| | |
|---|---|
| Project Number: | **CELTIC / CP7-011** |
| Project Title: | Mobile Networks Evolution for Individual Communications Experience – MEVICO |
| Document Type: | P |

| | |
|---|---|
| Document Identifier: | D 4.3.1 |
| Document Title: | **Intermediate Design of the MEVICO Traffic Engineering Architecture** |
| Source Activity: | WP4 |
| Main Editor: | Ece Saygun |
| Authors: | See the Authors section |
| Status / Version: | 1.00 |
| Date Last changes: | 27.8.2012 |
| File Name: | D 4.3.1 |

Abstract: In this document the intermediate traffic engineering architecture of MEVICO and the preliminary integration considerations of the analyzed traffic management solutions for EPC are introduced. The solutions presented here will be extended by D4.3.2 where the architecture and the integration framework will be finalized.

Keywords: QoS/QoE-aware traffic engineering, traffic management building blocks, microscopic and macroscopic traffic management, resource selection, caching.

| Document History: | |
|---|---|
| 01.02.2012 | Ericsson Turkey: Document created |
| 30.03.2012 | NSN: update section §5.4.2 |
| 01.04.2012 | Ericsson Turkey: Bulk Analysis of network data inserted |
| 20.04.2012 | ALU: Edition of §3.4.4 Traffic Engineered Handovers |
| 24.04.2012 | BME-MIK: added Offloading and NB-IFOM sections |
| 25.04.2012 | Turk Telekom: updated MCCS sections |
| 30.04.2012 | TUB: updated §3.1, 3.5.5, 3.7, 4.5 |
| 09.05.2012 | Avea: added §4.2 e2e QoS. |
| 25.05.2012 | ALU: updates in §3.4.3 wrt revision |
| 31.05.2012 | VTT: Reviewed §2, 3.1, and §3.2 |
| 01.06.2012 | CUT: added §3.3 and revised §6.5 |
| 06.06.2012 | BME-MIK: updated §6.5 |
| 08.06.2012 | ALU: updates on §5.1.4.5.3 |
| 13.06.2012 | DTAG: updates on §5.4 |
| 23.07.2012 | Ericsson Turkey: editorial work |
| 24.07.2012 | Montimage: §4.1.1 updated with more analysis of the results |
| 27.08.2012 | Final editorial work and document release |

# Table of Contents

## Authors

| Partner | Name | Phone / Fax / e-mail |
|---|---|---|
| Nokia Siemens Networks | | |
| | Christian Ruppelt | Phone: + 49 89 5159 39125<br>e-mail: christian.ruppelt@nsn.com |
| | Wolfgang Hahn | Phone :<br>e-mail: wolfgang.hahn@nsn.com |
| Technical University of Chemnitz | | |
| | Thomas Bauschert | Phone:<br>e-mail: thomas.bauschert@etit.tu-chemnitz.de |
| | Gerd Windisch | Phone:<br>e-mail: gerd.windisch@etit.tu-chemnitz.de |
| Deutsche Telekom | | |
| | Gerhard Hasslinger | Phone:<br>e-mail: Gerhard.Hasslinger@telekom.de |
| T-Systems | | |
| | Anne Schwahn | Phone:<br>e-mail: anne.schwahn@t-systems.com |

## Technical University Berlin

Łukasz Budzisz

Phone:

e-mail: lukasz.budzisz@tu-berlin.de

Sven Wiethoelter

Phone:

e-mail: wiethoel@tkn.tu-berlin.de

Berthold Rathke

Phone:

e-mail: rathke@tkn.tu-berlin.de

Tacettin Ayar

Phone:

e-mail: ayar@tkn.tu-berlin.de

## Budapest University of Technology and Economics – Mobile Innovation Centre

László Bokor

Phone:

e-mail: bokorl@hit.bme.hu

Zoltán Faigl

Phone:

e-mail: zfaigl@mik.bme.hu

Zoltán Németh

Phone:

e-mail: znemeth@mik.bme.hu

## University of Vienna

Florian Metzger

Phone:

e-mail: florian.metzger@univie.ac.at

Albert Rafetseder

Phone:

e-mail: albert.rafetseder@univie.ac.at

## Ericsson Turkey

Ece Saygun

Phone:

e-mail: ece.saygun@ericsson.com

## Turk Telekom

Ahmet Serdar Tan

Phone:<br>
e-mail: ahmetserdar.tan@turktelekom.com.tr

## VTT

Tapio Suihko

Phone:
e-mail: Tapio.Suihko@vtt.fi

## Alcatel Lucent

Sabine Randriamasy

Phone:
e-mail: Sabine.Randriamasy@alcatel-lucent.com

## Montimage

Bachar Wehbi

Phone:
e-mail: bachar.wehbi@montimage.com

## AVEA

Engin ZEYDAN

Phone:
e-mail: engin.zeydan@avea.com.tr

Çağatay EDEMEN

Phone:
e-mail: cagatay.edemen@avea.com.tr

Salih Ergüt

Phone:
e-mail: salih.ergut@avea.com.tr

# Executive Summary

As mobile and wireless communication networks move toward broadband converged networks and applications, the demands on the infrastructure will increase tremendously. The MEVICO project aims at identification of the technologies for the evolution of 3GPP LTE-mobile broadband network. The target is to innovate and develop new network concepts for meeting the future requirements of the evolving mobile networks and usage. The work related to this document encompasses the smart traffic management techniques on which the MEVICO partners have focused on. Each partner is focusing on the technologies according to their research plans and presenting the state-of-the-art in the corresponding technology mentioning the areas for improvement.

This document follows the categorization logic by grouping different techniques into building blocks and analyzing each mechanism separately.

Using these building blocks, various mechanisms can be categorized according to common functionality or solution space. Since this project focuses on innovative approaches on packet core, mobile backhaul and operator service domain, the radio management aspects were not taken into consideration. The same applies to traffic management, which is restricted to the application layer or is mainly about business modelling aspects. The four building blocks of highest interest to the project consortium are ‘microscopic traffic management’, ‘macroscopic traffic management’, ‘improved resource selection and caching’ and ‘deployment of new network resources’. ‘Deployment of new network resources’ has not been handled in this document.

In this document the preliminary traffic engineering architecture is proposed based on research completed. Next document on traffic engineering architecture will be D 4.3.2 where the proposed architecture will be finalized.

# List of acronyms and abbreviations

| | |
|---|---|
| 2G/3G/4G | 2nd/3rd/4th Generation Cellular Mobile Phone System (GSM, UMTS, LTE,…) |
| 3GPP | 3rd Generation Partnership Project, based on GSM Technology |
| 3GPP | 3rd Generation Partnership Project 2, based on GSM Technology |
| AC | Application Client |
| ALTO | Application Layer Traffic Optimization |
| AMBR | Aggregated Maximum Bit Rate |
| API | Application Programming Interface |
| APN | Access Point Name |
| ARP | Allocation and Retention Priority |
| ARQ | Automatic Repeat Request |
| AS | Autonomous System |
| AVP | Active Virtual Peer |
| BNG | Broadband Network Gateway |
| BS | Base Station |
| CAPEX | Capital Expenditure |
| CDN | Content Distribution Network |
| CIF | Common Intermediate Format |
| CSCF | Call Session Control Function |
| CSP | Communication Service Provider |
| DHT | Distributed Hash Table |
| DL | Downlink |
| DNS | Domain Name System |
| DoA | Direction of Arrival |
| DPI | Deep Packet Inspection |
| DSL | Digital Subscriber Line |
| DSMIPv6 | Dual Stack Mobile IPv6 |
| e2e | End to end |
| ECMP | Equal Cost Multi-Path |
| eNB | eNodeB |
| EPC | Evolved Packet Core |
| EPS | Evolved Packet System |
| E-UTRAN | Evolved UTRAN |
| FTP | File Transfer Protocol |
| GAN | Generic Access Network |
| GBR | Guaranteed Bit Rate |
| GGSN | Gateway GPRS Support Node |
| GPRS | General Packet Radio Service |
| GSM | Global System for Mobile Communications |
| GW | Gateway |
| HA | Home Agent |
| HeNB | Home eNodeB |
| HLR | Home Location Register |
| HO | Handover |
| HSS | Home Subscriber Server |
| HTML | Hypertext Markup Language |
| HTTP | Hyper Text Transfer Protocol |
| IETF | Internet Engineering Task Force |
| IFOM | IP Flow Mobility |
| IMPEX | Implementation Expenditure |

| | |
|---|---|
| IMS | IP Multimedia Subsystem |
| ISP | Internet Service Provider |
| LAN | Local Area Network |
| L-GW | Local Gateway |
| LIPA | Local IP Access |
| LTE | Long Term Evolution |
| LTE-A | LTE-Advanced |
| MAC | Medium Access Control |
| MacTM | Macroscopic Traffic Management |
| MAPCON | Multi-Access PDN CONnectivity |
| MBR | Maximum Bit Rate |
| MicTM | Microscopic Traffic Management |
| MIH | Media Independent Handover |
| MME | Mobility Management Entity |
| MNO | Mobile Network Operator |
| MP2MP | Multipoint-to-Multipoint |
| MPLS | Multiprotocol Label Switching |
| MPTCP | Multi-Path TCP |
| MT | Mobile Terminal |
| MWR | Microwave Radio |
| NAPTR | Naming Authority Pointer [Resource Record] |
| NAS | Non Access Stratum |
| NAT | Network Address Translation |
| NB-IFOM | Network Based IP Flow Mobility |
| NE | Network Element |
| NIC | Network Interface Card |
| OAM | Operation, Administration, and Maintenance |
| OPEX | Operational Expenditure |
| OSPF | Open Shortest Path First |
| P2P | Peer-to-Peer |
| P2MP | Point-to-Multipoint |
| P4P | Provider Portal for (P2P) Applications |
| PCC | Policy and Charging Control |
| PCE | Path Computation Element |
| PCP | Priority Code Point |
| PCRF | Policy and Charging Rules Function |
| PDN | Packet Data Network |
| PDP | Packet Data Protocol |
| P-GW | PDN Gateway |
| QoE | Quality of Experience |
| QoS | Quality of Service |
| RAN | Radio Access Network |
| RFC | Request For Comments |
| RNC | Radio Network Controller |
| RNL | Radio Network Layer |
| RRM | Radio Resource Management |
| RTSP | Real Time Streaming Protocol |
| SCADA | Supervisory Control And Data Acquisition |
| SCTP | Stream Control Transmission Protocol |
| SeGW | Security Gateway |
| SGSN | Serving GPRS Support Node |
| S-GW | Serving Gateway |

| | |
|---|---|
| SIP | Session Initiation Protocol |
| SIPTO | Selected IP Traffic Offload |
| SLA | Service Level Agreement |
| SMTP | Simple Mail Transfer Protocol |
| SRV | Service [Location Resource Record] |
| TAI | Tracking Area Identity |
| TCP | Transmission Control Protocol |
| TE | Traffic Engineering |
| TNL | Transport Network Layer |
| UDP | User Datagram Protocol |
| UE | User Equipment |
| UL | Uplink |
| UMTS | Universal Mobile Telecommunications System |
| URI | Uniform Resource Identifier |
| UTRAN | UMTS Terrestrial Radio Access Network |
| VLAN | Virtual LAN |
| VoIP | Voice over IP |
| VPN | Virtual Private Network |
| WLAN | Wireless Local Area Network |
| WMN | Wireless Mesh Network |

