iJOIN

INFSO-ICT-317941

# iJOIN

## IR4.1

## Analysis and identification of the network-level requirements and state-of-the-art review

| | |
|---|---|
| Editor: | Peter Rost, NEC |
| Deliverable nature: | Confidential |
| Suggested readers: | iJOIN GA |
| Due date: | April 30th, 2012 |
| Delivery date: | TBD |
| Version: | 0.6 |
| Total number of pages: | 50 |
| Reviewed by: | UC3M |
| Keywords: | iJOIN |
| Resources consumed | 20 PM |


**Abstract**

This document provides a comprehensive list of state-of-the-art technology, system requirements and assumptions, as well as first details of candidate technologies related to the work package "Network-layer solutions and system and operation management."

# List of authors

| Company | Author |
|---|---|
| **NEC** | Peter Rost, Daniel Camps Mur |
| **TUD** | Vinay Suryaprakash, Jens Bartelt |
| **UC3M** | Carlos J. Bernardos |
| **TID** | Ignacio Berberana |
| **IMC** | Umer Salim |
| **UNIS** | Yingli Sheng |

# History

| Modified by | Date | Version | Comments |
|---|---|---|---|
| Peter Rost | 2013-05-13 | 1.0 | Final version of IR4.1 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

 

# Table of Contents



 

# List of Figures



 

# Abbreviations

| | |
|---|---|
| 3GPP | 3rd Generation Partnership Programme |
| ANDSF | Access Network Discovery and Selection Function |
| API | Application Programming Interface |
| AQM | Active Queue Management |
| BGP | Border Gateway Protocol |
| BS | Base Station |
| CCO | Capacity and Coverage Optimization |
| CSG | Closed Subscriber Group |
| DGW | Distributed Gateway |
| DMM | Distributed Mobility Management |
| DNS | Domain Name Service |
| DSMIP | Dual Stack Mobile IP |
| ECN | Explicit Congestion Notification |
| eNB | Evolved Node B |
| EPS | Evolved Packet System |
| FBS | Femto Base Station |
| GGSN | Gateway GPRS Support Node |
| GPRS | General Packet Radio Service |
| GTP | GPRS Tunnelling Protocol |
| HSS | Home Subscriber Server |
| IETF | Internet Engineering Task Force |
| IFOM | IP Flow Mobility |
| iJOIN | Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks |
| IP | Internet Protocol |
| IT | Information Technology |
| LEDBAT | Low Extra Delay Background Transport |
| LGW | Local Gateway |
| LIMONET | LIPA Mobility and SIPTO at the Local Network |
| LIPA | Local IP Access |
| LMA | Local Mobility Anchor |
| LTE | Long Term Evolution |
| MAG | Mobility Access Gateway |
| MIMO | Multiple Input Multiple Output |
| MIPv6 | Mobile IP |
| MME | Mobility Management Entity |
| MN | Mobile Node |

| | |
|---|---|
| MPLS | Multiprotocol Label Switching |
| NAS | Non-Access Stratum |
| NB-IFOM | Network-Based IP Flow Mobility |
| NFV | Network Function Virtualization |
| NGMN | Next Generation Mobile Network Alliance |
| ONF | Open Network Forum |
| OPEX | Operational Expenditures |
| OSG | Open Subscriber Group |
| OSPF | Open Shortest Path First |
| PDN | Packet Data Network |
| P-GW | Packet Gateway |
| P2P | Peer to Peer |
| PMIP | Proxy Mobile IP |
| QoE | Quality of Experience |
| QoS | Quality of Service |
| RACH | Random Access Channel |
| RAN | Radio Access Network |
| RANaaS | RAN as a Service |
| RAP | Radio Access Point |
| RAT | Radio Access Technology |
| RED | Random Early Detection |
| RNL | Radio Network Layer |
| SDN | Software Defined Networks |
| SIPTO | Selected IP Traffic Offload |
| SLA | Service Level Agreement |
| SON | Self-Organised Networks |
| UE | User Equipment |
| WLAN | Wireless Local Area Network |

# 1 Introduction

The iJOIN project aims on tackling the increasing rate demands in future networks. Its approach incorporates two key concepts: very dense small-cell networks and (partly) centralized processing of RAN functionality in order to exploit centralization gain. Further to that, it considers the application of heterogeneous backhaul which requires a flexible assignment of functionality within the mobile network. The mobile network consists not only of the radio access part but also the backhaul network which connects radio access network and core network. Therefore, iJOIN pays special emphasis on the interworking and joint optimization of radio access and backhaul network. This joint operation will also require a joint management of networks in order to avoid contradicting measures applied to either one. Furthermore, iJOIN considers the partial remote execution of RAN functionality which will impose further requirements on the network management in order to guarantee minimum requirements for the centralization of functionality. Hence, the focus of this report is to analyse the network management and orchestration of the iJOIN system.

The unique attribute of a mobile network is the mobility of its users, i.e. users may connect from different physical locations. This implies temporal and spatial changes of the traffic distribution. As networks become denser in order to provide the required capacity, also the management of the network becomes more difficult which includes the mobility support (and its required signalling efforts), admission and congestion control, load balancing, traffic routing, and self-organisation in order to orchestrate the dense network efficiently. The future mobile network management may neither be fully centralized as networks become too complex nor fully decentralized as this it too inefficient and not effective. Hence, a suitable degree of multi-level coordination needs to be found. An example for such a network-layer functionality is routing where centralized solutions may easily become infeasible as an online routing on such a complex network is impossible. Full decentralization is not applicable as well because it is static or it requires significant signalling exchange between the individual nodes which easily leads to avalanche effects. Furthermore, routing needs to be application driven as different applications may require different path policies.

Beside the active network management after deployment (maintenance), also deployment and dimensioning aspects need to be taken into account. This includes the deployment density, connectivity, and performance of backhaul links which are required to satisfy the key performance indicators of iJOIN, i.e.

- Energy-efficiency: for instance, if eNBs are turned off, energy saving potential is exploited in the radio access network. However, this also impacts the backhaul network which needs to provide sufficient diversity in order to allow for exploiting energy saving potential in RAN and BH network.

- Cost-efficiency: for instance, deploying a very dense small cell network will also require a very dense backhaul network able to carry the required data rates. This requires a dimensioning of the backhaul network that is cost-efficient but still allows for high data rates.

- Utilization-efficiency: for instance, RAN and backhaul need to be utilized optimally instead of being dimensioned for peak-throughput. This requires methods that exploit diversity effects and efficient load balancing.

- Spectral efficiency: for instance, available backhaul resources should be used as efficiently as possible in order to reduce costs and increase system capacity.

In addition, future deployments will be very diverse, i.e. as much as services evolve, also the scenarios and use cases will further diversify. Therefore, iJOIN focuses on four different scenarios, i.e. dense hotspot in a stadium, dense hotspot on city plaza, wide area coverage with high data rates, and dense indoor deployments. All these deployment scenarios will require different network management functions and abilities, which are further outlined in this document.

 

# 2 Executive Summary

This document first introduces the state of the art within those topics that are of particular relevance for this work package, i.e. mobility management, load aware network architecture, self-organized networks, software defined networking, and traffic engineering. The analysis of the state of art not only points currently deployed protocols but also outlines their deficiencies with respect to the use cases considered by iJOIN.

Based on the analysis of the state of the art, this document introduces a set of candidate technologies which will be further investigated in the course of this project. This investigation will eventually lead to the selection of the most promising technology. Each technology candidate is introduced first with a brief motivation. Furthermore, the general assumptions and technical requirements which are imposed by each technology candidate are listed, and a brief description of each technology candidate is provided. For the moment, no results or detailed investigation is provided but rather a sketch of upcoming analysis within this work package. In particular, the following technology candidates are introduced:

- Distributed IP anchoring and mobility management
- Network-wide energy optimization
- Joint path management and topology control
- Routing and congestion control mechanisms
- Network wide scheduling and load balancing
- Backhaul analysis
- Software Defined Networking (SDN)

Finally, this report provides an overview of harmonized assumptions resulting from the individual technology candidates. In addition, the integration into the iJOIN architecture as described by work package 5 and the functional split are outlined. Both parts are particularly relevant to ensure consistent approaches for the envisioned iJOIN framework.

 

# 3 State of the Art

## 3.1 Mobility Management in 3GPP

Within 3GPP multiple mobility approaches are standardised which aim at coping with UEs. This covers both mobility within a given access network technology (e.g., intra-3GPP) and also between different radio access technologies (RAT).

### 3.1.1 IP Mobility

In order to integrate a heterogeneous set of access technologies, mobility can no longer be considered an intra-technology issue managed at link level. New network-layer mobility functions need to be introduced in order to extend mobility capabilities beyond the link-based procedures available at each access technology. The mobility solutions can be adopted at two different levels, depending on the terminal's degree of involvement in the mobility process: global or host-based mobility and local or network-based mobility.

Global or host-based mobility is based on the ability of the terminal to maintain a persistent and globally accessible address independent of its current point of attachment to the network. As its name indicates, it requires the terminal's awareness and involvement in the mobility process. Usually, a mobility management protocol is applied to map the persistent address to the temporary local address available at each moment. In further modifies accordingly the end-to-end routing of packets in order to maintain connectivity. This is addressed by Mobile IPv4/6 protocols [39][40] by IETF. Apart from host involvement, it also requires a Global Mobility Anchor (GMA) entity in the home network, where the permanent address of the terminal is attached and the mapping with the temporary address is kept.

Local or network-based mobility is intended to allow terminals to maintain connectivity when moving across a certain network area, known as local mobility domain. This type of mobility is provided by the network with no need of terminal intervention or awareness. Local mobility is based on the presence of two network entities: the Mobility Access Gateway (MAG) and the Local Mobility Anchor (LMA). The MAG is in charge of tracking terminal location and updating it on the LMA on behalf of the terminal by means of a mobility management protocol, whereas the LMA is devoted to keeping the terminal's persistent address and routing inbound and outbound traffic.

Current IP mobility protocols (GTP [1], PMIPv6 [2] and DSMIPv6 [3]) are centralised and require all user data traffic to traverse the mobile operators' core network. This implies several limitations [4]:

a) Sub-optimal routing: Mobile nodes are anchored at a central entity that leads to IP routing-paths that are generally longer than necessary.

b) Scalability problems: core networks are dimensioned to support peak data traffic.

c) Reliability: the central entity/core network is a single point of potential failure.

d) Lack of fine granularity of the mobility management service.

Within the IETF, a working group on Distributed Mobility Management (DMM)[1] has been chartered in March 2012 which works on a more generic framework for distributed mobility management. The group is currently working on identifying challenges and the scope of potential solutions [10][11]. Most of the solutions presented within the IETF community re-use existing definitions and operations specified for MIPv6 and PMIPv6, respectively for client- and network-based solutions. The draft [12] modifies the MIPv6 home agent in order to deploy it in each access network as default router for the mobile node (MN). This allows the MN to have several anchors, enabling better path establishment and improving handover latency. The authors in [13] introduce two possible approaches as network-based solutions, i.e., a partially distributed solution, where a central database supports mobility-featured access routers as a mobility sessions store, and a fully distributed solution, where the central mobility database is removed. In [14], the central session server is maintained, but the signalling is changed. These solutions provide dynamic mobility and fast handover management, but standard access routers need to be enriched with mobility functionalities inherited by the PMIPv6's local mobility anchor and mobile access gateway. A different approach is presented in [15], where

[1] http://datatracker.ietf.org/wg/dmm/

 

the main idea is to deploy many small PMIPv6 domains and to define a signalling protocol for LMA-to-LMA communications. Many benefits from the well-known PMIP protocol can be gained, but the design does not meet the flatness envisioned for future network architectures.

An overview of the DMM impact in standardization, both for IETF and 3GPP is given in [16], while other non-standard related, but more generic, solutions were proposed in [17]-[24]. The articles [17]-[19] explore similar solutions as those in [12]-[13], focusing on extensions and giving more attention to use case examples for deployment. In the documents [20] and [21], authors propose and evaluate a network-based mechanism for DMM without using dedicated signalling. Similarly to the drafts mentioned above, mobility capable access router can re-configure the routing in the network to grant session continuity for a moving terminal. The mechanism relies on inspecting the traffic generated by the terminal. Furthermore, peer-to-peer strategies are evaluated in [22]-[23]. In [22]-[23], access routers still need some mobility features to maintain the mobility database and to anchor IP flows. The signalling is derived from P2P technologies, such as employing distributed hash tables and creating a P2P overlay. The authors of [24] propose to handle mobility management relying on routing and DNS updates, based on iBGP, BGP and dynamic DNS protocols. Opposite to the previous designs, here routers do not need modifications, but the performance of the location update is bounded by the convergence time of the routing protocols.

While these solutions can provide some gains over existing centralised approaches, they do not take into account the particularities of the iJOIN scenario:

- Current DMM efforts ignore very dense deployments, and this might have a significant impact, for example in how to properly select an anchor point, because changes of attachment points will be more frequent in a small cell based environment.
- Jointly optimising access and backhaul networks is also out of scope of current DMM efforts.

These aspects, however, imply additional constraints and challenges to mobility-related procedures, such as access discovery and selection. The current ANDSF framework was not designed to operate well in these scenarios. Other solutions such as IEEE 802.21 [25] would also need extensions to properly operate when the density of potential attachment points increases significantly.

### 3.1.2 Mobility in 3GPP

The UE mobility state in 3GPP systems is classified into two states: idle mode and connected mode [53][54]. In idle mode [54], the cell selection and re-selection is performed for the mobility management of UE. When a UE is turned on, the UE searches for a suitable cell and chooses this cell to provide available services and tunes to its control channel. This is referred to as "camping on the cell." If the UE finds a more suitable cell, according to the cell re-selection criteria, it selects the more suitable cell and camps on it, which is referred to as cell re-selection. When a call is initiated, the idle mode is transited to the connected mode.

LTE utilizes a network-controlled handover procedure which is assisted by the UE in connected mode [53]. The UE measures the signal strength and sends a measurement report to the serving eNB. The serving eNB then performs the handover decisions based on the measurement reports. The handover procedure consists of three steps: handover preparation, handover execution and handover completion.

At present, operators are looking for more distributed approaches that are cheaper and more efficient. The 3GPP is working on approaches that permit offloading traffic from the operators' core network. These will result in two main solutions: Selected IP Traffic Offload (SIPTO) and Local IP Access (LIPA) [5]. SIPTO enables an operator to offload certain types of traffic at a network node close to that UE's point of attachment to the access network.

In case of 3GPP access, the Serving Gateway (S-GW) terminates the GPRS Tunneling Protocol (GTP) interface towards the 3GPP radio access networks, and therefore it is the extended equivalent of the Gateway GPRS Support Node (GGSN) of previous 3GPP releases. In addition, the S-GW provides IP routing features and takes the MAG role for the provision of network-based mobility. Conversely, the Packet Gateway (P-GW) provides the interface towards PDNs and takes the LMA role. Local mobility management between the S-GW and the P-GW may be based on GTP or on Proxy Mobile IPv6 (PMIPv6) [41], which is a local mobility implementation based on the MIPv6 protocol. There are also plans for global mobility support under 3GPP access, thus making the P-GW take also the GMA role and basing global mobility management between the User Equipment (UE) and the P-GW on Dual Stack MIPv6 (DSMIPv6) [42], which is a MIPv6 variant for dual stack (IPv4 + IPv6) hosts.

LIPA enables an IP capable UE that is connected by means of a femto-cell to access other IP capable entities in the same residential/enterprise IP network. In order to achieve this, a Local GW (LGW) that is collocated with the femto-cell is used. Both SIPTO and LIPA have very limited mobility support, especially in 3GPP specifications of up to Rel-10. In Rel-11, a work item on "LIPA Mobility and SIPTO at the Local Network (LIMONET)" [6] is studying how to provide additional, but still limited, mobility support to SIPTO and LIPA mechanisms, albeit mainly restricted to a localised area and requiring PDN connections to be deactivated and re-activated when not moving locally.

Furthermore, the research community has proposed extensions to current 3GPP mechanisms, such as [7] and [8]. These works mainly deal with relocating the P-GW in order to contrast the sub-optimal path that data packets need to take to traverse the P-GW. Authors of [9] propose instead to introduce an additional entity in the EPS called Distributed GW (D-GW), specifically designed to handle mobility in a distributed way, taking into account also non 3GPP access.

### 3.1.3 Mobility Management in Small Cell Networks

The authors of [43] propose two mobility management schemes applied to the Femto-GW at Radio Network Layer (RNL) for LTE Femto-to-Femto handover. The first proposal suggests that the Femto-GW could act as a mobility anchor which makes handover decisions. When the Femto-GW receives a handover request from the source cell, it checks the target cell ID. If the target cell is under its control, it will handle the handover directly. By contrast, the second proposal describes a Femto-GW which operates as a transparent node which simply forwards all handover messages between the Femto cell and MME. After handover completion, the S-GW is notified about the change of the attachment point. The first proposal is more suitable for enterprise use, because it reduces the signalling traffic within the core network. On the other hand, the second proposal is more suitable for home use, because more signalling messages are exchanged.

