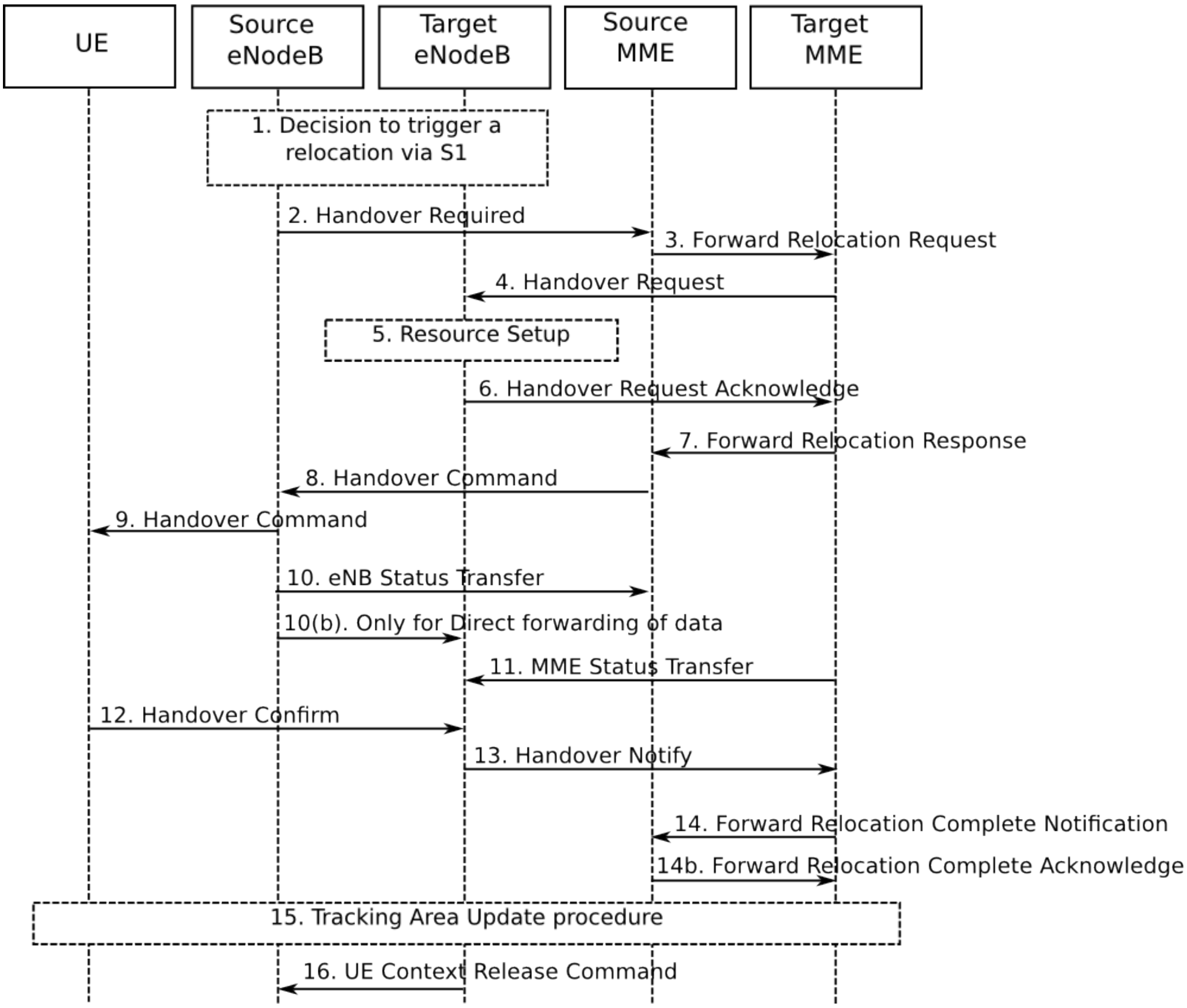

UE
Source eNodeB
Target eNodeB
Source MME
Target MME
1. Decision to trigger a relocation via S1
2. Handover Required
3. Forward Relocation Request
4. Handover Request
5. Resource Setup
6. Handover Request Acknowledge
7. Forward Relocation Response
8. Handover Command
9. Handover Command
10. eNB Status Transfer
10(b). Only for Direct forwarding of data
11. MME Status Transfer
12. Handover Confirm
13. Handover Notify
14. Forward Relocation Complete Notification
14b. Forward Relocation Complete Acknowledge
15. Tracking Area Update procedure
16. UE Context Release Command

Figure 2.7: S1-based handover

Entity (MME).

The source eNB indicates which bearers are subject to data forwarding. Direct Forwarding Path Availability indicates whether direct forwarding is available from the source eNB to the target eNB. This indication from source eNB can be based on e.g. the presence of X2. The target TAI is sent to MME to facilitate the selection of a suitable target MME.

Step 3: The source MME selects the target MME and if it has determined to relocate the MME, it sends a Forward Relocation Request message to the target MME. The target TAI is sent to the target MME to help it to determine whether S-GW relocation is needed (and, if needed, aid S-GW selection). If the MME has been relocated, the target MME verifies whether the source S-GW can continue to serve the UE. If not, it selects a new S-GW. If the MME has not been relocated, the source MME decides on this S-GW re-selection.

Step 4: The Target MME sends the Handover Request message to the target eNB.

Step 5: The reception of the previous message creates the UE context in the target eNB, including information about the bearers, and the security context. Handover Restriction List is sent if available in the Target MME. The Target MME shall include the CSG ID and CSG Membership Indication when provided by the source MME in the Forward Relocation Request message.

Step 6:The target eNB sends a Handover Request Acknowledge message to the target MME.

The EPS Bearer Setup list includes a list of addresses and Tunnel Endpoint Identifier (TEID)s allocated at the target eNB for downlink traffic on S1-U reference point (one TEID per bearer) and addresses and TEIDss for receiving forwarded data if necessary.

Step 7: If the MME has been relocated, the target MME sends a Forward Relocation Response message to the source MME.

Step 8: The source MME sends a Handover Command (Target to Source transparent container, Bearers subject to forwarding, Bearers to Release) message to the source eNB. The Bearers subject to forwarding includes list of addresses and TEIDs allocated for forwarding. The Bearers to Release includes the list of bearers to be released.

Step 9: The Handover Command is constructed using the Target to Source transparent container and is sent to the UE. Upon reception of this message the UE will remove any EPS bearers for which it did not receive the corresponding EPS radio bearers in the target cell.

Step 10: The source eNB sends the eNB Status Transfer message to the target eNB via the MME(s) to convey the Packet Data Convergence Protocol (PDCP) and Hyper Frame Number (HFN) status of the E-UTRAN Radio Access Bearers (E-RABs) for which PDCP status preservation applies. The source eNB may omit sending this message if none of the E-RABs of the UE shall be treated with PDCP status preservation.

Step 11: If there is an MME relocation the source MME sends this information to the target MME via the Forward Access Context Notification message which the target MME acknowledges. The source MME or, if the MME is relocated, the target MME, sends the information to the target eNB via the MME Status Transfer message.

Step 12: After the UE has successfully synchronised to the target cell, it sends a Handover Confirm message to the target eNB. Downlink packets forwarded from the source eNB can be sent to the UE. Also, uplink packets can be sent from the UE, which are forwarded to the target S-GW and on to the Packet Data Network (PDN)-GW.

Step 13: The target eNB sends a Handover Notify (TAI+ECGI) message to the target MME.

Step 14: If the MME has been relocated, the target MME sends a Forward Relocation Complete Notification message to the source MME. The source MME in response sends a Forward Relocation Complete Acknowledge message to the target MME. Regardless if MME has been relocated or not, a timer in source MME is started to supervise when resources in Source eNB and if the S-GW is relocated, also resources in Source S-GW shall be released. Upon receipt of the Forward Relocation Complete Acknowledge message the target MME starts a timer if the target MME allocated S-GW resources for indirect forwarding.

Step 15: The UE initiates a Tracking Area Update procedure when one of the conditions listed in clause *Triggers for tracking area* update applies. The target MME knows that it is a Handover procedure that has been performed for this UE as it received the bearer context(s) by handover messages and therefore the target MME performs only a subset of the Tracking Area (TA) update procedure, specifically it excludes the context transfer procedures between source MME and target MME.

Step 16: When the timer started in Step 14 expires the source MME sends an UE Context Release Command message to the source eNB. The source eNB releases its resources related to the UE and responds with a UE Context Release Complete message. When the timer started in Step 14 expires and if the source MME received the S-GW change indication in the Forward Relocation Response message, it deletes the EPS bearer resources by sending Delete Session Request (Cause, LBI) messages to the Source S-GW. Cause indicates to the Source S-GW that the S-GW changes and the Source S-GW shall not initiate a delete procedure towards the Packet Data Network GateWay (PDN-GW). The Source S-GW acknowledges with Delete Session Response messages. If Idle state Signalling Reduction (ISR) has been activated before this procedure, the cause also indicates to the Source S-GW that the Source S-GW shall delete the bearer resources on the other old Core Netweork (CN) node by sending Delete Bearer Request message(s) to that CN node.

2.2.1.2 X2-HO procedure

Handover through the X2 interface is triggered by default for intra-LTE mobility unless there is no X2 interface established or the source eNB is configured to use S1-handover instead. X2-handover procedure is illustrated in Fig. 2.8. Like S1-handover, it is also composed of a acquisition (steps 1 to 3) preparation phase (steps 4 to 6), an execution phase (steps 7 to 9) and a completion phase (after step 9).

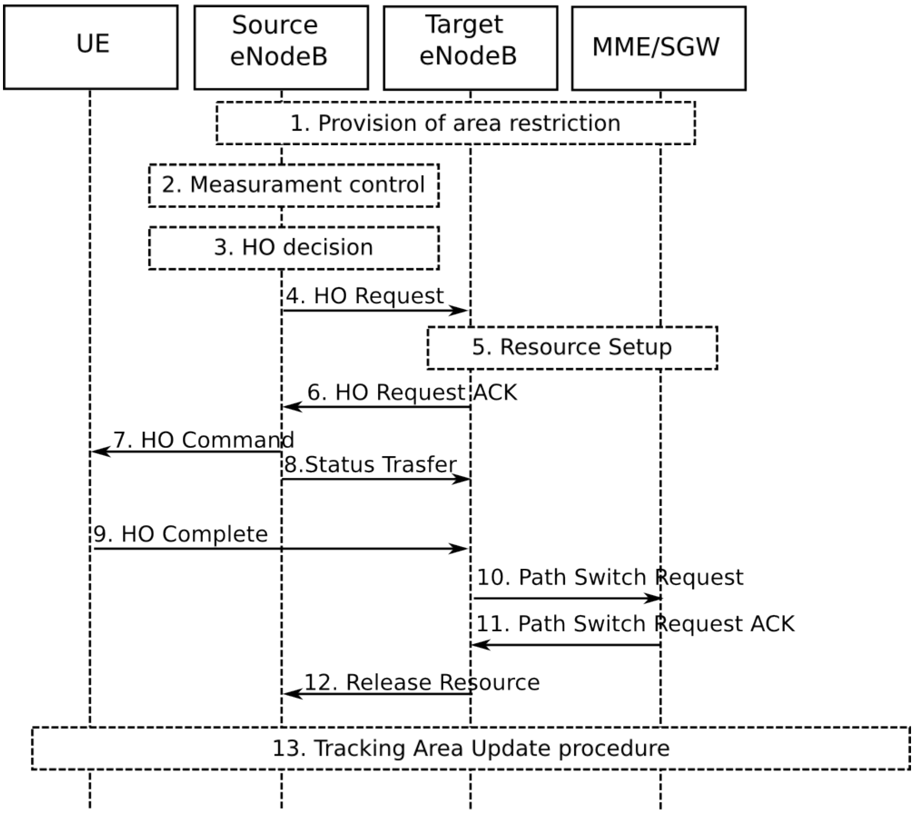


Figure 2.8: X2-based handover

The key features of X2-handover for intra-LTE handover are:

- The handover is directly performed between two eNBs, making the preparation phase quick.
- Data forwarding may be operated per bearer in order to minimise data loss.
- The MME is only informed at the end of the handover procedure when the handover is successful, in order to trigger the path switch.
- The release of resources at the source side is directly triggered from the target eNB.

Step 1-3: Analysis data to perform handover.

Step 4: If the source eNB selects the seamless mode for one bearer, it proposes to the target eNB in the Handover Request message to establish a GPRS Tunnel Protocol (GTP) tunnel to operate the downlink data forwarding.

Step 6: If the target eNB accepts, it indicates in the Handover Request ACK message the tunnel end-point where the forwarded data is expected to be received. The tunnel endpoint may be different from the one set up as the termination point of the new bearer established over the target S1.

Step 7: Upon receipt of the Handover Request ACK message, the source eNB can start forwarding the freshly arriving data over the source S1 path toward the indicated tunnel endpoint in parallel to sending the handover trigger to the UE over the radio interface. The forwarded data is thus available at the target eNB to be delivered to the UE as early as possible. When forwarding is in operation and in-sequence delivery of packets is required, the target eNB is assumed to first deliver the packets forwarded over X2 before delivering the ones received over the target S1 path, once the S1 path switch has been done. The end of the forwarding is signalled over X2 to the target eNB by the reception of special GTP packets that the S-GW has inserted over the source S1 path just before switching this S1 path; these are then forwarded by the source eNB over X2 like any other regular packets.

Step 8: the Status Transfer message provides the Sequence Number (SN) and the HFN that the target eNB should assign to the first packet with no SN yet assigned that it must deliver. This first packet can either be one received over the target S1 path or one received over X2, if data forwarding over X2 is used. When the source eNB sends the Status Transfer message, it freezes its transmitter/receiver status, that is, it stops assigning PDCP SNs to downlink packets and stops delivering uplink packets to the Evolved Packet Core (EPC).

Step 9-10: The Path Switch message is sent by the target eNB to the MME when the UE has successfully been transferred to the target cell (only when X2 initiated handovers). The target eNB sends Path Switch Request message to MME to inform that the UE has changed cell, including the TAI+Eutran Cell Global ID (ECGI) of the target cell and the list of EPS bearers to be switched. The MME determines that the S-GW can continue to serve the UE.

Step 11: The MME confirms the Path Switch Request message with the Path Switch Request Ack message.

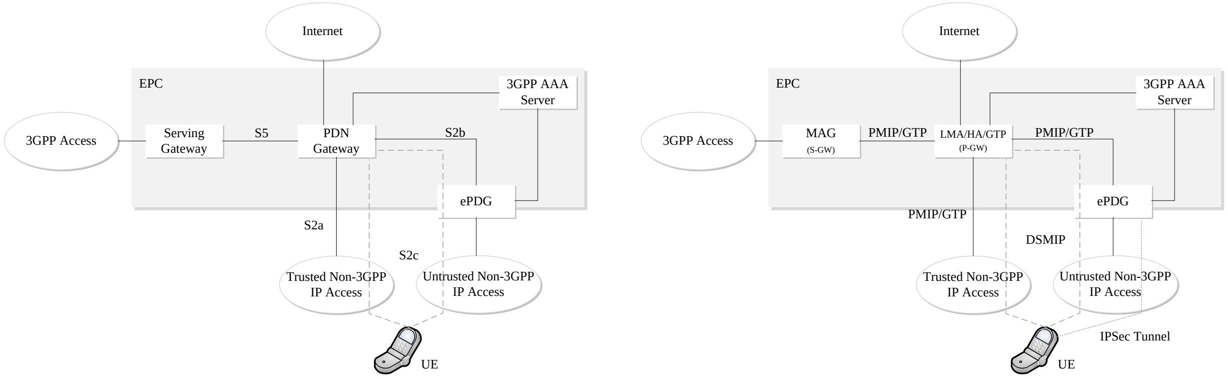
Step 12: By sending Release Resource the target eNB informs success of the handover to source eNB and triggers the release of resources.

Step 13: The UE initiates a Tracking Area Update procedure when one of the conditions listed in clause *Triggers for tracking area update* applies. If ISR is activated for the UE when the MME receives the Tracking Area Update Request, the MME should maintain ISR by indicating ISR Activated in the Tracking Area Update Accept message.

2.2.2 Connection of IEEE technologies to the EPC: Trusted and un-trusted non-3GPP networks

This section analyses the mechanisms defined by the 3GPP in order to connect non-3GPP access networks to the EPC. In order to incorporate non-3GPP access networks to the technologies available for the roaming of a terminal, the 3GPP has defined different mechanisms depending on the trust relation between the operators of each network. The level of trust put on a certain network does not depend on the technology used itself, but on the trust relationship defined between both operators. Such trust relations are normally defined as (legally binding) roaming agreements. If a roaming agreement between two operators is in place, the user is able to use another operator's access network, being authenticated either by the visiting or the home network. So, based on the roaming agreements signed by both operators, an access network may belong to two categories:

- Untrusted non-3GPP Networks: The non-3GPP network is completely untrusted by the UE and the 3GPP system.
- Trusted non-3GPP Networks: The non-3GPP network is completely trusted (such as a WLAN deployment done by the same operator) or some of the elements of the network may be trusted. For example, some servers at the non-3GPP network dealing with the authentication



(a) 3GPP modules and interfaces.

(b) Mapping to mobility protocols.

Figure 2.9: Interworking architecture for 3GPP and non-3GPP accesses.

and authorisation interworking with the 3GPP (e.g., Institute of Electrical and Electronics Engineers (IEEE) 802.1x mechanisms).

For a certain network, the fact of belonging to either category has a strong impact. The mechanisms used to connect the UE to the 3GPP depend on the trustiness of the access network. Fig. 2.9a shows a simplified scheme of the interworking architecture between 3GPP and non-3GPP networks to the EPC. Some entities involved in the interworking procedure, such as the Policy Charging and Rules Function (PCRF) and the Home Subscriber Server (HSS) have been deliberately left out to simplify the explanation. Fig. 2.9b shows the mapping of the entities and interfaces to the relevant mobility protocols (Proxy Mobile IPv6 (PMIPv6)/GTP/Dual Stack Mobile IP (DSMIP)). Although the 3GPP defines more options than the ones presented in Fig. 2.9b, we only show the options relevant to our discussion.

Focusing on the entities in charge of the UE mobility, in Fig. 2.9b the PDN Gateway (P-GW), defined in [5], is the user plane anchor for mobility between 3GPP and non-3GPP accesses. It includes the functionality of Localised Mobility Anchor (LMA), Home Agent (HA) and Generic Transport GPRS Tunnelling protocol end-point. Hence for all the different mobility protocols, it is in charge of defending the UE IP address and forwarding the packets to the current location of the UE. For 3GPP accesses, the Radio Access Network (RAN) is connected to the 3GPP through the S-GW, which is in charge of applying the different resource allocation policies to the 3GPP access and regarding mobility, acts as Mobility Anchor Gateway (MAG). Hence, the S-GW is responsible at the IP level of serving as DHCPv4 server and handling the Router Solicitation and Neighbour Advertisement used for IPv6 discovery. In addition, it forwards the different UE PDN connections to the P-GW through the appropriate PMIP/GTP tunnel. Before starting the discussion about the interworking with non-3GPP technologies, it is worth highlighting that there is a common solution that can be used in all cases, the usage of host mobility. As presented in Fig. 2.9a and Fig. 2.9b, the interface S2c maps to the use of DSMIP between the UE and the P-GW. In this case, the mobility is handled by the terminal and does not require any specific support by the network, hence the use of this solution is supported in trusted and untrusted 3GPP networks. The drawback of the solution is the requirement imposed on the terminal to implement the DSMIP protocol.

In the following sections we analyse the case of interconnecting the non-3GPP accesses to the EPC, focusing on the cases when the S2a and S2b interfaces are implemented with PMIPv6/GTP.

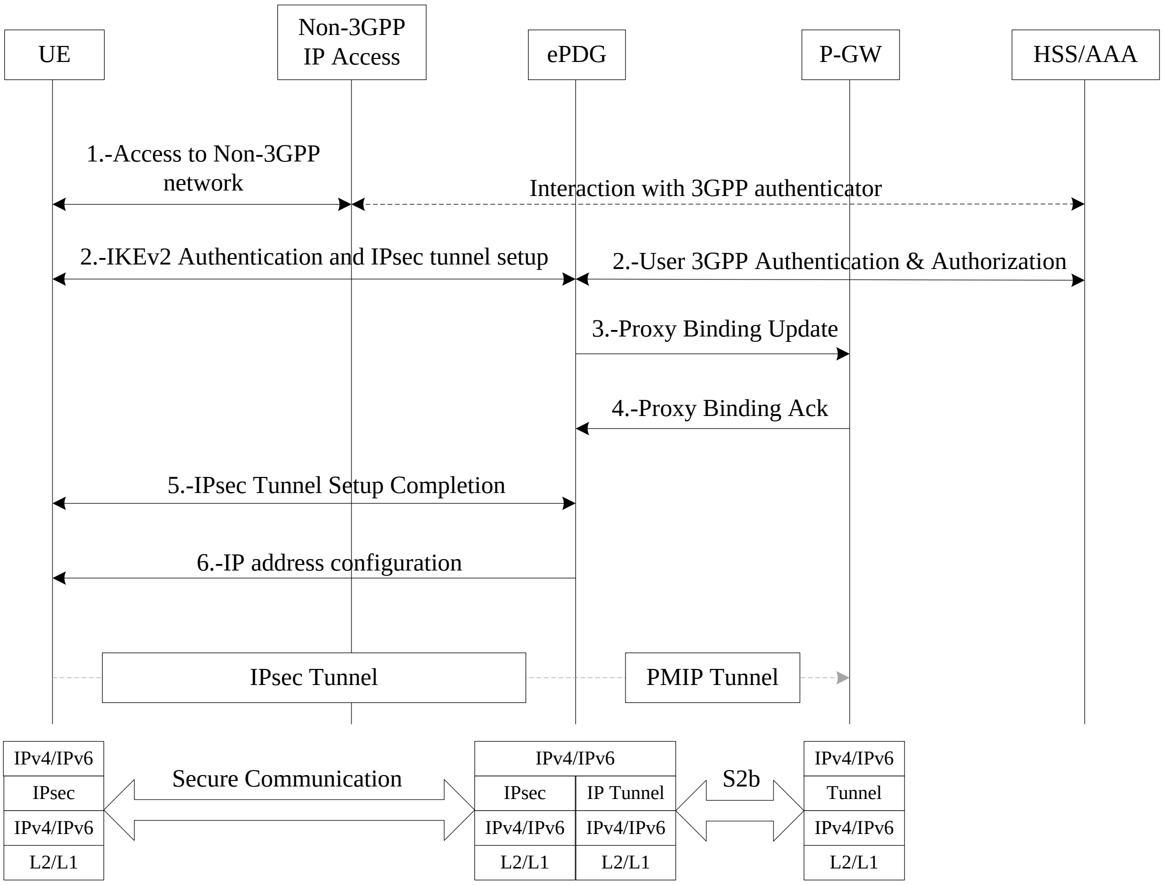


Figure 2.10: Untrusted non-3GPP procedures and user data protocol stack.

2.2.3 Un-trusted non-3GPP networks

Figure 2.9a shows the interworking architecture defined for the interconnection of untrusted non-3GPP access networks to the EPC. In this case, the non-3GPP network cannot be trusted and data from the UE should not travel un-protected through the non-3GPP network. The untrusted nature of the network also impacts on how it is connected to the EPC. In this case, there is no trusted element deployed on the non-3GPP network that can be used to interconnect to the EPC, hence an operator owned entity, called the Evolved Packet Data Gateway (ePDG), and located in the operator domain, is used. The ePDG connects to the P-GW through the S2b interface, which can implement PMIPv6/GTP. Fig. 2.10 shows the attachment procedures and the protocol stack used on the untrusted non-3GPP access network. As first step, the UE must gain access to the network through access specific mechanisms. At the end of this procedure, the UE must have acquired an IP address to be used in the next steps. Note that the non-3GPP access mechanisms are out of the scope of 3GPP and may or may not interact with the internal authentication and authorisation mechanisms of the operator. Once the IP address is acquired, the UE starts an IKEv2 tunnel establishment with the ePDG. The actual address of the ePDG to use is obtained through DNS interaction. Based on the APN selection provided by the UE, the ePDG selects the appropriate P-GW and starts the PMIPv6 binding procedure, by sending a Proxy Binding Update message to the P-GW, which answers back with the Proxy Binding Update (PBU) Ack. The reception of this last message triggers the completion of the IPsec tunnel setup, indication of which is transmitted

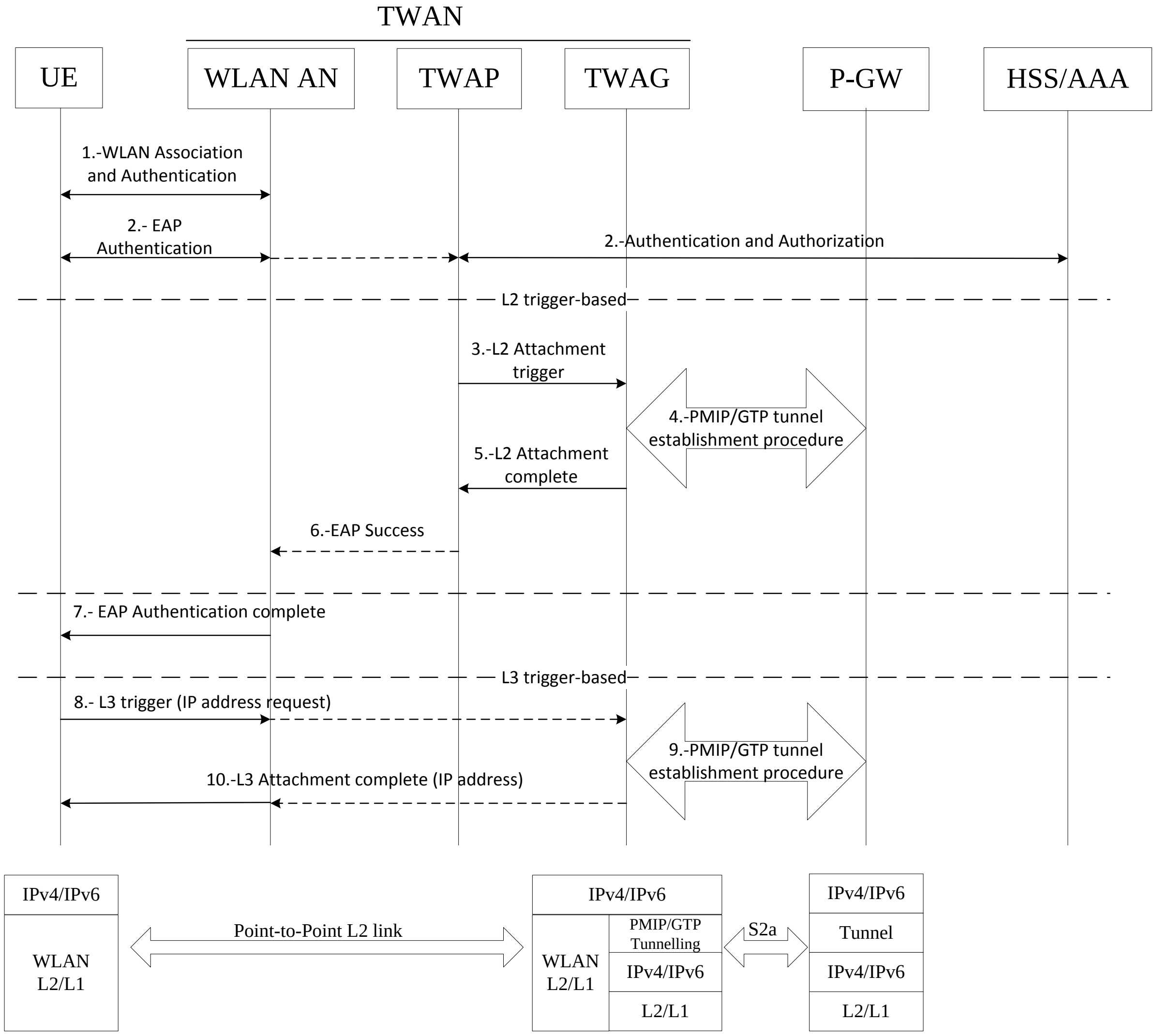

TWAN
UE
WLAN AN
TWAP
TWAG
P-GW
HSS/AAA
1.-WLAN Association and Authentication
2.- EAP Authentication
2.-Authentication and Authorization
L2 trigger-based
3.-L2 Attachment trigger
4.-PMIP/GTP tunnel establishment procedure
5.-L2 Attachment complete
6.-EAP Success
7.- EAP Authentication complete
L3 trigger-based
8.- L3 trigger (IP address request)
9.-PMIP/GTP tunnel establishment procedure
10.-L3 Attachment complete (IP address)
IPv4/IPv6
WLAN L2/L1
Point-to-Point L2 link
IPv4/IPv6
WLAN L2/L1
PMIP/GTP Tunnelling
IPv4/IPv6
L2/L1
S2a
IPv4/IPv6
Tunnel
IPv4/IPv6
L2/L1

Figure 2.11: Trusted WLAN procedures and user data protocol stack.

to the UE. Finally the ePDG, acting as MAG of the UE, assigns the IP address/prefix to the UE, using standard IPv4/IPv6 mechanisms. At this moment, the UE can start securely exchanging packets with the EPC. For the uplink direction, any packet is tunnelled to the ePDG using the IPsec tunnel. The ePDG then tunnels the packet to the PDN GW. For downlink packets arriving at the P-GW for the specific UE, the P-GW tunnels the packet based on the binding cache entry to the ePDG. The ePDG then tunnels the packet to the UE via the corresponding IPsec tunnel.