# 1 Introduction

The research project MEVICO investigates aspects of the 3GPP LTE-mobile broadband network for its evolution in the mid-term in 2011-2014 and beyond. The goal is to contribute to the technical drive and leadership of the Evolved Packet Core (EPC) network of the 3GPP, and thus support the European industry to maintain and extend its strong technical and market position in the mobile networks market. The project follows an end-to-end system approach on evolution of the EPC. The focus will be on the connectivity layers of the system, for example on the part of the future LTE network which provides the efficient packet transport and mobility support for the applications & end-user services accessed over the LTE and LTE-Advanced radio systems. The technical research areas of the project cover relevant topics in the areas of network architecture, mobility & routing, packet transport, traffic management, network management & engineering and techno-economic aspects. The project will include both conceptual research and demo/trial system implementations.

This is the second document of the end-to-end (e2e) QoS and Traffic Engineering Architecture task in MEVICO. Here traffic engineering techniques which fall only in the research interests of the partners are described. The state of the art is described in each topic and this is followed by proposed improvements and new technologies in line with the research progress.

This document is structured in the following way:

- Section 2 gives six main traffic engineering building blocks and the rest of the document deals with three of them,

- Sections 3, 4, and 5 give a description of a set of solutions in these three building blocks. These descriptions are based on: literature study, partners' expertise and in some cases on partners' trials and measurements.

- Section 6 suggests three architecture options (centralized, distributed, flat). These solutions will be finalized in the next deliverable.

# 2 Traffic engineering building blocks

The building blocks described in this section provide an overview about potential mechanisms introduced in the MEVICO deliverable D 4.2.1 to improve quality of experience for the user and to enable efficient usage of infrastructure and IT resources. For the latter there is a benefit for other stakeholders in the (mobile) communication business, such as communication service providers (e.g. MNO), content providers and CDN providers. We have identified six different categories, which can be used to assign the various building blocks and aspects. Figure 2-1 displays the principal building blocks for traffic management with some given examples. There are three blocks which may be correlated with each other similar like lower layer functions provide services to higher layer functions in a communication stack.

The traffic engineering mechanisms described in this document are classified according to these building blocks. The building blocks have been detailed in D 4.2.1. In this document, only the topics that fall in the MEVICO partners' research areas and where new methods are being suggested have been covered.

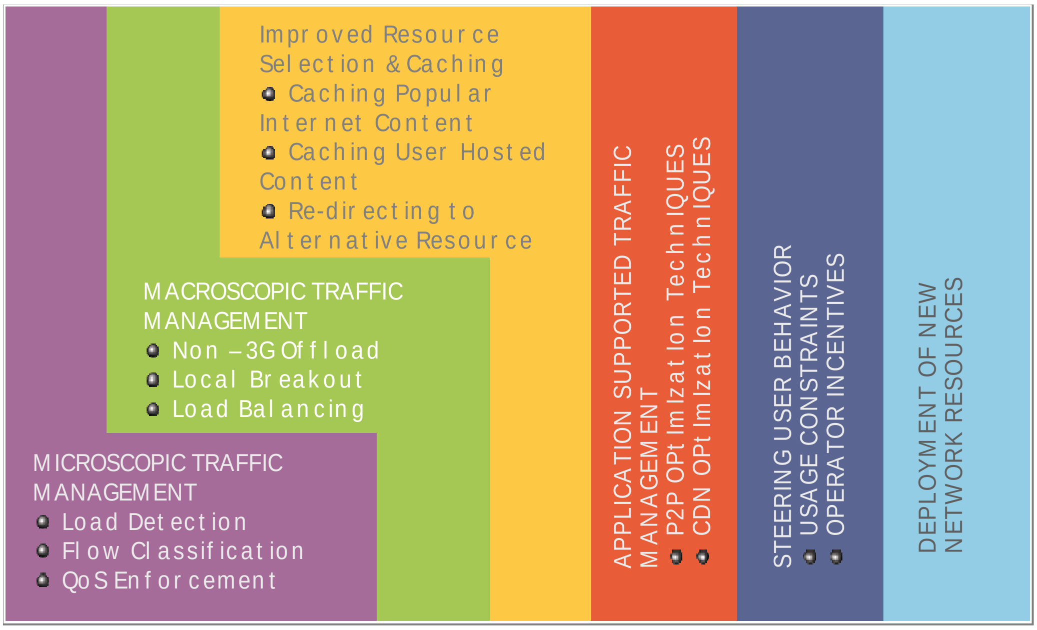

Improved Resource Selection & Caching
● Caching Popular Internet Content
● Caching User Hosted Content
● Re-directing to Alternative Resource
MACROSCOPIC TRAFFIC MANAGEMENT
● Non –3G Offload
● Local Breakout
● Load Balancing
MICROSCOPIC TRAFFIC MANAGEMENT
● Load Detection
● Flow Classification
● QoS Enforcement
APPLICATION SUPPORTED TRAFFIC MANAGEMENT
● P2P OPtImIzatIon TechnIQUES
● CDN OPtImIzatIon TechnIQUES
STEERING USER BEHAVIOR
● USAGE CONSTRAINTS
● OPERATOR INCENTIVES
DEPLOYMENT OF NEW NETWORK RESOURCES

**Figure 2-1:** ***Building blocks of traffic engineering mechanisms***

In the current case microscopic traffic management (MicTM) may provide services to macroscopic traffic management (MacTM). MicTM includes flow specific mechanisms, e.g. to improve quality of experience for the user. MacTM deals with the manipulation of routes through the networks, e.g. to improve efficient usage of resources. In addition to microscopic and macroscopic traffic management, a third group addresses the selection of resources in conjunction with caching, if necessary. This building block may rely on services of both microscopic and macroscopic traffic management. A resource in this context is associated with specific (multi-media) content, which is requested by users. All the above mentioned categories are associated with mechanisms that may require support from lower layers (below application). In addition we have identified two more building blocks which may require only little or no support from lower layers. These could be in place without dependence on other traffic management building blocks. On one hand there is application supported traffic management. There are many applications based on CDN and P2P, which try to optimize performance from end user perspective without getting support from network elements. Another building block has been identified, which is more relevant from business perspective without too many technical aspects. Mainly network operators but possibly other stake holders as well may influence user behaviour by defining certain constraints for usage of networks / services and certain incentive to comply with the usage constraints. Deployment of new network resources is identified as the last building block in this construction.

It should be noted that it is very hard to totally separate the above building blocks, there is certainly overlap among them. The most relevant overlaps within the MEVICO context have been addressed in Section 6, the Integration section.

For the purpose of scenario definition we want to introduce a simplified role model in order to understand the challenges, problems and requirements associated with the different stakeholders. These are the following:

- End user
- Communication service provider (CSP) – with a special viewpoint on mobile operators
- CDN Provider – usually a company with large amount of infrastructure to distribute content as close as possible to the end user. Customer is usually the content provider.
- Application / content provider – commercial or non-commercial entity, which inserts content for global or limited use in the Internet.

The roles of CSP, CDN Provider and content provider could be intertwined. For example, Google provides content but also has deployed a huge infrastructure to deliver it to the end user. On the other hand there are some operators, which want to make additional business by deploying CDN infrastructure or provide services via “walled-garden” models. This way end users should be motivated to connect to resources within the CSP or associated domains.

# 3 Macroscopic traffic management

This section proposes algorithms and mechanisms for macroscopic traffic management, in the areas of intra- and inter-technology cell (re)selection, selection of core network elements, traffic offloading, and cross-layer interference detection.

## 3.1 Access technology reselection

Today's access networks differ significantly with respect to coverage area, supported degree of mobility, and user data rates. Single WLAN hotspots can offer high data rates in small coverage areas without any mobility support. On the contrary, 3G networks support high mobility with large coverage but significantly lower data rates. The original design of WLANs has been driven by data service support, while that of cellular networks has targeted voice services. In course of the continuous evolution, both are targeting nowadays a complex mix of services. Furthermore, modern user devices such as smart-phones or tablets are equipped with multiple network interface cards (NICs) and are thus able to use either of several available access technologies.

Over the past years, an enormous growth regarding the traffic demand of mobile users has been reported. As a result, providers started to “offload” data traffic from 3G networks to WLAN hotspots or femto cells. For 2010, Cisco reports that the average amount of traffic per smart-phone doubled, while on a global level 31 percent of smart-phone traffic was offloaded. Predictions for 2015 assume that a smart-phone will generate as much as 1.3 GB traffic per month on average [CIS10] .