In [44] and [45], an adapted signalling flow is proposed for the three types of handovers in heterogeneous LTE networks, i.e. handover from macro cell to small cell, handover from small cell to macro cell, and handover between small cells. The proposed scheme considers the movement prediction mechanism as an additional parameter for handover decision, which effectively makes it a client-based handover. Reactive and proactive handover procedures are proposed to trigger the handover, because the handover procedure may be initiated by femto-cell, macro/micro-cell, and the UE. In reactive handover, the handover is trigged when the UE almost lost its serving cell signal or the most likely position of the UE can be predicted. Reactive handover aims to postpone the handover as long as possible to prevent frequent and unnecessary handover, and mitigate the generated overhead of handover. However, in proactive handover, the handover may occur at any time before the handover-condition is met, e.g. via estimate of the exact position of the UE. Proactive handover is expected to minimize packet loss and latency during handover.

In [46], a method for access and handover management for OFDMA femto-cell networks is proposed. In a Closed Subscriber Group (CSG) scenario, when a UE comes near a femto-cell, its serving macro-BS will check the UE's ID. If it is within the allowed list, the macro-BS informs the femto-BS to start the handover procedure. Otherwise, the macro-BS should notify the femto-BS to start a proactive interference management procedure. The authors also propose a hybrid access to the same scenario. After a non CSG UE enters a femto-cell, the cell measures the UE's signal strength and decides whether the potential interference caused by the UE is above the interference threshold or not. If so, the femto-BS will request a handover procedure from the serving macro-BS for the UE and indicates that this is done to avoid interference. The CSG scenario reduces unnecessary handovers and signalling load. However, in the hybrid scenario, the number of HOs is increased.

To solve the same problem as presented in [46], [47] proposed a pseudo handover based on the direct information exchange between base stations. This exchange includes sub-channel and power adaptation information in order to avoid excessive interference. The pseudo handover is executed in the Radio Access Network (RAN) and does not imply any signalling exchange with the MME in order to reduce signalling overhead significantly. When the UE tries to camp on a CSG FBS where it belongs to, the regular handover is triggered. Otherwise, the pseudo handover is triggered. The FBS will set up and maintain a table all non-CSG UEs that tried to camp on it (and executed a pseudo-handover).

 

Although there are many proposed solutions for handover management in small cells, most of the solutions have targeted only one or two parts of the handover procedure, such as handover preparation, handover decision parameter, handover signalling, macro cell to small cell hand-in algorithm. In [43], [44], [45], [48], [49], [50], [51], and [52], schemes are proposed for the signalling flow during the handover process with different additional parameters used to reduce the number of unnecessary handovers. For instance, [43] supports CSG and OSG scenarios for the handover between femto-cells. The scheme uses the user speed, QoS, and load balancing as additional parameters for the handover decision. A comprehensive handover procedure is required to achieve optimal handover performance.

## 3.2 Load Aware Network Architecture

Base stations in networks are dimensioned to be able to support peak data-traffic. Hence, they are underutilized during low traffic periods [26]-[28]. BSs that operate with low or zero loads are almost as expensive as those running at full capacity. Hence, mobile operators see an opportunity to reduce their OPEX by deploying smart algorithms that allow dynamically switching BS on and off to adapt to the current load without impacting users. So far, most research activities have focused on very simple algorithms, such as static planning of base stations that are switched off at night. Very dense deployments provide a cost-efficient solution to increase network capacity. However, they have not been considered in detail. Since 70% of the small cell power consumption is static (i.e., does not scale with resource utilization [29]), sleep mode is an essential tool to bound network wide energy consumption of small cell deployments during off-peak periods. An overview is provided in [30]. The ability to turn off base stations relies on a highly adaptable backhaul which is able to adjust itself to these variations in network topology. Therefore, investigations of the limits of current backhaul capabilities and an identification of their shortcomings in handling such situations is essential and worthy of pursuit. In the current state of the art technologies, the design and optimization of a highly adaptive backhaul is a topic which hasn’t been investigated in great detail. This aspect is strengthened by investigations into backhaul dimensioning, path management, and topology control. An examination of admission and congestion control mechanisms is also necessary in order to gauge the impact of turning off BSs in networks with dynamic variations in traffic. The results of some of these investigations might also require a re-examination of existing energy optimization methodologies.

## 3.3 SON in Backhaul Networks

Self-optimized Networking (SON) refers to the automated operation of networks, in particular mobile networks [31]. The objective of SON is to increase the automation of RAN operation including installation, maintenance, and fail recovery. Among the best known and explored SON functionality are automatic inventory, software download, neighbour relation, cell id assignment, CCO, mobility optimization, RACH optimization, and load balancing. SON can be divided into different phases:

- Network planning, topology management and documentation,
- Service provisioning and change management,
- Service testing and verification,
- Reporting and network maintenance, and
- Fault and availability management.

So far, SON only played a role to manage the radio access network efficiently but it was mostly disregarded for the management of the backhaul network. SON may be part of the solution to provide the required backhaul capacity for 4G networks, as backhaul networks become a significant bottleneck. SON functions for backhaul networks include [31]:

- Planning: Network configuration and backhaul dimensioning.
- Deployment: Configuration, testing and verification including side-effects, i.e. interaction of neighbouring network elements in RAN and backhaul network.
- Optimization: Adaptation to traffic characteristics including concepts which may anticipate bottlenecks based on pattern analysis.

 

- Maintenance: Fast problem identification and recovery due to "avalanche risk," i.e. failure of one backhaul node or link affects a set of base stations. Also fast integration of new technology is considered.

SON will be part of the solution to implement the convergence towards Ethernet/IP based backhaul networks, which is required for the RAN evolution towards 4G and later technology eras.

## 3.4 Traffic Management in Backhaul Networks

As networks become more complex and traffic diversity increases, there is the apparent need to manage the traffic carried by the network. Traffic management has the goal to maximize the utilization of the network by minimizing the maximum link utilization as well as the objective to avoid congestion in the network. One means is to automate path selection procedures and to optimize traffic utilization because only adding new capacity does not suffice but it is necessary to manage these resources efficiently.

The most widely used protocol for path selection is Open Shortest Path First (OSPF) which assigns weights to links and computes the shortest path across the network (or the path with lowest weight). Even small changes to link weights may have an avalanche effect causing significant impacts to the network [32]. Among others, the selection of the optimal path should take into account QoS as well as dynamic metrics such as available bandwidth, delay, reliability, jitter, and mobility aspects (such as proximity to nodes that can act as anchor/offloading points for certain types of flows).

An alternative to OSPF is Multiprotocol Label Switching (MPLS) which creates explicit paths through an IP network. Using small labels which identify the route through the network, MPLS allows for fast routing and is able to handle heterogeneous networks (which requires a tight integration of MPLS and the underlying physical network). Computing the optimal route is computational complex and may not be done online, e.g. using Genetic Algorithms [33]. Hence, a combination of offline and online algorithms may be used, as presented in [34]. The authors in [34] introduced an algorithm which uses the expected traffic matrix to derive an optimal network topology offline, and an online component applies small, local changes based on dynamic traffic requests. It is further possible to combine OSPF and MPLS as described in [35] where MPLS paths are used to distribute the traffic and OSPF is used locally which implies less frequent and less severe changes of link weights.

## 3.5 SDN in Backhaul Networks

SDN is defined as "a network architecture in which the network control plane is decoupled from the physical topology." But beyond this general declaration, several trends can be identified:

- Separation of hardware from software, i.e. choosing hardware based on necessary features and software based on protocol requirements.
- Logically centralized network control, which is considered to be more deterministic, efficient and fault tolerant.
- Automation: Separate monitoring, management, and operation.

SDN allows for making currently rather static networks more flexible by tailoring and optimizing specifically them for different use cases. This creates a dynamic network environment which is adapted to the needs of the applications running on top of it.

The SDN framework, as proposed by the Open Network Forum (and in a simplified way), is represented in Figure 3-1.

 

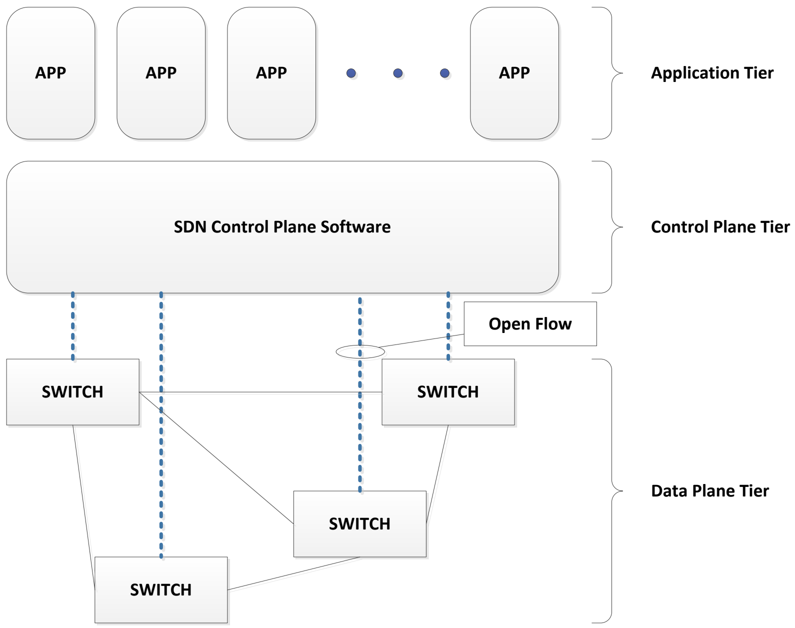


**Figure 3-1 SDN architecture as proposed by the Open Network Forum (ONF)**

The three tiers that compose the framework are the following:

- Application tier, e.g., virtual network overlays, network slicing (delegation), tenant-aware broadcast, application-aware path computation, integration with other software packages, policy, security, traffic engineering.
- Control plane tier, e.g., data plane resource marshalling, common libraries (e.g., topology, host metadata, state abstractions).
- Data plane tier, e.g., packet forwarding (as per flow table), packet manipulation (as per flow table), collection of statistics.

Between the application tier and the control plane tier a number of open APIs are defined, whose level of standardization is being explored by the Open Network Foundation (ONF). Between the data plane tier and the control plane tier the communications is carried out according to a standardized protocol, e.g. OpenFlow, which is promoted by ONF. OpenFlow allows direct access to and manipulation of the forwarding plane of network devices such as switches and routers, both physically and virtually (hypervisor-based). The protocol specifies basic primitives that can be used by an external software application to program the forwarding plane of network devices, just like the instruction set of a CPU would program a computer system.

Current access networks (last-mile backhaul networks) operate rather inefficiently which carries up to the aggregation network [36], i.e. the closer the aggregation network, the more efficient because diversity effects may be used (and need to be controlled). Furthermore, the move towards IP based backhaul networks also implies to move from fixed to non-deterministic bandwidth planning and management for continuously extending backhaul capacity.

In iJOIN, SDN is considered to tackle this challenge. SDN is applied to backhaul and radio access networks and allows for flexible adaptation of both, e.g. if different backhaul technologies are used and need to be controlled, or if physical links may be used for fronthaul and backhaul based on the functional split. This is in line with the backhaul requirements stated by NGMN in [37], which state that backhaul networks need a higher degree of configurability, e.g. granularity of information rates, resource sharing and prioritization of operators, traffic shaping, admission control, and load balancing. One could even consider SDN as part of RANaaS which applies the optimal functional split based on the underlying backhaul technology.

A complementary technology to SDN is Network Function Virtualization (NFV) [38] which refers to implementing and virtualizing network functionality on standard IT hardware. This is of particular interest for iJOIN where radio access network functionality is virtualized and should be accessible in a decentralized way on eNBs as well as in a centralized way in data centres. This functional split and flexible assignment requires virtualizing radio access network functionality, which is investigated within the iJOIN project.

# 4 Description of Technology Candidates

## 4.1 Distributed IP anchoring and Mobility Management

### 4.1.1 Motivation

This use case is motivated by the fact that in the past few years we have been witnessing an extraordinary data throughput explosion in cellular networks. Telecom operators have been carefully monitoring the disconnection between the average revenue per user (ARPU) and the associated cash costs per user (CCPU). Despite the remarkable volume increase of broadband data over mobile networks, mobile data revenue is falling fast.

This has some serious impacts on the dimensioning and planning of mobile networks. Specifically, we note that a) spectrum is limited and expensive, so available bandwidth for the access network cannot be easily increased; and b) deployed mobile core networks are highly hierarchical and centralized, which introduces serious scalability and reliability issues. The iJOIN project is tackling both, by looking at more densely deployed cells and enabling selected IP flows not to traverse the backhaul and mobile network operator's core. This is in line with several on-going 3GPP efforts, namely, the IP Flow Mobility and Seamless Offload (IFOM), the Local IP Access (LIPA), and the Selected IP Traffic Offload (SIPTO). It is also related to standardization activities within the IETF, namely the NETEXT working group.

As previously mentioned, highly hierarchical and cenralized mobile core networks introduce serious scalability and reliability issues. The iJOIN project also aims at mitigating this concern by investigating solutions that allow for distributing the data anchoring and mobility support. Note that this extends the previously described case, but with some key differences. In the case of local breakout and offload, the advantages come from the fact of selecting an offloading node for certain flows (unless there is an additional access technology deployed, such as WLAN, which is out of scope of the iJOIN project). This offloading node acts as anchor for the selected traffic, but it may happen that if the user moves, traffic cannot be seamlessly forwarded to the new location. However, in the case of a true distribution of the IP mobility management, the goal is to fully support user mobility.

Mobility Management in general is a set of tasks for controlling and supervising UEs in a wireless network to locate them for delivery services, as well as, to maintain their connections while they are on move. Mobility management is concerned with many aspects, such as Quality of Service (QoS), power management, location management, handoff management, and admission control. It is one of the most critical features in wireless communications due to the direct effect on user's QoE, network performance and power consumption. The core components of mobility management are location management, handoff management and the smart selection of network access.

The use of smaller cells is one of the approaches followed by iJOIN to increase the overall bandwidth capacity available to users. Whereas it is a well-known solution capable of providing significant enhancements, it also raises significant challenges in other areas, such as network selection and handover management. Decisions about which is the best radio access point to connect to are no longer mainly based on received signal strength, but they also need to take into account other multiple disparate aspects, such as: backhaul status, support for local breakout/offload, distributed anchoring, terminal/application specific aspects (e.g., mobility patterns, session lifetime, address continuity requirements, etc). A comprehensive solution is required to optimize the handover procedure, aiming at reducing packet loss, latency and minimizing signalling overhead as much as possible. Thus, fast and seamless user experience can be achieved. The mobility management design in iJOIN will go beyond improving network discovery and selection mechanisms, but improves signalling protocols. The availability of RANaaS will provide the possibility for enhanced mobility management. The available real-time load information and centralized control by RANaaS can enable more efficient load-aware handover management schemes, rather than the current 3GPP handover procedure. For example, handover decision can be optimized by taking into account the load information from neighbouring small cells.

### 4.1.2 Assumptions

 

Figure 4-1 shows a generic reference scenario for this use case (and other related IP mobility cases). With this regard, we consider the following general assumptions:

- The backhaul is a multi-hop IP network, meaning that there might be more than one node in the path between the radio head and the mobile operator's core, and that some (if not all) of these nodes are different IP hops. This, for instance, enables an easier integration of heterogeneous backhauling technologies. It will also allow for heterogeneous backhaul technologies, such as fibre and in/out-band connection between macro cell and small cells.
- Some nodes in the radio access network or in the backhaul might have local IP connectivity, which can even provide Internet access. These nodes can be used as offloading nodes for some selected flows, alleviating the load of the backhaul and core network.
- Some RAN and backhaul nodes are also able to provide mobility management (with low or none support from the core network). These nodes have control interfaces with the mobility entities in the mobile operator's core (e.g., MME, P-GW, HSS, etc.), i.e. including interface J2 between small cells for signalling exchange and J1 between small cells and RANaaS.