2.2.4 Trusted non-3GPP networks

The interconnection of trusted non-3GPP networks to the 3GPP is based on the existence of a trusted element within the non-3GPP access serving as point of union between both networks. Currently, although the specification covers undefined non-3GPP accesses, the most relevant access is Wireless LAN. For this technology, there is a specific working topic within the 3GPP that has gained some attraction lately, the so called S2a Mobility over GTP (SaMoG). The interconnection of a Trusted WLAN Access Network (TWAN) relies on the use of two entities, the Trusted WLAN

Access Gateway (TWAG) and the Trusted WLAN AAA Point (TWAP). The former one behaves as terminating point of the S2a interface, e.g., it acts as MAG and as default router for the UE on the access link, terminating Neighbour Discovery and DHCP signalling. The latter one, TWAP, is used as an AAA Proxy, relying the authentication signalling to the 3GPP AAA Server located in the core. It is also in charge of providing subscription information regarding the UE to the TWAG. Before the user attaches to the WLAN access, the UE needs first to discover the services provided by the network, e.g., if it supports Non-Seamless WLAN Offload (NSWO), multiple PDN connections, IP continuity etc.. Also, depending on the 3GPP release deployed, the UE may need to attach to an specific Basic Service Set ID (BSSID) in order to choose between the available services.

It is important to highlight that depending on the service the user is requesting, the requirements of the interconnection with the 3GPP are different. In this work we are focusing on three different services, NSWO, access to the 3GPP through one or several PDN connections (with the possibility of using different APNs) and IP continuity. For the first of these services, NSWO, the UE just requires an IP address attached to the trusted WLAN network, since the offloading is performed without considering mobility and the flow will not go through the EPC. For the two other services, the UE requires to have a connection to the EPC, hence it will use the S2a interface to communicate with it. In the following we focus on the services requiring 3GPP connection.

Fig. 2.11 presents the initial attachment procedure for a Trusted WLAN network implementing PMIPv6 or GTP in the interface between the TWAG and P-GW (S2a). The initial attachment procedure starts with the UE gaining connectivity at L2 with the WLAN network. In order to do so, standard IEEE 802.11 association and authorisation mechanisms must be performed. These procedures may require user interaction, e.g., captive portal. An important detail of the gained L2 connectivity is that the 3GPP is assuming a direct L2 connection between the UE and TWAG per PDN connection, i.e., no IP routers are transversed, even if several PDN connections are established between the same UE and APN.

Once the UE is able to start sending L2 frames, it starts (step number 2 in Fig. 2.11) EAP Authentication. In order to exchange the required EAP frames without IP connectivity, it is assumed the use of standard IEEE 802.1x. The authentication messages are carried towards the TWAP that is in charge of relying the message to the appropriate chain of AAA proxies, depending on the roaming scenario. Once the TWAP is able to contact the authoritative 3GPP AAA in the core, it downloads the list of capabilities the UE is authorised to request, such as IP continuity or local breakout capabilities. In addition, the TWAP needs to complement the information provided to the 3GPP AAA with some extra parameters such as the BSSID the UE is connected to. How the TWAP discovers this information is left out of the scope of the specification. At this point, depending on the specific interworking procedure implementation, the TWAP triggers the TWAG (step number 3 in Fig. 2.11) in order to start the S2a interface signalling (this is called L2 trigger since no IP connectivity has been provided yet to the UE), which in this example can be either GTP or PMIPv6 (step number 4 in Fig. 2.11). The exact trigger implementation is out of the scope of 3GPP and corresponds to signalling specific procedures between the TWAG and TWAP. In addition, the signalling required to setup the S2a tunnel, requires also some information provided by the UE, such as if this is a new attachment or a handover, the type of service required (IP continuity) or the IP address that requires handover survivability. Regardless of the use of the L2 trigger to start the S2a signalling, upon completion of the EAP 3GPP AAA exchange, the TWAP indicates the UE the result of the authentication procedure (step number 7 in Fig. 2.11). At this point of time, the UE is authenticated on the 3GPP and can request an IP address anchored at the operator's core. So, the UE requests an IP address through standard DHCPv4 or Neighbour discovery for the IPv6 case (both messages are named as L3 trigger in the successive). step number 8 in Fig. 2.11). If the S2a interface has not already setup the GTP or PMIPv6 tunnel, the reception of the L3 trigger starts the tunnel establishment (step number 9 in Fig. 2.11). Upon completion, the TWAG, acting as MAG or GTP end-point, conveys the 3GPP anchored IP address to the UE, through the DHCPv4/Neighbour Discovery response (step number 10 in Fig. 2.11).

In the case the UE desired to establish another PDN connection, the above procedure must be repeated per PDN connection, with some extra parameters to indicate some PDN specifics such as the APN and the credentials to be used.

Previous explanation has provided a simplified view of the process of interworking between the 3GPP and the non-3GPP technologies. In Chapter 3 we present the set of requirements that an interworking solution must meet in order to successfully provide all the functionality required for the above procedure and a solution landscape of the proposal being tackled at the 3GPP.

2.3 IP mobility management

This section is devoted to the critical analysis of the IP mobility management protocols. The reader is invited to refer to [7] for a historical introduction on the topic and the basic terminology used hereafter.

2.3.1 IPv6 mobility management

IPv6 mobility management has started with Mobile Internet Protocol version 6 (MIPv6) [8], which has been standardised by IETF in 2004. MIPv6 is a mobility protocol that provides global mobility support. Session continuity and reachability are maintained for the Mobile Node (MN) while undergoing IP handovers even when moving from one network to another. In order to achieve this, MIPv6 adopts a host-based approach where the MN participates in the mobility-related signalling. In MIPv6, each MN is always identified by its Home Address (HoA). While situated away from its home, the MN is also associated with a Care-of Address (CoA). The HA maintains the MN's bindings between HoA and CoA. It is also the MN's mobility anchor; all the data packets from/to the MN are tunnelled via the HA.

After MIPv6, discussions have been initiated on specifying a network-based mobility protocol designed/optimised for local mobility support in a single operational domain. For this objective, the network-based localised mobility management (NETLMM) working group was chartered at the IETF. Several years of work resulted in standardising a new protocol in 2008, namely PMIPv6 [9].

PMIPv6 inherits the concepts of MIPv6 and optimises them for providing mobility support in a localised domain. Compared to MIPv6, the main feature in PMIPv6 is to be a network-based mobility support protocol. This means that the network handles the mobility management on behalf of the MN. In order to achieve this, PMIPv6 introduces two network entities, namely the LMA and the MAG. The former is the HA for the MN in a PMIPv6-domain, where the mobility is managed by PMIPv6. The latter is a function on an Access Router (AR) that manages the mobility-related signalling on behalf of an MN that is attached to it. PMIPv6 manages all the sessions in the same way. All the data traffic of the MN passes through the LMA and a tunnel is established between the LMA and the MAG that is currently serving the MN.

2.3.2 DMM rationale

Current networks architectures are deployed in a hierarchical manner, relying on a centralised gateway. Thus, the existing IP mobility management protocols are generally deployed in a centralised manner. All the data traffic passes through a centralised mobility anchor, such as the HA in MIPv6 and the LMA in PMIPv6, and all the bindings are managed at this anchor as well. As the number of MNs and the volume of the mobile data traffic increase, such centralised architectures may encounter scalability issues (e.g. network bottlenecks, and single point of failure), security issues (e.g. attacks focused on the centralised anchor), as well as performance issues (e.g. centralised and non-optimal routing).

In addition, existing IP mobility protocols are designed to be always activated, managing all the services and all the traffic in the same way. They do not take into consideration that a given MN may not move during the use of a service (which is 60% of the cases [10] in operational networks) or that a service may not require mobility functions at all. Such approaches may thus lead to non-optimal routing and large overhead due to tunnelling mechanisms.

In order to cope with the rapid traffic explosion we are witnessing [11] IP mobility management protocols need to be adapted for such evolution. There is a need to define novel mobility management mechanisms that are both distributed and offered dynamically. They should be distributed in order to avoid any network bottleneck or single point of failure, and to provide better reliability. They should be activated/deactivated dynamically as needed, in order to globally reduce their signalling load and to increase the achieved performance.

Accordingly, the IETF chartered recently the DMM working group [12] in 2012. Various efforts from both industry and academia are being performed on specifying DMM schemes, e.g. [12, 13, 14, 15]. A common feature between different DMM schemes is distributing the mobility anchoring at the AR level. The MN changes dynamically its mobility anchor for new sessions, while keeping the previous anchors of ongoing sessions. When the sessions anchored at a specific mobility anchor are terminated, the MN deregisters from that anchor. Assuming that most of the sessions are relatively short, most of the data traffic is routed optimally without tunnelling [16].

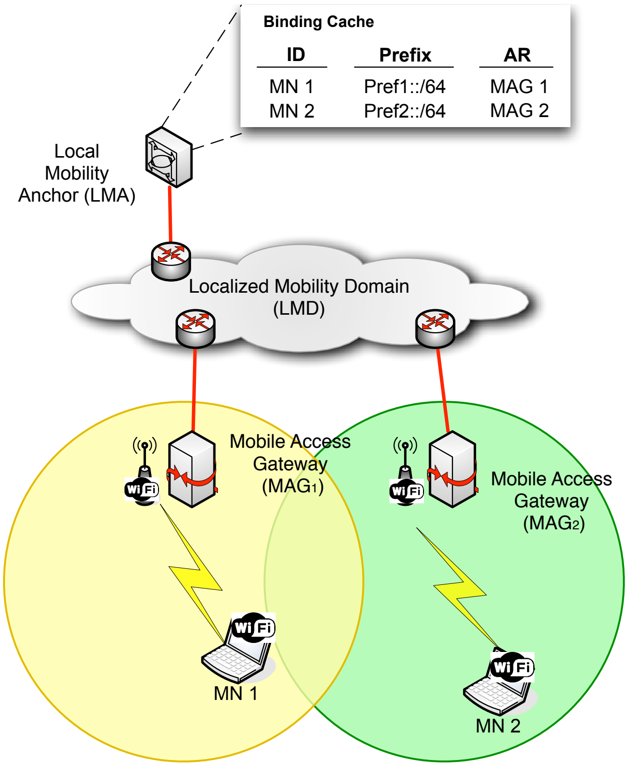
One of the DMM requirements [17] is to rely on the existing IP mobility protocols by extending and adapting them. This is in order to benefit such standardised protocols before specifying new ones, and also to facilitate the migration of networks architectures. Thus in the following, we consider two main approaches. The first is MIPv6-based, providing global as well as local mobility support for MNs that may move between several access networks. The second is PMIPv6-based, providing local mobility support for MNs moving in a single operational domain.

2.3.3 Mobile IP

MIPv6 [8] enables global reachability and session continuity by introducing the HA, an entity located at the Home Network of the MN which anchors the permanent IP address used by the MN, called the HoA. The home agent is in charge of defending the mobile device's home address when it is not at home and redirecting received traffic to the mobile's current location. When away from its home network, the mobile node acquires a temporal IP address from the visited network – called the CoA – and informs the HA about its current location by sending a Binding Update (BU) message. An IP bi-directional tunnel between the MN and the HA then redirects traffic to and from the mobile. In this way, the packets generated by the mobile node's communication peer – called the CN – and sent to the mobile's permanent address (i.e., its HoA) are tunnelled to the current MN location and arrive at the CoA. There is also optional support to avoid this suboptimal routing, which enables the MN to directly exchange traffic with a correspondent node without traversing the home network. This additional support is called Route Optimisation and allows the mobile to inform correspondent nodes about its current location.

2.3.4 Proxy Mobile IP

PMIPv6 [9] is a mobility-management protocol that allows legacy mobile terminals to perform handover operations across heterogeneous networks, without their involvement in the management of their own IP mobility signalling. As an example of operation (see Fig.2.12), consider a Localised Mobility Domain (LMD) scenario, where PMIPv6 provides mobility support, that comprises two MAGs, and an LMA. In addition to maintaining the state regarding the location of the MN in the



Binding Cache

ID	Prefix	AR
MN 1	Pref1::/64	MAG 1
MN 2	Pref2::/64	MAG 2

Figure 2.12: Proxy Mobile IPv6 (PMIPv6) scenario.

LMD, the LMA must maintain an IPv6-in-IPv6 tunnel with every MAG for forwarding the data traffic of their MNs. When an MN first arrives at the LMD, it attaches to an Access Point (AP) and sends a RS message requesting an IPv6 prefix. This message is received by the MAG, which asks the LMA for an IPv6 prefix for the MN through a PBU message. Next, the LMA replies to the MAG with a newly assigned IPv6 prefix for the MN through an Proxy Binding Acknowledgement (PBA) message and stores the mapping in its local lookup table, named Binding Cache (BC). Then, the MAG forwards the IPv6 prefix to the MN through an Router Advertisement (RA) message. Finally, the LMA uses the existing IPv6-in-IPv6 tunnel with the MAG (or creates a new one if there is none) for the data traffic exchanged by the MN with the network. When the MN moves to the coverage area of a second MAG, the process is repeated, but this time the LMA finds an existing entry in its BC for that MN, and therefore replies to the MAG with the same IPv6 prefix that the MN was using previously, updating the record for the MN and diverting its traffic to the new MAG tunnel. Thanks to the fact that the MAGs show the same layer-2 and IPv6 link local addresses to the MNs, these do not detect any layer-3 change while moving within the LMD.

In conclusion, thanks to PMIPv6, a Mobile Node may change from one layer-2 Point of Attachment (PoA) to another, but it always keeps the same IP address across the LMD managed by an LMA. It is also worth noticing that the operation of PMIPv6 does not require the MN to implement any modification or extra software in its layer-3 stack, although it may require the assistance of some layer-2 mechanisms to work more efficiently. These mechanisms are known as link-layer triggers, and are required to quickly detect a change of layer-2 PoA.

2.3.5 Distributed Mobility Management

Currently standardised IP mobility solutions come at the cost of handling operations at a central point – the mobility anchor – and burdening it with data forwarding and control mechanisms for a great amount of users. This central anchor point is in charge of tracking the location of the mobile and redirecting traffic towards its current topological location. While this way of addressing mobility management has been fully developed by the Mobile IP protocol family and its many extensions, it brings several limitations: *a)* sub-optimal routing, as traffic always traverses the central anchor, leading to paths that are, in general, longer than the direct one between the mobile node and its communication peer; *b)* scalability problems, as existing mobile networks have to be dimensioned to support all the traffic traversing the central anchors, and the anchor itself has to be powerful enough, and; *c)* reliability, as the central entity is a potential single point of failure.

In order to address such issues, a new paradigm of solution, the so-called Distributed Mobility Management (DMM), is currently being analysed by both academic and standards communities. DMM basically develops the concept of a flatter system, in which the mobility anchors are placed closer to the user, distributing the control and data infrastructures among the entities located at the edge of the access network [18].

The approaches proposed so far towards distributing the mobility management function can be divided into different categories, namely: *i)* clean-slate approaches, proposing novel network architectures tackling from the foundation the problems inherent to current IP mobility architectures, *ii)* architecture-dependent solutions, such as the different efforts initiated in the 3GPP (e.g., LIPA Mobility and SIPTO at the Local Network (LIMONET)), *iii)* peer-to-peer approaches, distributing the mobility management functionality across several nodes in the network. and *iv)* solutions based on or extending existing IETF protocols. Although the majority of the work has been focused on extending or modifying already existing IETF protocols, due to the benefits inherent to extending an already accepted protocol such as PMIPv6 or MIPv6 in 3GPP standards, there are several relevant works on each of the categories explained above, which we describe in the following.

The more relevant clean slate approach that has been discussed in the DMM related forums is the one presented in [19]. The solution proposed in [19] (which has been also submitted to the IETF as a draft) presents a novel approach that breaks with current trends on mobility management. It proposes the use of routing updates between routers to manage the mobility of the nodes within the mobility domain. It relies on DNS lookups to detect the prefix assigned to the node and BGP update messages to renew the information in the routing within the domain. Global roaming is also supported by issuing BGP primitives between several ASs. Although conceptually the proposal has been already discussed it still lacks of a deep analysis of its expected performance. Other existing clean slate approaches leverage the concept of identification/locator split to provide flatter architectures. In [20] the authors present a novel approach called HIMALIS (Heterogeneity Inclusion and Mobility Adaptation through Locator ID Separation) that advocates for a new architecture for mobility management built on top of the concept of locator and ID separation. End host traffic is routed through the optimal direct path through the change of the locators used in the communication, while the connection is not close since the identifiers are kept constant.

Regarding the second category, Architecture-dependent solutions, it is worth to mention the relevance of the DMM work being performed in the 3GPP. This organisation is currently looking for solutions specifically focused on providing enhanced mechanisms for local breakout, offloading and, in summary, enabling the optimisation of the mobility of the users, reducing its transit through the operator core network, hence alleviating the overload on the operator core network. Apart from already standardised solutions such as Selected IP Traffic Offload (SIPTO), Local IP Access (LIPTO), LIMONET, there are some works pushing for novel mechanisms for relocation of the GW allocated to the MN. Examples of these efforts are [21] and [22]. Both works provide complementary solutions for P-GW relocation within the 3GPP release 10 specification. Mainly both propose new mechanisms for detecting non optimal paths while being aware of the applications requirements.

One of the key aspects of the DMM concept is the distribution of the mobility management functionality across multiple entities and, ideally, as near as possible to the MN. Considering the idea of distributing the mobility management functionality, there have been several approaches that take advantage of new paradigms, such as Peer-to-Peer (P2P) technologies, to provide novel mechanisms for distributing the binding information. Relevant works using this approach are [23], [24] and [25]. In [23] the authors present m-Chord, a protocol used to distribute the HA and Foreign Agent (FA) functionality of MIPv4. The performance analysis performed indicates that in some cases the performance of the solution is even better than standard Mobile IP, although in the general case, there is a performance drawback from the use of the P2P technology. In [24], a new mobility management protocol based on Distributed Hash Table (DHT), Distributed IP Mobility Approach (DIMA) is presented. The protocol is similar to Mobile IP but the HA functionality is split and spread across different nodes that share a common binding distributed database. The data traffic towards the MN is intercepted by one of this nodes that acts as home agent, anchoring the HoA of the MN. The distributed mobility is achieved by relocating the nodes acting as distributed HA, closer to the MNs. Finally, [25] presents a solution for mobility management that distributes the functionality of the HA across multiple nodes through the use of a P2P approach. The protocol selected for the distribution of the information is Chord. The authors argue that one of the main drawbacks o using P2P overlays for mobility management is the lack of coherence between the overlay and the actual physical topology and features of the nodes. Hence they extend the P2P protocol to consider physical information through a Markov decision process, optimising the update and query performance.

There are several benefits inherent to the extension of already established protocols to support DMM. In fact, the 3GPP has already accepted as part of release 10 the use of PMIPv6 as an alternative to the well-known GTP protocol, and MIPv6 as the client-based mobility management protocol of choice. It is also a good practice within the IETF to first try to provide solutions based on already existing protocols, since these protocols have been already validated and discussed in the IETF. So, once the DMM charter was created in March 2012, there were several proposals aiming at extending these two protocols (PMIPv6 and MIPv6) to provide DMM support.

Based on the same mobility concepts as PMIPv6 and MIPv6, in [26] the authors present a generic solution for mobility management in Flat IP networks. The design encompass the use of Anchor Access nodes (AAN) and Visited Access nodes (VAN). The first ones maintain the mobility state of the node, while the data traffic is tunnelled between AAN and VAN when the node is roaming. The detection of the traffic and the AAN in charge of the mobility of a node is performed through packet inspection of uplink traffic. In order to overcome the situation when there is no uplink traffic the node is required to send uplink void packets when receiving downlink only flows, hence modifications to the end node are required. This work is extended in [27] to support a prefix relocation mechanism, which after some time is able to relocate the prefixes used by the MN to prefixes allocated to the serving AR. As in their previous work it requires modifications in the MN to indicate to the network the best moment to perform the relocation. Works focusing on MIPv6 based solutions are basically trying to reduce the impact of the triangular routing on the overall performance of the flows. In [28] the ADA (Asymmetric Double Agents) extension to MIPv6 is presented. The extension introduces two new entities in the network, the Local Mobile Proxy (LMP), that takes care of the functionality of the HA in MIPv6 but is located closer to the MN and the Correspondent Mobile Proxy (CMP), which located near the CN provides it with Route optimisation towards the LMP. A different approach to reducing the HA-MN delay is taken in [29]. This work proposes a solution that enables the use of multiple HAs distributed through the Internet, interconnected by high speed links and communicated through anycast routing. Hence these nodes can be always placed near the MN, in this way reducing all the problems of centralised deployments. Works levering on the PMIPv6 protocol, are mainly focused on providing route optimisation mechanisms between MAGs. In [30] the authors perform an analysis of the different mobility functionalities required in PMIPv6. Then, a solution is a design that splits these functionalities across several nodes in the network. Nevertheless, the proposed solution uses for the actual routing of the flows a centralised approach, not providing local breakout of the connections, hence no real distributed mobility is achieved. In [31] three mobility schemes are proposed: Signal-driven PMIPv6, Data-driven Distributed PMIPv6 and Signal-driven Distributed PMIPv6. The three mechanisms rely on control/data split and multicast or peer to peer communication to route the data towards the MN through the optimal path. Finally in [32] the authors present an extension for Proxy Mobile IP that enables the LMA to designate an entity to handle a certain flow. The anchoring function will follow the terminal as it roams across the mobility domain. The new entity in charge of performing route optimisation between the MAGs is called Intermediate Anchors (IAs). These entities are in charge of establishing tunnels with old and new MAGs, hence providing connectivity between them. The main problem of this solution is that it never provides the optimal path, but a close to optimality one.

2.3.6 SDN-based Mobility

SDN refers to the dynamic reconfiguration of the functions and capabilities of an operational network without changing the hardware and basic software of its elements. Under this general definition, SDN is being currently applied also to wireless and cellular networks (e.g., [33]), as investigated in other CROWD Work Packages (WPs). However, for historical reasons related to the initial efforts going under such name, it is often a synonym with data vs. control plane separation. In the following we adopt such terminology, which is very close to the CROWD WP4 theme of traffic and mobility management, which is discussed in this document. To highlight the importance and relevance of SDN for mobility, we mention that the latter is considered as the "gold" use case for the immediate application of very broad SDN concepts in [34].

The first widespread example of protocol for the configuration of network elements is OpenFlow (OF) [35], which has been endorsed and standardised by the Open Networking Foundation (ONF)[3] Initially OF attracted significant interest in universities, since it made it possible for researchers to experiment with new protocols without modifying the firmware of switches/routers, whose source code is typically not made available by the hardware vendors. More recently, OF is accepted by the industry at large and, as a matter of fact, it has become the reference protocol of the realisation of an SDN-enabled network architecture, since it has been integrated into a number of SDN frameworks with wider scope and objectives, including Network Function Virtualization (NFV) (e.g., OpenNaaS[4], Project Floodlight[5], OpenDaylight[6]).

However, it is worth noting that the separation of the data and control plane is neither new nor necessarily linked to SDN. For instance, the Multiprotocol Label Switching (MPLS) [36] natively distinguishes the forwarding elements and protocols from the control functions, which only reside on a very limited number of elements in the network, so as to achieve (mostly) high scalability at very high bit rates and to allow for a fast reconfiguration of the data paths. However, MPLS remains confined within a single domain and translation to/from Internet Protocol (IP) is necessary at its edges. Therefore, it is not an option as an end-to-end connectivity management solution. Another example is Forwarding and Control Element Separation (ForCES) [37], which is a charter in the IETF aimed at creating a framework, requirements, a solution protocol, a logical function block library, and other associated documents in support of the separation of the forwarding and

[3] `https://www.opennetworking.org/`.
[4] `http://www.opennaas.org/`.
[5] `http://www.projectfloodlight.org/`.
[6] `http://www.opendaylight.org/`.

control elements in a network. Even though it deals with timely and important topics, ForCES has not received major attention, possibly because of its abstract nature. The integration of both MPLS and ForCES with OF has been attempted, e.g., [38, 39, 40].

With specific reference to wireless and cellular networks, which are the only focus of CROWD, [41] can be considered a first attempt to implement a proof-of-concept instance of an SDN-driven approach to (campus) network design. In fact, vertical handover between WiFi and Worldwide Interoperability for Microwave Access (WiMAX) was implemented in a manner totally transparent for the users and using OF and other *open source* tools, including FlowVisor[7]. A more industry- and standard-oriented path is identified in [42], which illustrates in technical details the benefits of virtualising the EPC functions of a 3GPP core network into a data centre. Such vision is endorsed and extended in [43], where a full “software-defined mobile network” is defined in terms of its key requirements. A prototype implementation is also proposed, based on the concept of the MobileFlow Forwarding Engine (MFFE), which encompasses all the user plane protocols and functions, and the MobileFlow Controller (MFC), which is a logically centralised entity that configures dynamically the MFFEs. Virtualisation of the EPC functions to achieve higher scalability at a reduced cost [22, 21, 44] is also one of the objectives of the project Mobile Networks Evolution for Individual Communications Experience (MEVICO)[8], which has been funded under the Celtic Initiative.

Finally, a novel paradigm is explored in [45], which reports the experiences from the implementation of an SDN-based prototype of Follow-Me Cloud (FMC). FMC provides mobility features in a TCP/IP network for both users and services, maintaining all the ongoing network connections active, by exploiting the locator vs. identifier split: each MN is assigned an *identifier*, which is unique and persistent, while the *locator* is temporarily assigned depending on which network is currently visited by the MN. In the prototype, the IP address originally assigned to the MN in its home network is used as identifier and initial locator, and the latter is changed transparently to the MN by an underlying network of layer-2 OF switches. A user-centric view is also taken by the authors of [46], which illustrates the realisation of an Application Delivery Network (ADN) paradigm exploiting SDN concepts and the OF protocol. The solution is mostly intended for smart phones, which are also the target of [47], where the authors propose a preliminary solution for handling network functions, including mobility, through SDN, but only in *ad hoc* networks created by Device-to-Device (D2D) communications.

2.4 Access selection and vertical HO

The spreading popularity of smartphones, laptops and tablets is changing the way in which the wireless access network has been used recently and, therefore, it is also changing its architecture. In order to cope with the increasing bandwidth demand, macro-cells, femto-cells and pico-cells start to be smoothly integrated in the deployment of cellular networks, while WiFi and WiMAX are not only used in dedicated hotspots or to connect isolated regions to the Internet, but they are effectively used to offload the cellular network. Alongside the classic wireless access network, a wide set of other possibilities is also flourishing. For example, opportunistic relaying of the traffic by mobile devices through D2D communications [48] has also been investigated. Due to the heterogeneous nature assumed by the access network, multi-homed terminals are commonly available in the market too.

This new environment raises the issue of the selection of the best point of access, a.k.a. the access selection (see Fig. 2.13). If the access selection decisions are not fixed, but they are indeed periodically renewed in order to follow the dynamics of the environment, it comes straightforward

[7] `https://github.com/OPENNETWORKINGLAB/flowvisor/wiki.`

[8] `http://www.mevico.org/.`

that they may involve the switch of the mobile terminals to a different access technology, i.e. the access selection mechanisms may involve Vertical HandOver (VHO).

In this section we cover the most important concepts to consider when addressing this issue, i.e. we present the basic mechanisms underlying the access selection decision, the input parameters typically used and the objective that the state of the art solutions address through the access selection decision.