To allow data offloads to WLAN hotspots without requiring any actions from the end users, the Hotspot 2.0 Task Group in the Wi-Fi Alliance targets a quick-and-easy Wi-Fi authentication and roaming process on the basis of existing standards. With this, industry partners such as Cisco aim at changing the untrusted Wi-Fi into a trusted network leading to a user experience comparable to that of 3G [CIS11] .

Recently, “Smart Offload"” extensions have been proposed, where additionally all services of mobile operators (voice, SMS, MMS, etc.) may be offloaded to WLAN hotspots [Kineto10] . This enables the operator not only to maximize capacity and coverage, but also to offer free or low-budget solutions, such that the user does not move to competitors.

While today data offloading either happens on the discretion of the user or just simply anytime it is possible, our work goes a big step further. We target to select which data streams should conduct a handover so as to maximize the number of VoIP calls as well as the volume of data traffic that is offloaded to the WLAN hotspot. We have however to assure that the Quality of Service (QoS) level for each application must not undergo a certain minimal threshold. Thereby, we focus on traffic from users with low or no mobility as it is expected that more than 80 percent of the mobile data traffic will appear indoors in office or home scenarios [NokiaSiemensNetworks_2]. Among this static traffic, we claim to be able to make a proper selection that maximizes the WLAN utilization and thus allows higher volumes of offloaded traffic.

In the literature, several studies have identified the amount of traffic that could be offloaded today. Balasubramanian et al. [BMV10] report about 11 to 30 percent with immediate offloading considering a vehicle mobility pattern in city environments. Lee et al. [LRL+10] gain about 65 percent offloads for mobility and traffic patterns of every-day smart-phone users. In both studies, traffic was offloaded, if WLAN connectivity was available.

The approaches presented so far either apply just rough WLAN capacity estimations or implicitly assume that there is enough free capacity in WLAN. With the growing traffic demand, this premise becomes increasingly questionable. Therefore, it is necessary to select which traffic streams will be taken in WLANs. Wiethölter et al. [WW09] proposed a cost-function approach that evaluates the suitability of a traffic stream for a transport in an IEEE 802.11 hotspot with respect to its actual load on the wireless channel.

Now, this work extends the methodology of [WW09] as it tackles the following open issues: firstly, we demonstrate that, using the approach of Wiethölter et al., the maximum, offloadable number of traffic streams significantly improves as compared to Received Signal Strength Indication (RSSI) and random decisions. Secondly, we consider mixes of real-time and elastic traffic, i.e., VoIP and FTP streams, rather than only homogeneous traffic. Thirdly, as different decision schemes lead to different operational points of the WLAN network, our results allow operators to fine-tune the offloaded traffic mix such that it maximizes the utilization of WLANs.

This current work is not connected to any current 3GPP/IEEE standardisation activities. Proposed solution can be applied independently of a specific set of standards, making it therefore more compelling. In the following, Section 3.1.1 discusses the System Model, while the decision schemes are presented in Section 3.1.2. Then, Section 3.1.3 presents the methodology for the planned performance evaluation.

### 3.1.1 System Model

The system under consideration uses a cellular technology (CT) as basic means of communication enhanced by WLAN hotspots. One mobile operator owns both, CT and WLAN. The WLAN access cells are completely within the coverage area of CT. We assume that all terminals are equipped with a NIC for each access technology such that they can perform vertical handovers, which are conducted seamlessly, i.e., unnoticeable for the end users.

Throughout this work, we assume that highly mobile end users are served by only CT. In contrast, all stationary users within CT may be subject for consideration as handover candidates if they are within the coverage of one of the WLAN hotspots. The movement of streams happens, by the way, not necessarily one way. In order to accommodate the maximum number of streams at WLAN hotspots, some data streams may be pushed back from WLAN into CT.

We apply the following algorithm with which we decide for or against handovers from and to the WLAN hotspot. In our simulations, we evaluate at each interval $\Delta t_{HO}$ whether the QoS values of the served stations (STAs) in WLAN have been violated. If no violation has occurred, we accommodate an additional user in the WLAN cell by selecting him in a random-uniform fashion from all STAs being within WLAN coverage as well as being connected to CT. In case that QoS limits have been violated for at least one STA, we calculate the respective decision metric. Then, we handover the STA with the worst metric value from WLAN. To account for the situation that QoS limits of a STA are violated rather because of a bad channel instead of an overload situation, we introduced a QoS penalty. That is, if the QoS for a certain STA has been violated continuously for more than five $\Delta t_{HO}$, it is selected for a handover, too.

In principle, there exist three general concepts regarding the placement of the handover/offloading decision entity. This can be realized within the WLAN network, the CT network, or by a separate arbitration entity. We assume the signaling delay between such entities to be in the range of some tens of milliseconds leading to a negligible impact. As a result, for the topic investigated in this paper, it is not necessary to make any assumptions regarding the mapping of decisions on specific entities.

### 3.1.2 Decision schemes

We firstly describe the existing approaches for WLAN offloading, with which we are going to compare our novel solution. We refer to the former as “Comparative Schemes”. Afterwards, we summarize the design of our “Cost-Function Approach” presented in [WW09].

#### 3.1.2.1 Comparative Schemes

*3.1.2.1.1 RSSI*

For vertical handover decisions, the most common approach conducts a handover once the received signal strength of a terminal undergoes a certain threshold [YSN10]. Accordingly, the first selection scheme relies on RSSI measurements in WLAN: for each traffic stream, RSSI-values are collected on the receiver side(s) for each successfully received data frame and averaged over $\Delta t$. Terminals with the lowest RSSI-values are selected as handover candidates.

*3.1.2.1.2 Random Selection*

The simplest approach offloads traffic to WLAN once the connectivity is available. To capture this somewhat simple behavior also for the backward handover direction, we apply a decision scheme that selects terminals randomly from the 802.11 cell in a uniformly distributed fashion. In other words, each terminal in WLAN has the same probability of being selected for a handover to CT.

#### 3.1.2.2 Cost-Function Approach

In a WLAN cell, the capacity cannot simply be determined just by accounting transmitted bits per second. In particular, the achievable aggregated throughput highly depends on the number of contending terminals, terminals' positions, the channel quality, as well as on the transported traffic patterns. Thus, approaches with only rough throughput-based estimates regarding the remaining capacity (cf. the survey in [YSN10]) are not well suited for the maximization of the number of traffic streams in a WLAN cell. As a result, we follow our approach presented in [WW09], where the cost function of the WLAN access cell evaluates the load together with the effectiveness of occupied resources for each traffic stream *m*,

$$c_{\mathrm{WLAN}}(m) = \omega_1 \frac{t_a(m)}{\Delta t} + \omega_2 \frac{D(m)}{D_{\max}(m)}, \text{ with weights } \omega_1 + \omega_2 = 1. \qquad (1)$$

The ratio $t_a/\Delta t$ consists of the occupied airtime on the channel in relation to measurement interval $\Delta t$. The airtime $t_a$ represents the amount of time that the wireless medium has been occupied (or reserved, in case of inter-frame spaces and NAV settings) for the transmission of all packets *K* of traffic stream *m* during *Δt*:

$$t_a(m) = \sum_{j=1}^{K} t_{a_j}(m) = \sum_{j=1}^{K} \sum_{i=1}^{\text{trials}} t_{\text{IFS}} + t_d(R_{i,j}) + t_{\text{ack}} \,. \tag{2}$$

This includes the whole transmission sequence consisting of the inter-frame spaces DIFS or AIFS and SIFS ($t_{\text{IFS}}$), the duration $t_d$ of the complete data frame "on air", where the data part is encoded with a certain modulation scheme $R_{i,j}$ and the acknowledgment $t_{\text{ack}}$. The number of trials represents the transmission attempts that have been required to ensure the delivery of a single MSDU.

Further, *D* represents the inefficiency metric that evaluates the resource utilization on behalf of each traffic stream. For a complete discussion of the metric design, the reader is referred to [WW09] here we shortly summarize just the basic ideas. Overall, we target to identify certain users or traffic flows that contribute significantly to the load in the access network but benefit only marginally from these expenditures. Such behavior is denoted as "inefficiency" in the following.

Inefficiency *D* consists of two parts: the surcharge $\zeta$ and the overhead factor $\alpha$. While the surcharge is a measure for additional expenditures required for error control and correction, the overhead factor allows for an evaluation of different data packet sizes regarding their suitability in WLANs. In the following, both parts as well as their composition to the final inefficiency metric are discussed shortly.

#### *3.1.2.2.1 Surcharge*

This part is derived from the very basic definition of efficiency: In engineering, efficiency is usually defined as relation of system's output $\nu$ to the overall effort $\psi$ one has to invest:

$$\eta = \frac{\text{output}}{\text{effort}} = \frac{\nu}{\psi}. \tag{3}$$

The design rationale behind this part is to identify terminals with smallest efficiency values as handover candidates. However, it may be difficult to distinguish between two very small efficiency values near zero although the corresponding difference of effort values may be significant. Hence, the reciprocal is applied to enable comparability:

$$\text{surcharge} \quad \zeta = \eta^{-1} = \frac{\psi}{\nu} = \sum_{j=1}^{K} \frac{t_{a_j}}{\nu_j}. \tag{4}$$

As effort $\psi$, we consider the duration for the complete transmission sequences (Eq. 2). In contrast, we define system's output at MAC level as the smallest possible duration for each of these transmission sequences that would be required in case of an ideal error free channel:

$$\nu_j = \Delta t_{\text{opt}} = t_{\text{IFS}} = +t_d(R_{\max}) + t_{\text{ack}} \,. \tag{5}$$

The output definition includes the duration of the whole data frame when the data part is encoded with the highest modulation $R_{\max}$ for a single, successful transmission.