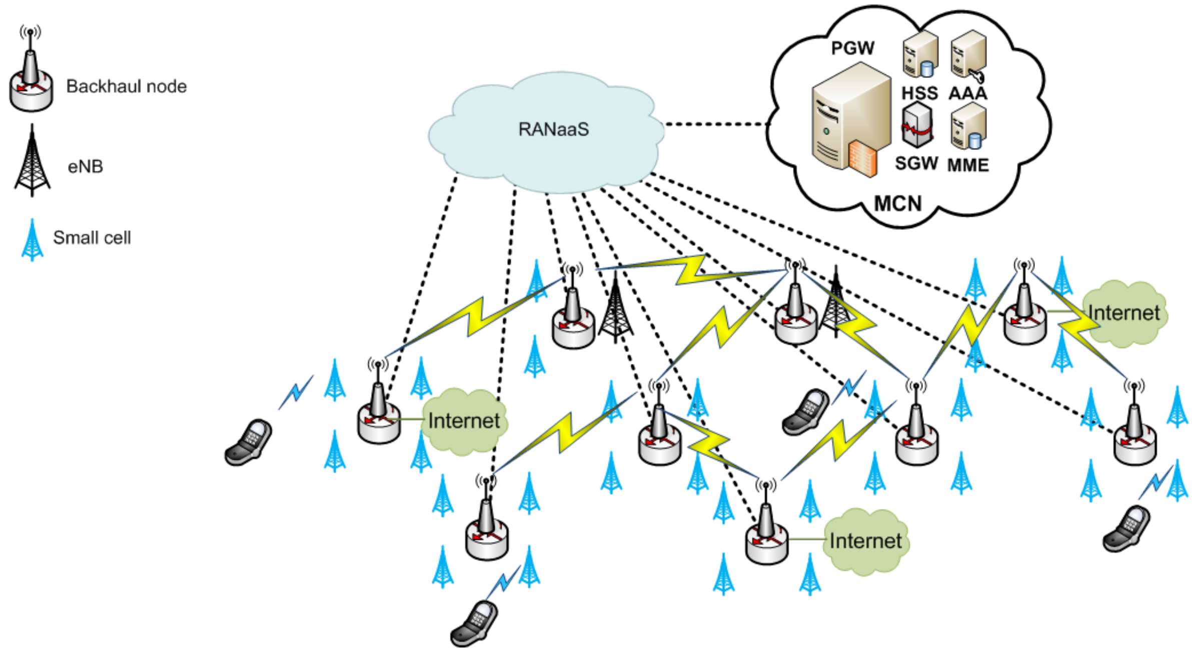


**Figure 4-1 Reference scenario for IP mobility related technologies**

### 4.1.3 Technical Requirements

Although there exist various works on IP offloading in the context of a generic network as well as 3GPP, the iJOIN architecture presents several particularities that require these solutions to be revisited or even reworked from scratch. Among these challenges, we can highlight the following: i) the backhaul network is dynamic and can self-configure itself to better adapt to the user traffic demand or to improve energy efficiency; ii) centralized processing in the cloud is available. Based on this, we list the main technical requirements of the solutions for this particular use case with respect to the network-layer:

- The solution should be IP based to allow its operation with different wireless backhauling technologies.
- The solution has to both consider the radio access and backhaul, and might need user terminal support. This would allow a certain user terminal for selecting the best radio access point based on whether local breakout is available (for the traffic the UE is sending/receiving).
- The designed mechanisms may interact with the backhaul routing function, so the path between the wireless radio head where the user is attached to and the offloading node can be dynamically set up and modified. This also allows performing energy optimizations.

 

- Backhauling capacity has to be taken into account, so the node selected to perform IP anchoring functionality can be selected based on the actual and current network load. Note that since the selected node has to remain playing the function of IP anchor while the user flow is alive, mobility considerations about where the UE might roam, should be also considered (and fed back to the routing function).
- The solution has to cooperate with the energy efficiency mechanisms as a decision of switching on/off a node might have an impact on mobility.
- The solution might be completely network-based, completely terminal-based or network-aided, user-based (hybrid approach).
- Not all the traffic might need mobility management (i.e. address continuity). Solutions should try to provide only mobility support to those that require it, and just offload (if possible) traffic that can survive an IP address change. This is actually related to the previous use case, as it might be the case that the networks prefers to switch certain flows to a different anchor/offloading node, because of backhauling capacity reasons, and it might be more efficient to do it for applications that can cope with an IP address change on their own.
- Handover preparation and execution mechanisms should be very fast in order to minimize packet loss.
- Solutions need to be applicable to overlay scenarios where macro and small cells co-exists and small-cells may leverage the support of macro-cells.
- Support of virtual cells as illustrated in Figure 4-2. The formation of a virtual cell, i.e., a cluster of cooperating and logically grouped small cells, appears to the UE as a single cell. In this case, handovers would occur only at the virtual cell boundaries. The handover within a virtual cell could be solved locally.
- Handover between virtual cells can be addressed using IP mobility mechanisms.
- Cooperation between virtual cell and RANaaS to enable load-aware handover decision.

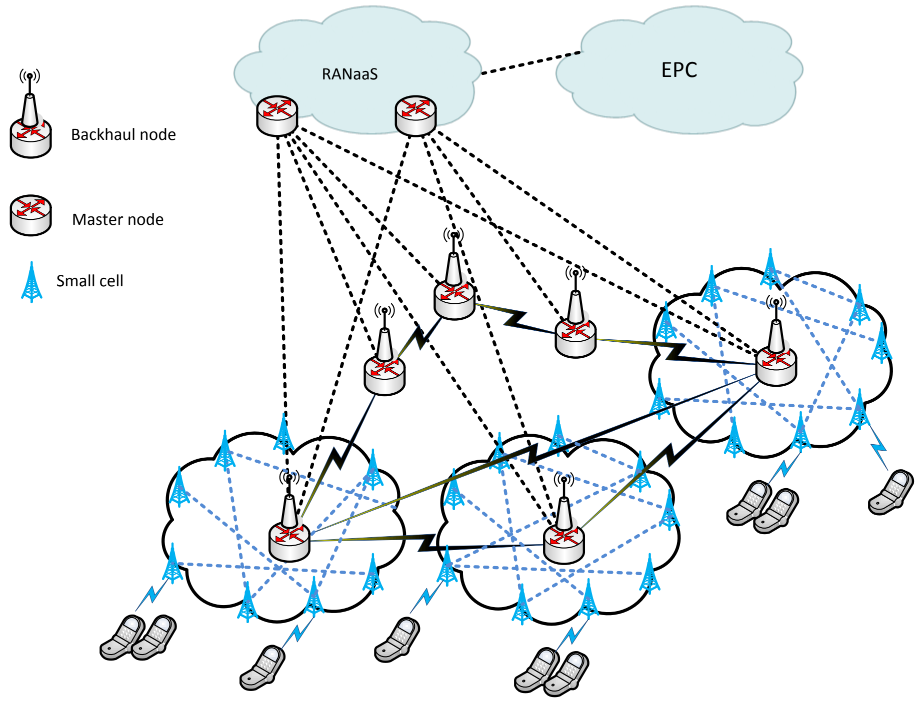


Figure 4-2 Mobility management: concept of virtual cells

 

### 4.1.4 Description of Technology Candidate

This technology candidate explores two paths which provide two different solutions which are, nevertheless, harmonized: Distributed Mobility Management and the Virtual Cell Concept. Both will be described in further detail in the following.

**Distributed Mobility Management**

Figure 4-3 shows an outline of the solution which will be further developed within the iJOIN project in order to provide a dynamic IP distributed mobility support with offloading support. The key aspects of the solution are:

- The network has multiple nodes in the access and in the backhaul that can perform the role of IP mobility anchors or offloading nodes. The difference between these is that the access network can provide additional mobility support if the UE moves away of its area of influence, while the backhaul network cannot.
- The UE and the network, upon initial start-up, and on an application basis, select the best radio access point of attachment and anchor for its traffic.
- If the UE moves and changes its serving cell (i.e. attaches to a different radio access point), some of the existing flows might need to be provided with mobility support. This is done by dynamically establishing/updating tunnels between the current radio access point and the original anchor of each IP flow. Note that this requires interaction with the backhaul routing function, to ensure that each flow is guaranteed the necessary quality of service. In some cases, tunnelling can be avoided if the routing function is capable of performing the required traffic redirection.
- The required control signalling to enable this dynamic and smart IP anchoring functions will benefit from the logically centralized cloud infrastructure to which all nodes have access to.

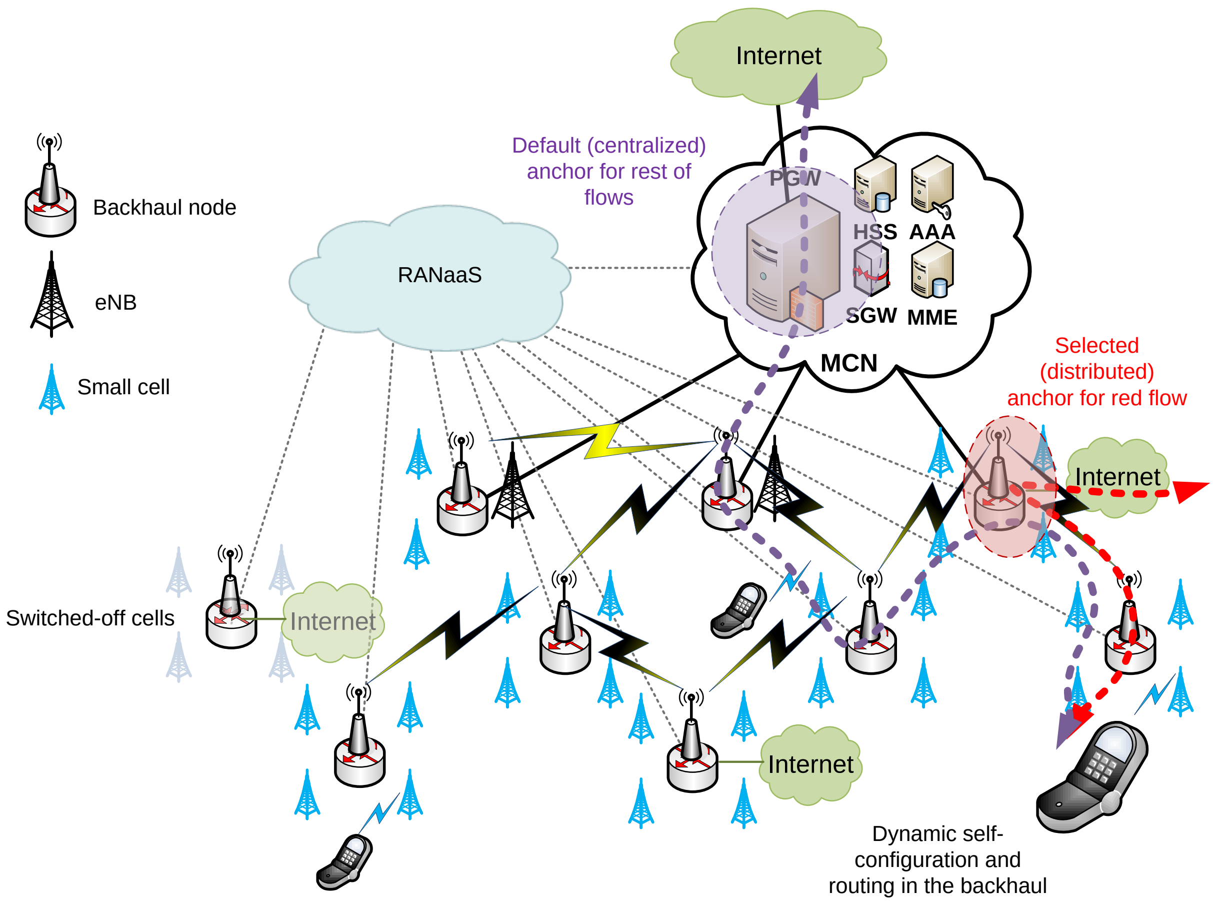


**Figure 4-3 Outline of the DMM solution**

 

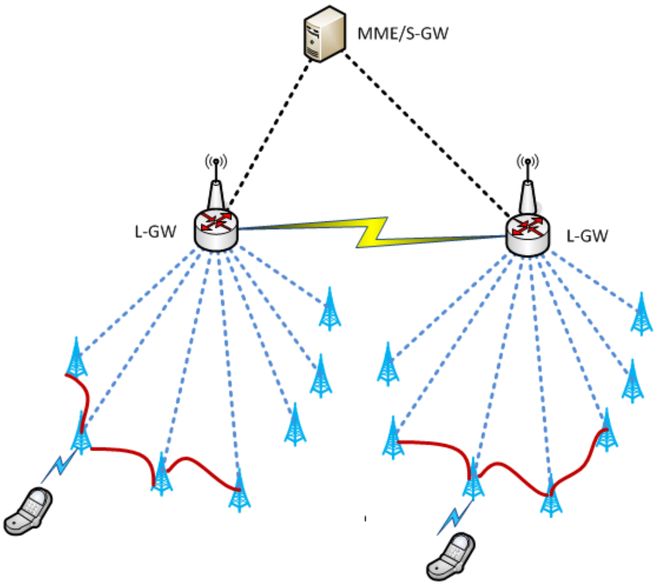


**Figure 4-4 Outline of the mobility management solution**

**Virtual Cell Concept**

Figure 4-4 shows an outline of the mobility management solution which will be further investigated. The key aspects of this solution are:

- Small cells are connected to the core network through multiple backhaul transport nodes. The introduction of L-GW provides the possibility of localized mobility management for handover between small cells within one L-GW domain. The L-GW is located between small cells and the mobile core network to act as the mobility anchor point for inter-small cell handover.
- The formation of virtual cells, a cluster of cooperating small cells, that appears to the user as a single cell. In this case, handovers would occur only at the virtual cell boundaries. The handover within a virtual cell could be solved locally. One virtual cell is managed by the L-GW accordingly.
- The solution investigates the handover signalling for three handover phases: handover preparation, handover execution and handover completion. To minimize the data lost during handover, the traffic forwarding scheme between small cells will be enabled. The traffic forwarding and path switch mechanism will be investigated in the handover execution phase and handover completion phase accordingly.
- The traffic forwarding scheme looks beyond the shortest path traffic forwarding and traffic forwarding with a threshold schemes, by taking into account traffic load conditions along the forwarding chain. Since the local traffic forwarding may increase the end-to-end communication latency and consume local resources, traffic load conditions along the forwarding chain should be considered to balance the trade-off between the path switch cost and traffic forwarding cost.

The solution investigates handover between small cells in two different scenarios: within one virtual cell where the handover can be handled locally by the L-GW and between virtual cells where the handover performance is enhanced by cooperation between neighbouring L-GWs.

## 4.2 Network Wide Energy Optimization

### 4.2.1 Motivation

Energy efficiency is both ecologically and commercially important to Information and Communication Technologies. Over 0.5% of the global energy consumption comes from wireless communication systems, mainly by outdoor cellular network BSs. A key challenge is to significantly reduce the energy consumption level whilst maintaining and even enhancing network capacity. Moreover, in order to improve competitiveness and the average revenue per UE, operators have to reduce OPEX of cellular networks.

 

Existing research on reducing the energy consumption of cellular networks has mainly focused on capacity improving transmission and RRM techniques, such as multi-user MIMO. Considering the total energy consumption of the Radio Access Network, the amount of energy saved by transmission and RRM techniques alone is fundamentally limited, while the energy saved by re-deployment can be much more significant.

According to a recent survey, nearly 80% of the energy consumption of a typical cellular network comes from the BSs. Furthermore, 70% of the BS energy consumption is caused by power amplifiers and air conditioning, which are used to keep the BS active even when there is no traffic in the cell. Hence, the optimization of BSs should have a large impact on the overall cellular energy efficiency. Energy efficiency can be improved from the following aspects:

- Offloading traffic from macro-cells to small-cells in order to be able to turn off macro-.
- Mechanisms that intelligently switch small-cells on/off for energy saving purpose. Energy saving for green networking is mainly realized by preventing cells from emitting at full power when there is no UE to serve.

The backhaul links will consume energy and some forwarding nodes on the backhaul links may be switched off based on the network utilization. SON techniques can be used to provide network-wide energy optimization. Each time the central server in centralized SON or the cell in distributed SON gathers fresh data, it can perform a new optimization if necessary.

### 4.2.2 Assumptions

Based on the previous introduction, we can state the following key assumptions for the technology candidate:

- Backhaul is an IP-based network
- Interfaces between small-cells and between small-cells and backhaul transport nodes,
- Availability of network topology and real-time energy consumption information for small cells and backhaul nodes
- Network topology and energy consumption (as a function of load) per node is known.
- Current utilization level of backhaul links and radio resource utilization at small cells are known.
- Spatio-temporal traffic profiles are known.
- SLA and users QoS requirements are known.

### 4.2.3 Technical Requirements

From the previous description, we can derive the key technical requirements which need to be fulfilled in order to implement the technology candidate:

- The small-cell BS and backhaul nodes should have the ability to invoke a low-power sleep mode when not required to serve any data traffic. The low power mode can be driven by small cells or core network.
- When a small cell is switched off or suspended to a low power sleep mode, the attached UEs need to be evenly assigned to neighbouring small cells. Both access and backhaul links should be able to support the newly assigned UEs.
- SON technologies for energy optimization, taking user mobility and traffic pattern into consideration.