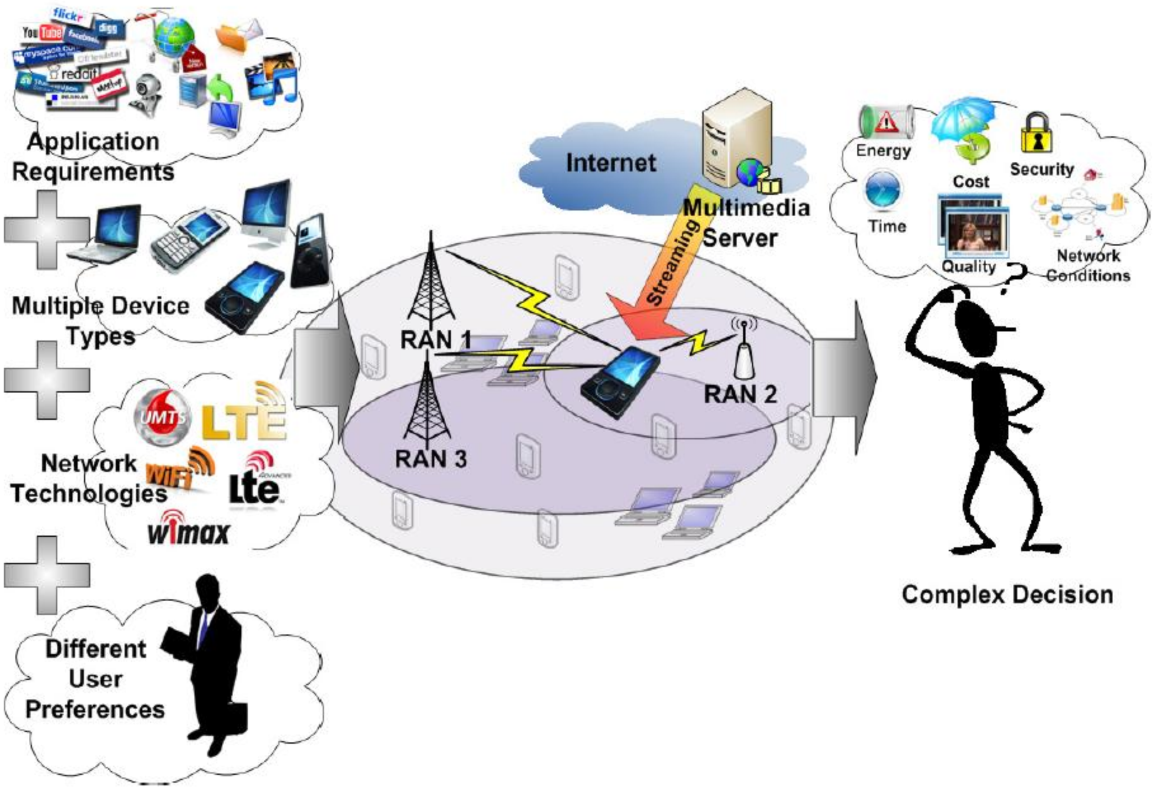


Figure 2.13: Heterogeneous Wireless Networks Environment (from [1]).

2.4.1 Overview over Access Selection/Vertical Handover mechanisms

The scenario where the access selection issue arises, is a heterogeneous wireless environment with terminals using multiple access interfaces and running non-real-time or real-time services. Depending on the goal of the access selection technique, the state of the art solutions envision different kind of environment set-up. The most conservative solutions suppose that a user could choose different technologies from the same operator [49]. Due to the fact that the operator has full control over his network, such environment choice typically implies that the access selection policies are optimised from the point of view of the network. Other solutions suppose that the user could attach to access networks of different operators thanks to economic agreements among them or between customers and operators [50]. The most daring ones envision a scenario with no contractual agreements where the users have a pool of potential points of attachment [51]. In such cases, the access selection techniques are usually thought to satisfy the needs and the requirements of the mobile users and mobile applications.

Whatever is the set-up scenario, it is always possible to identify three different phases characterising an access selection/vertical handover mechanism.

Information gathering. During the information gathering phase, the device responsible of the access selection decision, i.e., the mobile itself or a network element, collects all the data it needs to process. Depending on the decision algorithm, those data do not concern just network information, but also the rest of the components of the system, such as network properties, mobile devices, access points and user preferences. The pieces of information typically collected are the following:

- Availability of neighbouring network links by means of throughput, cost, packet loss ratio, Received Signal Strength (RSS), Signal to Noise Ratio (SNR), Carrier to Interference Ratio (CIR), Signal to Interference Ratio (SIR), Bit Error Rate (BER), distance, location or QoS parameters.
- Mobile devices' state, such as battery status, resources, interfaces or service class.
- User preferences, such as point of access priorities or service required.

Access selection/VHO decision. Based on the gathered information, this phase is in charge of deciding the first point of access of a mobile device or *when* and *where* to trigger a handover. In a homogeneous network environment the access selection and the handover decisions usually depend on RSS values. In heterogeneous networks the solutions of the same problems are instead quite complex. To make the best decisions, the information gathered from the different sources, i.e., network, mobile devices, and user preferences are typically weighted up and evaluated under a specific criterion that depends on the objective pursued through the access selection mechanism.

Access/Handover execution. This phase implements the access selection. In case of handover, this phase should also guarantee a smooth session transition process. To perform the VHO, different handover strategies cooperate with control signalling, and the IP management protocols. In order to allow such operations, this phase exploits a framework that ensure the functions being in charge of the mobility mechanism that will let the system to have seamless connectivity. Examples of this kind of architecture are IEEE 802.21, designed by the IEEE, and ANDSF, designed by 3GPP.

2.4.2 Scope and Optimisation goals addressed through Access selection/VHO

An important concept when developing an access selection mechanism is the access networks that are going to be compatible with it. In the state of the art solutions, several wireless technologies have been taken as candidates to be selected and, among them, a point of access belonging to the cellular network is typically present. Apparently, such solutions try to merge the cellular network with different other technologies in order to cope with the increasing demand of throughput. IEEE 802.11, 802.16, but also Bluetooth [52] or Digital video Broadcasting–Terrestrial (DVB-T) [53] in some specific solutions, are the other technologies taken into account. Surprisingly, also some wired technologies (as Ethernet [54]) have been considered. Furthermore, some solutions claim that the input parameters used and the optimisation goal they aim to reach are so general that any existing or future technology deployed would be a candidate for their access selection mechanisms [51, 55].

Through the use of an access selection mechanism, different goals could be accomplished. In the following we list the most important ones, together with a brief explanation of the used methodology.

Extend the battery lifetime. Along with the emergence of multi-purpose terminals, the gap between the energy requirements and the devices' battery capacity has progressively widened. For this reason, some of the solutions aim to reduce as much as possible the power consumed by the multiple radio interfaces equipped by the terminals. Each interface of the terminal can be in three different states: active, standby, or turned-off. When an interface is in active state, it can receive and transmit signals, or it can scan neighbouring cells. The battery is mainly consumed in this state. An interface in standby state is periodically activated just to scan new access networks or to monitor the link signal quality of neighbouring cells. In this state, the point of access typically exchanges control signalling messages with the terminals to track users for incoming communications. Finally, if the radio interface is turned off, no communication or measurement occurs and no energy is consumed. When there are not communications, typically all the interfaces are in standby state. However, mobile devices are most of the time in idle communication mode and consequently standby interfaces still consume a significant portion of the terminal's energy. In order to optimise the energy consumption of the terminals, it has been therefore proposed an access selection mechanism that always connects the devices just to the least power expensive technology among the ones in coverage [56]. The objective is to have only one interface active at a time except during vertical handover periods.

Throughput Maximisation. In case of a heterogeneous network, is not rare the case where a mobile device would have experienced a higher throughput if it had accessed the Internet through a different network technology. Indeed, it is always possible for a device to sit at the edge of the coverage area of a given point of access, or to be assigned to an overloaded technology. It is therefore possible to maximise through an appropriate access selection/VHO mechanism the overall throughput received by the user in the network, or, if the mechanism is user-driven, to choose the best access point among the available ones. Typically, the state of the art solutions, like [57], fall in this category, and the choice is given by an appropriate comparison of the different RSS received from all the available access technologies.

Energy Efficiency. Through the use of the access selection, it is also possible to achieve contemporary the maximisation of the throughput and the minimisation the power consumption of the terminal. Such a goal, is typically achieved through the optimisation of the energy efficiency of the whole system, or of the terminal where the access selection mechanism runs. Indeed, if we denote with P the energy consumed and with B the throughput achieved by the terminal where the access selection mechanism runs, the energy efficiency E is defined as:

$$E = \frac{B}{P}. \tag{2.1}$$

Since the energy efficiency maximisation is known to be NP-Hard, simplifications are typically introduced. In [58], for example, the energy efficiency is computed at the user side, where the required throughput is supposed to be known. The user then chooses the access technology that is able to serve his traffic, with the lowest energy consumption.

Load Balancing. From the point of view of the network, an interesting goal that could be achieved through the use of the access selection/VHO mechanism is the balancing of the load of the costumers among all the points of attachment. Indeed, the load balancing of the traffic ensures intrinsic robustness to peaks in the traffic demand or to failures of network's components. By the way, due to the complexity of the optimisation to be performed, also in this case basic assumptions are typically introduced. The work in [59] is an example of such a use of the access selection, where the load that the users offer to the network is again supposed to be known. Furthermore, starting from an initial assignment of the mobile nodes to the points of access, the optimisation of the load balancing is achieved by performing the access selection only for the user or for the set of users whose RSS goes below technology dependent thresholds. For this set of users, the access selection is performed so to maximise a cost function that is directly proportional to the average of the remaining bandwidth at each point of access.

Minimise Hand-Over Occurrence Frequency. Among all the possible input parameters that an access selection mechanism could have, there are also mobility-related parameters such as terminal velocity, moving pattern, moving history and location information. Through those pieces of information it is possible to estimate what would be the position of the users in the next future and, therefore, it is also possible to design an access selection mechanism that minimises the frequency of handover occurrence. The authors of [2], for example, given the mobility of devices in a heterogeneous scenario where cellular networks and hotspots technologies are present, evaluate whether it is worth performing the handover to hotspots whenever possible.

Reduce Monetary Cost. From the point of view of the user, another interesting goal that could be achieved through an intelligent choice of the point of access is the minimisation of the monetary costs a user has to afford in order to satisfy the QoS requested by the applications he/she is running. An example of an access selection done with this purpose is the "WiFi first" policy that devices commonly implement nowadays. Such a policy indeed minimises the monetary costs exploiting the fact that the WLAN contracts typically do not relate the traffic generated by the users to the final cost they have to afford, while instead the cellular data traffic contracts do so. Another example is given by [53]. In this case the access selection is performed by means of the evaluation of a function that takes into account the level of satisfaction of each application running in the device weighting the performance achievable and the monetary cost.

Multi-criteria Optimisation. When the access selection mechanism is not only used for a specific goal, but tries to achieve several objectives at the same time, the point of access is chosen taking into consideration a Multi-Criteria Decision-Making (MCDM) process. Due to the fact that the different objectives are typically competitive and the optimum for one of them does not necessarily represent the optimum for the other ones, the selection of the point of access represents the best possible trade-off. MCDM optimisations exploit a framework as the one shown in Fig. 2.14. In such framework the "best" network is obtained through the evaluation of different objectives categorised into four different groups: operator policies, e.g., load balancing; terminal properties, e.g., power consumption; customers preferences, e.g., monetary costs; QoS requirements, e.g., throughput or latency. Such objectives are then weighted and adjusted by some criteria, that could be both static or dynamic in time, in order to obtain a ranking among the available points of access. The ranking is achieved by evaluating a *distance* toward an ideal optimal solution, e.g., through the use of Grey Relational Analysis (GRA) or through the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), by gathering optimal weights through which to evaluate the different objectives, e.g., with Simple Additive Weighting (SAW) or with Multiplicative Weighting Exponent (MWE), by means of fuzzy logic or by modelling such competitive environment

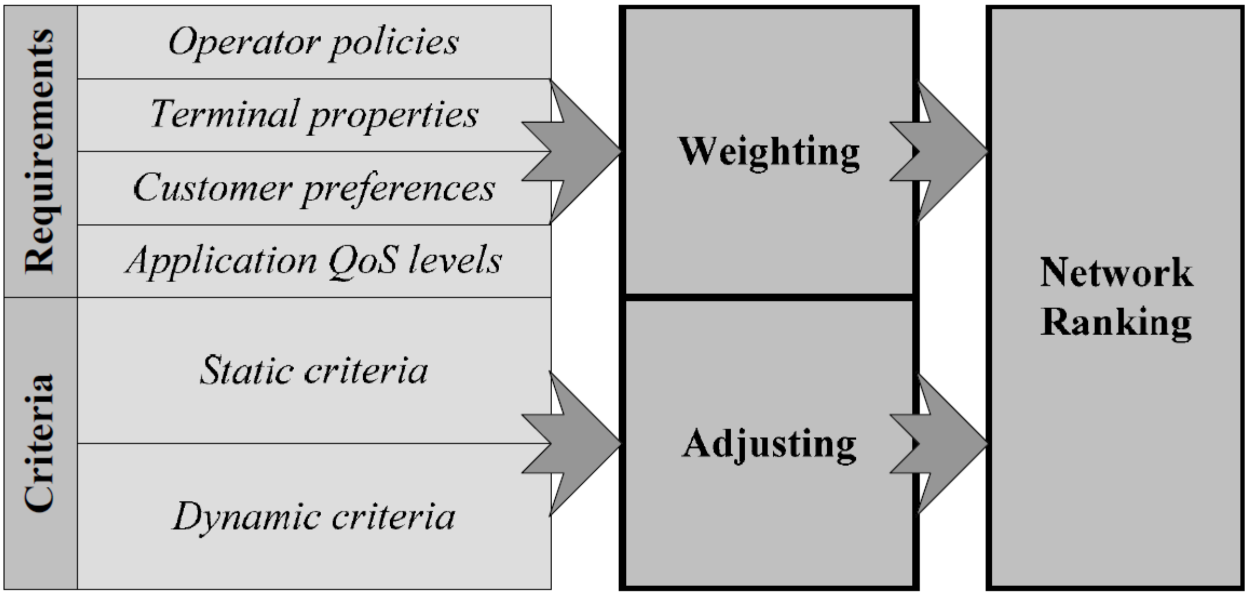


Figure 2.14: MCDM framework for access selection (from [2]).

through game theory. Solutions that follow these approaches are, e.g., the ones presented in [49, 51, 52].

2.5 IEEE 802.21

The main purpose of IEEE 802.21 [60] is to enable handovers between heterogeneous technologies (including IEEE 802 and cellular technologies) without service interruption, hence improving user experience of mobile terminals. Many functionalities required to provide session continuity depend on complex interactions that are specific to each particular technology. 802.21 provides a framework that allows higher levels to interact with lower layers to provide session continuity without dealing with the specifics of each technology. That is, the upcoming protocol can be seen as the glue between the IP centric world developed in IETF and the reference scenarios for future mobile networks currently being designed in 3GPP and 3GPP2 or other technology specific solutions. Additionally, while IETF does not cover specific layer-2 technologies, 3GPP/3GPP2 only addresses cellular technologies and how to integrate in them upcoming technologies such as WLAN. IEEE 802.21 provides the missing, technology-independent, abstraction layer able to provide a common interface to upper layers, thus hiding technology specific primitives. This abstraction can be exploited by the IP stack (or any other upper layer) to better interact with the underlaying technologies, ultimately leading to an improved handover performance.
To achieve these goals, IEEE 802.21 defines a media independent entity that provides a generic interface between the different link layer technologies and the upper layers. To handle the particularities of each technology, 802.21 maps this generic interface to a set of Media Dependent Service Access Points (SAPs) whose aim is to collect information and to control link behaviour during handovers. In addition, a set of remote interfaces terminal-network and network-network are defined to convey the information stored at the operator's network to the appropriate locations, e.g. to assist the terminal in handover decisions.

2.5.1 802.21 Objectives

Following the lines presented in the introduction, the contribution of the 802.21 standard is centred around the following three main elements:

i. A framework that enables seamless handover between heterogeneous technologies. This framework is based on a protocol stack implemented in all the devices involved in the handover. The defined protocol stack aims at providing the necessary interactions among devices for optimising handover decisions.

ii. The definition of a new link layer SAP that offers a common interface for link layer functions which is independent of the technology specifics. For each of the technologies considered in 802.21, this SAP is mapped to the corresponding technology-specific primitives. The standard draft includes some of these mappings.

iii. The definition of a set of handover enabling functions that provide the upper layers (like e.g. mobility management protocols such as Mobile IP [8]), with the required functionality for performing enhanced handovers. These functions trigger, via the 802.21 framework, the corresponding local or remote link layer primitives defined above.

Although the main purpose of IEEE 802.21 is to enable the handover between heterogeneous technologies, a set of secondary goals have also been defined. These secondary goals are:

- Service Continuity, defined as the continuation of the service during and after the handover procedure. One of the main goals of 802.21 is to avoid the need for restarting a session after a handover.

- Handover aware applications. The 802.21 framework provides applications with functions for participating in handover decisions. For instance, a voice application may decide to execute a handover during a silence period in order to minimise service disruption.

- QoS (Quality of Service) aware handovers. The 802.21 framework provides the necessary functions in order to take handover decisions based on QoS criteria. For instance, we may decide to handover to a new network that guarantees the desired QoS.

- Network discovery. This is an 802.21 feature that allows to provide users with information on the candidate neighbours for a handover.

- Network selection assistance. Network selection is the process of taking a handover decision based on several factors (such as QoS, throughput, policies or billing). In line with the above, the 802.21 framework only provides the necessary functions to assist network selection, but does not take handover decisions which are left to the upper layers.

- Power Management can also benefit from the information provided by 802.21. For instance, power consumption can be minimised if the user is informed of network coverage maps, optimal link parameters or 'sleep' or 'idle' modes.

2.5.2 IEEE 802.21 Architecture

In this section we present the general architecture of IEEE 802.21. We describe the different layers in the 802.21 protocol stack and their interaction, both at the node and network level. Fig. 2.15 shows the logical diagram of the general architecture of the different nodes in an 802.21 network. It shows a Mobile Node with an 802 interface and a 3GPP one, and that is currently connected to the network via the 802 interface. The figure shows the internal architecture of the Mobile Node, the 802 network, the 3GPP network and the Core Network. As it can be observed from the figure, all 802.21 compliant nodes have a common structure surrounding a central entity called Media Independent Handover Function (MIHF). The MIHF acts as intermediate layer between the upper and lower layers whose main function is to coordinate the exchange of information and commands between the different devices involved in taking handover decisions and executing the handovers. From the MIHF perspective, each node has a set of MIHF users, which will typically be mobility management protocols, that use the MIHF functionality to control and gain handover related information. The communications between the MIHF and the other functional entities such as the MIHF users and the lower layers are based on a number of defined service primitives that are grouped in SAPs. Currently, the following SAPs are included in the 802.21 standard draft (see Fig. 2.15):

- MIH_SAP: This interface allows communication between the MIHF layer and the higher layer MIHF users.

- MIH_LINK_SAP: This is the interface between the MIHF layer and the lower layers of the protocol stack.

- MIH_NET_SAP: This interface supports the exchange of information between remote MIHF entities.

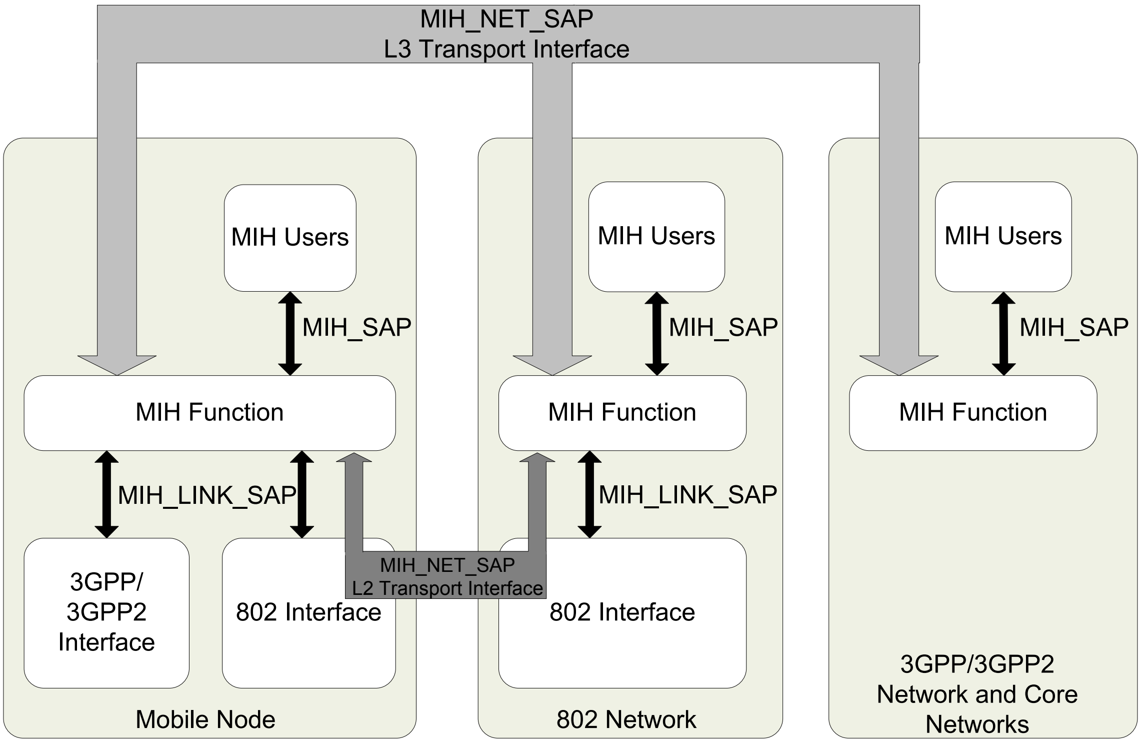


Figure 2.15: 802.21 General Architecture.

It is worth to notice that all communications between the MIHF and lower layers are done through the MIH_LINK_SAP. This SAP has been defined as a media independent interface common to all technologies, so that the MIHF layer can be designed independently of the technology specifics. However, these primitives are then mapped to technology specific primitives offered by the various technologies considered in 802.21. A table with the mapping of the primitives of the MIH_LINK_SAP interface to the link primitives of several technologies is included in the 802.21 draft.

The 802.21 architecture and reference model explained above, present a framework which support a complex exchange of information aiming at enabling seamless handover between heterogeneous technologies. 802.21 defines three different types of communications with different associated semantics, the so called Media Independent Handover (MIH) services:

- The Media Independent Event Service (MIES) provides event reporting of dynamic changes in link characteristics, status and quality.

- The Media Independent Command Service (MICS) enables MIH clients to manage and control the link behaviour related to handover and mobility.

- The Media Independent Information Service (MIIS) provides details about the characteristics and services provided by the serving and surrounding networks.

The MIH services can be delivered in an asynchronous or synchronous way. Events generated in link layers and transmitted to the MIHF or MIHF users are delivered by an asynchronous method, while commands and information, generated by a query/response mechanism are delivered in a synchronous way.

2.5.3 IEEE 802.21 amendments

The base IEEE 802.21 specification was published in 2008. Since then, the IEEE 802.21 WG has been working on several amendments extending the basic services provided by the specification or addressing specific cases which were not tackled in the base specification. Currently there are two approved amendments, IEEE 802.21a and IEEE 802.21b, while another two, IEEE 802.21c and IEEE 802.21d, are being finalised. Herein we provide a brief overview of each of these works:

- IEEE 802.21a: Security Extensions to Media Independent Handover Services and Protocol [61], provides mechanisms for securing the communication between an MN and a Point of Service (PoS). In addition, it provides pre-authentication mechanisms between an MN and a target PoA, in such a way that the terminal is pre-authenticated in the target network prior to the handover, reducing significantly the time associated to the handover.

- IEEE 802.21b: Extension for Supporting Handovers with Downlink Only Technologies [62], provides a novel mechanism to command a group of nodes attached to a downlink only technology, such as Digital video Broadcasting (DVB), to handover to a certain network. The base protocol defined in IEEE 802.21-2008 required a request/response exchange for commands, this specification extends this behaviour with a new indication message, without requiring confirmation that is not possible in such technologies without returning channel.

- IEEE 802.21c: Single radio handover [63], provides mechanisms to optimise the handover operation of single radio terminals. These terminals have multiple interfaces of different technologies but are not able to use them simultaneously, hence extensions to the base protocol to enable this use case are required.

- IEEE 802.21d: Multicast group communication [64], provides mechanisms for managing group of nodes, enabling multicast communication. Through this amendment it is possible to command handover operations to large networks of nodes, such as sensors, using multicast transport, hence reducing significantly the amount of signalling used.

2.6 Information gathering in dense networks

This section focuses on technologies enabling the terminal and the network to obtain information regarding the status and available services of the access network. Hence we study here the two most relevant technologies currently, the Application Layer Traffic Optimisation (ALTO) protocol (Section 2.6.1) and IEEE 802.11u (Section 2.6.2).

2.6.1 ALTO protocol

The ALTO [65] service provides network information to optimise the use of network resources and hence improve the performance of the application. For example, the ALTO client may request the cost to communicate with several content servers that all contain the same required data, then the ALTO server sends back the answer, by giving an ordered cost list with the cost associated to each content servers, allowing the ALTO client to choose the best one for the given cost.

The topology, the information about the network element, the computed cost are all given by the Internet Service Provider (ISP) and are managed and stored in the ALTO server information base.

The targeted application type are distributed applications where there is a choice between several sources disseminated over the network. Peer-to-peer (P2P) and content delivery networks (CDN) are the kind of applications that may use ALTO to optimise their communication.

Currently there is no way to be able to do the proper selection of the best information provider based on the network characteristics. Few information is provided to hint such choice based on optimisation of the communication between the application and its information provider. The current approach is either a fixed choice (given by the DNS for example), or a random choice. The first solution does not take advantage of the distribution of the information and creates a bottleneck, the second one is sub-optimal when the performance of the communication is not equivalent between all the information providers.

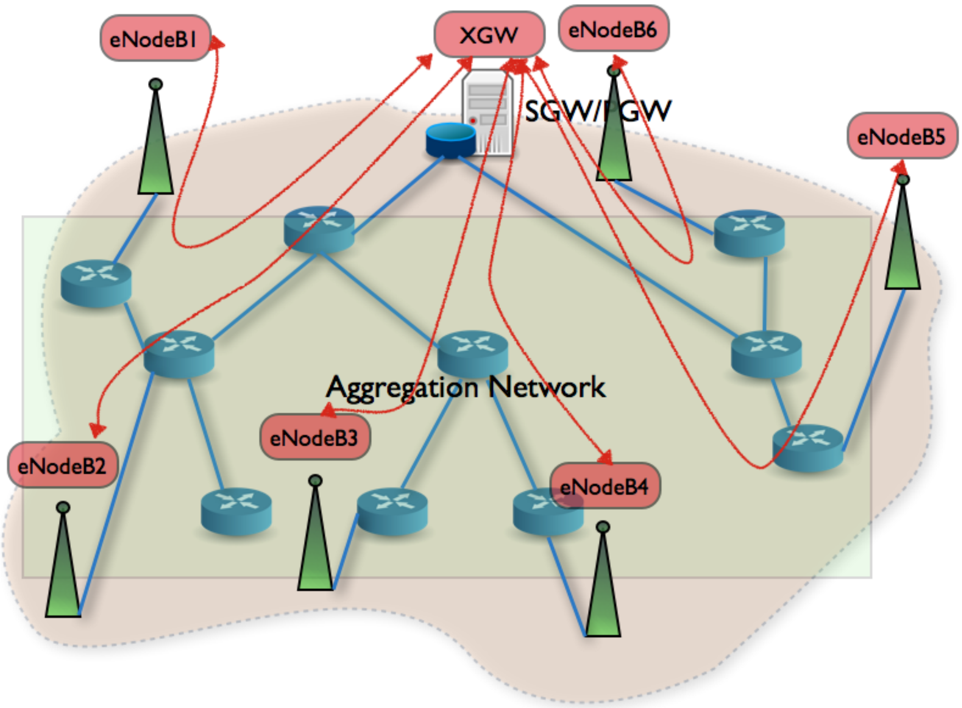


Figure 2.16: ALTO network abstract view.