#### *3.1.2.2.2 Overhead Factor*

IEEE 802.11 introduces a fixed amount of overhead (PHY framing, inter-frame spaces and immediate ACK) for one transmission regardless of the MSDU size. Thus, the smaller the MSDU size, the less optimally 802.11 becomes utilized. To accommodate this behavior, we introduce the overhead factor as

$$\alpha = 1 - \frac{\Delta t_{\text{MSDU}_{\text{opt}}}}{\Delta \text{t}_{\text{opt}}} = 1 - \frac{\Delta t_{\text{MSDU}_{\text{opt}}}}{\Delta \text{t}_{\text{MSDU}_{\text{opt}}} + \Delta t_{\text{oh}}} \tag{6}$$

While $\Delta t_{opt}$ is again the smallest possible duration for a frame exchange (Eq. 5), $\Delta t_{\text{MSDUopt}}$ represents the duration of the bare MSDU assuming the highest modulation scheme. $\Delta t_{oh}$ includes all necessary overheads due to framing, inter-frame spaces, and immediate ACK. The overhead factor starts for small MSDU sizes shortly beneath one and continuously decreases with larger MSDU sizes. We slightly changed the previous definition in [WW09] as small packets such as TCP ACKs were penalized too strongly.

#### *3.1.2.2.3 Metric Composition*

To allow the handover candidate selection among users with heterogeneous traffic patterns, both parts are combined to the inefficiency metric such that the overhead factor $\alpha$ serves as a penalty factor for the surcharge:

$$D = \alpha\zeta. \qquad (7)$$

Within this work, we calculate surcharge $\zeta$ over intervals $\Delta t$ of 100 ms, which is in the order of a WLAN beacon interval.

*3.1.2.2.4 Cost-Function Schemes*

Extending the previous work in [WW09], here we consider two flavors of the Cost-Function Approach, denoted as ”Inefficiency” and ”Equal Weight (EW)” decisions, with different weightings of their parameters. The first one only takes into account the Inefficiency of the wireless transmissions (i.e., $\omega_1 = 0$, $\omega_2 = 1$ in Eq. 1).

In order to penalize inefficient traffic streams evoking a high channel load, ”Equal Weight”' will consider the impact of a mix of Inefficiency and wireless channel load, measured by the occupied airtime, by setting $\omega_1 = 0.5$, $\omega_2 = 0.5$ in Eq. 1.

### 3.1.3 Outlook: Methodology for the Performance Evaluation

We will compare the performance of the four selected decision schemes regarding the number of VoIP flows and the volume of data traffic that can be accommodated by a WLAN cell thus unloading CT. To trade off both traffic types against each other, we will additionally consider the question how much FTP traffic one can transport at the costs of a reduced number of VoIP clients. For this, we will apply a two-stage process: we will firstly determine the capacity of the WLAN cell in terms of VoIP users which can be simultaneously served without violation of QoS constraints. In the second stage, we will consider VoIP and FTP traffic mixes in the WLAN cell.

## 3.2 Selection of core network elements

EPS supports resilience (through NE redundancy), optimized routing, balancing of data plane and control plane traffic load in the NEs, and sharing of eNBs among mobile operators. These features are enabled by

- Separation of control and data plane functions in different NEs: MMEs and S-GWs;
- Grouping of MMEs and S-GWs, which allows an eNB to be connected to one or more MMEs and S-GWs in a pool; and
- Flexibility in PDN GW selection.

### 3.2.1 Initial selection of core network elements

One fundamental change from 3GPP Release 7 to Release 8 was the definition of a new core network. The main difference of the so called evolved packet core (EPC) in Release 8 from the old packet core of Release 7 is the separation of user and control plane. With that, the control plane und the user plane can scale independently: control plane nodes and the user plane nodes can be placed and selected separately. This is convenient, because the user traffic is expected to grow faster than the control traffic.

In general the selection process of the core network elements is based on the domain name system (DNS) [3GPP_29.303]. The selection entity makes a DNS request to a DNS Server to obtain a list of possible network elements. This list is sorted with respect to various criteria and then the selection entity chooses the first entry in the sorted list. If this network element is not available, it takes the next entry until a network element responds or the list is exhausted.

**Control Plane**

The control plane consists of one element: the mobility management entity (MME). The MME is selected by the eNodeB during the initial attachment of a user. The main criterion for MME selection is to check, if the MME serves the eNodeB on which the UE attaches. In case more MMEs are serving the area (e.g. the MMEs belong to the same MME pool area), a weight factor could be used for statistical load balancing among the MMEs. The weight factor is based on the capacity of a MME compared to the other MMEs in the same pool area. The probability for a MME to be selected from an eNodeB is proportional to its weight factor. Another criterion for selection is to minimize possible MME relocation events, see [3GPP_23.401].

**User Plane**

The user plane consists of two network elements: the serving gateway (S-GW) and the packet data network gateway (PDN GW). These two gateways are selected by the MME during an initial attach of a user. Another PDN GW selection procedure is needed when the user establishes an additional PDN connection.

The MME uses information provided by the HSS or the DNS Server to select an appropriate PDN GW for a PDN connection. The address of a PDN GW used for a certain APN could be included in the subscription profile of a user in the HSS. In this case this address has to be used and the DNS aided selection by the MME is skipped. In all other cases the MME performs a DNS request (including the APN) to obtain a list of suitable PDN GWs. This list may also include weight factors for statistical load balancing and/or might be sorted, to indicate that a certain PDN GW should be preferred for the requested APN. This kind of sorting could be used for example to encode the distance of the PDN GWs to the service domain. The weight factor works similar to the weight factor of the MMEs.

The main criterion for selecting an S-GW is the Tracking Area Identity (TAI) or ID of the eNodeB serving the user. The MME performs a DNS request, including the TAI or eNB-ID to obtain a list of suitable S-GWs. This list may include weight factors for load balancing or may be sorted according to the distance of the S-GWs to the eNodeB. Another criterion mentioned in [3GPP_23.401] is to minimize the probability of S-GW reselections.

After the MME has obtained the list of S-GWs and PDN GWs it sorts these two lists so that collocated S-GW/PDN GWs or pairs of topologically close S-GWs/PDN GW are on top position. To sort the S-GW and PDN GW according to their topological proximity is an optional feature and has to be indicated through the DNS response. If one gateway had been already preselected (PDN GW through HSS information, S-GW during additional PDN connection establishment) only the list of the other gateway is reordered.

So far the following criteria could be considered for GW selection according to [3GPP_23.401] and [3GPP_29.303]:

- The used service class (known from the used APN), [3GPP_23.401]
- The area served by a S-GW [3GPP_23.401]
- The distance between S-GW and eNodeB (known from the ordered list in the DNS response) [3GPP_29.303]
- The distance between PDN GW and the service domain (known from the ordered PDN GW list in the DNS response) [3GPP_29.303]
- The load of the GWs [3GPP_23.401]
- The topological proximity of S-GW and PDN GW (known from the structure of the DNS name of the GWs) [3GPP_29.303]

To support traffic management also in distributed gateway scenarios it could be beneficial to include additional criteria into the GW selection process like:

- Load of the transport network
- Mobility behavior of users (e.g. low mobile, high mobile)
- Access networks supported by the GWs and the UEs
- A finer differentiation of the used services
- Other criteria (Applications, QoS)

These new parameters can be included either via DNS records or by changing the gateway selection procedure, performed by the MME. The selection procedure of the PDN GW can be also adjusted through changing the information held in the HSS.

### 3.2.2 Reselection of core network elements

Besides the initial selection of the core network element a reselection might be necessary during operation. Today the main reasons for reselecting network elements are mobility events like tracking area updates and handovers. In MEVICO it is intended to use the reselection functionality also for traffic management. In the following, only reselection mechanisms for user plane entities are considered, because the control plane traffic has a much lower share of the overall traffic compared to the user plane traffic.

S-GW reselection is standardized since 3GPP Release 8 (see [3GPP_23.401]). It can be performed mainly in reaction to user mobility causing either during a Handover event (if the UE is in active state), or during a Tracking Area Update (if the UE is in idle state). The MME decides if an S-GW change is necessary. This might be the case because the UE has moved into a tracking area which is not served by the old S-GW. Other reasons for S-GW reselection, like load balancing, are not specified in [3GPP_23.401] so far, but should be a subject of the investigations performed within MEVICO.

Since Release 10 PDN GW relocation is standardized as well (see [3GPP_23.401]). The current procedure is that the MME first triggers a PDN disconnection and then a PDN reactivation after the Tracking Area Update, such that the connection to an APN through a PDN GW is shut down and a new connection to the same APN through a new PDN GW is established afterwards. This method doesn't allow a seamless change of the connection, because the IP address in the UE changes as well. In general, this procedure can be performed at any time. However, it is convenient to perform the PDN GW reselection when the UE changes its state from idle to active.

## 3.3 Change of routing within backhaul

Mobile backhaul is the transport network and the equipment that is used for connecting together mobile system elements in the RAN and for connecting them to the corresponding elements in the mobile core network (in LTE, for connecting eNBs together and connecting them to MMEs and S-GWs). The transport functions integrated into the mobile system elements themselves are part of the mobile backhaul network. The transport service type needs to match with the connectivity requirements (non-optimal routing paths should be avoided to make efficient use of network resources and to minimize packet delay).

Routing changes in the backhaul involve the following considerations:

- The alterability in routing is constrained by the underlying physical network topologies;
- Traffic flows requiring different QoS may be routed via separate routing paths that offer distinct packet forwarding treatment;
- For network survivability, redundant routing paths need to be provisioned in order to recover from node and link failures;
- Dynamic re-routing of traffic may also involve re-allocation of transport capacity, if such flexibility is available;
- Routing changes within the backhaul should be coordinated with any load-balancing and offloading decisions among mobile NEs.

In the traditional tiered mobile access networks, the routing topology has been a tree rooted at the controller site. This has changed since the introduction of distributed and pooled mobile NEs and local breakouts. Besides, user plane traffic between eNBs is expected to increase in the future, beyond the temporary tunnelling over LTE's X2 interface in handovers. Therefore, multipoint-to-multipoint (MP2MP) transport service may be desired between mobile NEs instead of the traditional point-to-multipoint (P2MP) connectivity. A meshed physical topology allows freedom in route computation in the backhaul (e.g. consider load balancing in a last-mile microwave mesh).