  Small cell deployments and hierarchical deployments with overlay macro-cells may lead to a situation where many cells are barely loaded. In particular, this applies to situations where the load varies over different times of the day. In high load situations the best solution may be to provide coverage using many small cells, whereas in low load situations cells with only few users can be turned off by the network management. Self-organizing mechanisms and signalling protocols are required to detect traffic situations in order to redirect UEs and to adjust the network coverage.
- The solution needs to provide better energy efficiency performance and can therefore be possibly coupled with a mobility management approach. For example, when making a handover decision, the

 

small cells which are on low-power mode should be exempted from the neighbouring cell lists, thus they will not take any more UEs unless the small-cell is asked to return from low-power mode.

- Availability of a traffic profiling mechanism at RAPs.
- Fast and reliable handover mechanisms which ensure non-disruptive behaviour when RAPs are turned off.
- Traffic handover from one backhaul node to neighbouring backhaul nodes in the case that the energy saving algorithm indicates that a particular BN may be turned off.
- Traffic profiling mechanism might be improved with the support from UEs, i.e. UEs need to be capable of gathering statistics of their usage.
- Multi-tier energy saving algorithm, i.e. a backhaul node may decide to turn a RAP off in the case of low utilization.

### 4.2.4 Description of Technology Candidate

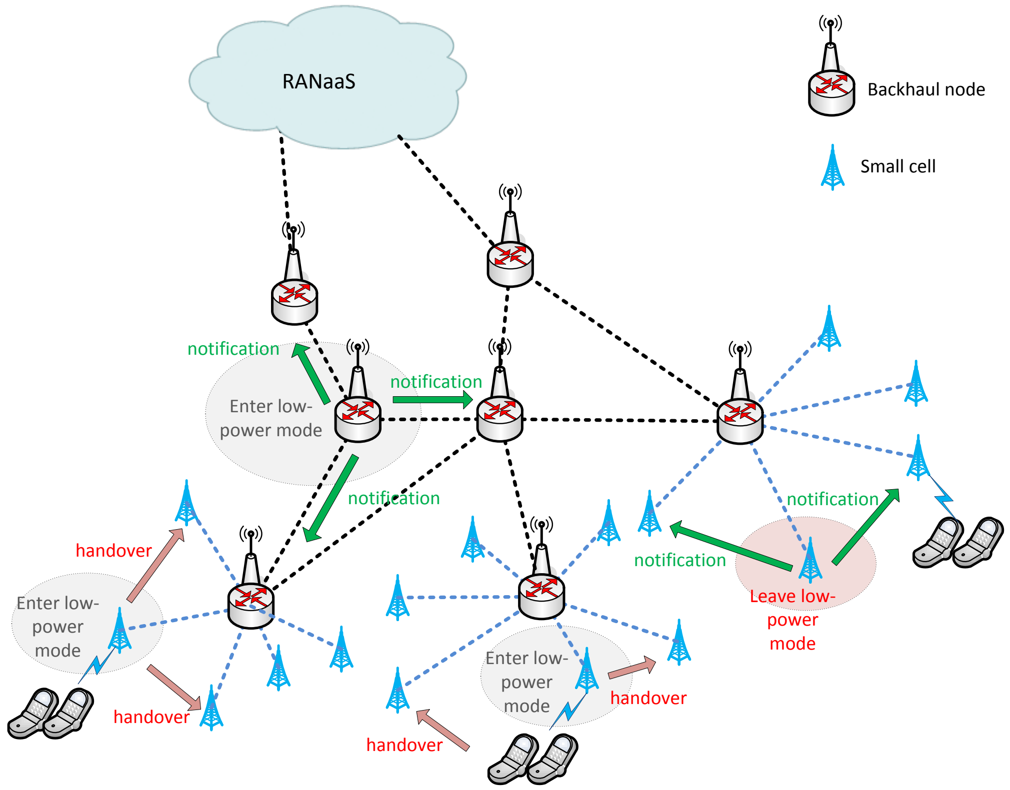


**Figure 4-5 Possible communication requirement of the energy optimization solution**

Figure 4-5 shows an outline of the energy efficiency solution to be further developed within the iJOIN project. In the following, we summarize the key features of the solution.

The energy efficiency solution for small cell deployment scenarios are based on cell/network load situation. The solution should guarantee user accessibility when a cell/backhaul node is transferred to low-power sleeping mode. The solution should not have negative impact on the UE power consumption.

 

**Energy Saving Procedures**

The energy saving procedures can be executed by different approaches: centralized approach, decentralized signalling approach and hybrid approach. In the centralized approach, small cells/backhaul nodes enter or leave low-power sleeping mode based on centralized decisions, which are made based on the real-time information obtained from the network, e.g. load information. The decisions can be either pre-configured or directly commuted to the small cells/backhaul nodes. If a small cell/backhaul node enters or leaves its low-power sleeping mode, the neighbouring small cells/backhaul nodes should be informed through signalling. In this case, the neighbouring cell list can be updated and routing decisions can be made correctly.

In the signalling approach, small cells/backhaul nodes may decide to enter low-power mode autonomously or based on information exchanged with neighbouring small cells/backhaul nodes. The small cells/backhaul nodes are aware of whether they are energy saving capable or not based on proprietary information, e.g., load information. When a small-cell/backhaul node decides to enter the low-power mode, it will initialise communication with the corresponding small cells/backhaul nodes, and related information may be included in the request message. The final decision is made after the signalling exchange. The enter/leave low-power mode decisions/requests will be based on information locally available in the small cells/backhaul nodes, including load information of the neighbouring nodes. Leaving low-power mode can be invoked based upon requests from the neighbouring cells/backhaul nodes, or the local policy available in the node, such as a pre-defined max switch off time. The neighbouring small cells/backhaul nodes should be informed after each on/off decision. And in order to perform energy efficiency in a more efficient way, some energy efficiency parameters might be exchanged between small cells/backhaul nodes if it is required, e.g., power consumption, traffic threshold and etc.

In the hybrid solution, the small cells/backhaul nodes are pre-configured by a centralized network entity, such as RANaaS. Also the RANaaS communicates to all small cells/backhaul nodes the values of some parameters that determine the behaviour of entering/leaving low-power mode.

**Exploiting Traffic Patterns**

In addition to the actual optimization approach, this technology candidate will also study different traffic patterns at RAPs which would allow for the definition of fine grain traffic profiles in the spatio-temporal domain. Such patterns would be estimated and identified at RAPs and backhaul nodes. Based upon the traffic profile, energy optimization algorithms may take the decision to switch off RAPs when they are underutilized and the shifting of traffic to neighbouring RAPs would not create bottlenecks. This would take into account the radio access resources at neighbouring RAPs and their backhaul connectivity. Such decisions may be taken by the energy saving algorithm running either at backhaul nodes or RANaaS, i.e. in this context we pay particular attention to a hybrid solution as explained before.

Similarly to scenarios with known traffic profiles, these algorithms may be equally applied to backhaul nodes. If a backhaul node is highly underutilized, energy management mechanism may switch it off under the condition that an alternative backhaul node with enough resources is available. This algorithm may be performed within the RANaaS. Finally, the solution will avoid ping-pong effects as a result of switching on/off RAPs and backhaul nodes.

## 4.3 Joint Path Management and Topology Control

### 4.3.1 Motivation

Dense small cell deployments are going to be subject to load demands that vary both in space and time. In addition, given the declining revenues experienced by mobile operators and the high number of small cells to be deployed, the small cell backhaul infrastructure should be as cost efficient as possible. Hence, dimensioning the small cell backhaul for peak traffic demands is neither a scalable nor a cost efficient approach. Instead, in iJOIN we envision that the infrastructure used to backhaul small cells should be able to adapt to varying capacity demands and to allocate backhaul resources where they are needed and when they are needed.

The iJOIN project has identified wireless backhauling as a critical enabler for small cell backhauling. Several wireless technologies are currently considered as viable candidates for small cell backhaul, both high capacity LOS technologies (e.g. 60GHz and E-Band) and lower capacity NLOS technologies (< 6GHz). Thus, the future small cell backhaul will likely be composed of a mix of heterogeneous technologies.

 

### 4.3.2 Assumptions

The following assumptions are imposed by this technology candidate:

- The leftover capacity in the current backhaul infrastructure used for the macro-network is reused to backhaul small cells. A typical current backhaul infrastructure for the macro network is composed of two to three levels of microwave aggregation trees connected to an optical ring.
- Additional fibre connection points might be deployed to backhaul small cells, i.e. some small cells may have a direct fibre connection.
- A centralized controller is available which is capable to manage the backhaul network and to modify the traffic flow on a per-node basis,
- In order to connect street level small cells to the macro cell backhaul (rooftop), a heterogeneous set of technologies may be used:
    - 60GHz LOS P2P links: 60GHz radios are installed in the rooftop macro-site and on street level together with small cells. This allows small cells with LOS to the macro cell, to connect to it directly. Other small cells may use multi-hop to the macro site. In order to achieve path diversity, a single small cell can be backhauled with more than one 60GHz P2P backhaul unit.
    - E-Band (70-90GHz) LOS P2P links which are licensed but achieve longer range than 60GHz links at similar data rates. E-Band links can be installed at rooftop level to connect macro-sites with each other.
    - NLOS technologies (< 6GHz) which offer lower bandwidth and some of them may be subject to interference, but may offer a solution to backhaul small cells with no LOS to other small cells and a macro-site.

### 4.3.3 Technical Requirements

The following technical requirements are set in order to implement the technology candidate:

- The backhaul network needs to provide sufficient path diversity in order to allow for traffic engineering solutions.
- The backhaul network allows for low configuration granularity, i.e., backhaul nodes may be configured to route traffic through different connections,
- Provision of long and short time statistics at eNB towards backhaul network, which requires an appropriate interface (possibly even on flow-level),
- Existing backhaul infrastructure may be reused and existing equipment may be upgraded to support the proposed technology candidate.

### 4.3.4 Description of Technology Candidate

Providing small cell backhaul with a certain degree of path diversity enables the application of traffic engineering techniques that dynamically adapt the transport paths used in the small cell backhaul to varying load demands. In the context of this technology candidate, we are going to study:

1. What are the key topological properties in the small cell backhaul that enable system-wide gains in a cost efficient way (e.g. tree or mesh topologies, number of fibre connection points, etc).
2. The system level design aspects required to appropriately manage dynamic load demands in the small cell backhaul (e.g. protocols that enable path reallocation, information available in RAN and backhaul that can be used to trigger path reconfiguration, etc),
3. The design of algorithms that dynamically manage the transport paths in the small cell backhaul.

In order to illustrate the previous concepts, Figure 4-6 contains an exemplary small cell deployment in a Manhattan type scenario. The figure explicitly depicts the backhaul links of both the small cell and macro cell networks, where we can see: i) rooftop level microwave links used to backhaul existent macro-cells, ii) attachment points to the metro fibre infrastructure collocated with some of the macro sites, iii) street level

 

60GHz point to point links used to backhaul small cells with each other (or with a macro site if there is LOS), and iv) rooftop level E-Band links that connect macro-sites with each other providing path diversity.

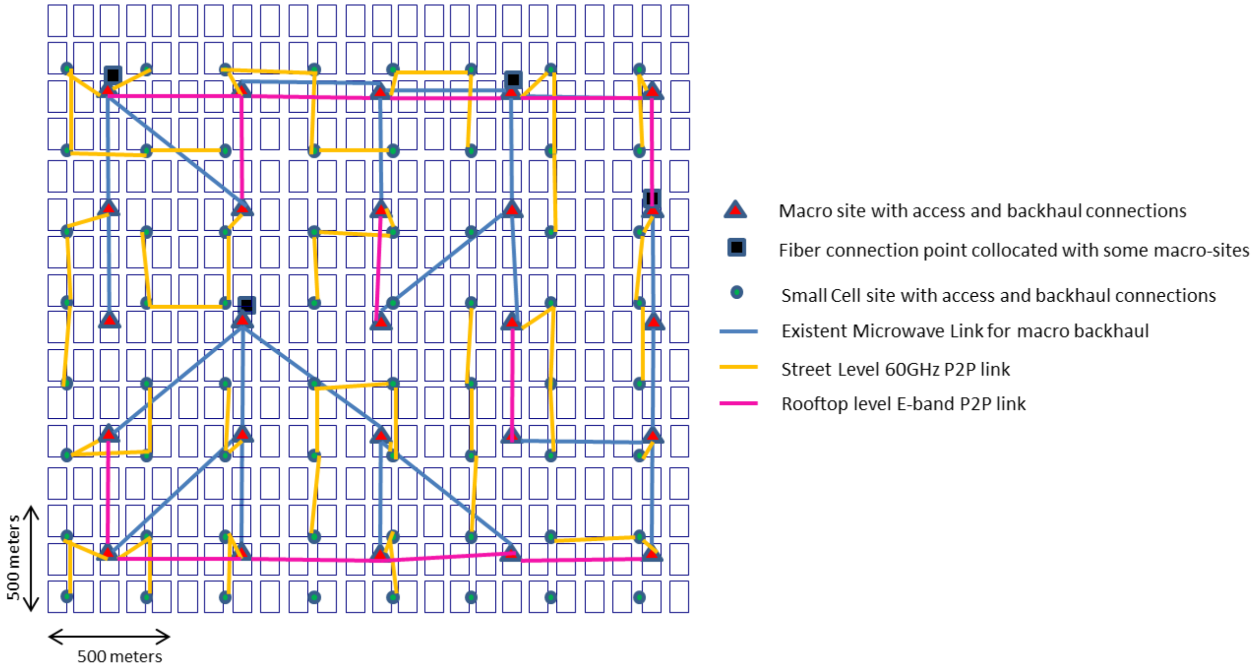


**Figure 4-6 Exemplary small-cell deployment and backhaul network**

In order to design effective traffic engineering solutions for the small cell backhaul we envision a two stage approach:

1. Study the topological properties that the small cell backhaul should fulfil. For that purpose there is a need to characterize realistic small cell deployments and realistic traffic models (with space and time dynamics). Given the previous models, we will consider the characteristics of wireless backhaul technologies (e.g. 60GHz, E-Band, NLOS) and study the system wide benefits of particular deployment strategies in the backhaul (e.g. tree, mesh, etc).
2. Given a small cell backhaul topology, we will study the applicability of the Software Defined Networking (SDN) paradigm to implement real-time traffic engineering in the small cell backhaul, considering:
   - A centralized controller that takes traffic engineering decisions for the small cell backhaul network (e.g. path allocation, load balancing, energy saving).
   - Small cell access and backhaul nodes that report relevant performance metrics to the centralized controller, so that this can take appropriate management solutions.

The following objectives will be particularly addressed by this technology candidate:

- Efficiently use leftover capacity in the existent macro backhaul, and any additional capacity especially deployed for the small cell network.
- Be able to adapt to traffic demands that vary both in time and space to avoid congestion in the backhaul.
- Be able to quickly restore backhaul links in case of failure.
- Support differentiated treatment of traffic aggregates in the backhaul.
- Interact with the RAN so that joint RAN and backhaul optimizations are possible (e.g. mapping of RAN QoS parameters to backhaul).
- Coexist and optimize the designed RAN mobility solutions (e.g. quick provisioning of X2 interface over the backhaul to prepare for small cell to small cell handovers).

 

- Enable application awareness, so that traffic engineering decisions can incorporate application knowledge.
- Enable virtualization and multi-tenancy, so that the small cell backhaul can be shared among different tenants.
- Gracefully degrade if not all small cell nodes are iJOIN capable.

## 4.4 Routing and Congestion Control Mechanisms

### 4.4.1 Motivation

The objective of congestion control is to allow network operators to simultaneously achieve high throughput and low average delay. This technology candidate addresses congestion control and analyses the impact of the support of advanced congestion control algorithms and their impact on the performance of the backhaul network. In this sense, we distinguish between active queue management (AQM) techniques deployed inside the network that, e.g., aim at reducing TCP sending rate by intentional packet drop, and low priority congestion control (LPCC) solutions, that intend to transfer at a lower priority by reacting faster to network congestion using indicators other than packet loss.

The deployment of small cells poses challenges on routing in the network layer, such as UEs performing handovers between small cells very frequently, traffic-aware portioning of small cells into either static or dynamic clusters causing congestions towards one central entity. Therefore, cooperation of small cells for an enhanced routing algorithm is required. Within this technology candidate, a novel routing algorithm will be investigated beyond the classic centralized/distributed routing algorithms. It is important to investigate the routing and admission/congestion control issues jointly to optimize the load distribution between small cells/backhaul nodes and provide seamless connectivity considering user's mobility between small cells. Routing algorithms should address the congestion control issues by avoiding traffic to be routed through the backhaul nodes which are already overloaded.