The basic information of ALTO is based on maps of an abstract network corresponding to a simplified view of the network. In Fig. 2.16 , the abstract view is drawn in red, the aggregation part of the mobile network has been removed, only the eNBs and the S-GW and the P-GW has been taken into account. It may be assumed that the relevant part of the network for the cost constraint are only on the radio accesses level and the LTE core network and its interconnection with other providers. The routing cost example, in Fig. 2.17 shows how a cost may differentiate several paths between the MN on the operator 1 network and other devices that are accessible through other operators. The internal routing cost for the operator 1 has been set to 3 and the cost between operator 1 and the other operators has been set to different values (15, 12, 10), representing the effective cost that the operator 1 will have to pay for a communication. As the operator 1 is not able to compute routing cost inside other operators networks, the routing cost is uniform for all devices being located in these networks. If there are some agreements between operators, it is possible to have a collaboration between ALTO servers from different operators, providing a more accurate answer. In the example, the ALTO service will reply with an ordered list of paths, the best one being the path between the MN and IP3, followed by the paths MN IP2 and MN IP1.

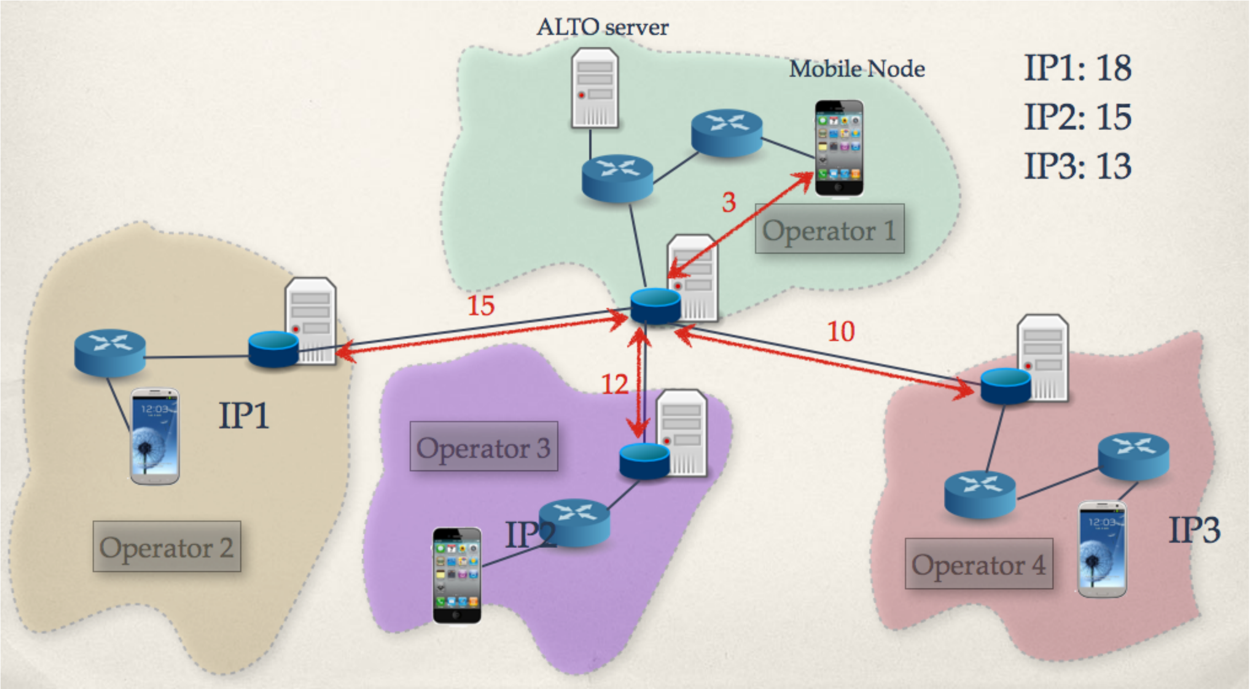


Figure 2.17: ALTO routing cost example.

The network map describes the topology of the network, this can be a partial description and then can be considered as an abstract view.

The cost map records cost for several properties between connected network locations, and allows to compute the cost for this property for the path between two endpoints (the one that corresponds to the ALTO client, and the one that will provide the information data).

The cost attributes are:

- Metric identifying what the cost represents (routing cost, hop count, ...).
- Mode identifying how the cost should be interpreted (expressing a numerical value which can be used in a computation to determine the best choice or a ranking which is just the given order for each alternative).

The attributes set defines a cost type.

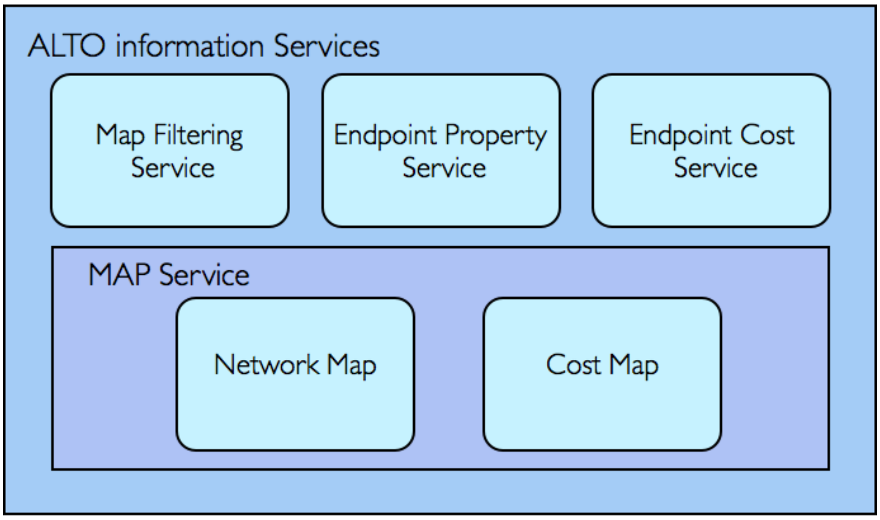


Figure 2.18: ALTO services.

The ALTO Protocol uses a Representational State Transfer Representational State Transfer (REST)-ful design. The communication between the ALTO client and the ALTO server uses standard Hypertext Transfer Protocol (HTTP). The HTTP data is encoded on the JavaScript Object Notation (JSON) language.

There is a version tag associated to the Network Map Information Resource that may be used as a way to give information about the modifications brought to the ALTO Server data.

2.6.2 802.11u

IEEE 802.11u [66] is an amendment to IEEE 802.11 to improve the ability of devices to discover, authenticate and use nearby WiFi APs. Its main goal is to provide information that will help to access WiFi APs in a way that is as easy as it is in the 3GPP world [67]. This is implemented via modifications of the following WiFi messages: the *Beacon*, the *Probe* and the *Association* messages.

The main changes are about the advertising protocol to support additional information, the roaming information, the QoS map and the Subscription Service Provider (SSP) interface.

IEEE 802.11u introduces the concept of a SSP which is the entity responsible for managing the user's subscription and associated credentials. Multiple SSPs will typically be accessible through a single AP, reflecting the various roaming relationships. This breaks the relationship between the Service Set ID (SSID) and the access credentials. Instead, devices dynamically query which SSPs are accessible via the AP, irrespective of the SSID that the AP is broadcasting.

IEEE 802.11u also introduces the concept of Homogeneous Extended Service Set ID (HESSID), which is a new identifier that is used to identify a set of access points that exhibit common networking behaviour.

Access Network Query Protocol (ANQP) allows the mobile node to retrieve information such as location, cellular network roaming, emergency services support, authentication, etc. And it is one of the technologies used in the HotSpot 2.0 initiative.

Compared to other protocols such as IEEE 802.21, ANDSF or ALTO that also gives information about network elements, IEEE 802.11u is inherently a distributed way to access to information such as roaming, the network element itself provides the requested information, there is no need to collect the data and then process it to make it available to all MNs. In IEEE 802.11u the network discovery process is done by the mobile through communicating with the APs and doing the selection with the MN's own criteria. The other protocols assume that the network discovery is done, and provide the list of APs that the MN should find on this location. To be relevant, services such as 802.21 Media Independent Information Service (MIIS) or ANDSF should have a view of the network that is recent enough to take into account the last changes for known APs. On the other hand, 802.11u eases the mobile scan process to select only valid (in the sense of APs the terminal can connect to) APs.

2.6.3 Protocol overlap and information lifetime

As seen above, the previous protocols share common information. For example roaming information are present in both 802.21 MIIS, ANDSF and 802.11u. Then the choice of which protocol should be used to get this kind of information arises. Since each protocol has its own benefits and disadvantages, the choice of protocol must be left to the operator to decide the best option for his deployment.

Another way is to get the information that has been updated the most recently, although mechanisms allowing versioning, to enable the creation of an ordered solution, are present only on 802.11u and ALTO protocols. Furthermore, the use of a protocol that is based on environment discovery such as 802.11u allows to get the real picture and to check the number of times the information required is not present, helping the evaluation of the information system.

3 Research activities

This section presents the research work performed in the area of Connectivity Management during the last 9 months of work in the CROWD project. Herein we focus on each of the components of the WP, analysing each of the modules required to provide IP-based mobility in a CROWD context. Hence, Section 3.1 is devoted to the analysis and dissection of the architecture of the next generation of connection managers, which must be smarter than current ones to be able to deal with the increasing complexity of wireless access networks. The section explores the different possibilities for information gathering and analysis the opportunities that the handover can present for power savings. Finally, Section 3.2 is focused on providing the requirements and tools which will be used in Chapter 4 to sketch the proposed CROWD mobility architecture.

3.1 Decision management

3.1.1 Analysis of the Connection Manager of different smartphone architectures

IP mobility has been a research topic for a long time and the different solutions designed to allow a user to freely roam across different points of attachment are a clear example of the evolution of the research on this topic, continuously adapting to the new requirements imposed by the operators. Although there are hundreds of different solutions, none of them has been a clear market success and none is massively deployed. So the current panorama on mobility management is somehow mixed, since mobility solutions have only been deployed within the cellular operator boundaries, e.g., a user can freely roam through the different access networks defined by 3GPP, but there is no solution for inter technology handover or IP mobility within non-3GPP technologies in the wide sense[1], due to the lack of support in the network and in the terminal. The lack of a common IP mobility solution implemented in the majority of smartphones and networks has as consequence the interdependence between the mobile user experience and the smartphone mobile services exposed to the applications. The smartphone operating system provides a set of mobility related functions which an application can benefit from in case it decides to handle mobility, as explained in section above. These functionalities depend on the operating system (e.g., Android, iOS, Windows Phone 8) and include connectivity events such as network down events, and highly variable network commands the application can use. Usually the terminal connectivity is decided by the service widely known as the Connection Manager, which is in charge of deciding what is the best connection for the terminal in a specific moment, and the application has to deal with these decisions. This section focuses on understanding the current state of the art on the mobility support at the Connection Manager in different terminals, providing a functional view of the differences between the major operating systems to identify the improvements that can be done in the future to optimise the mobile user experience. This section reports ongoing research work that will be completed during the next year of the project. The rest of the section is structured as follows: In Section 3.1.1.1 we present the experimental deployment and we evaluate the first results obtained on the experimental analysis for the management of network connectivity in Section 3.1.1.2. In this deliverable we report on findings regarding the initial attachment and default terminal configuration, while in subsequent

[1] Note that some technologies have their own mobility support at link layer but the connections at the terminal will not be able to survive an IP address change.

deliverables we will provide analysis on the differences while performing vertical and horizontal handover procedures.

3.1.1.1 Experimental setup

Table 3.1: Main characteristics of the analysed smartphones

	LG Nexus 4 E960	**iPhone 3GS**	**HTC 8S**
OS Version	Android 4.2.2	iOS 6.0	Windows Phone 8.0
Chipset	Qualcomm APQ8064 Snapdragon (S4 Pro)	Samsung APL0298C05	Qualcomm Snapdragon S4 MSM8227
CPU	Quad-core 1.5 GHz Krait	600 MHz Cortex-A8	Dual-core 1GHz Krait
RAM	2GB	256MB	512MB
WLAN	Atheros WCN3660 Murata SS2908001 Module (Qualcomm Prima)	Broadcom BCM4325	Atheros

This section describes the characteristics of the mobile terminals under test and also provides an overview of the different experiments performed while analysing the default Connection Manager of the terminal under analysis. Therefore, we study this element in several devices running the most representative OS – iOS, Windows Phone 8 and Android – in order to establish a comparison among them, and all the terminals are updated to their latest available version at the moment (subject to the terminal limitations). The characteristics of the smartphones in our experiments are collected in Table 3.1. For the sake of fairness we have avoided performing software modifications to the terminals (e.g., jailbreaking the iPhone) except for gaining root access permission in the Android device.

We intend to characterise the Connection Manager on the three devices under study, identifying the main strengths and weaknesses on the handling of network connectivity and comparing the behaviour of the three devices under different study cases. The set of experiments is composed of the following variations:

1. **Analysis of the default initial attachment procedure to an IEEE 802.11 network**. The goal of this experiment is twofold: *i)* to find out how the attachment to a WLAN is carried out by every device, analysing the differences among them, if any, and *ii)* study the network selection algorithm and the criteria that allow choosing to connect to a given access point when there are several wireless networks available.

2. **Analysis of the initial configuration of the protocol stack**. Once the mobile terminal has gained access to the Internet, we analyse the main steps and the protocols it uses to complete its networking stack configuration. We analyse this procedure both through the WLAN and the cellular interfaces. Through this experiment we intend to understand how the mobile terminals configure their interfaces and how they connect to the different services. Note that, if this procedure takes place entirely whenever there is a change in the point of attachment to the network, and not only as an initial configuration, it may enable potential optimisations for the handover process.

3. **Horizontal handover**. We examine the handover procedure between two IEEE 802.11 APs. We play with different network parameters to have a wide view of the performance for the different mobile terminals. Specifically, the current AP and the target one may have the same or different Extended Service Set ID (ESSID), operate in the same or different channels and

manage the same or different IP subnet, in which case the handover would imply also a L3 reconfiguration. This set of experiments allows us to know whether there are any dominant factors when the mobile device changes its point of attachment to the network and to what extent the different changes impact the configuration and connectivity management.

4. **Vertical handover**. We evaluate the handover procedure when it involves a change in the access technology. This test makes possible to know how the mobile devices under study handle the inter-technology handover, whether they can keep both technologies operative simultaneously and whether they handle the survival of ongoing connections. Characterising the inter-technology handover is essential to design potential optimisations and flow mobility solutions.

In addition to the mobile terminals under study, we deploy two IEEE 802.11 APs that provide Internet access while being under our complete control to keep track of the behaviour of the terminals attached to them and being able to modify network parameters, such as the ESSID, the wireless channel in which the AP operates and the IP subnet managed by the access router. The same APs were used for the vertical handover analysis.

3.1.1.2 Initial network attachment

WLAN Initial attachment

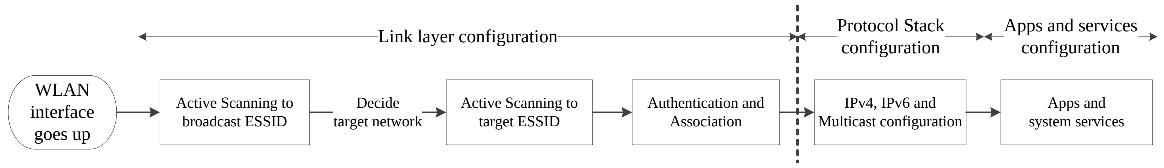

Link layer configuration
Protocol Stack configuration
Apps and services configuration
WLAN interface goes up
Active Scanning to broadcast ESSID
Decide target network
Active Scanning to target ESSID
Authentication and Association
IPv4, IPv6 and Multicast configuration
Apps and system services

Figure 3.1: Initial attachment to the WLAN.

Through this section we will explore the initial attachment to the IEEE 802.11 procedure followed by the three different terminals being analysed. Understanding this process allows to note the differences among the terminals and the order of operations performed to configure the complete protocol stack.

In order to test the default attachment to a WLAN, we deploy a wireless access point and run several repetition of the experiment for each smartphone. In each experiment, which lasts for 60 seconds, we monitor the traffic by means of a network analyser[2]. The monitor interface is located close to the access point, so we can likely capture all the frames involved in every exchange. Initially, the WLAN interface of the mobile terminal is down, so we do not miss any packet or reach misleading results because of having the device already connected to another network. We start our experiment by bringing up the interface and checking that the device actually connects to the AP under our control. This experiment describes the case in which the terminal finds an already known network and connects successfully. Note that to connect to a network for the first time the user must identify and select manually the network to connect to.

Fig. 3.1 presents a diagram synthesising our findings. In the following, each of the steps occurring in the initial attachment is explained in detail, highlighting the differences among terminals:

[2] http://www.wireshark.org/

Active scanning to broadcast address: When the WLAN interface of a mobile terminal goes up, it detects all the surrounding wireless networks available by initiating an active scanning procedure. The terminal sends Probe Request frames to a wildcard ESSID in every channel sequentially. Neighbouring APs will receive this Probe Request and answer with a Probe Response message, indicating their capabilities and providing synchronisation information. The three systems under study perform this active scanning phase in a different way. By monitoring the traffic in the different channels we measure the interval between consecutive Probe Requests sent in every channel. The results show that the Android phone scans every 10 seconds in every channel, sending a number of consecutive Probe Requests inter-spaced approximately 15 ms. However, the Windows phone is faster, elapsing only 6 seconds between consecutive scans in the same channel. The time lapse between consecutive Probe Request is similar to the Android one, which means that the Windows phone sends a smaller amount of frames. Lastly, the iPhone terminal presents an interval of approximately 9 seconds between the scan in every channel and it is the one that sends the smaller amount of frames, being the delay between consecutive Probe Requests around 20 ms. It is interesting to evaluate the behaviour of the three terminals in the channels 12, 13 and 14, which are not allowed in Europe, where we are based. The Android and Windows phones do not list the networks operating in those frequencies as available, but the iPhone does. However, the three terminals send Probe Request frames in every channel, including those out of the allowed frequency band.

The scanning policy followed by a terminal has several potential effects on the terminal. First, the less the number of messages sent the less the battery consumed. Second, the time required to finish this stage is longer as a higher number of probes are sent before going to the next step of the attachment, hence the attachment to a WLAN AP is delayed by this necessary procedure and this impacts the handover latency. Finally, a terminal can only obtain information regarding the signal level from the AP by the frames received, hence receiving more responses before deciding the target point of attachment can be beneficial.

Table 3.2: DNS queried for initial configuration of services on WLAN interface start-up

Service	Android	iOS	WP8
Initial connection to central servers	clients3.google.com (IPv4 and IPv6)	www.apple.com; e3191.c.akamaiedge.net	N/A
Network status indication	N/A	N/A	www.msftncsi.com (IPv4 and IPv6)

push notification service	mtalk.google.com (IPv4 and IPv6)	init-p01st.push.apple.com;
		a1441.g.akamai.net;
		36-courier.push.apple.com;
		36.courier-push-apple.com.akadns.net
			N/A

NTP	N/A	time.apple.com	N/A
Software updates	N/A	mesu.apple.com; a97.gi3.akamai.net	N/A

Target Network Decision: This step corresponds to the actual decision on the AP to connect to. Before performing this analysis it seems reasonable to guess that the terminals will perform some kind of fancy algorithm considering, for instance the signal level received from the different APs. We have discovered through our extensive tests that all three terminals use a simple rule to decide where to connect to. If the AP immediately previously used is available, the terminals will connect to it, no matter its signal level. In case the last visited AP is not available, the terminal connects to the one previous to the last one, and so on. In the case of secured and open networks are available, only iPhone terminal shows a preference for secured networks if the immediately previous connection is not possible. It is worth to note that the Android phone under study implements the last OS version as of the moment of writing this article (4.2.2 or Jelly Bean), which includes an option in the WiFi settings to connect to a different WLAN or to the cellular network if the signal is weak. However, even if this option is enabled, the terminal follows the same approach and disregards signal strength information to connect to an AP.

Active scanning to selected ESSID: In this step, the terminal addresses a Probe request message directly to the AP selected previously and indicates the target ESSID in the correspondent field in the frame. We argue that this procedure is done to ensure that the AP previously selected is available before performing the authentication and association processes, which depending on the security properties might be very time consuming.

Authentication and Association: These are the last two steps before being attached to a WLAN AP. Both of them are standard procedures and are equally executed in the three terminals. Security procedures are out of scope as the main focus in this work is on the Connection Manager.

IPv4, IPv6 and Multicast configuration: We are grouping these configuration steps because they are basically the same for the three systems being evaluated. The three of them configure an IPv4 address by means of DHCP. Then, they configure a local and a global IPv6 addresses. Finally, they configure the multicast membership and that is where they behave slightly different. For instance, Windows Phone makes use of LLMNR (Link Local Multicast Name Resolution) protocol in addition to IGMP and MLDv2 which are used in the iPhone signalling. The Android phone just makes use of MLDv2 messages.

Higher layer configuration: As soon as the mobile terminals gain Internet connectivity, they try to reach the central servers of their correspondent manufacturers to update their location, configure some services - time synchronisation or push notification service – and re-initiate some connections – as GTalk, in the case of Android. IP addresses on the server side are expected to change, so the terminal connects by hostname, issuing DNS queries rather than by IP address. Commonly, servers implement a load balancing scheme, so it is possible that the same query returns a different IP address for the same host name. For that reason, to identify the connections we track an IP address block, instead of a specific name or address. In addition we observe that many providers and applications use Content Distribution Network (CDN) nodes to offer their services, which complicates the identification of the service as the connection is hidden by the CDN. In the case of the Windows Phone system, the latest versions of the Microsoft's OS, including the mobile version, implement Network Connectivity Indicator (NCSI) service, which issues a DNS query to establish a TCP connection intended to send an HTTP GET method and retrieve a light-weighted web page, which is mainly void. This procedure serves to check the Internet connection at the device and prompt the users with a login form to introduce their credentials, if required by the WLAN administrator. The API in Android includes a way to access the information on whether the terminal is connected or connecting to the Internet and through which interface and register to event notifications in case of network status changes. In addition the application developers have also the option to try and reach a website from their application when it starts running, because the connection manager can only detect the status of the interface, but not if there is actual connectivity. Table 3.2 collects the service connections that are preceded by a DNS query when the three different terminals attach to a WLAN.

Cellular initial attachment:
In this section we present a similar analysis for the initial attachment through the cellular data interface. Our access to the cellular connection is much more restricted than to the WiFi network, so our analysis is quite limited as well. We extract our conclusions by monitoring traffic directly at the mobile terminal. For that reason, we have information from the Android and iPhone devices only. In this case, the mobile terminal accesses the network by means of a point-to-point connection, so the configuration and signalling overhead is reduced to multicast configuration and ARP querying

Table 3.3: Information types

information type	information source	network element provider
location (GPS)	GPS framework	mobile
network element map (eNodeB, AP)	ANDSF 802.21 MIIS	ANDSF server centralised MIIS server
roaming partners	802.11u 802.21 MIIS	WiFi AP centralised MIIS server
WiFi authentication optimisation	802.11u 802.21a	WiFi AP 802.21 POA

for the other end of the point-to-point link.

3.1.2 Information gathering and its use for connectivity management

3.1.2.1 Access Selection

Most of the current mobile devices are able to use several radio accesses, it is up to the connection manager to choose which radio accesses should be used. When the connection manager is choosing another radio access rather than the current one, it may have many reasons for doing that. The usual reason is that the current radio access does not fit the requested characteristics links for the current communication (radio link layer: bad Signal to Interference plus Noise Ratio Signal to Interference plus Noise Ratio (SINR), too many errors (Bit Error Rate BER too high) ...). However issues may be related to upper layers like the transport layer: packet loss, congestion, unacceptable delays. Such conditions may also trigger a change of the radio access. In this case, the origin of the problem may be located in another part along the path between the MN and its CN, and not at the radio accesses. Thus, improving the communication will imply to choose the best path between the MN and the CN, the choice of the radio accesses being only a consequence of the path choice as the next element of the MN along the path.

Several sources of information may be used to give accurate data to the connection manager in order to improve the radio access selection. The information can come from protocols such as 802.11u, 802.21, ANDSF for the radio accesses information or ALTO for the path information or various statistics given by the system on the MN.

The information type from these protocols that the connection manager may use has been summarised in the Table 3.3.

3.1.2.2 Obtaining radio access information: Operator managed information service vs mobile self discovery

The protocols that will provide radio access information use two different approaches:

- The mobile asks to a server an up-to-date snapshot of the network (or part of the network). The server will manage the information about all radio accesses.
  - Advantages: A global view that allows to make choice on large area, for example, determining if there is a continuous WiFi coverage for the foreseen MN movement.
  - Drawbacks: the multi operator case is not well taken into account, in this case this is not a roaming issue, but in the case of people that are working both at an enterprise and at home, they will have both the enterprise's operator and their own operator, which may have no roaming agreement between them. Furthermore the WiFi enterprise network

may not depend of an operator, and hence will not be recorded in server's information base. This information base may not be up-to-date, some proposed APs may be not available at this time and other available APs being not yet recorded on it.

- The mobile discovers itself its neighbourhood radio environment by scanning it from time to time.
  - Advantages: Autonomy, up-to-date information.
  - Drawbacks: It may be energy consuming to have to periodically scan the potential WiFi APs that the UE is able to connect. It has only a local view, if the WiFi coverage is a spotted one, the MN will do handover from WiFi to 3GPP network and back continuously. Association/connection process is costly in time, especially the authentication step if the authentication keys should be stored for each APs. Authentication protocols like Extensible Authentication Protocol - Subscriber Identity Module (EAP-SIM), Extensible Authentication Protocol - Authentication and Key Agreement (EAP-AKA) that use the credentials of the Subscriber Identity Module (SIM) card are a good solution to avoid having to store the needed information to access to each AP, making the association step to WiFi APs as easy as it is for 3GPP networks.

Operator managed information service:

An operator owns a lot of network elements, that will impact the management of the information base of the ANDSF server (or 802.21 MIIS server). In medium-sized countries (like Spain or Germany) an operator network will manage tens of thousand of macro base stations. With small cells (it depends on required deployment) the number may rise up to hundreds of thousands. For WiFi several millions of APs has to be managed. (Operator owned WiFi access points, and WiFi set-top boxes) For example, the number of events that should imply a modification of base stations that should be brought to the database each day (mainly for maintenance reason) are around tens per day (in our medium-size country case). For the WiFi case, keeping the same ratio, the number of modifications will rise to thousands. Even worse, for the WiFi access point that are not directly managed by the operator (such set-top boxes, the occurrence of such event will be more significant (WiFi set-top boxes are switched off from time to time without direct mechanism to inform the operator). A trade-off should be found, between the accuracy of the information given by the server and the frequency of updating the information on the mobile client side.

Protocols such as ANDSF or 802.21 do not define what should be the area of interest of the mobile, for searching the best radio access. Getting the information about the radio accesses over a large area, has several issues. First, if the data transferred by the MN is small, the cost of transmitted control data will be too hight compared to the transmitted user data. Furthermore, the stored data will be quickly outdated. Second, getting information for small areas is useless in case the mobile is on the move at great speed. The requests to the server database will happen too often and sometimes too late.

Mobile self discovery:

802.11u allows the MN to get information about the different APs, including roaming agreements or security requirements, before the association step.

An advantage of 802.11u is the ability to get some QoS measurements that will be used to check if the connection to the AP will improve the data transmission.

However today, the 802.11u is not widely deployed on APs, the decision algorithm should not assume that every APs will implement it. When the AP does not implement 802.11u, previous methods using 802.21 MIIS or ANDSF should be used as a backup.

It seems that a solution that will use both 802.11u and either ANDSF or 802.21 will help to get a global view that will give some guidelines about the use of available networks and on the other hand an up-to-date information about the existing radio accesses and how to connect to them as fast as possible.

3.1.2.3 Path information

As seen before, the ALTO protocol is used to determine among a set of paths between the MN and a peer, the best one depending on a given criteria. An ALTO request is usually done by an ALTO client that may be a user application, such as a video player.

The first network element that will follow the MN in a path is the radio access network element (AP or eNB, ...), if the communication bottleneck is on the radio access the ALTO server may give a response that will prioritise a path that goes through an overloaded radio access, and then inducing a wrong choice. A combination of information from the ALTO server and the connection manager will help to get the best of the two worlds.

The idea is to define a cost attribute related to the access point performance (cost_access for example) that will allow to order the path depending on the performance of the radio access that they will use and then use the multicost features [68] that allow to build a unique request with several different cost attributes. Then implement an ALTO proxy that will let all the request go through it except those with the cost_access cost attribute that will be computed locally, using the information given by the connection manager about the relative performance of all radio accesses present in the paths.