The mobile NEs, as customer equipment to the transport network, have only limited means for affecting routing within the backhaul (such that the re-routing of traffic flows would not be a consequence of re-selection of mobile NEs), especially if the mobile operator leases the transport service from a network service provider. If the NEs are multi-homed at the IP layer they can use alternative IP source addresses (i.e. alternative egress interfaces) or alternative IP destination addresses in Stream Control Transmission Protocol (SCTP) [RFC4960], in S1 control plane. In Ethernet transport, the NEs may tag frames with one or two Virtual LAN (VLAN) Ids and with Priority Code Points (PCPs). The VLAN Ids may designate different operators, traffic types, and/or service classes. In case the mobile access network utilizes MPLS in higher aggregation layers (or even in lower layers), these Ids, along with the PCPs, may then be mapped to separately routed Label Switched Paths (LSPs) (however, the VLAN tag assignment and their mapping to LSPs is normally static). For splitting of TCP flows across multiple paths, see section 4.3.2.

In addition to IP transport, many network providers have under-laid their core networks with MultiProtocol Label Switching (MPLS). MPLS establishes direct source-to-destination paths between the edge routers of an MPLS domain and enables basic traffic management functions, such as standard measurement per path and for the complete traffic matrix. Moreover, backup paths can be provided in MPLS for failure cases with fast rerouting mechanisms. Re-routing initiated within the transport network itself typically happens only in failure situations. A recovery operation moves all traffic on a path or a path segment (on a certain layer, e.g. link, LSP or pseudowire) to an alternative path that may be pre-computed (e.g. linear/ring protection or Fast Reroute in MPLS, see [RFC6372]) or computed on-demand (e.g. spanning tree formation in Ethernet or restoration in OSPF). An "idle" protection path may be allowed to carry best effort "extra traffic", which may be pre-empted in the protection switch. In MPLS networks, the constraint-based path (re-) computation may follow the Path Computation Element (PCE) – based architecture [RFC4655], which allows flexible distribution of route computation and signalling components.

In principle, a transport network may balance traffic load among "Equal Cost Multi-Paths" (ECMP), but this is problematic for Operation, Administration, and Maintenance (OAM) in connection-oriented network modes. Dynamic load balancing that would be adaptive to variation in traffic loads is still a research issue (e.g., see [Susitaival09]) especially when IP routing based on the shortest path first (SPF) principle is in use. Then the routing weights have to be manipulated in order to balance the traffic, which ends up in NP-hard optimization problems such that heuristics have to be applied. Without such complex optimization techniques, ECMP randomly splits up traffic paths leading to a slight but no strict traffic smoothing tendency.

MPLS networks allow for simpler TE algorithms, i.e. linear programming, to establish an optimized LSP path design for load balancing, since each path between edge nodes (LERs) can be chosen arbitrarily. As discussed previously, this solution has to include relevant failure scenarios. If load balancing in normal operation is applied to improve resource utilization this can have adverse effect in failure situations due to lack of resources for recovery. Therefore load balancing has to take into account all relevant failure cases such that the optimized solution assures enough throughput even in the worst of the considered failures. Network operators have developed traffic engineering tools for MPLS, e.g. TE-Scout [Haßlinger05], which support the complete traffic engineering cycle from monitoring, triggering of a re-optimized path design to reconfiguration of changes in LSPs.

In addition, fast failure response and situation awareness is mandatory, which again is well established in MPLS. On the other hand, mobile network technologies and self-organizing environments (SON) are often much more challenging with regard to traffic monitoring and engineering as compared to MPLS [Haßlinger11].

A Radio Network Layer's (RNL) traffic steering solution, e.g., initial selection and re-selection of core NEs (Section 3.2) or traffic engineered handover between base stations or across radio access technologies (Section 3.4.3), which selectively moves UE traffic flows to alternate serving mobile NEs, should be aware of the traffic load and the available capacity on the Transport Network Layer (TNL) [Zhao11]. Therefore, there is a need for coordination of re-routing actions by co-operation between the RNL and the TNL. Ideally, the cross-layer coupling should cover the whole backhaul, which in practice may need to be realized through a comprehensive network management system that covers both layers end to end.

One of the main success criteria for the solutions is the end users' QoE, which is hard to measure and which implies that traffic steering needs to be application-aware. To this end, the steering control algorithms may make use of deep packet inspection of user traffic.

## 3.4 Offloading techniques

The enormous growth of mobile data traffic has surprised mobile operators and vendors by choking existing communication networks a lot earlier than expected: bottlenecks of anchor nodes, sub-optimal routes and hierarchical/centralized architectures became hot R&D topics. As a result of the progressive innovation novel technologies have emerged quickly to manage routing optimization of data flows, divert user packets around network bottlenecks and critical entities of the architecture, and to redirect traffic directly to/from the Internet. All the techniques around this problem space are called data offloading and used for any kind of complementary network technologies for optimized delivery of data originally targeted for cellular networks. The set of rules aiming at triggering the offloading action can either be set by the mobile operator or the mobile subscriber and the code operating on the rules can run in the user terminal, in a core network entity or can be divided between different elements of the architecture. For the end users the gain of applying offloading schemes relies on extended data service cost control and the availability of higher bandwidth. From the operators' perspective the main purpose for introducing offloading in the network is to avoid congestion of the cellular architecture due to mobile data traffic evolution. Key technologies include pico- and femtocells, Wi-Fi, Content Delivery Networks, routing and media optimization, and Core Network Offloading as the most important schemes to reduce the congestion and load on the operator's network.

This section explores the main drivers and demands for the most promising mobile data offloading techniques, including a state-of-the-art presentation of the existing schemes. The focus of this section is on 3G and beyond systems; therefore it closely studies the technologies available to enhance, augment and extend existing networks. The main evolutionary steps leading to new operator oriented solutions and offloading strategies are highlighted, and detailed introduction of Direct Tunnelling (DT), Selective IP Traffic Offload (SIPTO), Local IP Access (LIPA) and IP Flow Mobility (IFOM) are provided to illustrate the potential and power of mobile data offloading in advanced traffic engineering.

### 3.4.1 Evolution of mobile data offloading

A typical cellular mobile network today consists of a large number of different but similarly expensive network elements that are maintained in a hierarchical fashion and operated in a highly centralized system model. This centralized and hierarchical build-up exists since almost the beginning of the mobile and wireless telecommunication industry: in 1972 author of [Joel_72] presented a centralized Base Station Controller (BSC) which allows paging through a number of cells aiming at locating mobile nodes while Base Stations are attached to the Mobile Switching Center (MSC) for wired line connection. The architecture of Global System for Mobile Communications (GSM) also use BSC and MSC in the same manner for call processing, mobility support, switching, data forwarding, and voice encoding functions.

The 3GPP network architecture specifications 03.02 [ETSI_96] and 23.002 [3GPP_TS_23.002] show the evolution of the 3GPP network from GSM Phase 1 published in 1995 until the Evolved Packet System (EPS) specified in Release 8 in 2010. The core part of EPS called Evolved Packet Core (EPC) is continuously extended with new features in Release 10 and 11. The main steps of the architecture evolution towards the core network centric mobile offloading techniques and a more flat and distributed architecture are summarized in the subsections below (Fig. 2 illustrates the evolution steps of the packet-switched domain) followed by the introduction of the LIPA, SIPTO and IFOM schemes.

#### 3.4.1.1 The creation of hierarchical structure of the GSM CS domain (Phase 1 and 2)

In Phase 1 (1995) the main entities of the GSM architecture have been introduced. Base Stations and Base Station Controllers form hierarchy to handle radio access tasks. The core network includes the Home and Visitor Location Registers that provide information about the subscriber and the current location (VLR area) of the mobile station, and the mobile switching centre (MSC), which constitutes the interface between the radio system and the fixed networks, and performs switching and call handling. In Phase 2 (1996) additional entities have been introduced like the Authentication Center, Equipment Identity Register, Gateway MSC, Interworking Function between the Public Land Mobile Network and fixed networks (ISDN, PSTN, PDNs).

The reason behind the hierarchization and centralization of the GSM architecture was both technical and economical. Primarily it offloaded the switching equipment (cross-bar switch or MSC). In parallel, existing ISDN switches could be re-used as MSCs only if special voice encoding entities were introduced below the MSCs, hence further strengthening the hierarchical structure of the network.

#### 3.4.1.2 The formation of the packet-switched domain (from Phase 2+ to Release 6)

The main driver to introduce packet switching was that it allows multiplexing hence resources can be utilized in a greater extent. Circuit switching does not share an allocated virtual circuit with other transmissions. In Phase 2+ (1997) the packet-switched (PS) domain is described, hence General Packet Radio Service (GPRS) support nodes are added to the network. In the PS domain the Serving GPRS Support Node (SGSN) stores subscription and location information to each subscriber registered to it. The Gateway GPRS Support Node (GGSN) stores subscription and routing information needed to tunnel data from the Packet Data Network (PDN) to the MS for subscribers with active packet data protocol (PDP; e.g., IP, X.25, FrameRelay) context on the GPRS Support Nodes. This architecture guarantees with interfaces the transit between CS and PS domains.

Release 1999 (2002) describes the well known UMTS architecture clearly separating the CS and PS domains. Seeing that UMTS was designed to be the successor of GSM, it is not strange that the central anchors remained in place in 3G and beyond.

Progress of mobile and wireless communication systems introduced some fundamental changes. The most drastic among them is that IP has become the unique access protocol for data networks and the continuously increasing future wireless traffic is also based on packet data (i.e., Internet communication). Due to the collateral effects of this change a convergence procedure started to introduce IP-based transport technology in the core and backhaul network. The evolution can be seen in Figure 3-1.