### 4.4.2 Assumptions

- AQM will be implemented by the operator as part of its backhaul infrastructure
- Backhaul is an IP-based network
- Small-cells are connected to the RANaaS through Backhaul Transport Nodes
- An interface is required between small cells for signalling exchange.
- Network topology and network utilization information for small cells and backhaul nodes are available upon request

### 4.4.3 Technical Requirements

- Delay measurements are performed in order to estimate the queuing delay. While the time synchronization itself is not relevant, the clock skew is of particular interest as it may imply that measurements are biased.
- Delays that are implied by scheduling need to be measured and taken into account.
- Delay measurements need to consider route changes and incorporate them.
- Mobile network congestion management should be configured in such a way that policy enforcement to overcome congestion is activated before end-to-end TCP congestion mechanisms begin to operate.
- ECN mechanisms should be enhanced to deal with the case of several operators sharing the backhaul network.
- Novel routing approach for small-cell deployments

The approach will exploit a cooperative routing algorithm for small cells deployments. The classic centralized routing algorithm relies on a central entity and can easily cause congestion towards the central entity. Improvements have been made in distributed routing algorithms to distribute resource consumption amongst network nodes whenever it is required. By proposing the novel routing approach, the backhaul nodes can cooperate with each other to optimize the routing decisions, by selecting the less congested routes. The routing algorithm is designed to avoid the heavily loaded backhaul nodes to improve congestion control.

### 4.4.4 Description of Technology Candidates

This technology candidate considers two parts which regard LPCC based congestion control as well as joint routing and congestion control. Both parts are separately explained but are meant to interwork, which is detailed in the course of the project.

**LPCC based congestion control**

This part of the investigation will focus on LEDBAT although the requirements derived can be extended to other LPCC mechanisms. According to IETF, Low Extra Delay Background Transport (LEDBAT) is an experimental delay-based congestion control algorithm that seeks to utilize the available bandwidth on an end-to-end path while limiting the consequent increase in queuing delay on that path. LEDBAT uses changes in one-way delay measurements to limit congestion that the flow itself induces in the network. LEDBAT is designed for use by background bulk-transfer applications to be no more aggressive than standard TCP congestion control (as specified in RFC5681) and to yield in the presence of competing flows, thus limiting interference with the network performance of competing flows.

LEDBAT employs one-way delay measurements to estimate the queuing delay which may indicate that a link is in congestion. End-to-end delay can be decomposed into transmission (or serialization) delay, propagation (or speed-of-light) delay, queuing delay, and processing delay. On any given path, barring some noise, all delay components except for queuing delay are constant. To observe an increase in the queuing delay in the network, a LEDBAT sender separates the queuing delay component from the rest of the end-to-end delay. The latter constitutes the base delay, which is the minimum delay that can be observed on the end-to-end path.

To respond to true changes in the base delay, as can be caused by a route change, LEDBAT uses only recent measurements in estimating the base delay. The duration of the observation window itself is a trade-off between robustness of measurement and responsiveness to change, i.e. a larger observation window increases the chances that the true base delay will be detected (as long as the true base delay is unchanged), whereas a smaller observation window results in faster response to true changes in the base delay.

Assuming that the base delay is constant (in the absence of any route changes), the queuing delay is represented by the variable component of the measured end-to-end delay. LEDBAT measures queuing delay as simply the difference between an end-to-end delay measurement and the current estimate of base delay. The queuing delay should be filtered (depending on the usage scenario) to eliminate noise in the delay estimation, such as due to spikes in processing delay at a node on the path.

LEDBAT can be used as part of a transport protocol or as part of an application, as long as the data transmission mechanisms are capable of carrying timestamps and acknowledging data frequently. LEDBAT can be used with TCP, Stream Control Transmission Protocol (SCTP), and Datagram Congestion Control Protocol (DCCP) with appropriate extensions where necessary. It can be further used with proprietary application protocols such as those built on top of UDP for peer-to-peer (P2P) applications.

It seems reasonable to assume that, although LEDBAT should be implemented at the edges of the connection, its main impact will be observed at the bottleneck link in the connection. And it can be assumed that, in a high percentage of cases, this bottleneck link will be either the radio interface or the last mile backhaul link to the eNB.

The interaction of the LEDBAT congestion control with the radio interface radio management functionalities is very difficult to assess. RRM mechanisms are expected to act at a different, much shorter time scale than end-to-end congestion control mechanisms. Furthermore, the implementation of some RRM functionalities is vendor dependent, which makes it complicated to determine the impact that they may have on LEDBAT performance.

For LPCC mechanisms, the technical solution requires to implement mechanisms that keep baseline delay as constant as possible when there are no congestion issues to be solved. They may also modify them in order to activate the congestion control mechanisms in a preventive way.

In iJOIN, an LPCC based solution is investigated which builds upon LEDBAT. The proposed solution will be implemented in the iJOIN Transport Nodes (iTNs). The basic functional blocks of this solution are:

 

- Detection: mechanisms for detecting which flows are considered as low priority and may be affected by LPCC mechanisms. It should be based on the relationship between the variable delay and the bit rate of the flow.
- Baseline delay equalization: mechanisms to keep the baseline delay as constant as possible. This can be done by packet inspection, looking at timestamps, or by means of leaky bucket type of mechanism.
- Baseline delay modification, in such a way that it can compensate events like handovers or activate congestion control mechanisms in order to prevent congestion situations. This can be achieved by properly modifying the parameters of the baseline delay equalization mechanism.

Active Queue Management (AQM) schemes like RED [1] or REM [2] randomly drop or mark packets before the buffer of a network nodes becomes full. Hence, TCP senders can be notified to avoid excessive growth of queues in buffers. AQM is meant to be a general mechanism using one of several alternatives for congestion indication, but in the absence of ECN, AQM is restricted to using packet drops as a mechanism for congestion indication. In this case, AQM drops packets based on the average queue length exceeding a threshold, rather than only when the queue overflows.

ECN based AQM schemes use the two least significant (right-most) bits of the DiffServ field in the IPv4 or IPv6 header to encode four different codepoints:

- 00: Non ECN-Capable Transport — Non-ECT
- 10: ECN Capable Transport — ECT(0)
- 01: ECN Capable Transport — ECT(1)
- 11: Congestion Encountered — CE

If both endpoints support ECN, they mark their packets with ECT(0) or ECT(1). If the packet traverses an AQM queue that is experiencing congestion and the corresponding router supports ECN, it may change the codepoint to CE instead of dropping the packet. This process is referred to as "marking" and its purpose is to inform the receiving endpoint of a potential congestion. At the receiving endpoint, this congestion indication is handled by the upper layer protocol (transport layer protocol) and needs to be echoed back to the transmitting node in order to signal it to reduce its transmission rate.

The use of ECN based AQM for mobile networks it is being explored in the context of the 3GPP Release-12 Study Item UPCON (User Plane Congestion) [4]. The objective of UPCON is to improve resource efficiency in the network and to increase the number of active users while maintaining good user experience (QoE).

The way that the ECN procedures would be supported in a mobile network is illustrated in Figure 4-7.

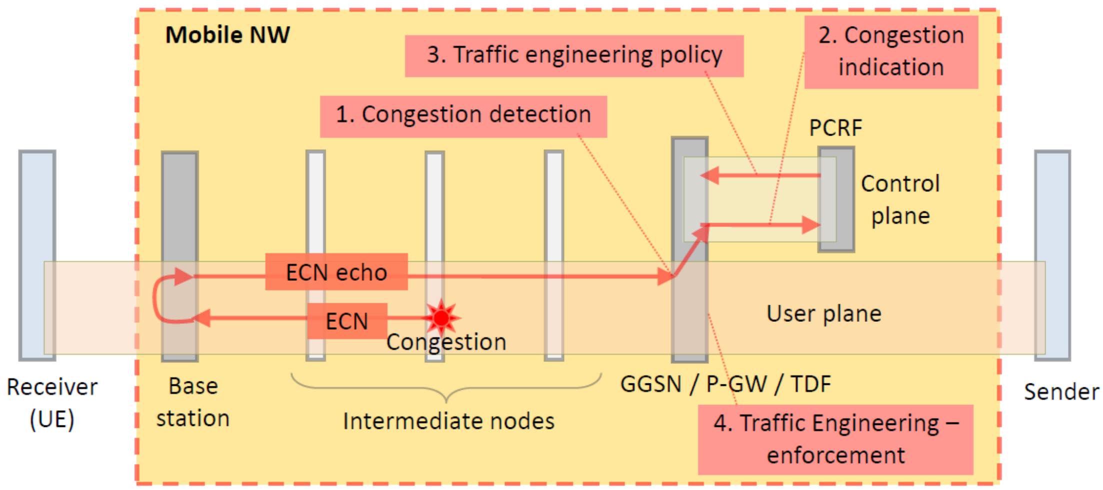


**Figure 4-7 Congestion control in a mobile network**

 

This represents a different approach to the standard one, where ECN functionalities are implemented by the edge devices. In the mobile network they are most likely to be implemented in the eNodeB and the P-GW. On top of this, as the LTE network may use two tunnelling protocols, Proxy-Mobile-IP (PMIP) and GPRS Tunneling Protocol (GTP), Congestion Indication should be based on GTP/PMIP-level ECN-ECHO.

There are also other differences: whilst the traditional approach in IP networks is to allow TCP congestion control mechanisms to deal with congestion situations, in the approach supported in UPCON congestion is dealt by means of traffic engineering enforcement procedures in the P-GW (based on the policy established by the PCRF). This approach has the advantage of taking into account not only the congestion level, but also the subscriber's profile, when implementing remedial procedures.

As indicated above, different problems should be tackled with depending on the kind of congestion protocol mechanism that is being implemented.

In iJOIN, a solution based on AQM is investigated which is based on traffic policing and assuming that the proposed iJOIN architecture enables more sophisticated congestion control in the backhaul network. Upon detection of congestion there are a number of alternative actions that may be pursued, such as the use of alternative routes to offload the congested node, using, for example, multicast TCP/IP as illustrated in Figure 4-8, the use of SLAs to prioritize traffic flows, and the activation of mechanisms that may reduce overhead, for example, header compression, use of different security mechanisms (MACsec instead of IPsec), packet aggregation of no delay sensitive flows, etc.

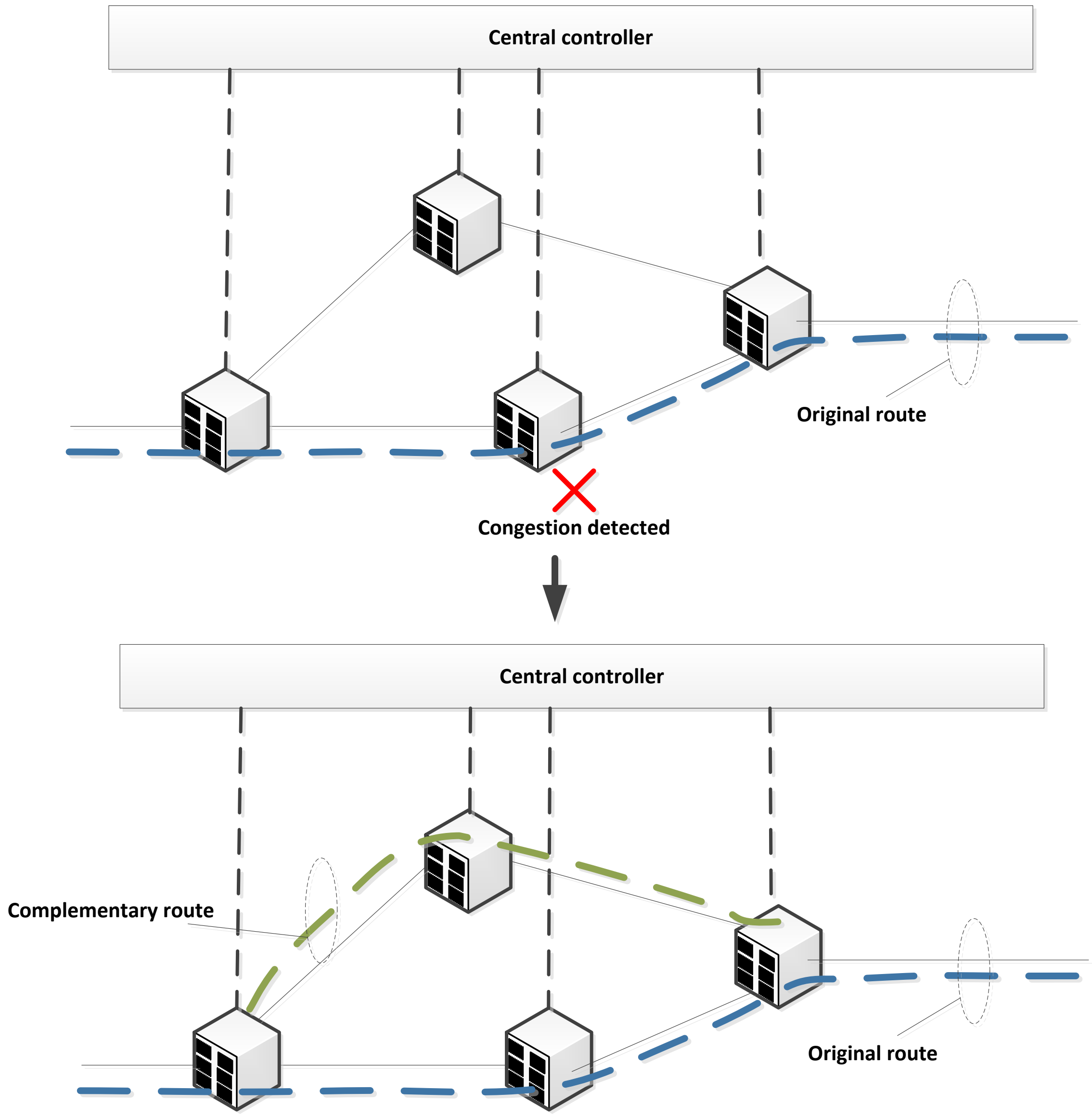


**Figure 4-8 Use of alternative routes for off-loading congested nodes**

 

**Joint routing and congestion control**

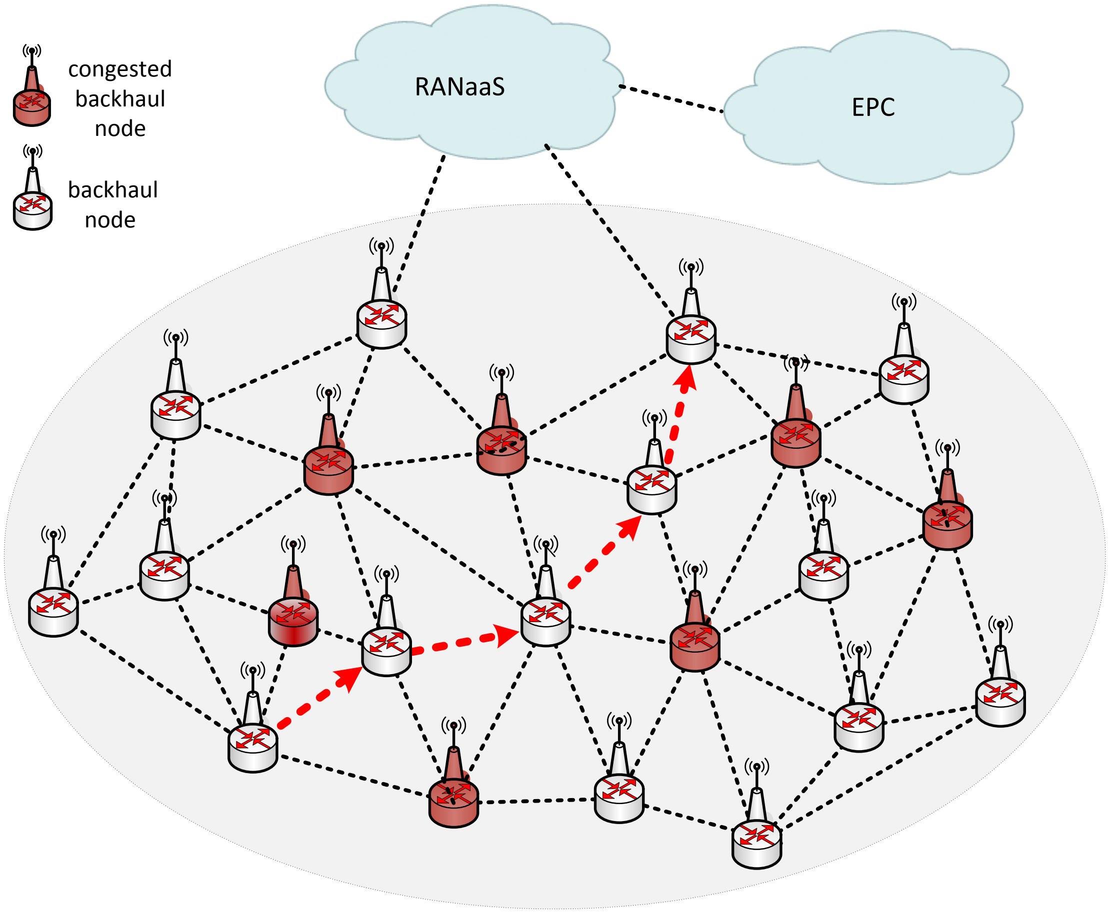


**Figure 4-9 Use Outline of joint routing and admission/congestion control**

Figure 4-9 shows an outline of the joint routing and admission/congestion control solution which is further investigated within the iJOIN project. We summarize the key features of the solution as below:

- It can be considered a mesh base approach, indicating that multiple backhaul nodes and paths are available between small cells and the EPC. However, the proposed solution goes beyond existing mesh-based approaches exploiting queue length and geographic information.
- The backhaul nodes cooperate with each other by exchange of signalling information. The information of current network utilization conditions of neighbouring backhaul nodes should be available locally within every backhaul node. This can be achieved either by centralized approach provided by RANaaS or by initiating requests to neighbouring backhaul nodes. In the centralized approach, the RANaaS should be able to learn the network topology and load information for backhaul nodes and commute relevant information to backhaul nodes accordingly. Otherwise, every backhaul nodes can request the information by signalling the neighbouring nodes.
- The solution selects the neighbour based on the principle of minimizing the congestion towards the EPC. When taking forwarding decisions at a given node, instead of just considering the shortest distance vector as in traditional routing algorithm, the load information in the neighbouring backhaul nodes is also taken into consideration. As shown in Figure 4-9, the optimized route is selected so that the congested backhaul node is avoided.
- The solution might be a proactive routing approach. All backhaul nodes maintain a routing table that contains separate entries for all the possible destinations, which need to be periodically updated. There might be a scalability problem if there are a lot of backhaul nodes towards the EPC. Therefore,

 

it is important to find a trade-off between the cooperative routing approach and the traditional IP routing/Mobile IP routing.