3.1.2.4 Information Gathering

All the standards that manage an information base and that have been used in this section (802.21 MIIS, ALTO and ANDSF) do not describe in detail how the information base is created or modified. Neither they specify the frequency of update, even if most of them specify that the service is not a real-time one. The client should assume that the information got in the response may not be valid anymore. The changes that are brought to the information server should be easily tracked by the client. In CROWD, some discovery mechanisms will be proposed to fill the information base.

3.1.3 ANDSF

The ANDSF is a function of the 3GPP EPC that has been introduced in the Release 8. The ANDSF is a policy server communicating with the UE over the S14 interface, which is essentially a declination of an OMA-Device Management (DM) Management Object (MO) specific to ANDSF.

The purpose of the ANDSF is to assist the UE to discover non-3GPP access networks, such as WiFi, that can be used for data communications and to provide the UE with rules policing the connection to these networks. The ANDSF can provide the following information to a UE, based on operator configuration:

- *Discovery information*: a list of networks that may be available in the vicinity of the UE and information assisting the UE to expedite the connection to these networks.
- *Inter-System Mobility Policy (ISMP)*: network selections rules for a UE with no more than one active access network connection (e.g., either LTE or WiFi);

- *Inter-System Routing Policy (ISRP)*: routing policies for a UE with potentially more than one active access network connectivity (e.g., both LTE and WiFi). Basically, the ISRP indicates the most appropriate interface per IP flow.

When requesting the ANDSF server, the UE indicates its capabilities in term of multiple interfaces management. Being aware of UE capabilities, the ANDSF server can adjust its behaviour in terms of policies (ISMP or ISRP) to be sent to the UE. Then, the UE has different ways to take benefit from multiple interfaces, it may either map certain IP flows to given interfaces or make decision to leverage on IP flow mobility; consisting in moving an IP flow from one interface to another, independently from other ongoing IP flows.

In 3GPP Release 10, the ANDSF framework has been enhanced with the introduction of ISRP, i.e. network selection rules for a UE with potentially more than one active access network connection (e.g., both LTE and WiFi). More than one active access network connection can be provided to the UE in different manners as specified in 3GPP: as illustrated in Fig. 3.2, a UE may employ:

- *Multiple-access PDN connectivity (MAPCON)*: the UE is able to simultaneously maintain different active PDN connections[3] through different access networks (i.e. on both a 3GPP access and a non-3GPP access).

- *IP Flow Mobility (IFOM)*: the UE is able to route different IP flows to the same PDN connection through different access networks [69]. This allows to choose on a per IP flow basis on which access each flow should be routed and to move them seamlessly between accesses.

- *NSWO*: a UE is able to route specific IP flows via a WLAN access without traversing the EPC (hence without support for session continuity).

According to 3GPP, the UE may employ IFOM, MAPCON and NSWO operator policy and user preferences.

3GPP standard only specifies the UE-ANDSF interface and the model object (i.e., policies format); however it does not give hints to compute policies. Policies can be static but, ideally, the ANDSF should take into account the network environment, so building policies relying on back-end functions (Fig. 3.3) providing this knowledge. Examples of information that may be stored at the ANDSF are given in the following:

- *WiFi cartography*: this function maps the user-location (e.g., 3GPP cell-ID) with WiFi access networks.

- *Network monitoring*: this function gives information regarding the quality, e.g., network load, on the WiFi access networks. Network monitoring could be implemented through an 802.11u/ANQP server for Hot-spots 2.0.

- *User profile*: this database stores information related to the user (subscription rights,..).

3.1.4 Energy efficiency optimisation through access selection/vertical HO in presence of dense/heterogeneous scenarios and relay nodes

Due to the increasing processing capacity and the multi-purpose nature of the usage of the modern mobile devices, the average battery lifetime of such terminals is progressively shortening. Therefore, reducing the power they consume is of absolute importance.

[3] an PDN connection corresponds to the association between a UE represented by one IPv4 address and/or one IPv6 prefix and a PDN, identified by an APN. From the UE side, a PDN connection corresponds to a virtual network interface with an IP address/prefix that belongs to the PDN.

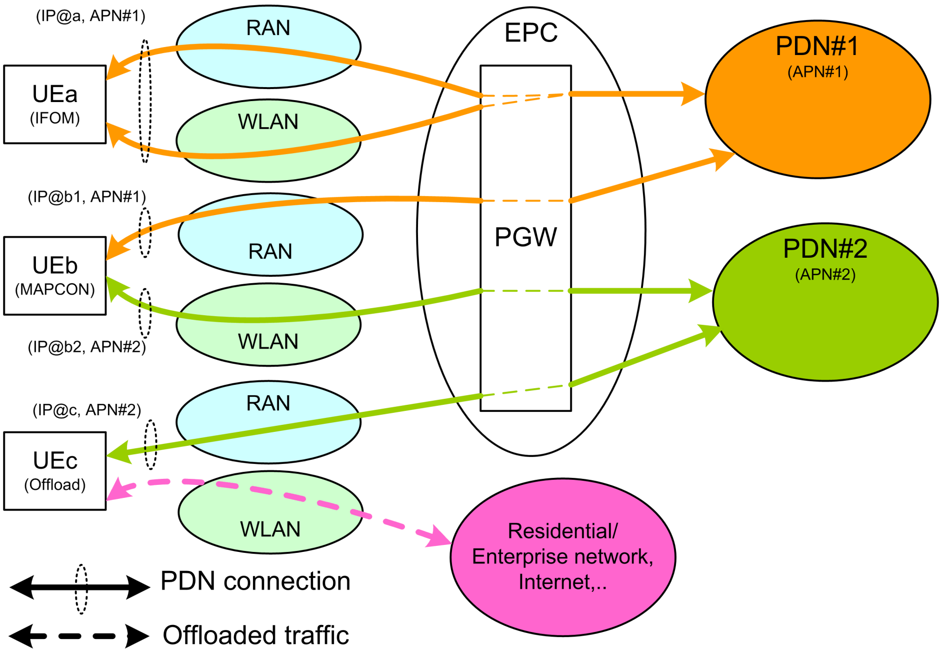


Figure 3.2: IFOM, MAPCON and NSWO.

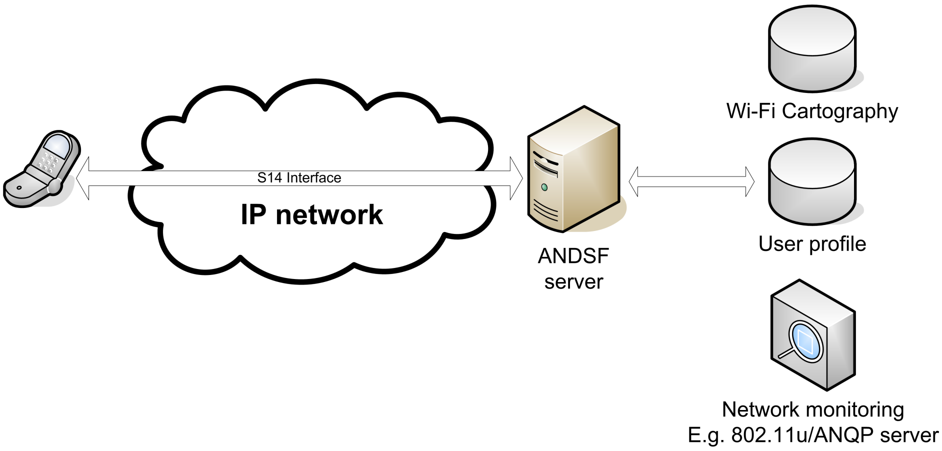


Figure 3.3: ANDSF network support.

One of the possible solutions to this problem is to improve the energy efficiency of the multiple radio interfaces (i.e., LTE, WiFi,...) present in the majority of the current devices. From the point of view of the operators, improving the efficiency of the power consumed by the terminals is of extreme importance too. Among all the possible ways in which the energy efficiency could be improved, a network-driven access selection mechanism, designed conveniently for this case, is the solution that leverages to his best the architecture we introduced in the CROWD project. The presence of local and regional controllers allows the operator to have a complete picture of the state of their heterogeneous access network at each moment and to find at the system level the optimal trade-off among power consumed and throughput achieved.

In the following, we first show the scenario where the access selection mechanism is likely to be used, then we show what is missing in the state of the art solutions in such a case, we sketch the mentioned access selection mechanism and, finally, we present the future research directions we intend to explore.

Scenario of application. Based on the actual evolution of the cellular access network, alongside the normal coverage of macrocells base stations, we envision a scenario with operators deploying short-range, low-cost access points, referred to as femtocells. Although femtocells typically support few costumers [70], they embody the functionality of a regular base station which operates in the mobile operator's licensed band. Even if it requires a careful interference management, the deployment of femtocells improves, from the mobile operator perspective, the licensed spectrum spatial reuse and decongests nearby macrocell base stations.

Furthermore, we also envision the possibility that operators offloads their cellular access network exploiting their wired access network by means of customer-deployed WiFi hotspots. In order to cope with costumers experiencing bad quality channels, we finally envision operators allowing devices to form WiFi-Direct groups so that the traffic of the whole group is sent/received to/from the cellular access network just by/to the Group Owner (GO). Adding such a packet relay option is a key factor to enhance the energy efficiency of the whole system. Indeed, if the access selection is performed correctly, the cellular base stations are connected to a set of users whose channel quality allows to use highly efficient Modulation and Coding Schemes (MCSs). In this sense, [48] is showing first promising results. Fig. 3.4 summarises the scenario of application.

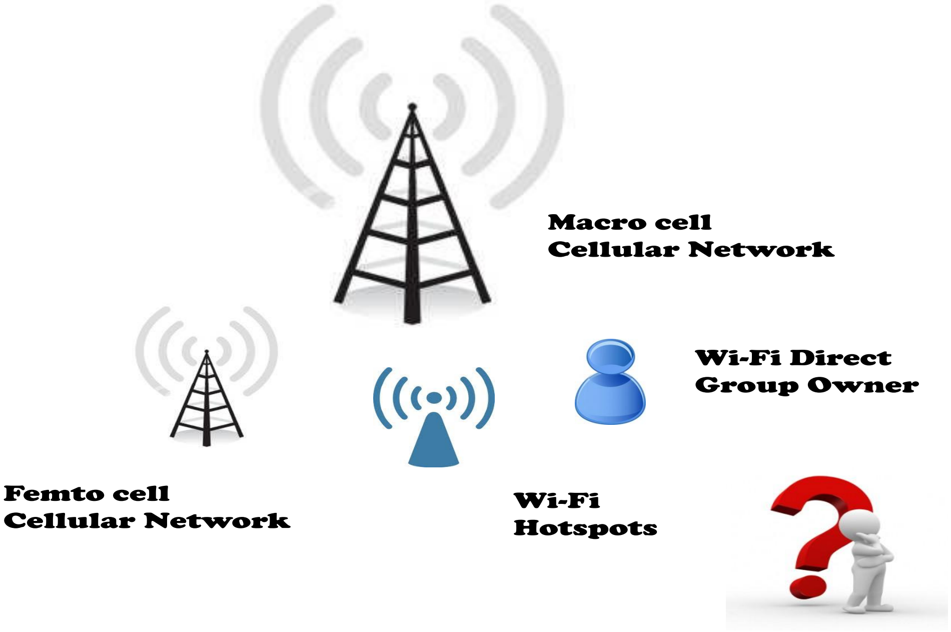


Figure 3.4: Application scenario of the proposed access selection mechanism.

Problems in available State of the Art Solutions. Achieving energy efficiency by means of access selection mechanisms in heterogeneous environment has already been proposed. By the way, all the solutions that tried to face such a problem suffers of severe limitations. The majority of them propose a user-driven access selection mechanism. In those cases the users choose the best suited point of access among the ones available without having a complete knowledge of what is happening in the system. The users typically rank the APs by uplink and downlink power required to serve the required loads. It is easy to see that, when they are experiencing similar conditions, the users rank the available point of access in the same way, and that one of the available technologies would be rapidly overloaded. Furthermore, even if the users could know what is the actual load of the point of access they actually chose and react accordingly, allowing to take the decisions to devices having a partial visibility of the state of the system would create a ping-pong effect of the users among technologies and continue vertical handovers that deteriorates the overall QoS experienced.

Even if the access selection mechanism is network-driven, the state of the art solutions appear to be extremely simplified with respect to the complexity of the problem. For instance, the basic assumption underlying each piece of work of this kind is that the load of any of the devices accessing the network is known and fixed in time. The access selection mechanism finds just the best assignment users/PoA such that the load is served while the power consumed is minimised. Moreover, ad-hoc modes are rarely taken into account when network-driven access selection mechanisms are designed, and, if present, they are just used to enlarge the coverage area of the access network [59].

A New Access selection/VHO mechanism. In CROWD, we propose a new network-driven access selection mechanism that leverages the architecture introduced by the CROWD project. Thanks to the presence of local and regional controllers, operators have the complete view of their access network, even if it spans different technologies. Among all the pieces of information the operator could exploit, we assume that it has the complete knowledge of the bit-efficiency each device would have if connected to a PoA, either if it is a cellular base station or a WiFi-Direct GO. Such assumption is not restrictive, indeed those technologies already provide this kind of information by means of control message feedbacks. Finally, even if the access selection mechanism is presented by means of a static configuration, we assume that the access selection mechanism is repeated over time in order to track the dynamics of the system.

Thanks to those assumptions, it is possible to restrict the set of possible assignment devices/PoA to a huge but finite set and to compute for each of them the downlink and the uplink throughput the devices would experience. In the following we present such downlink throughput for a specific device k for each of the possible ways it can access the network, i.e. *i)* the device connects to an eNB and does not retransmit the traffic from other devices in a WiFi-Direct group, *ii)* the device connects to an eNB and it becomes the GO of a WiFi-Direct group, *iii)* the device receives his downlink traffic from the GO of the WiFi-Direct group it belongs to and, finally, *iv)* the device connects to a WiFi hotspot. Same arguments could be used to compute the uplink throughput.

- *Device connected to an eNB, no WiFi group.* Let us assume that the device k has been scheduled at the LTE eNBs t with other N_t^{LTE} devices and that b_i, with $i \in 1, ..., N_t^{LTE}$, is the bit efficiency they could achieve through the MCS they use. Let us also imagine that b_k is the bit efficiency of device k and that the eNB node is using a throughput fair scheduler. Then, it can be easily proved that the downlink throughput S_k^{LTE} is (in saturation conditions):

$$S_k^{LTE} = \frac{N_{RB} \cdot N_{SC} \cdot N_{SY}}{T \cdot (\frac{1}{b_1} + ... + \frac{1}{b_{N_t^{LTE}}} + \frac{1}{b_k})}, \tag{3.1}$$

  where N_{RB} are the resource blocks for time slot the eNBs transmits, N_{SC} are the sub-carriers per resource block, N_{SY} is the number of OFDM symbols per sub-carrier in each resource block and T is the time slot duration. In case the eNB uses a different scheduler, S_k^{LTE} can be adapted accordingly.

- *WiFi GO connected to an eNB.* In this case, not all the downlink traffic the GO receives from the eNB is actually directed to it. Indeed, a fraction of it is relied by the GO into the WiFi Direct group. The GO acts as a normal device in the WiFi cluster, then it contends the medium with the other stations who are trying to send their uplink tp the GO. If S_k^{LTE} is the throughput the device receives from the LTE network, the fraction of throughput $S_k^{WiFi-GO} \leq S_k^{LTE}$ belonging to other devices can be gathered by [71]. Then, the actual downlink throughput S_k of device k is:

$$S_k = S_k^{LTE} - S_k^{WiFi-GO}. \tag{3.2}$$

- *Node attached to a WiFi Group.* A device attached to the WiFi Group run by GO t receives as downlink throughput a fraction of the traffic transmitted by t into the cluster. If we assume that a fair scheduler is again used, and if we denote the number of devices belonging to the cluster with $N_t^{WiFi-GO}$ users, then the downlink throughput of mobile device k is:

$$S_k = \frac{S_t^{WiFi-GO}}{N_t^{WiFi-GO}}. \tag{3.3}$$

- *Node attached to a WiFi Hotspot.* In this case, the device shares the medium with all the other devices attached to the same WiFi hotspot. The downlink throughput S_k^{WIFI} can be again gathered through [71].

Given the throughput computed as above, an estimation of the energy consumed by the network interfaces is possible thanks to available energy models present in the state of the art ([72] and [73] respectively for WiFi and LTE). Note that for the WiFi GO, the energy consumption of the LTE interface has to take into account the whole traffic received and sent toward the eNB the device it is connected to.

For each of the possible assignment of devices to access points, it is therefore possible to compute the energy efficiency of the system and select the best one among all the available ones.

Ongoing research directions. When applied to dense scenarios, with high number of heterogeneous access points and devices, the computation complexity of the proposed access selection mechanism is barely feasible. Therefore, the main focus of the ongoing research is to find a heuristic so to efficiently approximate its optimal solution. Another crucial factor of the proposed access selection mechanism is the reliability of the power consumption model used. In order to improve the gain effectively achieved by means of the access selection, another research direction we pursue is the design of a new energy model that takes into account also the power saving modes that the devices normally employ nowadays. Finally, we are also evaluating the possibility of jointly use a power control mechanism together with the access selection mechanism we propose.

3.1.5 Assessment of the Energy consumption savings with 3G offload

The design of an efficient mechanism to have the cellular and the WLAN connections sharing the traffic load benefits both the network operators and the final users. By offloading the traffic from the cellular network to a WLAN, a network operator can reduce the load on its network and reuse the freed resources for users that cannot handoff their traffic. On the other side, offloading and flow mobility in general can also improve the user experience, although usually this part is overlooked since the main driver of it is to alleviate the operator's problems. For example, the mobile users can experience a better quality due to the higher bandwidth that a WLAN can offer compared to 3G, or even use both interfaces at the same time in order to achieve a higher available bandwidth.

Although 3G offload can be used to just refer to the simple handover of all IP flows from one interface, e.g., 3G to a secondary one, e.g., WiFi, the opportunities that this technology enables are maximised when a fine-grained flow selection is allowed. For example, an operator might prefer not to offload Voice over IP (VoIP) flows, due to the inherited difficulties in providing QoS guarantees on an unmanaged WiFi access, while video traffic might be always offloaded to a technology providing higher bandwidth.

3.1.5.1 Energy consumption assessment

Enabling flow mobility implies benefits both for the operator, that can save resources in the radio access, and also for the user, that is able to take advantage of an increased available bandwidth. However, in order to fully assess the suitability of this mechanism, it is essential to evaluate it in terms of complexity and of another factor which is usually forgotten, its energy consumption. Energy consumption is specially critical for mobile devices and smartphones, which already suffer from battery-drain issues due to continuous and exhaustive use along the day. Despite the fact that 3G connection is heavily consuming the battery of the device, it is generally configured to be the default access connection and is almost always on. Therefore, in order to implement a flow mobility solution, it is reasonable to assume that in addition to this intensive usage of the 3G interface, we will need to add the energy consumption corresponding to additional network interfaces. Nevertheless, our experimental results show that the energy consumed by the 3G interface is higher than the one consumed by the WLAN interface, so offloading the 3G connection helps reducing this consumption. Through this section we provide experimental results supporting the claim that flow mobility can also be beneficial for the user in terms of achievable energy consumption savings.

Modern terminals such as Android or iPhone smartphones do not allow by default the simultaneous use of 3G and WiFi interfaces. To overcome this issue and perform an experimental assessment of the energy cost derived from enabling IP flow mobility (i.e., use of multiple network interfaces at the same time) we perform real power consumption measurements on a multi-mode device, equipped with a WLAN IEEE 802.11a/b/g and a 3G Universal Mobile Telecommunication System (UMTS) (High Speed Packet Access (HSPA) capable) interface. In order to be able to control as much as possible the used devices, capture traffic sent and received at the network interfaces, as well as closely monitor the device, we decided to use a small residential router, Asus WL-500GP v1.0, based on a Linux firmware. The measurements provided through this procedure are later validated by the analysis of the battery lifetime of an android smartphone while using 3G and WLAN interfaces separately. Finally we derive our main conclusions through a synthetic use case that allows us to provide quantitative gains on the percentage of battery spent through the use of the proposed flow mobility mechanism.

Assessing the power consumption of the joint operation of 3G and WiFi

The following section is devoted to perform the experimental assessment of the energy consumption associated to our flow mobility solution. To measure the power consumption of each technology we have chosen a small residential router: the Asus WL-500g Premium. This router is equipped with a 266 MHz processor, an IEEE 802.11b/g WLAN interface and an IEEE 802.3 Ethernet interface connected to a Virtual LAN (VLAN) capable 5-port switch. This version of the router has a mini-PCI slot that allows changing the original wireless card. We replaced the original Broadcom card by an Atheros based 802.11a/b/g (Alfa Networks AWPCI085S) one, which is supported by the Madwifi[4] driver. In order to mitigate as much as possible the impact of collisions and interference in the power consumption measurements, we avoided the 2.4GHz band (IEEE 802.11b/g) – which is very crowded in our lab, as reported in [74] – and configured the WLAN interface in 802.11a mode.

We replaced the original firmware of the router by installing a lightweight Linux-based version, which gives us more flexibility in the configuration. We choose the distribution Kamikaze 8.09.2 of OpenWRT[5] with a Linux-2.6 kernel and this allows the support of a 3G USB modem. In our tests, we used a Huawei E160 HSPA USB stick[6].

Power consumption was measured using a PCE-PA 6000 power analyser[7]. Power measurements were carried out using a PCE-PA-ADP current adapter where the power supply of the router was

[4] http://www.madwifi.org/
[5] http://www.openwrt.org/
[6] http://www.huawei.com/mobileweb/en/products/view.do?id=1960
[7] http://www.industrial-needs.com/technical-data/power-analyser-PCE-PA-6000.htm

Table 3.4: Power consumption results

3G ON		WLAN ON	
WLAN OFF	1.80 ± 0.10 W	3G OFF	1.03 ± 0.08 W
WLAN IDLE	1.86 ± 0.08 W	3G IDLE	1.21 ± 0.16 W
WLAN ON	2.16 ± 0.13 W	3G ON	2.16 ± 0.13 W

plugged in. Measurement data was transferred from the power analyser to a computer via an RS-232 interface for its processing.

Using this setup we performed the measurements described next. We first calibrated the power analyser by measuring the consumption when both the WLAN and 3G interfaces are switched off. All reported results are relative to this level. For the actual measurements, we are interested in the power consumption when the network interfaces are in the following states:

- OFF: the interface is switched off.
- IDLE: the interface is on but it does not send or receive any data traffic. For the case of WLAN, this means that the card is associated to an access point (so the card is receiving beacon frames) without sending or receiving any user data traffic. For the case of 3G, this means that the interface is up, a Packet Data Protocol (PDP) context has been activated and a Point-to-Point Protocol (PPP) interface has been set up, but no data is exchanged.
- ON: the interface is on and engaged in a data traffic exchange. In our tests, this means that a file is downloaded from a server using HTTP. By using TCP, the card is receiving at the maximum available rate, and traffic is sent in both directions (downlink: mostly data segments, uplink: mostly TCP acknowledgments).

We measured the power consumption for the different states of the WLAN and 3G interfaces. Table 3.4 shows the mean and 95% confidence interval of the results extracted from five 300-second experiments. We focus on the scenarios where at least one of the interfaces is actively sending or receiving traffic, as those are the cases in which it is important to evaluate the energy cost associated with having a second active interface. This second interface may be either receiving or sending traffic or just idle, ready to operate. Results show that the 3G interface consumes more power than the WLAN interface, but the difference between using only the 3G interface and using simultaneously the 3G and the WLAN interfaces is only of 20%. Note that this additional cost is only incurred when both interfaces are actively engaged in a data transfer, and that by using them simultaneously, the time required to send a given amount of data via WLAN would be shorter – since the throughput obtained via a WLAN network is typically higher than the one that can be obtained via a 3G network – and this would also contribute to a lower power consumption. The extra power consumption caused by activating the WLAN interface (IDLE state) is just around 3%, which besides would only be needed when the mobile is sending or receiving traffic, as it is then when the network operator and the user may benefit from offloading traffic from the 3G infrastructure to a WLAN hotspot, if available. It is important to highlight that the actual values of energy consumption of the device are not directly comparable with the results we would obtain with a smartphone, since the level of integration provided in such a platform is much higher, allowing further improvements in the energy consumed by the device. Due to this, and to be able to compare, we only focus on the relative difference between the 3G and WiFi consumption profiles which, as we will see in the next section, follow the same trend in both device families.

Energy consumption profiles of 3G and WiFi in an Android device

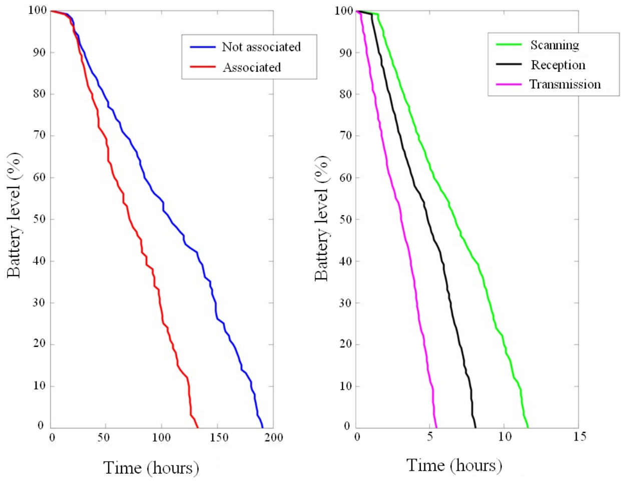


Figure 3.5: Battery drainage for the different WiFi states.

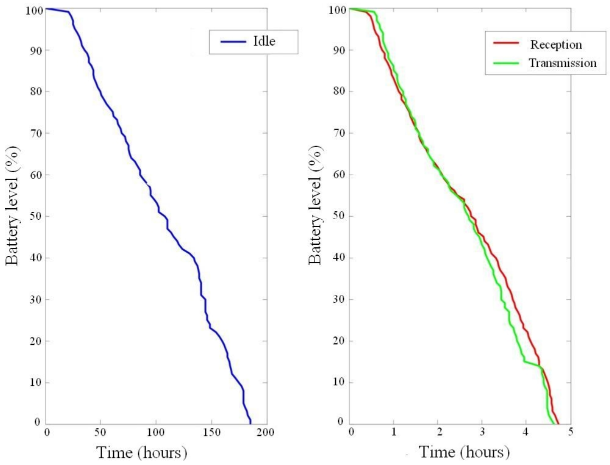


Figure 3.6: Battery drainage for the different 3G states.

In order to confirm the results obtained in previous section, we measured the battery duration of an HTC Legend device. The operating system of the device under test is Android 2.1 (Eclair). Apart from the device, we used a desktop to monitor and configure the parameters of interest in the mobile terminal. We also configured the WLAN interface of this desktop as Access Point to which the mobile terminal would associate. To measure the energy consumption we developed an application running in the background to monitor the battery continuously, keeping track of the voltage level in order to compute the power consumed. As this application is a service running in the background, it consumes negligible CPU resources, minimising the impact on the energy consumption measurements. In addition, no interaction with the user (or the tester in this case) is required, as all the information is saved to a text file. All the measurements have been performed keeping one element active and the rest inactive, in order to isolate the contribution of each individual element to the total power consumption of the device.

Figs. 3.5 and 3.6 show the battery drainage curves for WiFi and 3G respectively. Fig. 3.5 presents results for each of the possible states of the WiFi interface, while Fig. 3.6 considers only the states we can control on the 3G interface, namely: transmission, reception and disconnection states. The results from these measurements show that the battery life of the device is much shorter when the

3G interface is on than when the WLAN interface is at its maximum battery consumption task, which is transmitting packets. The battery of the mobile terminal can last 200 hours when the 3G interface is inactive, against a duration of less than 5 hours when there is incoming or outgoing traffic. However, in the case of the WLAN interface, the difference between the transmission and reception states are much more evident than those of the 3G interface, shrinking from 8 hours to 5 hours, respectively.

In addition, through the analysis of the slopes of each curve, we can obtain the relative difference between the cases of a single active interface (3G) and the case when both interfaces (3G+WiFi) are simultaneously used. Supposing a transmission and reception cycle of a 50%, the overall difference between both cases is approximately 15%. As expected, the relative difference between both cases is lower for the smartphone device, we argue that this difference is due to a higher integration of the components in the smartphone compared to the router.