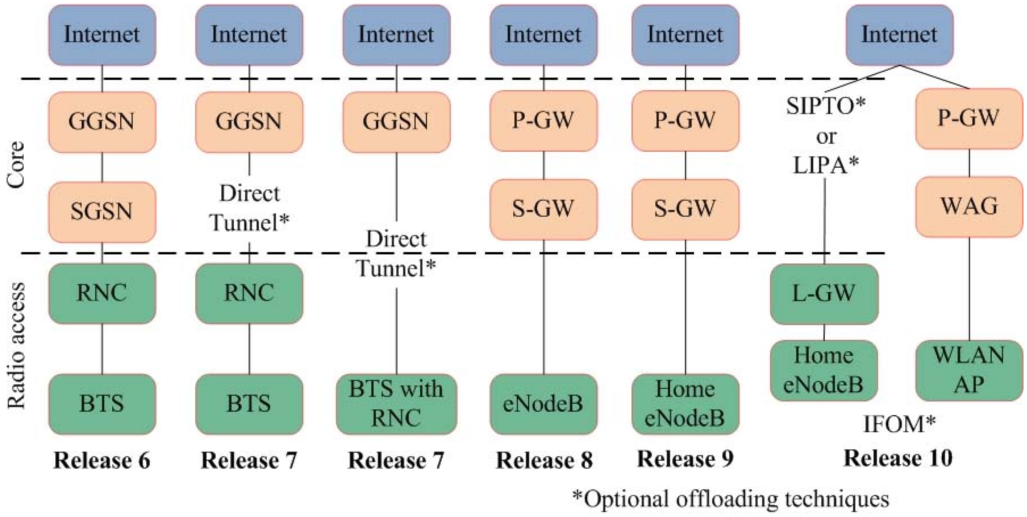


**Figure 3-1:** ***Evolutionary steps of the 3GPP packet-switched domain***

#### 3.4.1.3 Service convergence

In Release 4 (2003) the Media gateway function is specified for the CS domain to support media conversion, vocoding, payload processing and to translate between CS bearers, PS media streams. It is the premonitory sign of the introduction of the IMS core network functions in Release 5 (2003) for provision of IP services delivered over the PS domain. IMS provides a common service layer for IP based multimedia services, and aids the fixed mobile convergence.

Release 6 (2005) mainly introduces WLAN interworking and Multimedia Broadcast Multicast Service related entities for point-to-multipoint communication in the 3GPP architecture. Note that the latter scheme can be considered unsuccessful due to lack of practical deployment.

#### 3.4.1.4 The appearance of offloading techniques: transition towards a flat network architecture (Release 7)

The enormous increase of IP-based data traffic offloading and flattening of hierarchical and centralized functions became the main driving forces in the evolution of 3GPP network architectures.

Release 7 (also called Internet HSPA, 2008) supports the integration of the RNC with the NodeB providing a one node based radio access network. The address space of RNCs had to be extended to enable more than 4096 sites [Holma_074]. This architectural change was necessary to avoid performance bottlenecks caused by RNCs facing increasing IP traffic demands. Particularly the highly increasing smart phone generated traffic and Machine-to-Machine (M2M) demands represent challenges to network RNCs and SGSNs, because they generate high volume of tiny packets which highly utilize the central servers.

Another architectural enhancement of this release is the elaboration of Direct Tunnel service [3GPP_TR_23.919], [3GPP_TS_23.401] which can be considered as the first mobile data offloading technique standardized in 3GPP. Direct Tunnel allows offloading user traffic from SGSN by bypassing it. The Direct Tunnel enabled SGSNs can initiate the reactivation of the PDP context to tunnel user traffic directly from the RNC to the GGSN or to the Serving GW introduced in Release 8. This mechanism tries to reduce the number of user-plane traffic anchors. However it also adds complexity in charging inter-PS traffic because SGSNs can not account the traffic passing in direct tunnels. When Direct Tunnel is enabled, SGSNs still handle signalling traffic, i.e., keep track of the location of mobile devices and participate in GTP signalling between the GGSN and RNC. Removing RNC and SGSNs from the architecture enables I-HSPA to achieve ten times greater simultaneous throughput than WCDMA alone.

Release 7 specifies the Policy and Charging Control (PCRF) function which integrates QoS policy control and charging that were previously two separated functions. This makes possible more granular, IP flow-based charging and QoS enforcement for session-based applications in the PS domain, independently from the IP connectivity access network between the UE and the IMS entities.

#### 3.4.1.5 The Evolved Packet Core: a flat PS domain of the 3GPP architecture (Release 8)

Release 8 (2010) introduces a new PS domain, i.e., the Evolved Packet Core (EPC). Compared to four main GPRS PS domain entities of Release 6, i.e. the base station (called NodeB), RNC, SGSN, GGSN, this architecture has one integrated radio access node (called eNodeB) containing the precious base station and the radio network control functions, and three main functional entities in the core, i.e. the Mobility Management Entity (MME), the Serving GW (S-GW) and the Packet data Network GW (PDN GW).

The MME is the control plane entity supporting NAS signalling and security, inter core network node signalling for mobility between 3GPP access networks, PDN GW and Serving GW selection, SGSN selection for handovers to 2G or 3G 3GPP access networks, roaming, bearer management including dedicated bearer establishment, lawful interception of signalling, support for handovers to 3GPP2 access networks. The Serving GW is the user plane interface towards Evolved UTRAN. It is the local mobility anchor for inter-eNodeB and inter-3GPP handover. It provides lawful packet interception, accounting, and event reporting to PCRF. The PDN GW is the gateway towards the PDN. Its functions include per-user based packet filtering with deep packet inspection, lawful interception, UE IP address allocation, mobility management, DHCPv4 and v6 functions for configuration, transport level packet marking, uplink and downlink service level charging and rate enforcement. Please read Section 3.2.1 for more specific details on the above architectural elements of 3GPP release 8.

#### 3.4.1.6 Femtocells: offloading the macro-cellular radio access network (Release 9)

Release 9 (2010) introduces the definition of Home (e)NodeB Subsystem. These systems allow unmanaged deployment of femtocells at indoor sites, providing almost perfect broadband radio coverage in residential and working areas, and offloading the managed, pre-panned macro-cell network. A study [FemtoForum_10] describes that network capacity is growing at 29% per year, and demand is currently growing at 108% per year, which is a significant difference needing architectural innovation, i.e., offloading the macro-cell network. Informa released a report on mobile access at home, and they forecast that by the year 2012 55% of all mobile data usage will occur at home and 26% in the office [Informa_08]. Subscribers whose data is offloaded are not average users but the most demanding users because they are sitting behind radio-wave absorbing walls. Adding 1 Mbps of capacity using a femtocell is approximately 1/200th of the cost of adding the same capacity using a macro-cell infrastructure [Neu_Mobile_09]. Femtocells represent many technical challenges which include resource allocation, timing/synchronization, backhaul capacity with appropriate QoS, interference management with macrocells and femtocells, handoffs, mobility, emergency service provision facing that users may switch off HNodeBs, open or restricted access to a group of subscribers or the combination of both, and securely bridging the femtocell with the core network [Chandrasekhar_08]. Several hardware manufacturers are dealing with femtocell solution development nowadays. On one hand there are the traditional vendors in the telecommunications industry like Motorola, Alcatel Lucent, Huawei, Samsung, and Nokia Siemens Networks. On the other hand specialized companies focusing strictly on femtocells also do exist: Radioframe, Ubiquisys are good examples. Airave by Sprint USA can be considered as the biggest commercial femtocell deployment so far. The official launch date was back in September 2007 making the service available in the Sprint Nationwide network [Fiercewireless_10]. Existing problems and technical challenges are not really solved, which has been pointed out lately by the Vodafone femtocell hack even letting intruders listen to on-going calls [wikithc]. Low-power base stations can eliminate coverage holes in the macro-only system; however these systems require smarter radio resource coordination and interference management among base stations. As the spectral efficiency per link is approaching theoretical limits with the introduction of new radio access technologies (LTE Advanced), this generation of technology is about improving spectral efficiency per unit area, deploying pico-, femtocells and relays. This also requires smarter radio resource coordination and distributed optimization algorithms communicating over the backhaul network, e.g., the X2 interface of the eNodeBs [Khandekar_10].

#### 3.4.1.7 Selected IP traffic offload, Local IP Access and IP Flow Mobility (Release 10)

In Release 10 (2010) Selective IP Traffic Offload (SIPTO), Local IP Access (LIPA) and IP Flow Mobility services have been published [3GPP_TS_23.401], [3GPP_TR_23.829]. These enable local breakout of certain IP traffic from the macro-cellular network or the H(e)NodeB subsystems, in order to offload the network elements in the PS and EPC PS domain.

The LIPA function enables an IP capable UE connected via Home(e)NodeB to access other IP capable entities in the same residential/enterprise IP network without the user plane traversing the core network entities. LIPA is established by the UE requesting a new PDN connection to an APN for which LIPA is enabled, and the network selecting the Local GW associated with the H(e)NodeB, and enabling direct user path between the Local GW and the H(e)NodeB. Mobility of LIPA PDN connections is still not supported in Release 10.

SIPTO enables per APN and/or per IP flow class based traffic offload towards a defined IP network close to the UE's point of attachment to the access network. In order to avoid SGSN/S-GW from the path, Direct Tunnel mode should be used. Technical challenges of traffic offload are traffic engineering, mitigating security compromises on third party broadband access network, charging, and policy enforcement of the mobile operator. Yi-Leng et al. have demonstrated a prototype for offloading Ethernet-based cellular backhaul interfaces [Yi-Neng_Lin_10].

In Release 10 also the IP Flow Mobility has been specified [3GPP_TR_23.861], [3GPP_TS_23.261] in order to provide an efficient toolset for selective assignments of different UE traffic flows to separate radio access networks implemented by different technologies (e.g., LTE and WLAN). The scheme enables seamless IP flow mobility and allows the operator to indicate how different IP flows of UEs are routed through the access systems and to selectively offload some traffic (i.e., specific flows) to the WLAN segment while using UTRAN/E-UTRAN for other traffic.

### *3.4.2 Promising offloading techniques*

This section provides more details about the most promising offloading techniques in the literature.