## 4.5 Network Wide Scheduling and Load Balancing

### 4.5.1 Motivation

Figure 4-10 illustrates a multi-hop small cell network which utilizes heterogeneous backhaul. In this scenario, packets might need to traverse multiple hops with different capacities and nodes with different queue handling capabilities. This situation can severely impact end-to-end user experience when one or multiple links are congested. It would cause the QoS to degrade significantly. Such cases might require packet multiplexing for various traffic classes with smart priority queuing. Better user experience can be realized with fine grain queue management, intelligent resource scheduling of backhaul resources and proper radio resource assignment.

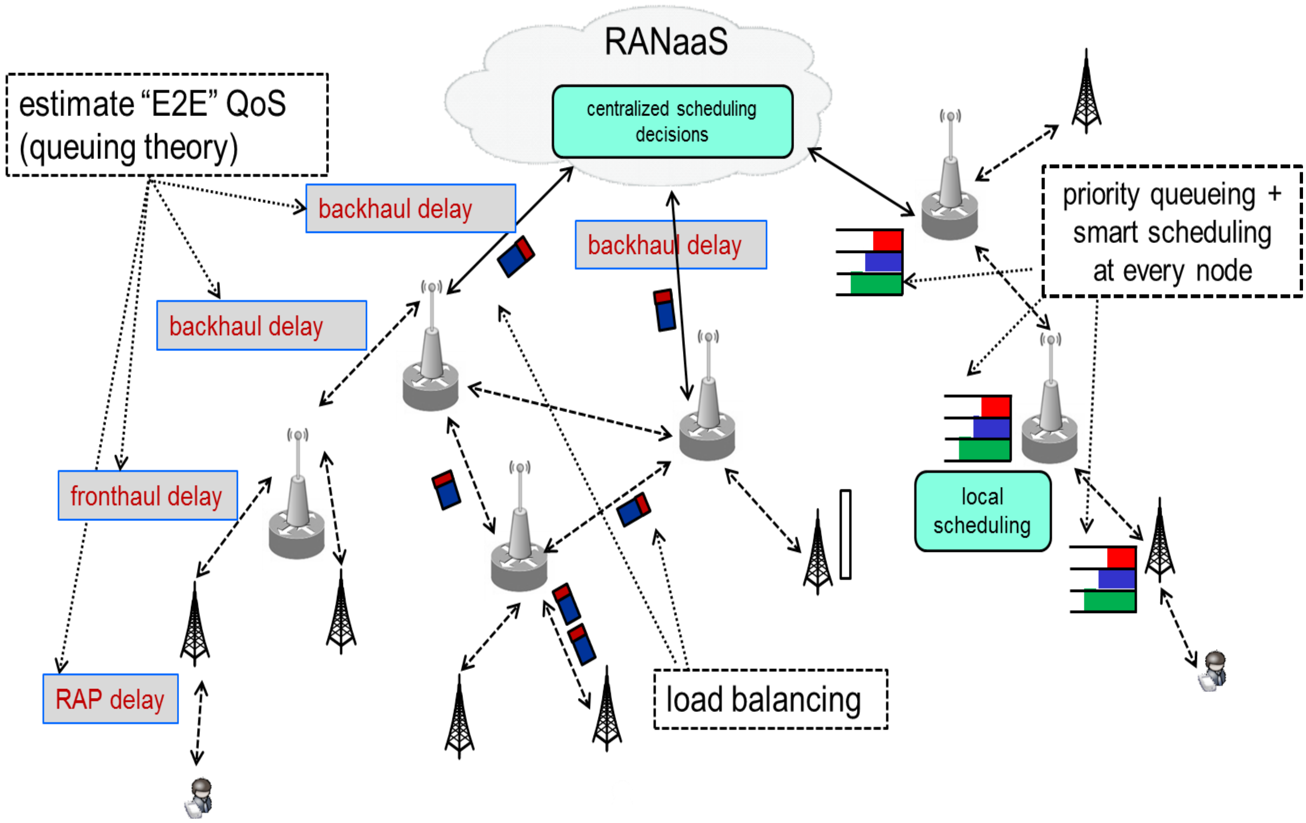


**Figure 4-10 Network wide scheduling and load balancing**

### 4.5.2 Assumptions

This technology imposes the following general assumptions:

- Network topology is known.
- Capacity of various backhaul links and for the nodes buffer (queue) handling capacity is known.
- The current availability of radio resources for RAPs is known.
- Application requirements and constraints of different traffic types are known.

### 4.5.3 Technical Requirements

The technical requirements for this candidate technology are listed below:

- Fast and reliable handover mechanisms which ensure non-disruptive behaviour when RAPs are turned off.

 

- Traffic handover from one backhaul node to neighbouring backhaul nodes in the case that the energy saving algorithm indicates that a particular BN may be turned off.
- Joint optimization of backhaul and radio access network energy consumption, i.e. ability to turn off backhaul nodes based on the RAN load.
- Users' demands for QoS are known.

### 4.5.4 Description of Technology Candidate

This technology candidate focuses on small cell networks which have been connected to the core network through multi-hop backhaul nodes. In the bid to provide better user experience and minimal end-to-end delay of services and applications, this proposal envisions accurate modelling of system wide end-to-end delays. Then multiple load balancing techniques can be applied and delays evaluated to optimize the user experience. There will be local scheduling algorithms running on backhaul nodes and a central algorithm running on RANaaS entity.

When backhaul links are becoming the bottleneck for traffic flows, capacity isolation for different traffic classes may be highly sub-optimal. In such cases, traffic multiplexing along all possible links combined with smart priority queuing and scheduling may be preferable.

This scheduling mechanism also explores the possibilities of topology control by evaluating the performance limits for a given topology. If needed, it may apply additional measures such as local break-out in order to improve the QoS.

## 4.6 Backhaul Analysis based on Viable Metrics and "Cost" Functions using Stochastic Geometry

### 4.6.1 Motivation

Heterogeneity in wireless networks implies increased randomness in base station deployment. This in turn leads to a scenario where the backhaul for such networks has to be highly adaptable. Design and deployment of such backhaul infrastructures depends on our ability to analyse networks of today and assess various methods of improvement using a common metric or standard. Stochastic geometry provides one such method of obtaining a metric that can be used as a benchmark for comparison. This approach deals with base stations as points of a point process, wherein system parameters such as transmit power, fading, path-loss, etc. are treated as functionals (or attributes) of each of these points. This can then be used to analyse the probabilities of coverage, spatially averaged rate or spectral efficiency. In such models, the backhaul can be described as a higher layer of points (distributed according to another point process) which is superimposed upon the base station layer and imposes certain restrictions on the layer below (i.e. base station layer), e.g., a throughput limit. The major advantage of using such a model is the fact that such an analysis observes the "average" behaviour of various system parameters by taking an expectation over infinitely many realizations of the point process. This implies that every network topology that can exist has been implicitly included in the observations. This is also the reason why such models can prove to be effective benchmarks against which other models can be compared.

### 4.6.2 Assumptions

- The backhaul network is considered to be the topmost layer of a multi-layered network with a point process describing the components (BSs or users) of each layer.
- This layer imposes constraints on the "cost functions" of interest that need to be evaluated. These cost functions could be energy consumption parameters or CAPEX/OPEX.

### 4.6.3 Technical Requirements

The following technical requirements are imposed by the above described analysis:

- The mathematical model requires that users aren't allowed to connect to the backhaul layer.
- For energy efficiency analysis, there exists a mechanism which enables an effective turn on and off of base stations.

 

- Users are served by the base station closest to it. If the base station is turned off, users are automatically served by the next nearest base station.

### 4.6.4 Description of Approach

This approach treats base stations and UEs as points of a point process in the Euclidean plane. Each point of the base station process has some functionals attributed to it. These functionals are typical system parameters such as transmit power, path loss, and fading. They further determine how the given area is divided or tessellated around each base station. The UEs (points of another point process) are assumed to connect to the base station from which the highest power is received (usually the nearest one). Based on such a framework, various performance indicators such as coverage probabilities and spectral efficiency can be observed over many such (theoretically infinite) realizations of the point processes. An expectation over these realizations can result in a description of the average behaviour of the network for a given number of UEs and base stations. With this framework, the backhaul can now be considered to be another layer containing points of a point process which is superimposed on the layer consisting of base stations and UEs. The backhaul layer imposes certain restrictions on the base station point process in the layer below, thereby limiting the performance and altering the performance indicators mentioned above. The interactions between the points of the backhaul layer can be modelled as functionals which vary based on the type of backhaul considered, i.e. wired or wireless backhaul. This framework now results in a relationship between the performance indicators, base stations, backhaul, and the UEs, which can then be used as a constraint in an optimization problem to evaluate the effectiveness of the backhaul in terms of CAPEX, OPEX, and energy consumption.

## 4.7 Use of Software Defined Networking in the iJOIN Network

### 4.7.1 Motivation and Assumptions

In order to assess the impact of an SDN architecture on the RANaaS concept, we focus on two main characteristics of RANaaS that can be enabled by SDN:

- The support of functional mobility between network elements, with different degrees of centralization and distribution depending on the scenario to be supported.
- The ability to support over the same infrastructure backhauling and fronthauling requirements associated with the Cloud RAN concept.

The support of functional mobility by an SDN architecture can be considered from two different viewpoints: the mobility of the functions that are supported directly by the backhaul infrastructure, and the mobility of the functions of the nodes that use the backhaul infrastructure. With respect to the first viewpoint, SDN may be used to support the distribution of the following backhaul functions:

- Transport service to be provided: MPLS, MPLS-TP, VLAN, IP, and related protocols.
- Routing functionalities for network sharing and reliability.
- Basic backhaul functionalities such as security (IPSec, MACSec 802.1AE, IDP/IPS), physical layer synchronization (frequency, phase), and reliability.
- Added value functionalities such as caching, transcoding, traffic engineering, and data collection.

With respect to the mobility of functions, SDN may be used to locate in different nodes the different processing elements. It is possible to distinguish among them:

- Data processing: baseband processing, scheduling
- Control processing: mobility support, load balancing.
- Non-Access Stratum (NAS) signalling processing

### 4.7.2 Technical Requirements

Apparently, a number of technical requirements on the backhaul nodes need to be imposed:

- Support for PDCP layer functionalities (header compression, ciphering).
- Support for RLC layer functionalities (segmentation, ARQ).

 

- Support for MAC layer functionalities (multiplexing, Hybrid ARQ).
- Phase synchronization between eNBs and backhaul nodes.

For other processing elements, different technical requirements will be derived.

### 4.7.3 Description of Technology Candidate

In order to explain the technology candidate, downlink common scheduling functionality is used as an example. Figure 4-11 illustrates the downlink scheduling process in LTE. The downlink scheduler controls which user terminals are served in a particular timeslot and which set of resource blocks of the Downlink Shared Channel (DL-SCH) should be occupied. In addition, it controls the transport-format selection (selection of transport-block size, modulation scheme, and antenna mapping) and logical-channel multiplexing for downlink transmissions. As a consequence, the RLC segmentation and MAC multiplexing will be affected by the scheduling decision.

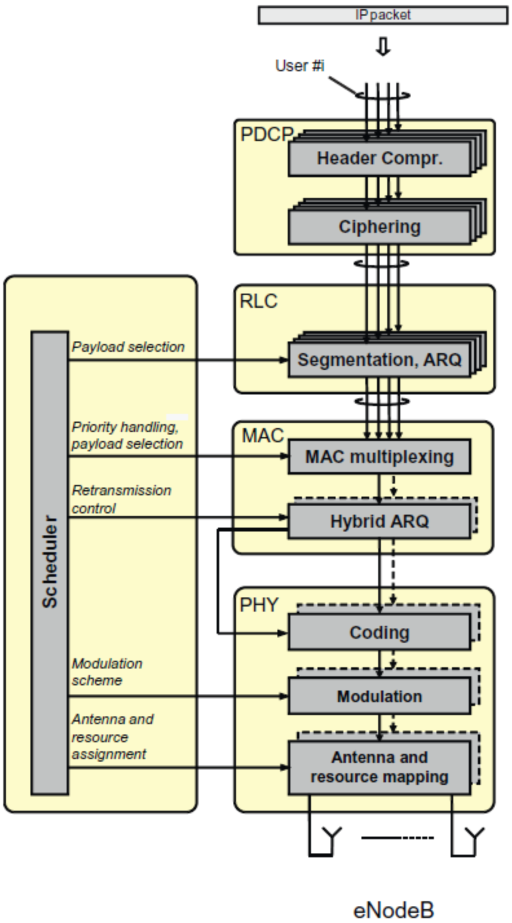


**Figure 4-11 Downlink scheduling process in LTE**

Common scheduling, as illustrated in Figure 4-12, allows for the selection of resources in different base stations such that inter-cell interference is minimized with the objective to maintain the scheduling gain in each cell. One way to implement this is to have one dedicated base station which controls the scheduling process. Another possibility is a decentralized implementation across all involved base stations which exchange the corresponding scheduling information in order to achieve a consistent decision.

 

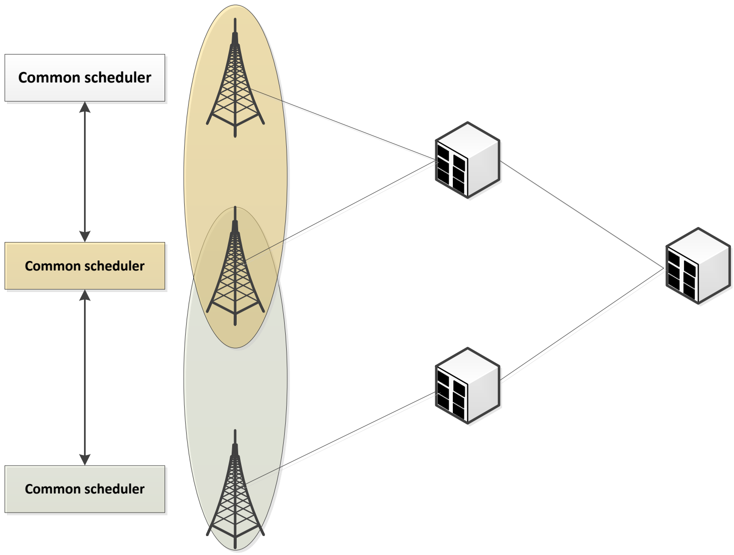


**Figure 4-12 Common scheduling across base stations**

A third option is illustrated in Figure 4-12. It involves an external node that is connected to the cooperating base stations. This node can perform the common scheduling process. It receives IP packets from the core network and sends transport blocks to the cells with an indication of where (i.e., in which resource element) they are transmitted. This division should be performed such that the scheduler is able to communicate directly with the cooperating cells.

This option would require to implement the common scheduling functionality in several nodes of the backhaul network. Hence, it implies some potential drawbacks: increased complexity of the nodes, need for coordination in order to select the most adequate node to support the scheduling function[2], impact of the introduction of new base stations and changes in the topology.