Finally, and in order to conclude this analysis, let us make a synthetic example of the energy saving that such an approach would provide to the end user. First let us consider some assumptions, for the sake of the simplicity of this analysis, which aims at assessing if a typical mobile user could afford the additional power consumption introduced by the use of flow mobility extensions. Several studies, such as [75], point out that users of smart hand-held devices download an average of 20 MBytes per day via 3G. Considering Fig. 3.6, and an average 3G speed of 1 Mbps, the download would take 160 seconds and consume around an 0.8% of the battery.[8] In case a flow mobility solution was deployed and the terminal was able to use WiFi to download the same amount of information, it would use the WiFi interface for approximately 6.4 seconds (assuming IEEE 802.11a extended rates and a real throughput of approximately 25 Mbps). During this time, the terminal will use a 15-20% more energy compared to the case of using only 3G, but the overall time would be highly reduced. This implies that the terminal would have spent less than a 0.1% of the battery downloading the file. This simple analysis does not aim at providing rigorous and precise figures, but just at roughly assessing if a flow mobility solution is affordable from the perspective of power consumption. Based on the obtained results, we can conclude that selectively using more than one network interface results in an reduced energy consumption.

From these experimental results we can derive that the use of the WLAN interface is considerably more efficient in terms of energy consumption than the use of the cellular 3G connection. In addition, the throughput and the achievable bandwidth by using a WLAN access are also higher than the ones that the 3G connection can offer. Therefore, we can take advantage from the higher bandwidth offered by the IEEE 802.11 access network, offloading the cellular connection and freeing resources for other users while reducing the energy consumption of our devices.

3.2 Mobility management

The CROWD vision relies on IEEE 802.11 and LTE technologies as radio access networks, tightly integrating both technologies within the same operator core. Due to the extreme density of the network, the IP mobility mechanism used in a CROWD network cannot be centralised, due to reasons that will be explained later, so an approach based on DMM concepts will be designed. This section then is devoted to: *i)* identify the different requirements that CROWD must consider, in order to interconnect the WiFi and LTE RANs, *ii)* the limitations already identified on tradition mobility management mechanisms and *iii)* to explain how an approach based on SDN can be used to provide a flexible and scalable mobility solution.

We start this discussion by understanding the different functional requirements currently imposed by operators while connecting non-3GPP technologies to their cores. Through this section

[8] This value matches perfectly with a real measurement of the power consumed by an iPhone 3GS downloading a 20 MBytes file.

we assume the technology to interconnect is trusted by the operator. The interworking procedure depicted in Fig. 2.11 corresponds to a simplified version of the mechanism used for the initial attachment of a UE to the trusted WLAN network. One could think that the mechanism proposed can be currently implemented without further modifications to the different protocols/technology stacks. In reality, the implementation of such mechanisms requires of the modification of several protocols and the introduction of solutions which are not standardised, hence yielding to interoperability problems. In the following we detail the different requirements in terms of functionality provided by the non-3GPP (in case of Fig. 2.11 an IEEE 802.11 Access Network (AN)) that need to be implemented for the correct operation of both networks. The following functional requirements have been extracted from the 3GPP-Trusted WLAN interworking specification [76]:

- L2/L3 Attachment/detachment triggers: Current solution defines two disjoint mechanisms to trigger the PMIPv6/GTP tunnel setup by the TWAG. These mechanisms are based on either a proprietary message between TWAP and TWAG once the Extensible Authentication Protocol (EAP) authentication is performed (L2 trigger) or the use of IP address configuration mechanisms (DHCP or Neighbour Discovery) to trigger the action. From the former approaches, the so called L2 trigger corresponds to vendor-specific mechanisms, which are not a good option to deploy for an operator, while the second one suffers from incompatibilities from some terminals (e.g., sending a Router Solicitation is only performed on power cycling of the interface) and performance issues (e.g., there is a delay between the L2 association and the IP stack detection of such event). Hence, the usage of real link layer events, which are generated once the link is setup, may improve the performance of the solution on an already implemented way. Within CROWD we address the detection of attachment/detachment of the user in WLAN through the use of standard bridging messages that are made available through the use of SDN (OpenFlow). Basically, the APs deployed in the WiFi district behave as standard OpenFlow switches with a wireless interface. Once a terminal attaches to an AP, the AP generates a standard Logical Link Control (LLC) frame containing the Medium Access Control (MAC) address of the new terminal. This behaviour must be extended for the LTE RAN.

- Point-to-point L2-based transport through the TWAN: The use of PMIPv6/GTP protocol on the S2a interface, mandates that the link between the UE and the MAG (TWAG) behaves as a point to point link. This fact has several implications, mainly because an IEEE 802.11 link by nature is broadcast and it does not behave as a point to point link by default. In addition, this point to point link must be encrypted and isolated between different nodes. The implementation of such a link increases in complexity when multiple PDN connections are setup, since packets belonging to different connections may originate at the same UE and terminate in the same TWAG, hence the MAC addresses in the L2 frame are the same and discerning PDN connections just by MAC addresses is no longer possible. This fact also introduces some added complexity in the TWAG, since it requires to map the different PDN connections to the appropriate S2a tunnel. In addition, different PDN connections must be isolated and secured even for the same users. Different solutions to implement the correct behaviour between UE and TWAP are discussed in the next section. This requirement is easily addressed by the SDN approach followed in CROWD. The idea behind this requirement is twofold, on the one hand this requirement implies that the traffic per user and PDN connection is isolated. On the other hand, the use of a point to point L2 transport implies that the operator does not care about the actual topology at L3 of the wireless access. Through the use of SDN, we aim at providing data path modification at L2, hence we meet both objectives.

- Authentication based on 3GPP protocols (EAP+802.1x): The authentication and authorisation mechanism defined for the interworking with 3GPP is based on EAP, specifically in the EAP-AKA and EAP-SIM variants. Although there has been an already strong effort on providing 3GPP-compatible security mechanisms such as the HotSpot 2.0 specification, there is still way to go in order to define a common authentication framework that can be used between 3GPP and non-3GPP technologies. In the area of IEEE 802, it is envisioned to use IEEE 802.1x to encapsulate EAP messages up to the node in the network acting as AAA proxy. Finally, we cannot forget, that the 3GPP authentication is performed just after the node is granted access to the non-3GPP network, hence an integrated mechanism for joining both the AN and the EPC core is desirable. Through the use of data path control at L2 (such as the one provided by OpenFlow), data packets from the terminal, such as the ones used for authentication, can be routed to the appropriate AAA proxy-server for the operator the terminal belongs to. In this way, the authentication and access to the LTE or WLAN RANs can be implemented in a common way.

- Service discovery mechanisms: Once the UE detects a neighbouring non-3GPP network, in order to be able to connect to it, it requires to know what services are available in the network. For example, the UE needs to know if non-seamless offloading is supported by the network or if the network allows the EPC access. In the case of the 3GPP Release 11, this requirement is even stronger, since it requires the UE to use different BSSIDs to access different services. Hence the UE needs to know what are the services offered by each BSSID, to choose the appropriate one at attachment time. In Release 12, the fact of allowing multiple PDN connections and IP continuity forces the non-3GPP PoA (i.e., WiFi APs) to provide information regarding the reachable APNs. In the CROWD connectivity management architecture, this is addressed through the use of several information repositories, such as ANDSF or IEEE MIIS, which are complemented with information provided by the ALTO protocol.

- Protocol for information exchange between UE and TWAP/TWAG: In addition to the L2/L3 triggers, the TWAG/TWAP needs of some extra-information to signal to the EPC. Specifically, it requires an indication of the type of attachment (if it is a handover or a new attachment), the service required (e.g., Non seamless handover or access to the EPC) and the BSSID/ESSID the UE is attached to. In order to provide this information, it is required the implementation of some kind of protocol allowing the exchange of information between the UE and the TWAP/TWAG. In addition, if multiple PDN connections are desired, the UE needs to signal also the APN it wants to connect to and the credentials for the authentication. In the CROWD architecture this is provided by the extended connectivity manager, which is able to interact with the network in order to configure the desired services.

Current IP mobility management solutions, as previously explained in sections 2.3.3 and 2.3.4, pose several well known problems [77], which are exacerbated due to the extreme dense nature of the CROWD network:

- Sub-optimal routing. Since the (home) address used by a mobile node is anchored at the home link, traffic always traverses the central anchor, leading to paths that are, in general, longer than the direct one between the mobile node and its communication peer. This is exacerbated with the current trend in which content providers push their data to the edge of the network, as close as possible to the users, as for example deploying CDNs. With centralised mobility management approaches, user traffic will always need to go first to the home network and then to the actual content source, sometimes adding unnecessary delay and wasting operator's resources. Through SDN approaches enabling the control of the data path by a central controller, the different variants of IP mobility management protocols can be implemented. Even more, although the key concepts are the same, the possibility of been able to change the data path enables the controller to modify the destination of the flows so they are delivered to the actual location of the terminal, without the need of going through a central node, such as the HA or LMA.

- Scalability problems. Existing mobile networks have to be dimensioned to support all the traffic traversing the central anchors. This poses several scalability and network design problems, as central mobility anchors need to have enough processing and routing capabilities to be able to deal with all the users' traffic simultaneously. Additionally, the entire operator's network needs to be dimensioned to be able to cope with all the users' traffic. In the CROWD approach, the control and data paths are split, hence the scalability issues are diminished.

- Reliability. Centralised solutions share the problem of being more prone to reliability problems, as the central entity is potentially a single point of failure. For our architecture, and due to the need of having a central controller in charge of managing the network, this problem is still present.

The rest of this section is structured as follows: Section 3.2.1 provides an overview of the two key solutions for DMM protocols that have been proposed by CROWD members to the IETF. Then Section 3.2.2 and Section 3.2.3 provide an overview of the SDN tools and mechanisms to be used in the CROWD DMM solution that is presented in Chapter 4.

3.2.1 DMM landscape

3.2.1.1 DMM concepts

DMM aims to adapt existing IP mobility protocols such as MIPv6 and PMIPv6 to the emerging flat IPv6 mobile networks architectures. It intends to distribute and confine the mobility support functions (or part of them) at the ARs level, keeping the rest of the network unaware of the mobility events and their support. For example, the mobility anchoring functionalities are distributed at each AR.

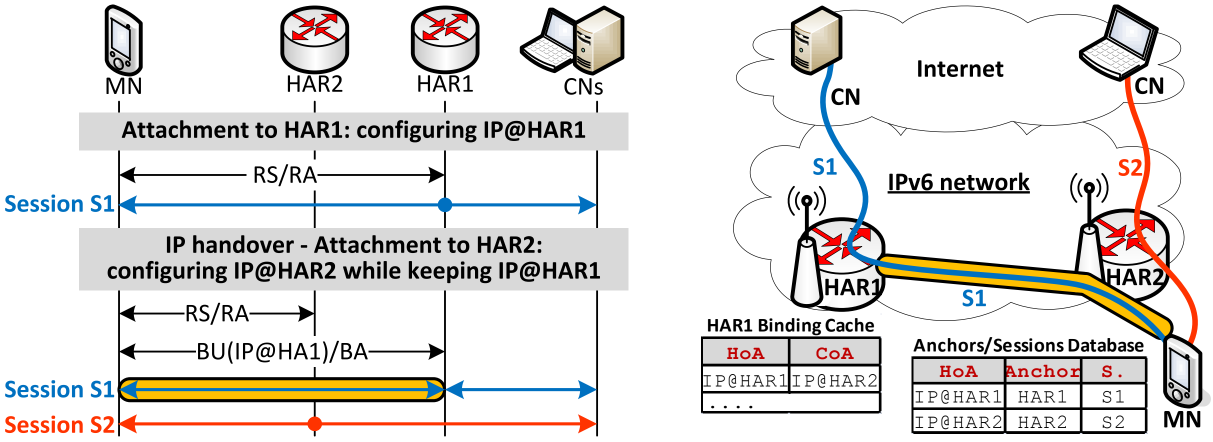
In addition, DMM intends to activate dynamically the mobility support only when needed, i.e. when the MN actually undergoes an IP handover. Compared to MIPv6/PMIPv6 where all the data traffic is anchored at the same entity i.e. the HA/LMA, dynamic mobility management aims at changing the anchor for new sessions. In particular, new sessions in DMM are anchored at the current mobility anchor (deployed at the AR) and initiated using the current IPv6 address. This is achieved thanks to the IPv6 feature allowing an MN to use several IPv6 addresses simultaneously.

As a result, the data traffic of the new sessions is routed optimally (without tunnelling) between the MN and the CN, until the MN undergoes an IP handover. If the MN undergoes one or more IP handovers before the end of a session, then the data traffic of this session is routed via the mobility anchor thanks to tunnelling mechanisms.

DMM is expected to optimise the routing path for a significant percentage of the data traffic [16]. This is due to the fact that at a given time more than 60% of the users in operational networks are non-mobile. In addition, DMM is expected to reduce the tunnelling overhead and global network signalling loads, as a consequence of the point mentioned above as well as because the mobility anchors are closer topologically to the MN. Finally, the single point of failure issues in MIPv6/PMIPv6 are expected to be eliminated with DMM since the bindings' management is distributed at the ARs level instead of being managed at the same entity.

3.2.1.2 MIPv6-based DMM for Global/Local Mobility Support

The MIPv6-based DMM approach (e.g. [15]) is required to support global mobility besides local mobility. The MN may move between different access networks and operational domains. Thus, it


MN
HAR2
HAR1
CNs
Attachment to HAR1: configuring IP@HAR1
RS/RA
Session S1
IP handover - Attachment to HAR2:
configuring IP@HAR2 while keeping IP@HAR1
RS/RA
BU(IP@HA1)/BA
Session S1
Session S2
Internet
CN
CN
S1
S2
IPv6 network
HAR1
HAR2
S1
MN

HAR1 Binding Cache

HoA	CoA
IP@HAR1	IP@HAR2
....	

Anchors/Sessions Database

HoA	Anchor	S.
IP@HAR1	HAR1	S1
IP@HAR2	HAR2	S2

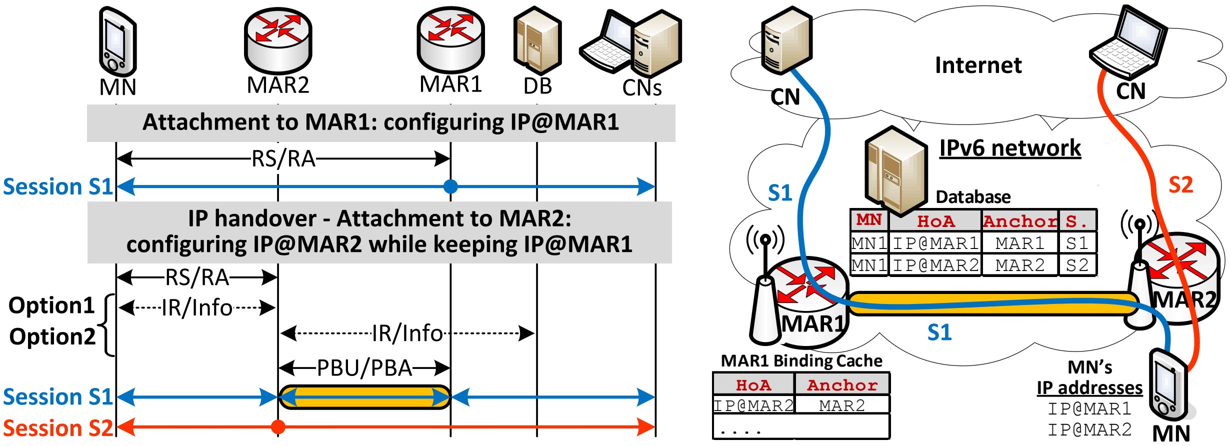
Figure 3.7: Mobility management in MIP-based DMM.

is not guaranteed that the MN is always attached to an HA; the MN may attach to a classical AR that cannot act as a mobility anchor. In addition, the MN may attach to a legacy MIPv6 network. This pushes towards not modifying the HA functionalities in the MIPv6-based DMM approach in order to be compatible with those legacy MIPv6 networks. As a result, the MIPv6-based DMM has to exploit the DMM concepts taking into consideration the different cases and conditions mentioned above. Hence, we distinguish two different use-case scenarios as follows.

In the first scenario, the MN moves in a zone where the HA functionalities are distributed at the ARs level. A concrete example on this scenario is the operational domain of an operator implementing DMM. Then, each AR is an HA (to be named Home Access Router (HAR)) and can play the role of mobility anchoring. This enables the MN to always anchor its new sessions at its current HAR. Upon attaching to an HAR, the MN configures a new IP address. The MN uses this IP address as a source address to initiate new sessions. The data traffic of these sessions is then routed optimally without any need for tunnelling. If these sessions are terminated before undergoing any IP handover, then there is no need to keep the HAR as an anchor. If the MN undergoes one or more IP handovers before terminating these sessions, then the MN manages the mobility by sending a BU to the HAR which replies with a Binding Acknowledgement (BA). A tunnel is then established between the MN and the HAR for the data packets of these sessions. After attaching to a new HAR, the MN uses its new IP address to initiate new sessions. Fig. 3.7 illustrates an example of mobility management in this scenario of MIPv6-based DMM.

In the second scenario, the MN moves in a zone where some ARs implement the HA functionalities but not the others. In order to know if it is attached to an HAR or an AR, the MN checks during the attachment phase the "H bit" [8] in the router advertisement (RA) message. If the "H bit" is set to one, then the MN is attached to an HAR and hence can use it as an anchor for new sessions (as in the previous scenario). Otherwise, it is attached to a classical AR and hence should select one of the existing anchoring HARs as an anchor for the new sessions; new sessions are initiated using the MN's HoA associated to the selected HAR and their traffic is tunnelled via this HAR (with the CoA set to the current IPv6 address).

As mentioned above, the MN sends BU messages to its anchoring HARs upon an IP handover. In order to perform this, the MN is required to create an anchors/sessions database. In this Data base (DB), each entry includes: an HoA, the address of the associated anchoring HAR, and the list of active sessions initiated using this HoA. For each entry, the MN sends upon an IP handover a BU to the registered anchoring HAR using the associated HoA as the HoA and the current IPv6 address as the CoA. The MN updates this DB upon configuring a new HoA when attaching to a new HAR, as well as upon initiating or terminating a session. The MN keeps an entry as long as



MN	HoA	Anchor	S.
MN1	IP@MAR1	MAR1	S1
MN1	IP@MAR2	MAR2	S2

HoA	Anchor
IP@MAR2	MAR2
....	

Figure 3.8: Mobility management in PMIP-based DMM.

it has some active session(s) associated to this entry. The MN deletes an entry when all the active sessions in this entry are terminated. However, even when the MN becomes idle, it keeps at least the most recent entry in order to avoid being without any anchor as required for global mobility support.

3.2.1.3 PMIPv6-based DMM for Local Mobility Support

Compared to the MIPv6-based DMM approach, the PMIPv6-based DMM approach (e.g. [13, 14]) is required to support the local mobility only. The MN moves in a single operational domain. The operator of this domain is required to implement the DMM requirements at each AR, i.e. LMA and MAG functionalities. Hence, it is guaranteed that the MN is always attached to an AR that can act as a mobility anchor (to be named Mobile Anchor Router (MAR)). This allows the MN to always initiate new sessions using the current IP address. The data traffic is then routed optimally without tunnelling. If these sessions are terminated before undergoing an IP handover, then there is no need to keep the MAR where they where initiated as an anchor. On the other hand, if the MN undergoes an IP handover before terminating these sessions, then the mobility management is needed and activated dynamically. Since the MN moves inside a single operational domain, the network is able to manage the mobility. This releases the MN from some additional requirements (compared to the MIPv6-based DMM). The new MAR sends on behalf of the MN a PBU to the previous MAR, which replies by a PBA. A tunnel is then established between the new MAR and previous MAR of the MN. The data traffic of the sessions that are initiated at the previous MAR is then routed via this tunnel. However, the MN initiates new sessions using the new IP address thanks to the IPv6 feature allowing to use several IP addresses simultaneously. The data traffic of these new sessions is routed optimally. Fig. 3.8 illustrates an example of mobility management in PMIPv6-based DMM.

In order to be able to send PBU(s), the new MAR needs to know some additional information. This information includes the IP addresses of the MN's previous MAR(s) used for anchoring the MN's previous sessions, and the associated active IP addresses of the MN. Hereafter are two different options that allow the new MAR to know such information.

Option 1: Relying on the MN: The MN knows all the information about its IP addresses, its sessions and associated anchors. The MN creates an anchors/sessions database and updates it upon initiating or terminating a session (as described in the MIPv6-based DMM approach). The new MAR retrieves the needed information from the MN by sending to it an Infor-Retrieval (IR) request, and then sends the PBU(s) on behalf of the MN.

This option does not introduce any new network entity but the MN is no longer agnostic of the mobility support. Since the MN participates even partially in the mobility-related signalling, this option is considered host-based and the approach is no longer network-based.

Option 2: Relying on a Database: In order to keep the PMIPv6-based DMM approach network-based, another option is to rely on a centralised database [14] in the network. This DB is required to store similar information as in the anchors/sessions database described above, but for all the MNs. For each MN, the DB stores its active IP addresses and associated anchors' IP addresses. The DB is expected to be updated each time an MN configures or releases an IP address. When the MN attaches to a new MAR, this MAR retrieves the needed information from the DB through an IR request, and then sends the PBU(s) on behalf of the MN.

This option does not affect the MN and keeps the approach network-based. However, it introduces a centralised entity in the control plane of the network. The approach is then considered partially distributed and no longer fully distributed.

3.2.2 Optimised mobility management in the IEEE 802.11 RAN

This section presents the tools used in Section 4.2 to build the CROWD Mobility Management solution for IEEE 802.11-based RANs. The WLAN districts considered by CROWD consist mainly in a set of dense packed IEEE 802.11 Access Points joint together by an IEEE 802-based switched network, which is denominated as backhaul within this deliverable, and connects the different APs to one or several CROWD-enhanced IP routers denominated Distributed Mobility Management Gateway (DMM-GW)s. The mobility management of terminals within this network is divided on Intra-district mobility and Inter-district mobility as explained in Section 4.2. Although separated, both solutions use the same key concepts to handle the mobility of the user; the use of the SDN paradigm to split the control and data plane, obtaining higher degrees of scalability and flexibility.

In order to manage the data path within such scenario, in CROWD we have decided to use the OpenFlow protocol. This protocol is specific to IEEE 802 switches at the moment, although extensions enabling its use in other technologies are under investigation. The protocol works by defining a centralised entity (the CROWD Local Controller (CLC) in our case) which is able to manage the data path of the different flows within the network. A data path is basically a set of switching table entries defining a certain match for the packets, an action to perform and an output port. Hence, the OpenFlow protocol defines the required communication between the applications controlling the data-path, in our case the mobility management applications running at the CLC, and the specific implementation of the switch forwarding tables, implemented in the actual hardware.

Hence, the simplified behaviour (suited for our purposes) of a switch supporting the OpenFlow specification is as follows:

- When a packet arrives at any of the OpenFlow controlled ports, the switch checks if there is any match corresponding to the packet. If there is a match, then the appropriate action is taken and the packet is either forwarded to a certain port (which can be a multicast port), dropped or sent to a secondary switching table within the router for further processing.

- If the OpenFlow switch does not have any match corresponding to the packet, the default behaviour of the router is to send the packet to the controller, through a separate outband or inband connection.

- Once the packet arrives at the controller, the application running on the controller takes care of running any required algorithm to implement the specific solution. As output the packet can be modified and returned back to the controller, a new matching rule and corresponding actions and output ports can be installed on the switch, the packet might be dropped or a different packet can be returned to the switch. This enables the controller to implement a great diversity of behaviours, providing a great flexibility to the network operator. In our case, apart of authorising the user, the controller is in charge of selecting the DMM-GW to be used by the terminal and of computing and installing the data path that will carry the packets to the DMM-GW. In case of a mobility event, the controller will react, modifying the data path to convey the packets to the new location of the terminal.

- Once a matching rule is installed, unless it is explicitly set on the action parameter, packets matching the rule are not sent back to the controller. Hence the number of processing events at the controller is bounded.

It is worth noting that a wireless interface can be set as OpenFlow controlled, behaving in exactly the same way as any other port, although it is not possible to differentiate links among the wireless stations connected to the wireless port.

As described above, it is clear that the implementation of different mobility management protocols, algorithms or even the customisation of behaviours according to operator policies can be very easily implemented in the CROWD architecture.

In addition to the OpenFlow capabilities within the network, for the mobility solutions presented in Section 4.2 to work, some other SDN-based interfaces are required. Specifically, it is needed some mechanism to be able to control an IP router through the controller, in a similar manner as the OpenFlow specification can control the switching table of the backhaul elements. This is specially needed in order to create tunnels and to setup routes and prefixes on the router interfaces. One of the objectives of this work package is the definition of such an interface, and will be subject of investigation in the following months.

3.2.3 Optimised mobility management in LTE

CROWD is focused on those scenarios with an extreme density of base stations, we therefore advocate that an incremental optimisation of the existing architecture would not yield the required level of efficiency, and hence put forward a radical evolution of the mobility management itself based on the concepts of DMM and SDN-based mobility illustrated in Chapter 2.

In this section we illustrate our initial results within WP4 of CROWD with specific reference to the LTE technology. It is worth noting that it is beyond the scope of the project to define a fully interoperable evolution of the existing architecture, elements, and protocols. Rather, we aim at proving the feasibility and effectiveness of a number of selected and focused solutions for improved mobility management that are consistent with the overall project architecture under definition as part of the WP1 activities (an initial sketch of which has been described in [4]). Therefore, we focus only of the CROWD scenarios, i.e., consisting of extremely dense and heterogeneous wireless networks, well understanding that real deployments will match this description only in small islands (e.g., crowded public places, very populated residential areas) and incrementally over time. Furthermore, we do not consider important system features that are affected by densification or do not have significant impact on the KPIs in which we are interested, e.g., security considerations, lawful interception, emergency calls.

The current system architecture of LTE is illustrated in Fig. 3.9, which shows the network elements that are most affected by the mobility management procedures. The figure also reports the standard interfaces used for both control and data transmission.

The proposed network architecture is illustrated in Fig. 3.10, which has the following differences with respect to Fig. 3.9:

- The MME is substituted by the CLC. In LTE the MME is a key control element with the following responsibilities:

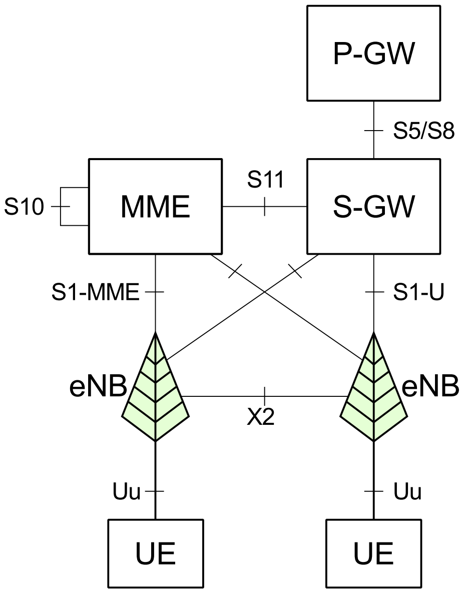


Figure 3.9: LTE system architecture (edge only).

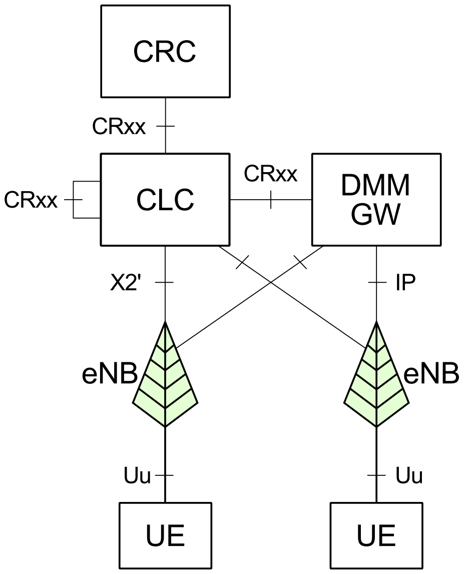


Figure 3.10: Proposed CROWD system architecture.