#### 3.4.2.1 Direct Tunnel (DT)

The Direct Tunnel is an optional feature in the User Plane operation of 3GPP standards and allows the SGSN to create a direct tunnel between RAN and GGSN (in case of GGSN connectivity through Gn/Gp) or S-GW (in case of S-GW connectivity through S4) for user plane communication in the packet switched domain. The scheme is introduced in TS 23.060 [3GPP_TS_23.060] by defining new functions in the SGSN to handle the control plane signalling and also to decide when to create a Direct Tunnel. Deployment issues are analysed in TR 23.919 [3GPP_TR_23.919].

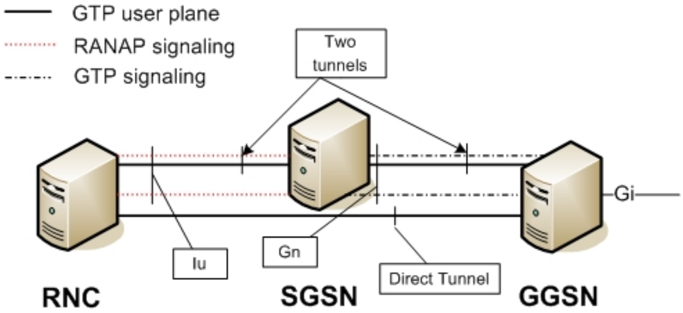


**Figure 3-2:** ***The main concept of the Direct Tunnel scheme***

In order to make the Direct Tunnel establishment and maintenance possible, the SGSN supplies the RAN with the Tunnel Endpoint Identifier (TEID) and user plane address of the GGSN, and also supplies the GGSN with the TEID and user plane address of the RAN. When the Radio Access Bearer (RAB) assigned for a particular PDP context is released but the PDP context remains still alive, the GTP-U tunnel will be (re)-established between the GGSN (in case of GGSN connectivity through Gn/Gp) and SGSN in order to be able to handle the downlink packets.

Direct Tunnel provides possibilities to offload user traffic from SGSN by bypassing it and delivering user plane traffic right between the RNC and the GGSN (Figure 3-2). However, the scheme also has some drawbacks, mostly regarding to increments in the signalling load. When informing the GGSN of the IP address of the RNC and the TEID according to the active PDP context, the SGSN sends an Update PDP Context Request to the GGSN during every PDP Context Activation for which a Direct Tunnel is created. This signalling increases the control plane load on both the SGSN and GGSN entities. In a Direct Tunnel system, release and re-establishment procedures of RAB are necessarily visible to the GGSN which also increases the volume of control plane messages between the SGSN and GGSN. In cases when the radio-link quality is low, the frequent RAB release and re-establishment will result also in increased signalling load on the GSN nodes. Also there are some Intra-SGSN procedures (e.g., intra-SGSN inter-RNC procedures), which are not visible for the GGSN in a legacy (i.e., two-tunnel) system. These procedures will become visible to the GGSN when applying Direct Tunnel and such will further increase the volume of signalling messages on the GSN entities.

If the system operates an active Direct Tunnel, the SGSN will not be able to measure data traffic volumes belonging to the PDP context for which the particular Direct Tunnel is created. It means that activating a Direct Tunnel for a PDP context will result in inconsistency in the data traffic volumes measured by the SGSN and GGSN for that particular PDP context. Therefore, if a Direct Tunnel is active, traffic volume based charging in the packet switched domain cannot be operated only based on SGSN charging information. Direct Tunnel also limits the operation of the system’s Lawful Interception (LI) capabilities: when Direct Tunnel is applied, LI can only gather reliable user plane communication information from the GGSN. However, when some LI functions in the GGSN are employed, certain control plane information (e.g., SMS and mobility management related information) can only be gathered from the SGSN entity.

Real-life deployment demands an important architectural requirement: inside the network area where Direct Tunnel is to be applied, the Iu-PS and Gn transport network segments must be assembled and configured in a way where these two transport networks are visible to each other (Figure 3-3). However this architectural necessity helps to eliminate a user plane anchor from the system, it also comes with a minor security drawback: user plane traffic will reach the Gn network directly from the RNC. This problem could be solved by introducing a firewall between Iu-PS and Gn network segments, therefore only traffic to predefined ports/IP-addresses from predefined ports/IP-addresses would be allowed to transfer between Iu-PS and Gn segments. A firewall further enhancing the network security can be deployed between the RNC and the IP network. Of course the introduction of such security appliances instead of a simple routers will likely impact user-plane latency.

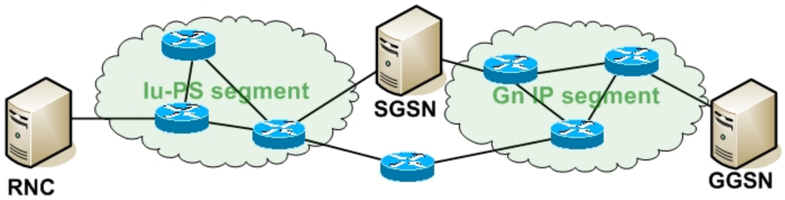


**Figure 3-3:** ***Connectivity between Iu-PS and Gn transport segments***

#### 3.4.2.2 Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO)

Traffic offloading schemes and particularly LIPA/SIPTO are analysed and started to be standardized within the 3GPP SA2 working group. LIPA is designed for residential and/or corporate environments in order to provide local network access as introduced in [3GPP_TR_23.829][3GPP_TS_22.220]. LIPA allows a mobile terminal employing IP communication via a H(e)NB to directly connect to and communicate with other IP-enabled devices and services residing in the local network by bypassing the operator's core network. SIPTO offloads selective IP traffic to the Internet at home and enterprise environments as well as at macro-cellular access networks. The main objective of these 3GPP efforts is to identify potential architectural enhancements and functionalities in order to support LIPA for the H(e)NB subsystem and SIPTO for the H(e)NB subsystem as well as for the macro cellular network.

As LIPA provides access for IP-enabled user equipment connected via a H(e)NB to other IP-enabled devices in the same residential/enterprise IP network, the offloaded user plane traffic (i.e., LIPA traffic) is expected to bypass the complete mobile operator's network except the components in the access segment located at the residential/enterprise premises (Figure 3-4). However, signaling traffic will continue to pass over the whole operator's network infrastructure. The LIPA breakout is executed at the local breakout point called the Local GW (LGW) deployed in the local network.

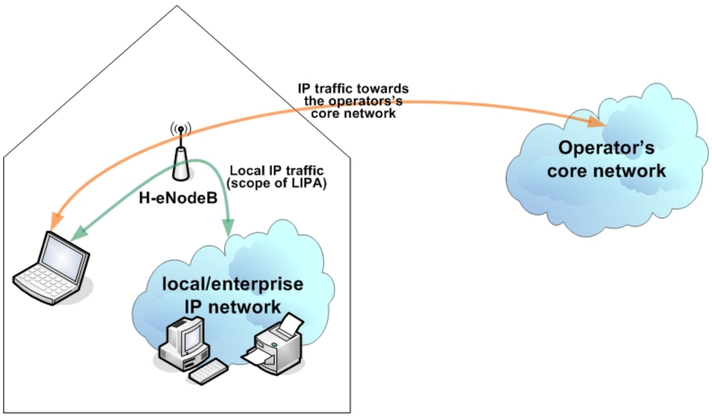


**Figure 3-4:** ***LIPA architecture***

According to the 3GPP specifications related to H(e)NB subsystems, LIPA schemes should fulfil the following main requirements:

- UEs shall be able to communicate with other network entities located within the residential/enterprise premises via the H(e)NB.
- Traffic offloaded by LIPA is expected to remain local.
- UEs shall be able to access both local IP network resources and the operator's core network simultaneously.
- LIPA is subject to subscription: according to specific roaming agreements, UEs may be able to use LIPA offloading services also in a visited network.
- Service continuity for UEs moving between different H(e)NBs within the same residential/enterprise network is advisable. Note that this functionality has not been standardized yet: it is handled by 3GPP TR 23.859 (LIPA Mobility and SIPTO at the Local Network) within the work item LIMONET.

SIPTO is a traffic offloading solution designed to offload Internet destined traffic from the 3G mobile network as early as possible [3GPP_TR_23.829][ 3GPP_TS_22.220]. The operation of SIPTO is two-fold, two main types of breakout architectures can be distinguished. On one hand architectures with breakout point "at or above RAN" can be created. This scenario covers macro and some H(e)NodeB SIPTO use cases. On the other hand, SIPTO also allows the breakout point to be located either in the residential/enterprise network (similarly to LIPA), or "above" H(e)NodeB (e.g.,. in the backhaul segment or in the H(e)NodeB-GW entity). This scenario mainly covers LIPA and some H(e)NodeB SIPTO use cases.

In the most common scenario the traffic offloading takes place on the data path between the UTRAN/LTE and the SGSN/SGW (Figure 3-5). The scheme allows the traffic offloading functions to transparently monitor the Radio Access Network Application Part (RANAP) control traffic and based on this information it becomes possible to detect and drive Internet destined communication sessions. Detection and intervention can be performed based e.g., on the used APN or on the IMEI, as several

Internet-only devices, as HSPA enabled laptops are identifiable in this way. E.g. 3GPP TR 23.829 proposes a new SIPTO_enabled flag associated with each APN in the user's subscription indicating whether the connection to that APN is enabled/disabled for SIPTO. The same scheme is followed in case of LIPA.

When the scheme detects sessions bound to inside operator services, the traffic offloading will not take any action and user data traffic will travel the standard path, without the shortcut of the traffic offloading function. In case of sessions bound to the Internet, SIPTO functions alter the traffic path in order to bypass both the SGSN (SGW) and GGSN (PGW) and forwarding it directly to the nearest Internet peering point. The main control plane functions (e.g., location updates to HLR) remain in the SGSN/GGSN or SGW/PGW entities and will not be affected by SIPTO. Applying fine grained traffic selection policy schemes, SIPTO can identify and direct specific Internet destined user traffic to the operator's CDN or other locally deployed service in case of need.