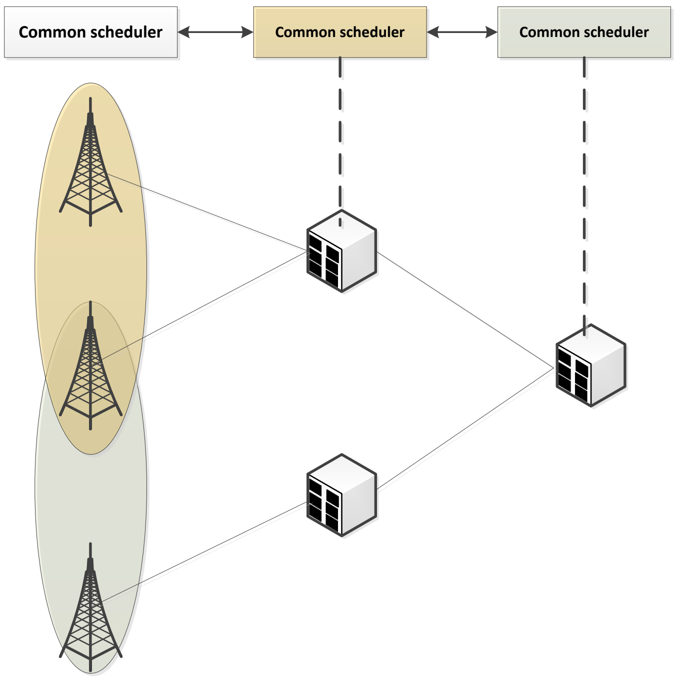


**Figure 4-13 Scheduling through an external node**

One potential way to address these challenges is illustrated in Figure 4-14. The solution may be to implement a SDN-like architecture for the backhaul network with a centralized control that supports the common

[2] This would require to implement a signalling mechanism that right now is not contemplated in the standards.

 

scheduling functionality (alongside with other functionality which is more related to backhaul such as explained earlier).

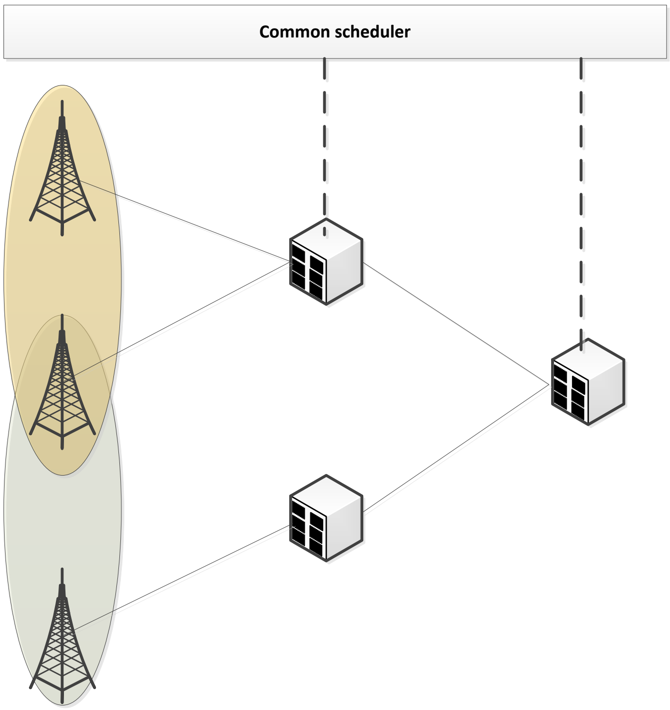


**Figure 4-14 SDN architecture for common scheduling**

The common scheduler should reside in a node that is accessible to all the nodes in the backhaul network. It would have two basic functions:

- Selection of the appropriate modulation and coding scheme, resource elements, and antennas based on CSI information reported by the cooperating cells and the buffer status.
- Repackaging the downlink IP packets into transport blocks consistent with the scheduling decision.

One of the most critical challenges is the potential imperfectness of CSI due to delays. The impact of this imperfectness will depend on the UE mobility and interference characteristics. However, under normal operating conditions this delay should not exceed a few milliseconds.

 

# 5 Harmonized Assumptions and Requirements

The general iJOIN architectural assumptions are described in IR5.1 [55]. In this section we introduce the mapping of these assumptions to the WP4 technologies candidates. This list will be used to derive the preliminary iJOIN architecture, and after that the specific requirements will be investigated and derived based on this architecture. We list next the considered technologies candidates:

- 4.1: Distributed IP anchoring and mobility management
- 4.2: Network-wide energy optimization
- 4.3: Joint path management and topology control
- 4.4: Routing and congestion control mechanisms
- 4.5: Network wide scheduling and load balancing
- 4.6: Backhaul analysis
- 4.7: Software Defined Networking (SDN)

| Assumption | Description | 4.1 | 4.2 | 4.3 | 4.4 | 4.5 | 4.6 | 4.7 |
|---|---|---|---|---|---|---|---|---|
| A.1 | Large number of iSCs in local area | x | x | x | x | x | | X |
| A.2 | Availability of macro BS in same frequency band | | | | | | | |
| A.3 | J1 interface between <u>all</u> iSCs and RANaaS with known parameters | o | | | | | | |
| A.4 | J1 interface between <u>some</u> iSCs and RANaaS with known parameters | x | | | | | | |
| A.5 | J2 interface for interconnections of all iSCs | x | | | | | | |
| A.6 | J2 interface for interconnections of some iSCs (direct neighbours, selection) | | | | x | x | x | |
| A.7 | Wired inter-node links between iSCs (fibre) | o | | o | | | | |
| A.8 | Wireless inter-node links between iSCs (60GHz) | * | | x | | | | |
| A.9 | Wired connection of iSCs to RANaaS (fibre) | o | | o | | | | |
| A.10 | Wireless connection of iSCs to RANaaS (60GHz) | * | | x | | | | |
| A.11 | Multiple Tx/Rx antennas at iSC | | | | | | | |
| A.12 | Availability of a logical controller for the joint RAN/BH optimization | x | x | x | o | o | | x |

**Table 5-1. Mapping of iJOIN architectural assumptions to WP4 candidate technologies**

<u>Legend</u>

| | |
|---|---|
| “x” | mandatory assumption |
| “*“ | optional choices for implementation candidates |
| “o” | optional assumption; this not-mandatory feature may lead to improvements |
| “ “ | not assumed for the TC |

As shown in Table 5-1, most of WP4 candidate technologies assume the existence of a large number of small cells, basically allowing to benefit from the mobility, dynamic resource management (routing, scheduling, congestion control) and network wide energy optimizations. While there are no strong assumptions in terms of the backhaul connectivity, the existence of wired links will of course allow for performance improvements. The mobility CT assumes the existence of the J2 interface among iJOIN small cells, as well as the routing and scheduling ones. A critical assumption for WP4 is the availability of a logical controller for the joint RAN/BH optimization (this logical entity is called iJOIN network controller and it is defined in IR5.1).

The following table summarises the WP4 implementation assumptions per technology candidate describing the fundamental framework of the investigations. This table is actually a summary of the assumptions elaborated while describing each of the technology candidates in Section 4.

| Assumption | Description | 4.1 | 4.2 | 4.3 | 4.4 | 4.5 | 4.6 | 4.7 |
|---|---|---|---|---|---|---|---|---|
| A.4.1 | The leftover capacity in the current backhaul infrastructure used for the macro-network is reused to backhaul small cells | | | x | | | | |
| A.4.2 | Heterogeneous backhaul (fiber and heterogeneous wireless) | | | x | | | | |
| A.4.3 | The backhaul is a multi-hop IP network | x | o | | x | | | |
| A.4.4 | Some nodes in the radio access network or in the backhaul might have local IP connectivity | x | o | | | | | |
| A.4.5 | Nodes providing local breakout have control interfaces with the mobility entities in the core | x | | | | | | |
| A.4.7 | Small-cells are connected to the cloud through Backhaul Nodes (BNs) | x | x | x | x | x | x | |
| A.4.8 | Path diversity within the backhaul network | o | x | x | x | x | X | |
| A.4.9 | Reusability of existing backhaul technology | x | | | | | | |

**Table 5-2 Mapping of iJOIN implementation assumptions to WP4 candidate technologies**

The following table lists the preliminary technical requirements for each technology candidate, and is up to changes during the progress of the project. This table is actually a summary of the requirements elaborated while describing each of the technology candidates in Section 4.

| Technical requiremen t | Description | 4.1 | 4.2 | 4.3 | 4.4 | 4.5 | 4.6 | 4.7 |
|---|---|---|---|---|---|---|---|---|
| R.4.1 | The iJOIN Transport Nodes will be remotely configurable by a centralized entity and will be able to report measurements to this centralized entity | | | x | x | | | |
| R.4.2 | The iJOIN Transport Nodes will have an interface towards the iSC to collect short and long time scale statistics about the RAN (possibly even on flow-level) | | | x | x | | | |
| R.4.3 | The solutions should work on an IP-based network | x | | | | | | |
| R.4.4 | Solutions have to both consider and interact | x | | x | x | | | |

 

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | with the radio access and backhaul | | | | | | | |
| R.4.5 | Solution might benefit from user terminal support | x | | | | | | |
| R.4.6 | Solutions have to cooperate with the energy efficiency mechanisms | x | | o | | | | |
| R.4.7 | Mobility support has to be provided on a per-flow basis only to applications that require address continuity | x | | | | | | |
| R.4.8 | Handover preparation and execution mechanisms should be fast enough to support real time communications, and should minimise packet loss. | x | | | | | x | |
| R.4.9 | Coexistence of macro and small cells should be considered. Support from macro cells as part of intra-small cell mobility might be considered. | x | | | | | | |
| R.4.10 | Transport nodes have support for PDCP layer functionalities (header compression, ciphering) | | | | | | | x |
| R.4.11 | Transport nodes have for RLC layer functionalities (segmentation, ARQ) | | | | | | | x |
| R.4.12 | Transport nodes have support for MAC layer functionalities (multiplexing, Hybrid ARQ) | | | | | | | x |
| R.4.13 | There is phase synchronization between eNBs and backhaul nodes | | | | | | | x |
| R.4.14 | The formation of virtual cell (i.e., a cluster of cooperating and logically grouped small cells that appears to the user as a single cell) will be considered for mobility purposes | x | | | | | | |
| R.4.15 | Cooperation between small cells within one virtual cell, and cooperation between virtual cells should be considered to enable load-aware handover decisions | x | | | | | | |
| R.4.16 | Small cells are connected to the master node in cloud through multiple BNs | x | | | | | | |
| R.4.17 | Ability of iJOIN small cells to go to low-power sleep mode when not serving any user traffic | | x | o | | | | |
| R.4.18 | Existence of traffic profiling mechanism in place at small cells | | | x | | | | |
| R.4.19 | Users demands for QoS are known | | | x | | | x | |
| R.4.20 | Mobility requirements of (at least some of the) running applications are known | x | | | | | | |

**Table 5-3 Mapping of iJOIN requirements to WP4 candidate technologies**

 

| Legend | |
|---|---|
| "x" | mandatory assumption |
| "*" | optional choices for implementation candidates |
| "o" | optional assumption; this not-mandatory feature may lead to improvements |
| " " | not assumed for the TC |

# 6 Integration in iJOIN Architecture and Functional Split

## 6.1 Integration of Technology Candidates in iJOIN Architecture

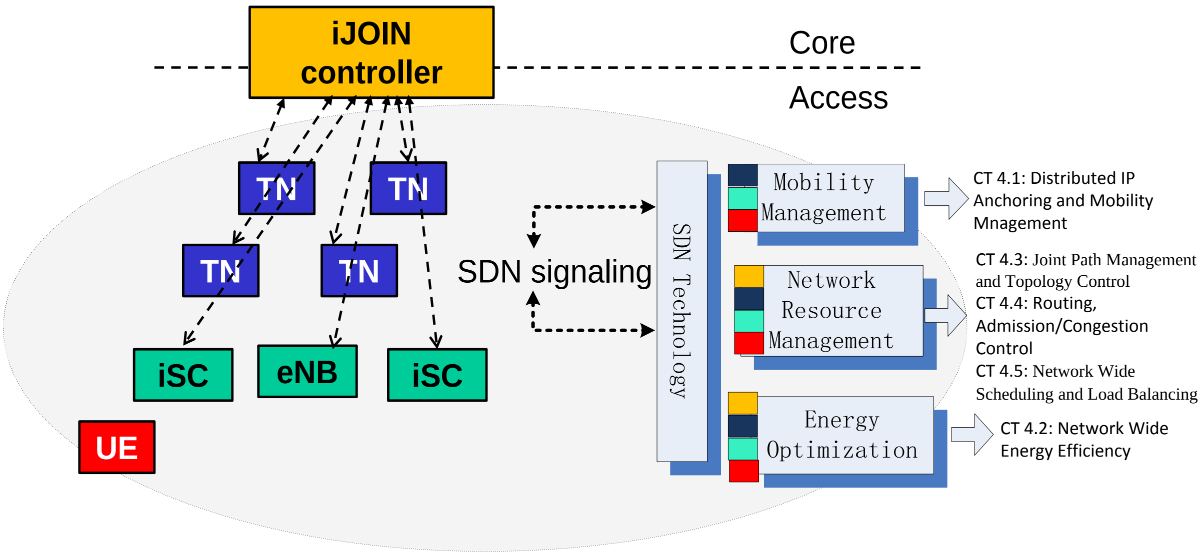


**Figure 6-1 iJOIN Architecture and WP4 Technology Candidate**

Figure 6-1 shows the integration of the iJOIN architecture and the candidate technologies. The technology candidates can be classified into three categories: mobility management, network resource management and energy optimization.

- Mobility management: CT 4.1 (Distributed IP anchoring and mobility management). This function will be implemented in iJOIN transport node, iJOIN small cell and possibly UE assistance is required.

- Network Resource Management: CT 4.3 (Joint path management and topology control), CT 4.4 (Routing, admission/congestion control), CT4.5 (Network wide scheduling and load balancing). This function will be implemented in the iJOIN transport node, the iJOIN small cell, and may be supported by UEs. The iJOIN controller, which provides a global view of the network utilization condition, is beneficial to the network resource management design.

- Energy Optimization: CT 4.2 (Network wide energy efficiency). This function will be implemented in the iJOIN transport node and the iJOIN small cell. The iJOIN controller, which provides a global view of the network topology and load information, is beneficial to the energy optimization design.

SDN technology provides a framework to integrate the technology candidates vertically in the iJOIN architecture.

## 6.2 Interaction of Technology Candidates

This section provides an overview of how the different technology candidates interact in order to achieve the WP4 goals. We present the WP4 architecture by introducing the different modules and the logical entities where they are implemented. The foreseen interfaces between these different modules are also identified in this section. This is a preliminary definition which will be further refined in the course of the project.