- Updating the TA of UEs. Among the other functions, this is necessary so that the UE can be reached even after it is not in a connected state anymore. Such function is partially delegated to the CLC and CROWD Regional Controller (CRC) while the UE is in a connected state, while it can be fully retained while the UE is not connected, since the impact on performance will be relatively small in this case.
- Selecting the best next S-GW when the UE moves outside of the area covered by the current S-GW. This function is fully delegated to the CLC and CRC, which also reconfigure the backhaul so as to optimise the data path from the eNB to the DMM-GW.
- Other functions which are beyond the scope of our work, as mentioned above, including terminating the authentication/ciphering end-points, enabling lawful interception, etc.

- The S-GW and P-GW are substituted by the DMM-GW. In LTE the S-GW and P-GW are in charge of managing the data path from the eNB to the CN. They do so via the establishment and maintenance of GTP tunnels encapsulating an EPS bearer, which is a set of packets directed to/coming from the UE which are expected to enjoy the same QoS.

- The S1-U interface is substituted by plain IP, so as to enable uniform addressing/routing with heterogeneous technologies within a district.
- The S1-MME and X2 interfaces are substituted by a X2' interface, which is a variation of the X2 protocols between eNBs that will be defined as part of the WP2 activities in CROWD for the interconnection of the eNBs to their CLC.

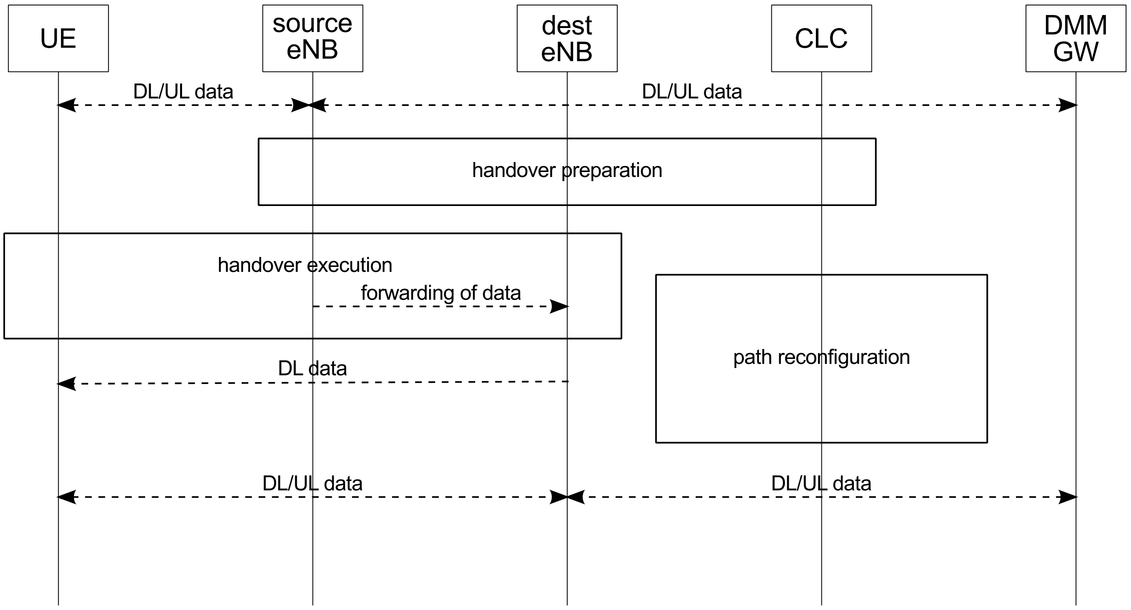


Figure 3.11: Procedure for intra-district LTE handover.

As an example, we report in Fig. 3.11 the procedure for intra-district LTE handover (see Section 2.2.1.1 for the S1-based procedure and Section 2.2.1.2 for the X2-based procedure in the legacy system). At the time of writing, complete and more elaborate procedures are being defined within WP4 in CROWD, in cooperation with the other WPs so as to achieve a uniform and consistent view on the proposed network architecture, functions, and protocols. In the figure we assume that a UE is in a connected state and has established data sessions with a DMM-GW in a district. At some time a handover preparation procedure is triggered. In the legacy system this is triggered by the source eNB based on measurements reported by the UE for its neighbouring cells. On the other hand, in CROWD this will be subject to the CLC, as currently investigated in WP2. Anyway, a handover execution phase begins after the handover has been triggered and prepared. This might require a forward to the destination eNB of the downlink data received by the source eNB but not yet delivered to the UE. Meanwhile, unlike in the legacy system, the CLC can immediately activate the procedure for the path reconfiguration in the backhaul network, which in this case does not require an interaction with the DMM-GW since we are under the assumption that the UE is not leaving the district. After such reconfiguration is complete, downlink and uplink data immediately flows between the destination eNB and the DMM-GW.

4 Functional architecture

The functional architectural of WP4 was already presented in D1.1 [4], herein we first provide a summary of the key concepts and models of our architecture, focusing on the most relevant changes in the WP4 architecture with respect to D1.1. Then we continue with the functional requirements and the IP mobility management solution specifically designed for CROWD.

Fig. 4.1 shows a representation of the different modules composing WP4 and the functional intra work package interfaces connecting them.

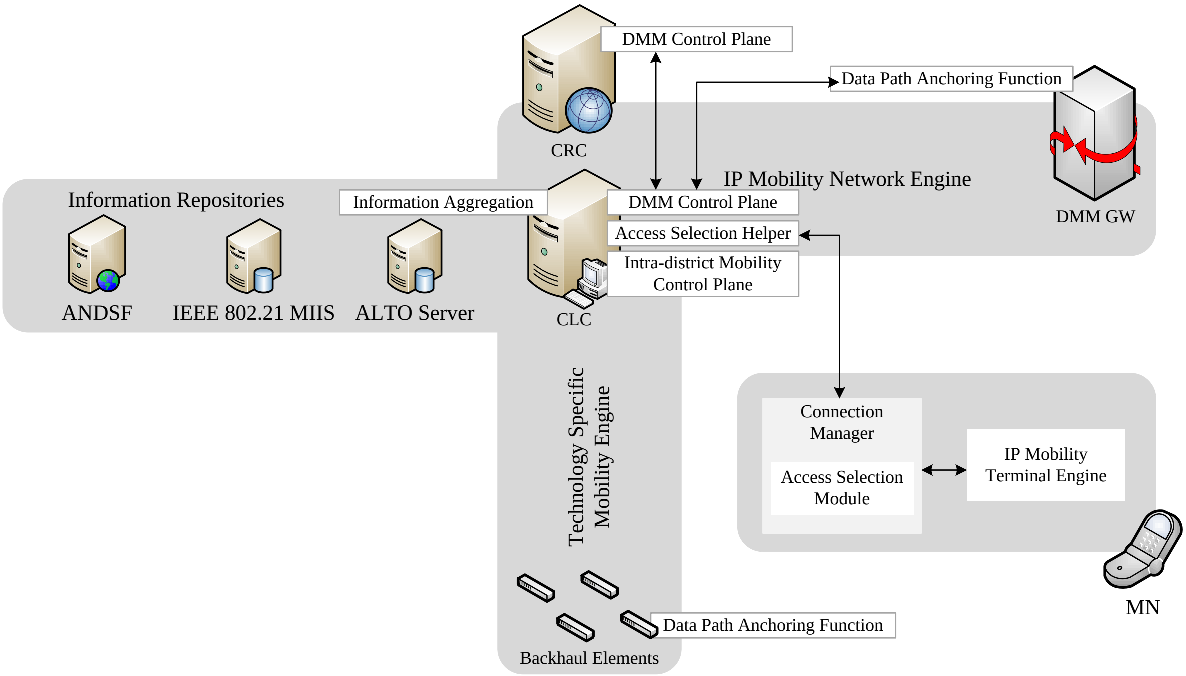


Figure 4.1: WP4 module view and functional internal-interfaces.

According to the general CROWD architecture presented in D1.1, the Connectivity Management modules are split between different locations, depending on the specific functionality provided. In this way, the Connectivity Management functionality is divided in 3 different locations: *i)* the district, including one or several DMM-GWs and a CLC functionality, *ii)* the terminal and *iii)* the network core. Note that in Fig. 4.1, we present a refined view of the architecture presented in D1.1, where we place each module on the appropriate entity. In the following we detail each of these locations and the functional description of the modules residing on them.

A CROWD district includes at least one DMM-GW which implements the data-path functions required to handle mobility at IP layer. Following the SDN approach of CROWD, the control plane of the mobility management is handled by interfacing with the CLC functionality, using a North-Bound (NB) interface to control the data path used by the terminal flows. The WP4 modules explanation and location, as shown in Fig. 4.1 are the following:

- IP Mobility Network Engine: This submodule is in charge of implementing the network-based variant of DMM to be used in the CROWD network. In order to do so, we have identified three functionalities to be deployed within the IP Mobility Network Engine:

1. DMM Control Plane Function: This function is in charge of performing the network-based signalling part for DMM, and it is located at the CLC and CRC (further explanations of the relation between CLC and CRC can be found in Section 4.2).
2. DMM Data Path Anchoring Function: This module is in charge of managing the data-path. Its main functionalities consist in managing the tunnelling and encapsulation of data packets between DMM-GWs. This functionality resides on the DMM-GW and is triggered by signalling from the CLC.
3. Access Selection Helper: This module is able to interface with the terminal in order to provide explicit instructions for access selection, bypassing the information repository. In case it is needed, this module is able to select a target network and inform the terminal of this selection and in addition, it may be able to prepare the hand-over by reserving resources in the target networks.

- Technology Specific Mobility Engine: This module represents the control part of the solution for intra-district mobility that is presented in Section 4.2. This module resides in the CLC.

- Data Path Anchoring Function: Residing in the equipment at the backhaul, this module is in charge of actually forwarding the packets to the appropriate destination as defined by the Technology Specific Mobility Engine.

- Information Aggregator: This module resides at the CLC and uses the aggregation functionalities provided by the CLC to provide a unified vision of the available resources in the district to the Information Repositories located in the core of the operator.

4.1 Functional requirements

This section gives the functional requirements that the CROWD mobility architecture shall meet to address global objectives stated in CROWD deliverable D1.1 [4]. Different solutions shall be defined to meet these requirements, and Section 4.2 gives solution hints to address some of these requirements, mainly requirements #1 and #2. CROWD mobility framework will be refined in further release of the WP4 deliverables (i.e. D4.2 and D4.3) to finally address the following requirements.

1. Distributed Mobility Management: Deliverable D1.1 identifiers the distribution of the network functions as a basic to address scalability and performances issues, raised by CROWD environment. It also applies to mobility management so that traffic does not need to traverse centrally deployed mobility management entities and thereby avoid non-optimal routes and single point of failure. Distribution should apply to both the data and control planes.

2. Fast data path reconfiguration: CROWD solutions must provide transparent mobility support to the user and the application [4]. However, a dense radio environment leads to more frequent handovers than experienced in current mobile network. In such difficult environment, faster data path reconfiguration is necessary to not degrade the user QoE during mobility.

3. Dynamic Mobility management: CROWD mobility management should come into play only when it is necessary. CROWD solutions must provide transparent mobility support to the application [4]; this transparency may lead to provide a stable IPv4 address or IPv6 prefix within the whole CROWD network area; typically, upon change of PoA to the network, an application flow cannot cope with a change in the IP address. However, the IP address

preservation is not always necessary since some applications can handle the handover. Also, a mobile node does not necessary change its PoA, it can initiate and terminate IP session still being attached to the same access point. CROWD solution should be able to deal with this particular situation to avoid unnecessary mobility management operations.

4. Intra and inter technology mobility: The mobile node must be able to move between different access networks technologies (e.g. WiFi and 3GPP mobile network).

5. Global Mobility Management: CROWD solution should work across different networks and be able to co-exist with network deployments that do not support CROWD, e.g., when the mobile node moves from a CROWD capable WiFi network to a legacy 3GPP mobile network. Another example is given by a mobile could initiate a communication within a CROWD WiFi area (city centre with dense WiFi), then move to a non-CROWD WiFi area (suburban area).

6. Keep separate connection management functions: the architecture design must not mix the connection management functions (execution, decision, initiation; see Section 2.1). So that a function can be updated without impacting the others.

7. Access selection taking into account the network conditions: By definition, dense radio environment leads to rapidly changing network conditions (see problem statement in D1.1). The access selection should thus be able to monitor the current access to trigger re-selection before quality degradation leads to communication break. The access selection should also be able to take into account the network condition of the candidate access network to avoid selecting an access with poor quality.

8. Handover upon network events and without user mobility: the terminal may change the current point of attachment because of network events. For example, this occurs when the CROWD controller makes decision to switch-off the current base station or if the quality of communication degrades because of interference.

9. Discovery of CROWD entities in distributed environment: Distributed architecture, as per CROWD, brings discovery issues since many network entities (i.e. IP anchoring gateways and CROWD controllers), providing the same service may come into play and regular discovery mechanisms used in centralised architecture (e.g. DNS) cannot apply.

10. Client or Network based mobility management: With network based mobility management, CROWD mobility management should not require specific protocol and mobility operations on the mobile side. Client based mobility does not have this constraint and the mobile node can be involved in mobility management operations. In any case, mobility management should take benefit from MAC layer enhancements developed in other CROWD WPs.

11. Multihoming: CROWD solution should allow a mobile node to maintain, and use, simultaneous attachments to more than one interface (e.g. WiFi and 3GPP interfaces).

12. Security: CROWD solution must not introduce new security risks or amplify existing security risks, described either in Request for Comments (RFC) 4832 [78] or in RFC 6275 [8], when CROWD relies respectively either on network based, or client based mobility management.

4.2 CROWD Mobility Management proposed solution

This section is devoted to the definition of the SDN-based DMM mobility solution provided by CROWD. It is important to note that in this deliverable we focus on the functional aspects only, without providing implementation or technology-specific details. Without loss of generality we explicitly refer to OpenFlow as the reference protocol for the reconfiguration of the backhaul network elements and we consider IEEE 802.11 APs only (LTE specifics will be reported in next deliverable). However, our approach is general enough to encompass much broader usage scenarios, including the cases in which:

- backhaul devices are reconfigured via non-OpenFlow protocols, e.g., via Interface to the Routing System (I2RS) or vendor-proprietary extensions;
- LTE base stations are present in the network, see Section 3.2.3;
- the mobility management logic is executed by a control application, either local or remote to the CLC, which accesses the latter via a set of NB APIs, currently under definition within CROWD.

The rest of the section is structured as follows. First, the mobility mechanisms to be used within a district will be defined. This mechanisms follows a network-based, terminal transparent mobility solution, very similar to Proxy Mobile IP, but with the benefits in terms of flexibility and scalability provided by SDN approaches. Second, we will focus on the network-based DMM approach for inter-district mobility. Finally, we sketch the solution for the case of global mobility, in which an MN moves to a non-CROWD domain.

4.2.1 Intra-district Mobility

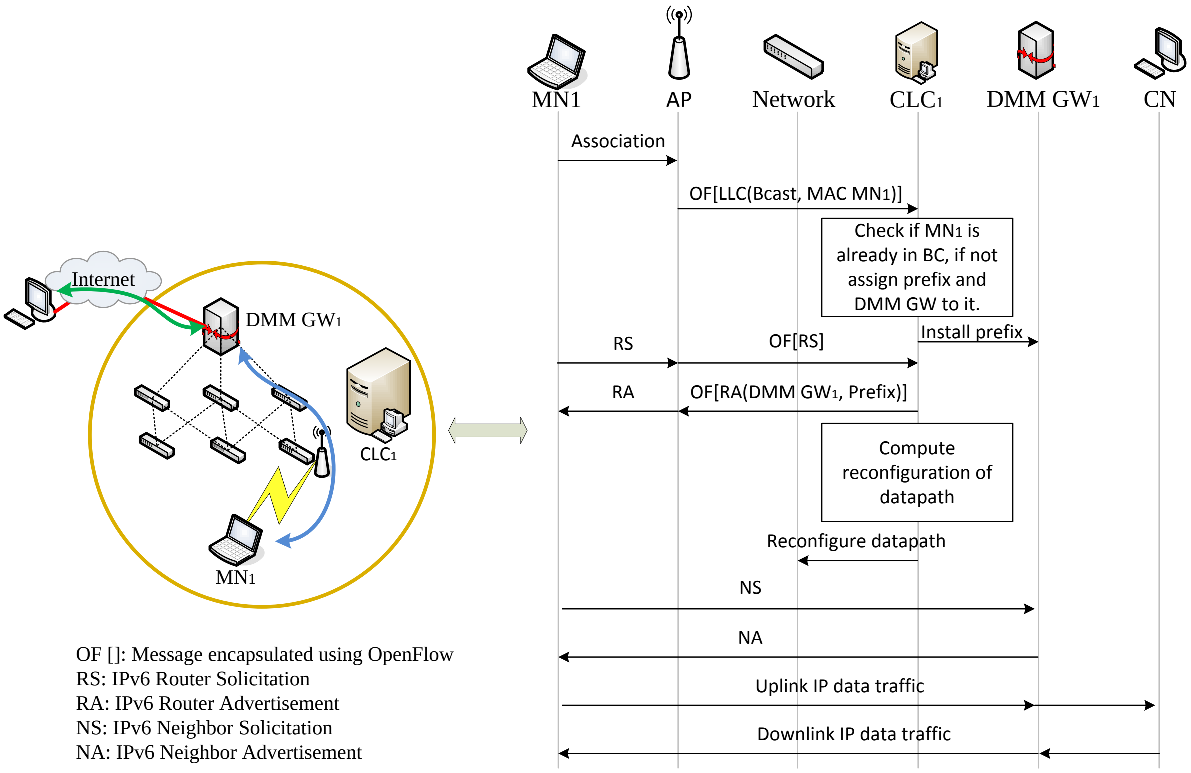

Internet
DMM GW1
CLC1
MN1
MN1
AP
Network
CLC1
DMM GW1
CN
Association
OF[LLC(Bcast, MAC MN1)]
Check if MN1 is already in BC, if not assign prefix and DMM GW to it.
Install prefix
RS
OF[RS]
RA
OF[RA(DMM GW1, Prefix)]
Compute reconfiguration of datapath
Reconfigure datapath
NS
NA
Uplink IP data traffic
Downlink IP data traffic
OF []: Message encapsulated using OpenFlow
RS: IPv6 Router Solicitation
RA: IPv6 Router Advertisement
NS: IPv6 Neighbor Solicitation
NA: IPv6 Neighbor Advertisement

Figure 4.2: Intra-district Mobility: Initial attachment.

Fig. 4.2, presents the initial connection of a terminal to a CROWD district. In this case, we are assuming the district is composed of IEEE 802.11 Access Points, an OpenFlow capable backhaul connecting the APs to the DMM-GWs and a CLC. Upon attachment of the terminal (MN_1 in Fig. 4.2), the AP generates an LLC message[1] that serves as mechanism to update the forwarding table in the switches connecting the AP to the gateway. In this case, as all the network is OpenFlow capable, this LLC message is encapsulated in an OpenFlow message and sent to CLC_1. The LLC message contains the MAC address of the terminal and the MAC address of the AP, hence the CLC can use this message to trigger the detection of mobility or to trigger the mechanisms required for a node to join the network. In the case depicted in Fig. 4.2, CLC_1 does not have any previous entry of the terminal on its BC, hence it assumes a new terminal has attached to the network. Then, the CLC assigns an IPv6 prefix and a DMM-GW to be used by the terminal and stores this information on its BC. The standard procedure followed by a terminal after successful attachment to a new network includes the sending of a RS, in order to configure its IP address using IPv6 Stateless auto-configuration (SLAC). As in the case of the LLC message, the network encapsulates the RS message and sends it to the CLC. The CLC uses an RA message to answer the RS, providing the prefix and default router ($DMM\ GW_1$) selected before. Hence, through the mediation of the CLC, highjacking the RA functionality of the network, we are able to control the IP level attachment of the terminal within the CROWD network. At this point in time, the CLC is able to compute the required match rules and data path modifications required to forward the terminal's packets to the selected DMM-GW. These modifications are configured into the network through the OpenFlow protocol, requiring several message exchanges among CLC_1 and the different switches conforming the path between the terminal and $DMM\ GW_1$. Once the data path is configured, packets originated at the terminal with layer 2 destination the DMM-GW are transparently forwarded at layer 2. This behaviour is completely transparent to the layer 3 stack of the terminal, which sees the path terminal – DMM-GW as a single hop. Finally, the terminal after performing a Neighbour Discovery procedure, is able to exchange packets with any CN through $DMM\ GW_1$.

The case of handover is illustrated in Fig. 4.3, where the terminal attaches to a second AP within the same district. For the case of mobility within a district, the only required change is to modify the data-path, so packets are forwarded between the new AP and the DMM-GW assigned to the terminal. The process followed corresponds to the flow diagram depicted in Fig. 4.3. Upon attachment to the new AP, the AP generates an LLC message including the MAC address of the terminal and the MAC address of the AP it has connected to. Upon reception of this message, the CLC is able to identify that the terminal has moved, since it stores on its BC the MAC address of the AP the terminal was connected to (note this is an example of ID, and any other terminal ID could be used). Once this information is obtained, the CLC is able to compute the required modifications to the data-path used by the terminal's packets, in such a way that, after modifying accordingly the OpenFlow configuration on the backhaul, the packets will flow from the new AP to the old DMM-GW ($DMM\ GW_1$).

4.2.2 Inter-district Mobility

The CROWD solution for inter-district mobility works on top of the intra-district mobility solution, and relies in two main entities, the CLC and the CRC. Since the solution works on layer 3, it is required the definition of new APIs to control the IP layer configuration of the DMM-GW. Specifically, we need two new APIs, one to convey information about the node, such as the IPv6 prefix and DMM-GW assigned to it and a second one to configure the IP layer of the DMM-GW,

[1] Note that this message is present in IEEE technologies, a suitable counterpart is required for the LTE districts, currently we are working on the design of this functionality.

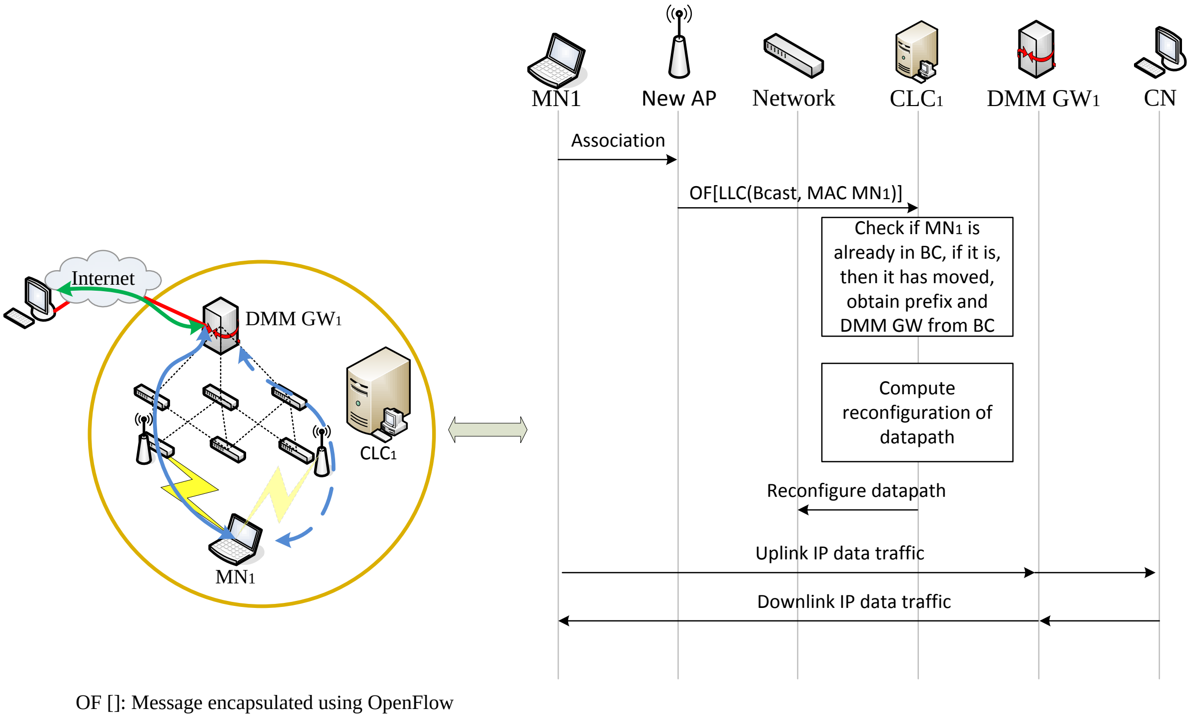
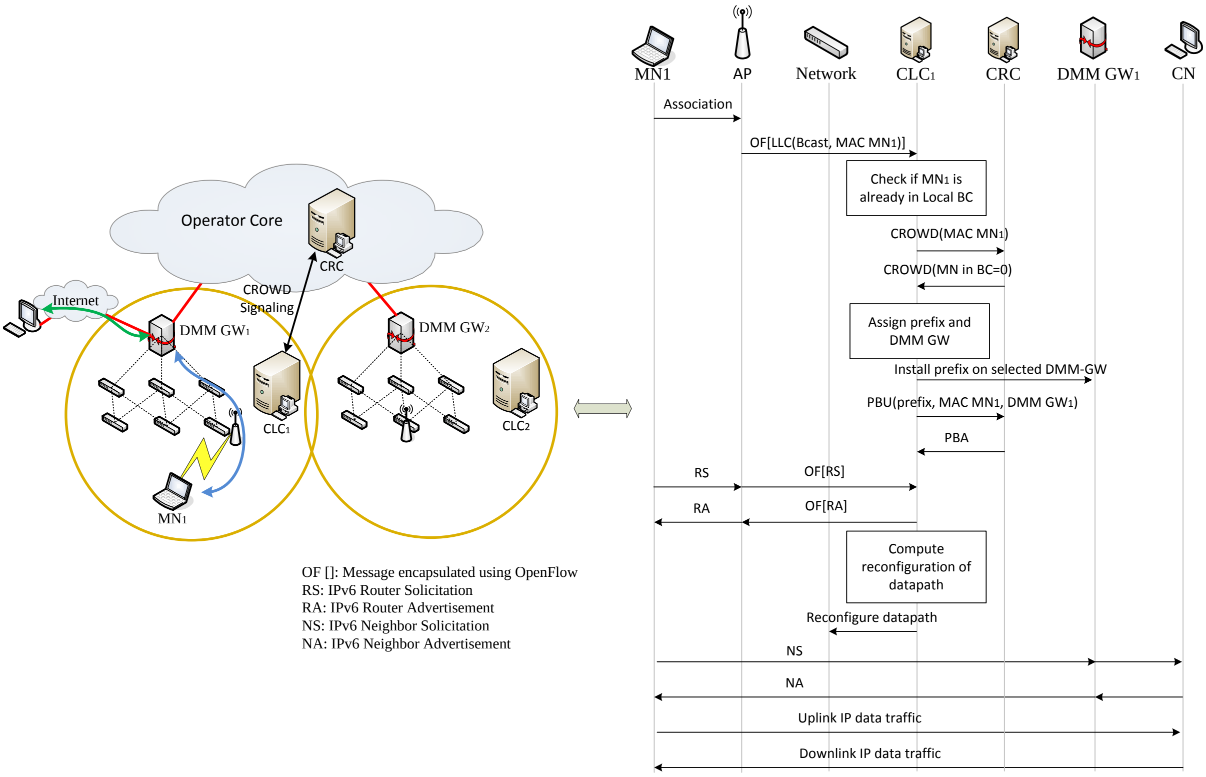

Internet
DMM GW1
CLC1
MN1
MN1
New AP
Network
CLC1
DMM GW1
CN
Association
OF[LLC(Bcast, MAC MN1)]
Check if MN1 is already in BC, if it is, then it has moved, obtain prefix and DMM GW from BC
Compute reconfiguration of datapath
Reconfigure datapath
Uplink IP data traffic
Downlink IP data traffic
OF []: Message encapsulated using OpenFlow

Figure 4.3: Intra-district Mobility: Handover.

the prefixes reachable through the interfaces and to setup an IP in IP tunnel. Fig. 4.4, shows the initial attachment phase of the CROWD inter-district mobility solution. The attachment to the network begins with the terminal associating to a point of attachment belonging to the district. As in the case of the intra-district mobility, this event triggers an LLC frame which is encapsulated in an OpenFlow message and forwarded to CLC_1. Through this message, the CLC is able to check on its local BC if the node is already attached to the district. If it is not the case, then it will contact the CRC, in order to check if the node is already registered in a previous district, and inter-district mobility is required. In the example depicted in Fig. 4.4, the terminal has not been attached previously to any CROWD network, so CLC_1 is free to assign it any of the available IPv6 prefixes in the district. Once the CLC has decided the prefix and DMM-GW to be assigned to the terminal, it proceeds to install the prefix in the DMM-GW ($DMM\ GW_1$). In this way, the CLC has an extra degree of flexibility, being able to assign arbitrarily the prefix to the selected gateway[2]. In addition, the prefix, DMM-GW and terminal identification (MAC address) are notified to the CRC, that is able to keep track of the previous attachments of the terminal. Once this process finalises, the rest of the procedure is exactly the same as the one performed for the intra-district mobility, depicted in Fig. 4.2. Fig. 4.5, presents the procedure of a handover between two different CROWD districts. The procedure assumes that the initial attachment process has been carried on as presented above. The procedure relies on the communication among the CLCs being orchestrated by the CRC. Basically, the CRC behaves as a data-base containing the list of previous DMM-GWs to be considered while performing handover. The configuration of the IP layer on the DMM-GWs and the tunnel setup among them is handled locally by the CLCs on each district. The procedure starts by the terminal attaching to an AP in a different district. As in previous cases, this event

[2] Subject to the routing constraints imposed by the operator



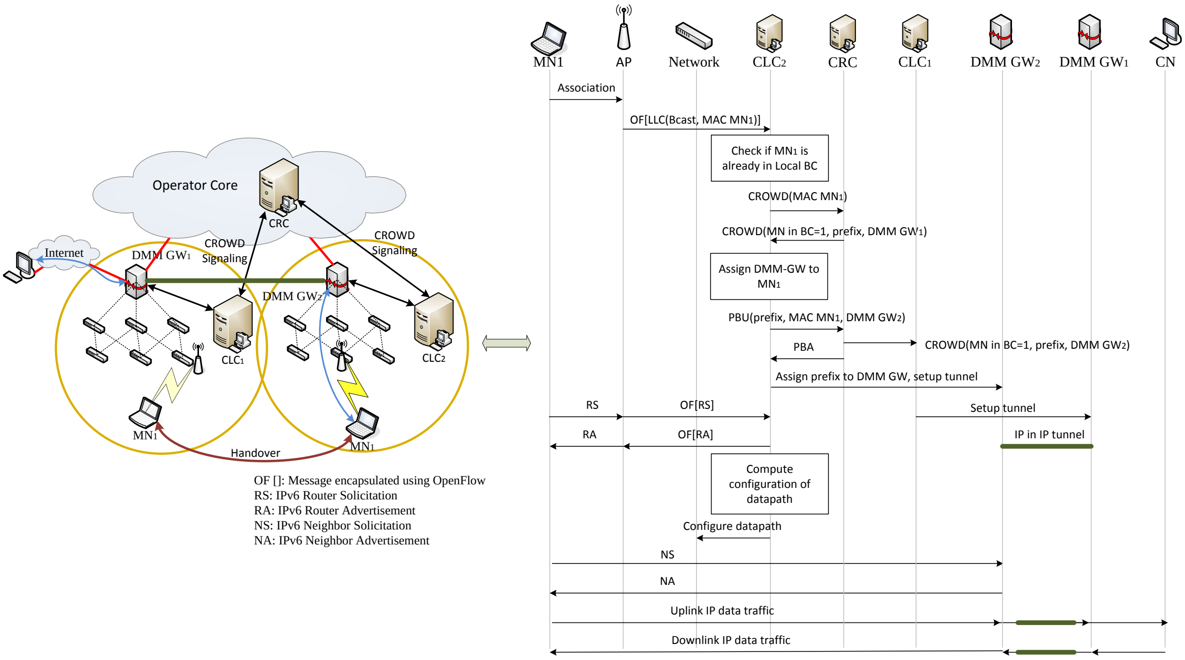
OF []: Message encapsulated using OpenFlow
RS: IPv6 Router Solicitation
RA: IPv6 Router Advertisement
NS: IPv6 Neighbor Solicitation
NA: IPv6 Neighbor Advertisement

Figure 4.4: Inter-district Mobility: Initial attachment.

triggers the CLC (CLC_2) to check if the node is registered on its internal BC. As this is the first time the terminal attaches to the district, the CLC asks the CRC for previous registrations. In this case, the CRC has information regarding the terminal, informing the CLC of the prior connection of the terminal to $DMM\ GW_1$ and the prefix used. With this information, CLC_2 is able to decide the DMM-GW ($DMM\ GW_2$) to be used within this district and provides the CRC with this information. The CRC store this information on its local BC for future reference. At this point in time several procedures are performed in parallel. First, the CRC informs CLC_1 of the new location of the MN; with this information, CLC_1 can configure $DMM\ GW_1$ with an IP in IP tunnel connection with $DMM\ GW_2$ and change the routes at $DMM\ GW_1$ so the prefix used by the terminal is routed through the tunnel. In parallel, CLC_2 configures the new prefix in $DMM\ GW_2$ and setups the IP in IP tunnel towards $DMM\ GW_1$. Once the tunnel is established, the configuration of the data path in the new network is performed as in previous cases. When all the procedure is complete, packets between the CN and MN_1 are forwarded first to $DMM\ GW_1$, which tunnels them to $DMM\ GW_2$. After $DMM\ GW_2$, the OpenFlow configured data path takes care of forwarding the packets to the appropriate location of MN_1 within the district.

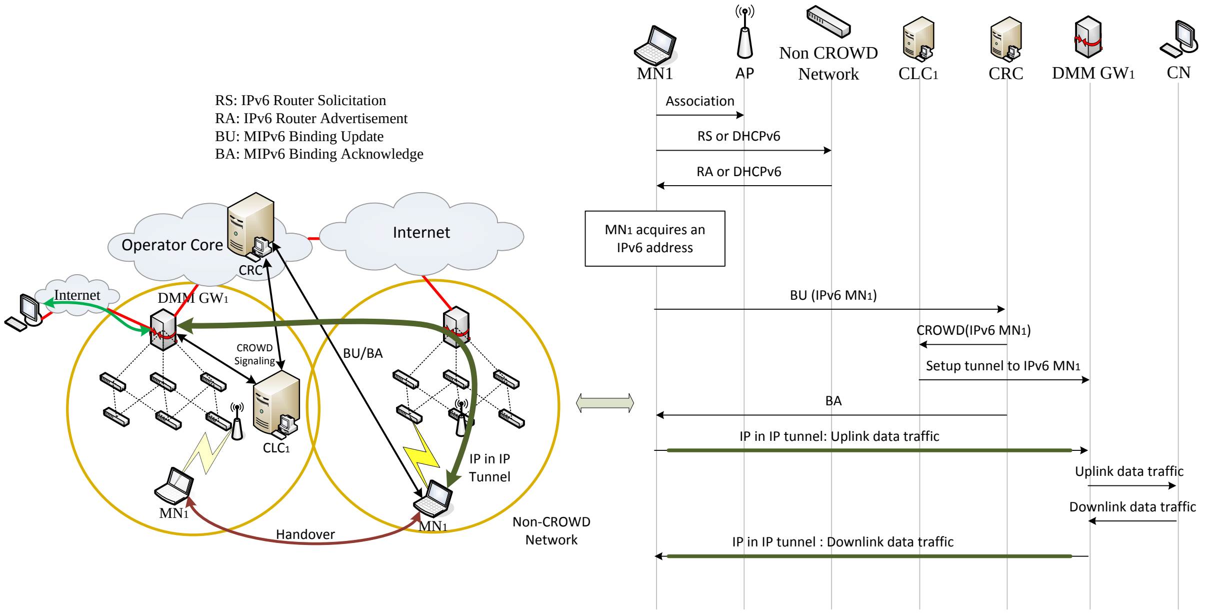
4.2.3 Host-based Mobility

This section presents the global mobility solution adopted by CROWD. The idea behind this solution is to enable the terminal to maintain connectivity when it has roamed to a non-CROWD network, which does not have any of the supporting technologies provided by CROWD. Fig. 4.6, presents the proposed solution. As the network-based proposals presented in above sections, this solution relies on the split of control and data planes, reusing concepts from traditional IP mobility solutions. In this case, we have chosen Mobile IPv6 as the basis of the proposal. Following the



OF []: Message encapsulated using OpenFlow
RS: IPv6 Router Solicitation
RA: IPv6 Router Advertisement
NS: IPv6 Neighbor Solicitation
NA: IPv6 Neighbor Advertisement

Figure 4.5: Inter-district Mobility: Handover.



RS: IPv6 Router Solicitation
RA: IPv6 Router Advertisement
BU: MIPv6 Binding Update
BA: MIPv6 Binding Acknowledge

Figure 4.6: Host-based Mobility CROWD solution

diagram in Fig. 4.6, the procedure starts with MN_1 roaming to a point of attachment not belonging to a CROWD network. Upon attachment, the terminal performs standard IPv6 procedures, soliciting a prefix or a DHCP lease. Once the terminal has acquired an IPv6 address, it can check that the IP address provided by the network does not correspond to any of the IP addresses that are currently been in use by the terminal, so it decides that the Host-based Mobility solution must be used. In order to regain IP connectivity, the terminal uses a MIPv6 BU message, which contains the newly obtained IPv6 address as CoA. This message is sent to the CRC, that in this case acts as HA for the terminal. Upon reception of the message, the CRC contacts the CLCs in charge of managing this terminal, e.g., the CLCs that control the prefix the terminal is using as HA. Once the CLC is aware of the terminal needs, it commands the previous DMM-GW serving the terminal, to create an IP in IP tunnel with destination the terminal's new IPv6 address (i.e., the CoA). Through this solution, packets sent by the CN arrive at the DMM-GW that was previously serving MN_1 ($DMM\ GW_1$ in Fig. 4.6), to be forwarded through the tunnel to the terminal. The same behaviour is enforced for uplink packets.

4.3 Interface definition

This section presents a high level definition of the interfaces used by WP4. We present a high level definition of the interfaces, providing the main information that will be exchanged among WP4 modules. Detailed specification will be provided in D4.2.

Currently the Connectivity Management WP has defined two set of interfaces: *i)* Interfaces used for gathering and distributing information and *ii)* Interfaces for managing the terminal mobility. In the following we briefly describe the two sets of interfaces:

1. Information gathering and distribution interfaces:
   - Information management at the Controllers:
     - **WP4_CLC_CRC_Inf_Aggr**: This interface is used to convey the information obtained by the Information Aggregation at the CLC, to the CRC which in turn will use the WP4_CRC_Inf_Rep interface to distribute it to the different Information Repositories.
     - **WP4_CRC_Inf_Rep**: This interface is defined between the Information Repositories (ALTO, ANDSF and 802.21 MIIS) and the CRC to get the information of the district managed by a certain CLC.
   - Information management at the Connection Manager:
     - **WP4_UE_ALTO_Inf**: Interface between ALTO server and mainly the UEs (JSON ALTO) to get cost map.
     - **WP4_UE_IEEE80221_Inf**: Interface between the UE and the IEEE 802.21 MIIS server for information about the network and the UE and a PoA to get information about an access node. This interface implements the standard IEEE 802.21 protocol.
     - **WP4_UE_ANQP_Inf**: Internal UE interface to get the information of the IEEE 802.11u capable APs, through the use of the Generic Advertisement Service (GAS)/ANQP protocol.
     - **WP4_UE_ANDSF_Inf**: Interface between ANDSF server and UEs (Open Mobile Alliance - Device Management (OMA-DM) ANDSF).
2. Mobility Management Interfaces:
   - Mobility Management control signalling between controllers:
     - **WP4_CLC_CRC_MM**: This interface provides communication capabilities between the CLC and CRC to exchange mobility management related information such as the existence of an MN in the CRC Binding Cache and the prefix and DMM-GW associated to it.
   - Mobility Management control signalling within a district:

- **WP4_CLC_DP**: Interface between the CLC and the different elements in the data path, used to control the forwarding of UE packets through the network. This interface will be defined in two separated ways; one for IEEE-based technologies, which will use OpenFlow based signalling and a second one based on the X2 interface for 3GPP networks.
- **WP4_CLC_GW_DMM**: Interface enabling the communication between the CLC and the DMM-GW. Used to setup tunnels and configure routes.

5 Conclusion

This deliverable focuses on connectivity management in a CROWD environment (i.e. dense access network context). In this document, and in WP4 in general, the Ariadnés thread is the clear separation of the connectivity management functions, which are *initiation*, *decision* and *execution*. Even if these functions are closely related, we believe that they remain independent for the sake of design efficiency. So, after going through the state of the art with respect to each of these functions, the document presents first thoughts regarding the application of these concepts to the CROWD context. Following this trend, we have analysed first the module in charge of initiating the handover, usually known as Connection Manager, analysing possible architectures been discussed at the IETF. As a second stage we have also studied the target network decision, considering how information repositories can influence this decision and how the target network technology can affect the energy consumption. Finally, in Chapter 4 we provide, along with the functional requirements of the architecture, the DMM-based mobility protocol designed for the ”execution phase” of the CROWD Connectivity Management solution.

Although we think that the architecture and designs provided in this deliverable are a major milestone in the project, we aim at refining them in the following months, focusing in the following areas for future work:

- Confirm the mobility trigger capabilities at the terminal side.
- Design an access selection algorithm optimising network consumption.
- Detail specification of the interfaces among CROWD entities.
- The impact on current standard (3GPP, BBF, IETF, IEEE) will also be studied and extensions, if any, will be defined and proposed as change request in the appropriate standardisation body.

Lastly, we want to highlight that SDN based mobility is a novelty in the mobility management area and the corner stone of the CROWD WP4; as such, the SDN based CROWD architecture will be challenged by existing routing and mobility protocols. We aim at answering this challenge providing solutions for the next generation of networks.

Bibliography

[1] R. Trestian, O. Ormond, and G. Muntean. Game Theory-Based Network Selection: Solutions and Challenges. In *IEEE Communications Surveys and Tutorials*, pages 1212–1231, 2012.

[2] L. Wang and D. Binet. Mobility-Based Network Selection Scheme in Heterogeneous Wireless Networks. In *VTC Spring*, 2009.

[3] D. Liu, T. Lemon, and Z. Cao. MIF API consideration. Internet-Draft (work in progress), draft-ietf-mif-api-extension, 2012.

[4] C. Cicconetti, S. Auroux, E. Bizouarn, A. de la Oliva, M. Draexler, V. Mancuso, A. Morelli, R. Gupta, I. Sanchez, V. Sciancalepore, P. Seite, and P. Serrano. Preliminary architecture design. Project deliverable D1.1, CROWD, 2013.

[5] 3GPP. General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access. *3GPP Std. TS 23.401*, 2013.

[6] K. Samdanis, T. Taleb, and S. Schmid. Traffic Offload Enhancements for eUTRAN. *Communications Surveys Tutorials, IEEE*, 14(3):884–896, 2012.

[7] D. Le, X. Fu, and D. Hogrefe. A review of mobility support paradigms for the internet. *Communications Surveys Tutorials, IEEE*, 8(1):38–51, 2006.

[8] C. Perkins, D. Johnson, and J. Arkko. Mobility Support in IPv6. RFC 6275 (Proposed Standard), July 2011.

[9] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil. Proxy Mobile IPv6. RFC 5213, August 2008.

[10] Cisco Systems Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2009-2014. Technical report, Cisco, 2010.

[11] Wireless Intelligence (operators data): https://wirelessintelligence.com .

[12] IETF DMM WG: http://datatracker.ietf.org/wg/dmm/charter/ .

[13] P. Seite. Dynamic Mobility Anchoring. Internet-Draft (work in progress), May 2010.

[14] Seite, P. and Bertin, P. Distributed Mobility Anchoring. Internet-Draft draft-seite-dmm-dma-06 (work in progress), IETF, 2013.

[15] H. Ali-Ahmad, M. Ouzzif, P. Bertin, and X. Lagrange. Distributed dynamic mobile IPv6: Design and evaluation. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 2166–2171, 2013.

[16] H. Ali-Ahmad, M. Ouzzif, P. Bertin, and X. Lagrange. Comparative performance analysis on dynamic mobility anchoring and proxy mobile IPv6. In *Wireless Personal Multimedia Communications (WPMC), 2012 15th International Symposium on*, pages 653–657. IEEE, 2012.

[17] H.A. Chan, H. Yokota, J. Xie, P. Seite, and D. Liu. Distributed and Dynamic Mobility Management in Mobile Internet: Current Approaches and Issues. *Journal of Communications*, 6(1), 2011.

[18] P. Bertin, S. Bonjour, and J.M. Bonnin. Distributed or centralized mobility? In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6. IEEE, 2009.

[19] P.J. McCann. Design of a flat wireless Internet Service Provider network. In *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*, pages 1 –5, oct. 2011.

[20] V.P. Kafle, Y. Kobari, and M. Inoue. A distributed mobility management scheme for future networks. In *Kaleidoscope 2011: The Fully Networked Human? - Innovations for Future Networks and Services (K-2011), Proceedings of ITU*, pages 1 –7, dec. 2011.

[21] W. Hahn. 3GPP Evolved Packet Core support for distributed mobility anchors: Control enhancements for GW relocation. In *ITS Telecommunications (ITST), 2011 11th International Conference on*, pages 264–267. IEEE, 2011.

[22] W. Hahn. Flat 3GPP Evolved Packet Core. In *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*, pages 1 –5, oct. 2011.

[23] R. Farha, K. Khavari, N. Abji, and A. Leon-Garcia. Peer-to-peer mobility management for all-ip networks. In *Communications, 2006. ICC'06. IEEE International Conference on*, volume 5, pages 1946–1952. IEEE, 2006.

[24] M. Fischer, F.-U. Andersen, A. Kopsel, G. Schafer, and M. Schlager. A Distributed IP Mobility Approach. In *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pages 1 –6, sept. 2008.

[25] Y. Zhai, Y. Wang, I. You, J. Yuan, Y. Ren, and X. Shan. A DHT and MDP-based Mobility Management Scheme for Large-Scale Mobile Internet. In *Proc. IEEE INFOCOM Workshop on MobiWorld*, pages 1–6, 2011.

[26] P. Bertin, S. Bonjour, and J.-M. Bonnin. A Distributed Dynamic Mobility Management Scheme Designed for Flat IP Architectures. In *New Technologies, Mobility and Security, 2008. NTMS '08.*, pages 1 –5, nov. 2008.

[27] P. Louin and P. Bertin. Network and host based distributed mobility. In *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*, pages 1 –5, oct. 2011.

[28] M. Liu, X. Guo, A. Zhou, S. Wang, Z. Li, and E. Dutkiewicz. Low latency IP mobility management: protocol and analysis. *EURASIP Journal on Wireless Communications and Networking*, 2011(1):1–16, 2011.

[29] R. Wakikawa, G. Valadon, and J. Murai. Migrating home agents towards internet-scale mobility deployments. In *Proceedings of the 2006 ACM CoNEXT conference*, page 10. ACM, 2006.

[30] H.A. Chan. Proxy mobile IP with distributed mobility anchors. In *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, pages 16 –20, dec. 2010.

[31] H. Jung, M. Gohar, J.I. Kim, and S.J. Koh. Distributed mobility control in proxy mobile IPv6 networks. *IEICE Transactions on Communications*, 94(8):2216, 2011.

[32] M. Boc, A. Petrescu, and C. Janneteau. Anchor-based routing optimization extension for Proxy Mobile IPv6 in flat architectures. In *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*, pages 1 –5, oct. 2011.

[33] Li. Erran, L. Morley, and J. Rexford. Toward Software-Defined Cellular Networks. In *European Workshop on Software Defined Networking (EWSDN)*, Darmstadt, Germany, October 2012.

[34] G. Hampel, M. Steiner, and T. Bu. Applying software-defined networking to the telecom domain. In *INFOCOM, 2013 Proceedings IEEE*, pages 3339–3344, 2013.

[35] S.J. Vaughan-Nichols. OpenFlow: The Next Generation of the Network? *Computer*, 44(8):13–15, 2011.

[36] X. Xiao and L.M. Ni. Internet QoS: a big picture. *Network, IEEE*, 13(2):8–18, 1999.

[37] L. Yang, R. Dantu, T. Anderson, and R. Gopal. Forwarding and Control Element Separation (ForCES) Framework. RFC 3746, IETF, 2004.

[38] E. Haleplidis, C. Tranoris, S. Denazis, and O. Koufopavlou. Adopting software engineering practices to network processor devices introducing the Domain Specific Modeling paradigm to the ForCES Framework. In *Network and Service Management (CNSM), 2010 International Conference on*, pages 366–369, 2010.

[39] O. El Ferkouss, S. Correia, R. Ben Ali, Y. Lemieux, M. Julien, M. Tatipamula, and O. Cherkaoui. On the Flexibility of MPLS Applications over an OpenFlow-Enabled Network. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–6, 2011.

[40] S. Ishida and I. Nishioka. Software-defined packet optical transport networks offering multiple services. In *Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), 2013*, pages 1–3, 2013.

[41] K. Yap, R. Sherwood, M. Kobayashi, T. Huang, M. Chan, N. Handigol, N. McKeown, and G. Parulkar. Blueprint for introducing innovation into wireless mobile networks. In *Proceedings of the second ACM SIGCOMM workshop on Virtualized infrastructure systems and architectures*, VISA '10, pages 25–32, New York, NY, USA, 2010. ACM.

[42] J. Kempf, B. Johansson, S. Pettersson, H. Luning, and T. Nilsson. Moving the mobile Evolved Packet Core to the cloud. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2012 IEEE 8th International Conference on*, pages 784–791, 2012.

[43] K. Pentikousis, Yan Wang, and Weihua Hu. Mobileflow: Toward software-defined mobile networks. *Communications Magazine, IEEE*, 51(7):–, 2013.

[44] N. Varis, J. Manner, and J. Heinonen. A layer-2 approach for mobility and transport in the mobile backhaul. In *ITS Telecommunications (ITST), 2011 11th International Conference on*, pages 268–273, 2011.

[45] R. Bifulco, R. Canonico, M. Brunner, P. Hasselmeyer, and F. Mir. A Practical Experience in Designing an OpenFlow Controller. In *Software Defined Networking (EWSDN), 2012 European Workshop on*, pages 61–66, 2012.

[46] S. Paul and R. Jain. OpenADN: Mobile apps on global clouds using OpenFlow and Software Defined Networking. In *Globecom Workshops (GC Wkshps), 2012 IEEE*, pages 719–723, 2012.

[47] P. Baskett, Yi S., Wenjun Z., and B. Guttersohn. SDNAN: Software-defined networking in ad hoc networks of smartphones. In *Consumer Communications and Networking Conference (CCNC), 2013 IEEE*, pages 861–862, 2013.

[48] A. Asadi and V. Mancuso. Energy efficient opportunistic uplink packet forwarding in hybrid wireless networks. In *e-Energy*, pages 261–262, 2013.

[49] Qingyang Song and A. Jamalipour. Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques. *Wireless Commun.*, 12(3):42–48, June 2005.

[50] E. Gustafsson and A. Jonsson. Always best connected. *Wireless Commun.*, 10(1):49–55, February 2003.

[51] R. Trestian, O. Ormond, and G. Muntean. [reputation-based network selection mechanism using game theory. In *Physical Communication*, pages 156–171, 2011.

[52] B. Xing and N. Venkatasubramanian. Multi-Constraint Dynamic Access Selection in Always Best Connected Networks. In *MobiQuitous*, pages 56–64, 2005.

[53] E. Adamopoulou, K. Demestichas, A. Koutsorodi, and M. Theologou. Intelligent Access Network Selection in Heterogeneous Networks - Simulation Results. In *ISWCS 2005, 2nd International Symposium on Wireless Communications Systems*, pages 279–283. IEEE Communications Society, September 2005.

[54] M. Buddhikot, G. Chandranmenon, S. Han, Y. Lee, S. Miller, and L. Salgarelli. Integration of 802.11 and Third-Generation Wireless Data Networks. In *INFOCOM*, 2003.

[55] P. TalebiFard and V. Leung. A Data Fusion Approach to Context-Aware Service Delivery in Heterogeneous Network Environments. In *ANT/MobiWIS*, pages 312–319, 2011.

[56] Q. Nguyen-Vuong, N. Agoulmine, and Y. Ghamri-Doudane. ”A user-centric and context-aware solution to interface management and access network selection in heterogeneous wireless environments ”. *Computer Networks*, 52(18):3358 – 3372, 2008.

[57] M. Ylianttila, M. Pande, J. Mäkelä, and P. Mähönen. Optimization Scheme for Mobile Users Performing Vertical Handoffs Between IEEE 802.11 and GPRS/EDGE Networks. In Arthur Henley and Booker Tyron, editors, *GLOBECOM 2001, IEEE Global Telecommunications Conference*, volume 6, pages 3439–3443. IEEE Communications Society, November 2001.

[58] C. Desset, N. Ahmed, and A. Dejonghe. Energy Savings for Wireless Terminals through Smart Vertical Handover. In *ICC*, pages 1–5. IEEE, 2009.

[59] S. Lee, K. Sriram, K. Kim, J. Lee, Y. Kim, and N. Golmie. Vertical Handoff Decision Algorithm Providing Optimized Performance in Heterogeneous Wireless Networks. In *GLOBECOM*, pages 5164–5169, 2007.

[60] LAN/MAN Committee of the IEEE Computer Society. IEEE Std 802.21-2008, Standards for Local and Metropolitan Area - Part 21: Media Independent Handover Services, 2008.

[61] LAN/MAN Committee of the IEEE Computer Society. IEEE Standard for Local and Metropolitan Area Networks- Part 21: Amendment 1: Security Extensions to Media Independent Handover Services and Protocol, 2012.

[62] LAN/MAN Committee of the IEEE Computer Society. IEEE Standard for Local and Metropolitan Area Networks- Part 21: Amendment 2: Extension for Supporting Handovers with Downlink Only Technologies, 2012.

[63] LAN/MAN Committee of the IEEE Computer Society. IEEE Draft Standard for Local and Metropolitan Area Networks- Part 21: Amendment 3: Single Radio Extensions, 2013.

[64] LAN/MAN Committee of the IEEE Computer Society. IEEE Draft Standard for Local and Metropolitan Area Networks- Part 21: Amendment 4: Multicast Group Management, 2013.

[65] R. Alimi et all. ALTO Protocol. Internet-Draft (work in progress), draft-ietf-alto-protocol-07.txt, March 2011.

[66] LAN/MAN Committee of the IEEE Computer Society. IEEE Standard for Local and Metropolitan Area Networks- Part 11: Amendment 9: Interworking with External Networks, 2011.

[67] B. Orlandi and F. Scahill. Wifi Roaming - Building on ANDSF and Hotspot 2.0, 2012.

[68] S. Randriamasy, W. Roome, and N. Schwan. Multi-Cost ALTO. Internet-Draft (work in progress), draft-randriamsy-alto-multi-cost-07, October 2012.

[69] 3GPP. IP flow mobility and seamless Wireless Local Area Network (WLAN) offload; Stage 2. *3GPP Std. TS 23.261*, 2012.

[70] A. Galindo-Serrano, L. Giupponi, and M. Dohler. Cognition and Docition in OFDMA-Based Femtocell Networks. In *GLOBECOM*, pages 1–6, 2010.

[71] A. de la Oliva, A. Banchs, and P. Serrano. Throughput and energy-aware routing for 802.11 based mesh networks. *Computer Communications*, 35(12):1433–1446, 2012.

[72] A. Garcia-Saavedra, P. Serrano, A. Banchs, and G. Bianchi. Energy consumption anatomy of 802.11 devices and its implication on modeling and design. In *CoNEXT*, pages 169–180, 2012.

[73] J. Huang, F. Qian, A. Gerber, J. Morley-Mao, S. Sen, and O. Spatscheck. A close examination of performance and power characteristics of 4G LTE networks. In *MobiSys*, pages 225–238, 2012.

[74] P. Serrano, C. J. Bernardos, A. de la Oliva, A. Banchs, I. Soto, and M. Zink. FloorNet: Deployment and Evaluation of a Multihop Wireless 802.11 Testbed. *EURASIP Journal on Wireless Communications and Networking*, 2010, 2010.

[75] P. Rysavy. Mobile Broadband Capacity Constraints and the Need for Optimization, February 2010.

[76] Technical report 23.852 (samog). 3GPP TR 23.852 v1.4.0, 2013.

[77] Chan, H. and Liu, D. and Seite, P. and Yokota, H. and Korhonen, J. Requirements for Distributed Mobility Management. Internet-Draft draft-ietf-dmm-requirements-07 (work in progress), IETF, 2013.

[78] J. Kempf and C. Vogt. Security threats to network-based localized mobility management (netlmm). RFC 4832, 2007.