 

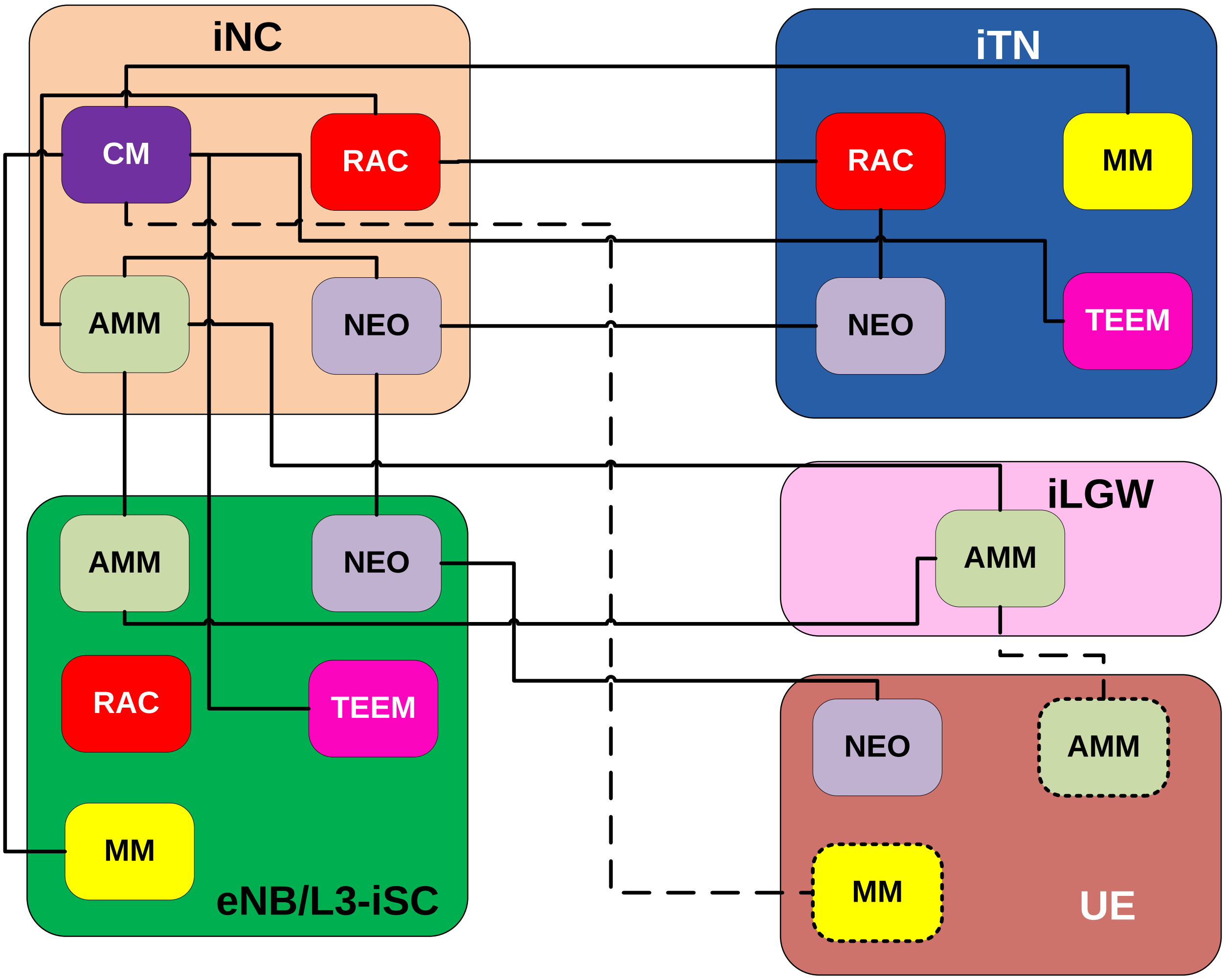


**Figure 6-2 WP4 preliminary architecture**

Figure 6-2 shows the preliminary WP4 architecture, the different modules, in which logical entities they are deployed, and the different interfaces. Note that a dashed line denotes an optional module/interface. We next list and briefly describe ach of the modules (this will be further detailed in D4.1):

- Anchor & Mobility Management (AMM). This module, defined by the CT 4.1 (Distributed IP Anchoring and Mobility Management), is in charge of providing and managing IP addresses to the UEs, as well as ensuring that those addresses used by applications which cannot handle an address change are provided with mobility support. This module is located in the iNC, iLGW, eNB/L3-iSC and optionally on the UE (to benefit from terminal-aided support). Note that CT 4.1 is in charge not only of providing mobility support on an address (application) basis, but also to ensure that resources are optimally exploited, both in the backhaul and in the access. This is achieved by selecting and using an anchor closer to the UE. This does not mean that for some flows legacy EPS Rel-10 mobility mechanisms (and anchors, i.e., the PGW) are not used but that they are actually complemented by the iJOIN solutions.

  The AMM interacts with the RAC and NEO modules.

- Routing, Admission/Congestion Control (RAC). This module, defined by the CT 4.4 (Routing, Admission/Congestion Control) is in charge of properly configuring the layer-3 routing in the backhaul, considering the status of RAN, as well as the UE traffic requirements. This module is deployed in the iNC, iTN and eNB/L3-iSC.

  The RAC interacts with the AMM and NEO modules.

 

- Network wide Energy Optimizer (NEO). This module is defined by the CT 4.2 (Network wide Energy Efficiency). It is in charge of taking network wide decisions about switching on/off physical nodes, as well as ensuring that UE traffic is still properly routed by the nodes that are running at each time. To do so, the NEO module is foreseen to interface with RAC and AMM modules. The NEO module is deployed in the iNC, iTN, eNB/L3-iSC and the UE.
- Measurement Module (MM). This module is defined by CT4.3 (Joint path management and topology control). It is deployed in the iTN, eNB/L3-iSC and UE and its task is to measure performance metrics as indicated by the Controller Module (CM) residing in the iNC. Envisioned metrics to be reported by the MM comprise: locally experienced congestion, available neighbours, available data rates, and number of connected UEs. The MM will support several reporting modes, e.g.: asynchronous, periodic or event based.
- Traffic Engineering Enforcement Module (TEEM). This module is defined by CT4.3 (Joint path management and topology control). It is deployed in the iTN, eNB/L3-iSC and UE. This module offers a programmable API to the Controller Module (CM) residing in the iNC that can be used by this entity to engineer the backhaul and access networks. Envisioned actions to be supported by the TEEM are a programmable forwarding plane in the transport nodes, programmable traffic classification/prioritization, and programmable per-flow rate control.
- Controller Module (CM). This module is defined by CT4.3 (Joint path management and topology control), and it is deployed in the iNC. The task of this module is to configure the MMs in the iTN/eNB and UE entities under its control, to gather the measurements reported by the configured MMs, and to configure the TEEMs in the iTN/eNB and UE entities under its control based on the collected measurements.

## 6.3 Interaction of technology candidates and RANaaS

As RANaaS focuses on the flexible centralization of RAN functionalities towards a cloud platform, it will have to interact with the management of the backhaul network and the support of mobility procedures in multiple ways:

- Technologies such as SDN may be enablers for the realization of RANaaS' flexible functional split. Requirements on these enabling technologies should be identified in order to assess whether they can support them or not. Extensions of southbound protocols, e.g. OpenFlow, may be required, as well as the implementation of northbound applications.
- Requirements on the backhaul infrastructure, mainly in terms of capacity and latency, can differ significantly depending on the functional partition. The feasibility of the proposed scheme will require the interaction with a centralized path control mechanism.
- Mobility and load balancing procedures may result into a different functional split, for example, the UE handovers from an iSC to a conventional macro-cell. The RANaaS supporting infrastructure should support this functional transfer in an optimized way.
- Congestion control mechanisms may have a different behaviour depending on the functional split because parts of the network may not be visible to them.
- The combined optimization of access and backhaul network in a RANaaS environment has also implications. Decisions that are taken in a centralized way (e.g., scheduling) may have an impact on the performance of the backhaul network (e.g., may lead to congestion in the backhaul network). Therefore, feedback mechanisms are required.
- For other use cases such as network sharing it could be necessary to ensure that the options adopted in RANaaS and the WP4 technology candidates are compatible with the requirements associated to them.

 

# Acknowledgements and Disclaimer

This work was partially funded by the European Commission within the 7th Framework Program in the context of the ICT project iJOIN (Grant Agreement No. 317941). The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the iJOIN project or the European Commission.

# References

[1] 3GPP, "TS 29.274: Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C)," September 2011.

[2] S. Gundavelli, K. Leung, V. Devarpalli, K. Chowdhury and B. Patil, *Proxy Mobile IPv6, RFC 5213 (Proposed Standard), Internet Engineering Task Force, August 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5213.txt.*

[3] H. Soliman, "Mobile IPv6 Support for Dual Stack Hosts and Routers, RFC 5555 (Proposed Standard), Internet Engineering Task Force," June 2009. [Online]. Available: http://www.ietf.org/rfc/rfc5555.txt.

[4] H. Chan, "Problem statement for distributed and dynamic mobility management," March 2011.

[5] 3GPP, "TS 23.401: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," September 2011.

[6] 3GPP, "TR 23.859: LIPA Mobility and SIPTO at the Local Network," July 2011.

[7] W. Hahn, "Flat 3GPP evolved packet core: Improvement for multiple net-work connections," in *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Brest, France, October 2011.

[8] W. Hahn, "3GPP Evolved Packet Core support for distributed mobility anchors: Control enhancements for GW relocation," in *International Conference on ITS Telecommunications (ITST)*, St. Petersburg, Russia, August 2011.

[9] C. J. Bernardos, J. C. Zúñiga, A. Reznik, "Towards flat and distributed mobility management: A 3GPP evolved network design," *2012 IEEE International Conference on Communications (ICC)*, pp.6855-6861, June 2012.

[10] H. Chan *et al.*, "Requirements for Distributed Mobility Management," Internet Engineering Task Force, draft-ietf-dmm-requirements-03.txt (work-in-progress), December 2012.

[11] D. Liu *et al.*, "Distributed Mobility Management: Current practices and gap analysis," Internet Engineering Task Force, draft-ietf-dmm-best-practices-gap-analysis-00.txt (work-in-progress), February 2013.

[12] C. J. Bernardos, A. de la Oliva, F. Giust, “A IPv6 Distributed Client Mobility Management approach using existing mechanisms,” Internet Engineering Task Force, draft-bernardos-mext-dmm-cmip-00.txt (work-in-progress), March 2011.

[13] C. J. Bernardos, A. de la Oliva, F. Giust, T. Melia, R. Costa, “A PMIPv6-based solution for Distributed Mobility Management,” Internet Engineering Task Force, draft-bernardos-dmm-pmip-01.txt (work-in-progress), March 2012.

[14] P. Seite, P. Bertin, J. H. Lee, "Dynamic Mobility Anchoring," Internet Engineering Task Force, draft-seite-dmm-dma-06.txt (work-in-progress), January 2013.

[15] H. Chan, F. Xia, J. Xiang and H. Ahmed, "Distributed Local Mobility Anchors," Internet Engineering Task Force, draft-chan-netext-distributed-lma-03.txt (work-in-progress), March 2010.

 

[16] J. C. Zúñiga, C. J. Bernardos, T. Melia, A. de la Oliva, R. Costa, A. Reznik, “Distributed Mobility Management: a Standards Landscape”, accepted for publication in *IEEE Communications Magazine, issue on Telecommunications Standards*, March 2013.

[17] F. Giust, C. Bernardos, S. Figueiredo, P. Neves and T. Melia, "A Hybrid MIPv6 and PMIPv6 Distributed Mobility Management: the MEDIEVAL approach," in *IEEE Workshop on multiMedia Applications over Wireless Networks (MediaWiN), organized in association with the IEEE Symposium on Computers and Communications (ISCC 2011)*, Kerkyra, Corfu, Greece, June 2011.

[18] F. Giust, A. de la Oliva and C. Bernardos, "Flat Access and Mobility Architecture: an IPv6 Distributed Client Mobility Management Solution," in *3rd IEEE International Workshop on Mobility Management in the Networks of the Future World (Mobiworld 2011), in conjunction with The 30th IEEE International Conference on Computer Communications (IEEE INFOCOM 2011)*, Shanghai, China, April 2011.

[19] F. Giust, A. de la Oliva, C. J. Bernardos and R. P. Ferreira Da Costa, "A Network-based Localized Mobility Solution for Distributed Mobility Management," in *International Workshop on Mobility Management for Flat Networks (MMFN 2011), collocated with WPMC 2011*, October 2011.

[20] 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 (NTMS) '08*, November 2008.

[21] P. Bertin, S. Bonjour and J.-M. Bonnin, "An Evaluation of Dynamic Mobility Anchoring," in *2009 IEEE 70th Vehicular Technology Conference Fall (VTC 2009-Fall)*, September 2009.

[22] M. Fischer, F.-U. Andersen, A. Kopsel, G. Schafer and M. Schlager, "A Distributed IP Mobility Approach for 3G SAE," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2008.

[23] L. Yu, Z. Zhijun, L. Tao and T. Hui, "Distributed mobility management based on flat network architecture," in *2010 The 5th Annual ICSTWireless Internet Conference (WICON)*, March 2010.

[24] P. McCann, "Design of a flat wireless Internet Service Provider network," *2011 14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pp.1-5, October 2011.

[25] IEEE, "IEEE P802.21/D14 Draft Standard for Local and Metropolitan Area Networks: Media Independent Handover Services," IEEE, 2008.

[26] M. A. Marsan, L. Chiaraviglio, D. Ciullo and M. Meo, "Optimal energy savings in cellular access networks," in *IEEE ICC Workshop*, Dresden, Germany, June 2009.

[27] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *GLOBECOM*, December 2010.

[28] V. Mancuso and S. Alouf, "'Reducing costs and pollution in cellular networks," *IEEE Communications Magazine, vol. 49, no. 8,* pp. 63-71, August 2011.

 

[29] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, D. Sabella, M. Gonzalez, C. Desset and O. Blume, "Cellular Energy Efficiency Evaluation Framework," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, Yokohama, Japan, May 2011.

[30] Ashraf, F. Boccardi and L. Ho, "SLEEP Mode Techniques for Small Cell Deployments," *IEEECommunications Magazine, vol. 49, no. 8,* pp. 72-79, August 2011.

[31] NGMN, "Next Generation Mobile Networks Use Cases related to Self Organising Network, Overall Description," Deliverable, NGMN, December 2008.

[32] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World," *IEEE Journal on Selected Areas of Communication,* vol. 20, no 4, May 2002.

[33] A. Oliveira and G.R. Mateus, "Using Genetic Algorithms to LSP Setup in MPLS Networks," Proc. XXIV Simpósio Brasileiro de Redes de Computadores, pp. 705-721 2006.

[34] G. Conte, P. Iovanna, R. Sabella, M. Settembre, L. Valentini, "A Traffic Engineering Solution for GMPLS Network: A Hybrid Approach Based on off-line and on-line routing methods," *ONDM Conference,* February 2003.

[35] M. Zhang, B. Liu and B. Zhang, „Multi-Commodity Flow Traffic Engineering with Hybrid MPLS/OSPF Routing," *IEEE Globecom 2009,* December 2009.

[36] NEC, "Mobile backhaul evolution and convergence," White Paper, January 2010.

[37] NGMN, "A requirement specification by the NGMN alliance, NGMN optimised backhaul requirements," August 2008.

[38] Chiosi et al. "Network Functions Virtualisation," *SDN and OpenFlow World Congress,* Darmstadt, Germany, October 2012.

[39] C. Perkins, "IP Mobility Support for IPv4", RFC 5944, IETF, November 2010.

[40] D. Johnson, C. Perkins and J. Arkko, "Mobility Support in IPv6", RFC 3775, IETF, June 2004.

[41] S. Gundavelli et al., "Proxy Mobile IPv6", RFC 5213, IETF, August 2008.

[42] H. Soliman, "Mobile IPv6 Support for Dual Stack Hosts and Routers (DSMIPv6)", RFC 5555, IETF, June 2009.

[43] L. Wang, Y. Zhang, and Z. Wei, "Mobility Management Schemes at Radio Network Layer for LTE Femtocells", *IEEE 69th Vehicular Technology Conference (VTC Spring),* pp. 1 –5, April 2009.

[44] A. Ulvan, R. Bestak, and M. Ulvan, "The Study of Handover Procedure in LTE-based Femtocell Ne twork", in Wireless *and Mobile Networking Conference (WMNC), Third Joint IFIP,* pp. 1-6, October 2010.

[45] A. Ulvan, R. Bestak, and M. Ulvan, "Handover Scenario and Procedure in LTE-based Femtocell Networks", the *Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM),* October 2010.

[46] Z. Fan and Y. Sun, "Access and Handover Management for Femtocell Systems", IEEE 71st Vehicular Technology Conference (VTC Spring), 2010, pp. 1 –5, May 2010.

[47] H. Li, X. Xu, D. Hu, X. Chen, X. Tao, and P. Zhang, "Pseudo-Handover based Power and Subchannel Adaptation for Two-tier Femtocell Networks", *Wireless Communications and Networking Conference (WCNC)*, IEEE, pp. 980 –985, March 2011.

 

[48] H. Zhang, X. Wen, B. Wang, W. Zheng, and Y. Sun, "A Novel Handover Mechanism Between Femtocell and Macrocell for LTE Based Networks", *Second International Conference on Communication Software and Networks ( ICCSN)*, pp. 228 –231, February 2010.

[49] Z. Chowdhury, W. Ryu, E. Rhee, and Y. Jang, "Handover between Macrocell and Femtocell for UMTS based Networks", *11th International Conference on Advanced Communication Technology (ICACT),* February 2009.

[50] M. Chowdhury and Y. Jang, "Handover Control for WCDMA Femtocell Networks", *the Journal of Korea Information and Communication Society,* vol. abs/1009.3779, September 2010.

[51] J. Kim and T. Lee, "Handover in UMTS Networks with Hybrid Access Femtocells", *the 12th International Conference on Advanced Communication Technology (ICACT),* vol. 1, pp. 904 –908, February 2010.

[52] S. Wu, "A New Handover Strategy between Femtocell and Macrocell for LTE-Based Network", *4th International Conference on Ubi-Media Computing (U-Media),* pp. 203 –208, July 2011.

[53] 3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10)", TS36.300, v10.9.0, December 2012.

[54] Generation Partnership Project, "Technical Specification Group Radio Access Network; User Equipment (UE) procedures in idle mode and procedures for cell reselection in connected mode (Release 10)", TS 25.304, v10.7.0, December 2012.

[55] iJOIN Project, "IR 5.1 Preliminary definition of requirements and scenarios for the iJOIN architecture," April 2012.

 