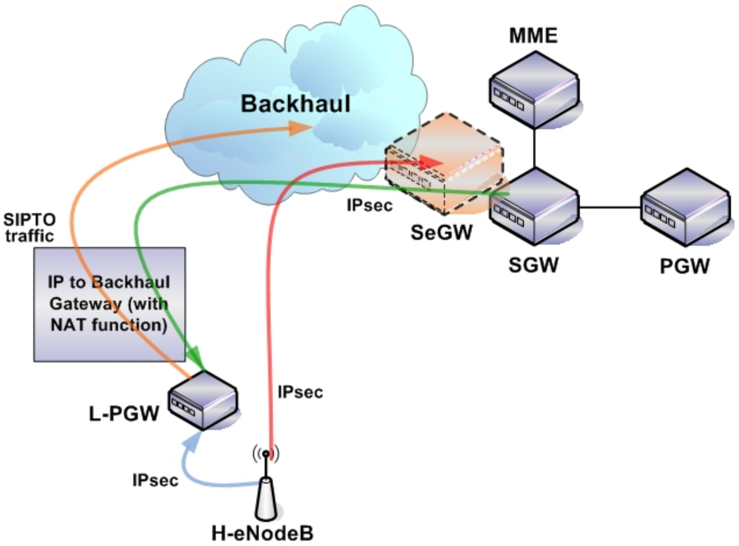


**Figure 3-5:** ***A potential SIPTO deployment scenario***

Similarly to LIPA, SIPTO schemes also require a local the breakout point. The entity is called LGW (Local Gateway), which should either be collocated with the H(e)NB or form a standalone element connected directly to the H(e)NB subsystem. SIPTO offloading for macro cellular network takes place at or above the RAN. By performing offloading for selected IP traffic closer to the users (i.e., the breakout point is located close to the edge of the network), operators could significantly reduce the load on their network resources such PDN and/or Serving gateways. Also suboptimal routing in the mobile backhaul segment may be eliminated.

The connection with the operator's core network is achieved through the IP backhaul by spanning a secure tunnel between the H(e)NB and a Security Gateway (SeGW) in the core (Figure 3-5). IP traffic from the local network is tunnelled through the Home Router and the IP backhaul towards the SeGW. By using this solution security can be ensured, which is crucial considering that the IP backhaul could be owned by a different operator. It is also important to note that the SeGW entity could also address the need of data aggregation from different heterogeneous sources.

In LIPA when accessing a local network resource, a user is preferred to have control on selecting the available LIPA service. However, in case of SIPTO the traffic offloading service should be completely transparent to the users. It could be problematic as handling separate dedicated PDN connections will be needed for this, but such a function is only applicable to UEs supporting such a feature, and this surely will not be the case for several legacy end-user terminals. Therefore, in case of common UEs with a single PDN connection, the implementation of a Network Address Translation function at H(e)NBs, HNB-GWs and at eNBs will be required in order to translate the address of the UE into an IP address exclusively assigned SIPTO operation. This feature will expect packet inspection in the affected entities, introducing higher processing load.

Considering the above, the following two main different LIPA/SIPTO architectures are to be defined:

1. LIPA/SIPTO based on a dedicated offload PDN connection, in which separate PDN connections are used for the offloaded and the non-offloaded traffic.
2. LIPA/SIPTO based on a specific traffic offload function and NAT, where a single PDN connection is used for both offloaded and the non-offloaded traffic, and where the breakout point including the NAT function is located within the H(e)NB.

Actual standardization efforts regarding SIPTO services mainly focus on locating the functionality on the operator's network and on providing methods for distinguishing the SIPTO related traffic. Lawful Interception (LI) functionalities are actual research topics and under discussion at 3GPP SA1 working group and at the Femto Forum. In case of LIPA the H(e)NB subsystem is a promising candidate to perform LI. However, the case of SIPTO is not such an easy question. The main problem in SIPTO is that user traffic may need to travel inside the operator's core network for getting intercepted, which is totally against the main objective of SIPTO. Researches also consider mobility management of SIPTO sessions. For example authors in [Taleb_11] enhance the existing methods by proposing a DNS-based scheme for dealing with SIPTO traffic and simultaneously consider service continuity for SIPTO traffic during mobility events. The service continuity of this scheme is provided by enforcing both downlink and uplink traffic to traverse the LGW at the anchoring (H)eNB.

#### 3.4.2.3 IP Flow Mobility (IFOM)

IP Flow Mobility (IFOM) is a special, 3GPP standardized [3GPP_TR_23.861], [3GPP_TS_23.261] extension of IP mobility, a well known IETF communication protocol family containing MIP, PMIP, HMIP, DSMIP, NEMO, etc. It allows moving selected ongoing communication flows from one network to another, without any interruptions of the modified flows while keeping the other flows on their actual network.

3GPP TS 23.261 [3GPP_TS_23.261] defines the above goals and the regarding mechanisms for mobile cellular architectures: based on this specification UEs are able to simultaneously connect to 3GPP and WLAN accesses and are allowed to exchange different IP flows belonging to the same PDN connection through different access networks. The mechanisms of [3GPP_TS_23.261] enable seamless IP flow mobility between 3GPP and WLAN accesses with IP flows lying in the same or different applications. The solution introduced in [3GPP_TS_23.261] is based on DSMIPv6 [RFC5555], and using the fine grained (i.e., flow based) mobility management of IP sessions it makes applicable for network operators to modify how the flows are travelling through the available access network. The scheme is available for both the Evolved Packet System and the I-WLAN architecture, and thanks to the DSMIPv6 origins, IP address preservation and session continuity is achievable during movements of IP flows between access systems.

The standard level of granularity in case of access system connectivity and inter system mobility is defined in TS 23.402 [3GPP_TS_23.402] and TS 23.327. These technical specifications introduce a per PDN connection basis for user equipments, meaning that in case of handovers all the IP flows of a particular PDN connection will be moved between the source and target access systems. However, IFOM functions make a much finer granularity available in access system connectivity and inter system mobility: handover procedures are not bond to PDN connection level, they can be used to single or multiple IP flows belonging to a particular PDN connection. This also means that some IP flows of one PDN connection with pre-defined parameters and characteristics can be routed through one access network while other IP flows of the same PDN connection can be routed through another access network. Therefore, the scheme makes possible to selectively offload some traffic (such as best effort traffic) to e.g., the WLAN segment while using UTRAN or E-UTRAN access systems for other traffic (such as traffic with higher QoS requirements).

To provide this fine grained mobility the inter-system mobility signalling is improved aiming to transport special routing filters inside the operator's network. The routing filer transporting extensions to the mobility signalling framework of DSMIPv6 for multihoming situations are specified in RFC 6089 [RFC6089] and are applicable to both H1 and S2c interfaces.

Aiming to allow the mobile operator to indicate to the UE through which access network IP flows are assumed to be routed, inter system routing policies are introduced in TS 23.402 [3GPP_TS_23.402]. These policies can be created and used "per APN, per IP flow class under any APN or per IP flow class under a specific APN" and the UE can be informed about such policies either through ANDSF or based on static pre-configuration. In case of IP flows which are routed via WLAN, the inter system routing policies also define if the traffic supposed to be routed via the Home Agent or directly through the WLAN access. This means that in order to apply routing policies for IFOM, the inter-system routing policy must include one or more filter rules. These filter rules are special formulas identifying a prioritised list of potentially available access systems to which the UE should be connected to when available in order to route traffic that matches specific IP filters on a particular APN or on any APN. Such a filter rule also identifies which access networks are not usable or limited for the traffic matching specific IP filters (e.g., WLAN is prohibited for RTP/RTCP traffic flows on APN-X).

In the mobility procedures of [3GPP_TS_23.261] the UE is assumed to be simultaneously connected to a WLAN and a 3GPP access and it is supposed that the UE uses both the access systems for the same PDN connection. According to this, the UE is the entity that initiates flow mobility operations using DSMIPv6 messages (i.e., the UE adds/modifies/deletes/moves flows between accesses). However, there are use-cases, where Home Agent initiated flow operations could also be beneficial.

#### *3.4.2.3.1A special IFOM application: Home Agent Initiated Flow Bindings*

There are cases in multihoming Mobile IPv6 environments when flow mobility (or flow binding) is initiated by a central entity, such as the always available Home Agent. Operations like network-controlled flow binding revoking, moving, or provisioning are equally possible with this mechanism; making it possible to revoke an existing flow binding in case of an error, or move a flow from one interface to another on the MN side, or simply provide default flow settings for newly connected Mobile Nodes. The approach is not mutually exclusive with the MN initiated flow binding described in RFC 6089 [RFC6089], it merely extends the mobility features it provides, meaning that flow bindings are not always initiated by the HA. There are drafts (e.g., [Yokota_11]) which introduce a new Mobility Header and signalling messages based on the flow binding protocol implemented in RFC 6089.

Possible application use cases of HA Initiated Flow Bindings are Default Flow Binding Provisioning, Traffic Offloading and Flow Binding Revocation. Default Flow Binding Provisioning is used for example in an environment where a central entity wants to force Service Level Agreements (SLA) to a customer, e.g., forcing P2P traffic through Wi-Fi while allowing 3GPP access for HTTP traffic. The Traffic Offloading technique makes it possible to move certain data flows from one interface to another, e.g., in case of increasing traffic load in 3G segment move video streams to the Wi-Fi segment. Policies can be much complex based on the fact that the core network entities know about their actual traffic conditions. Flow Binding Revocation is useful when due to an administrative decision; a certain flow binding is no longer valid for the MN.

The main protocol operation is based on RFC 6089: the multiple care-of address registration extension to the Mobile IPv6 protocol makes it possible to use multiple egress interfaces and operate policy based routing using these interfaces by the flow binding mechanism specified in RFC 6089. The document describes that in order to initiate a flow binding operation a valid Mobile IPv6 binding is required. Similarly to that technique a HA initiated flow binding operates via the Flow Binding Indication (FBI) and the Flow Binding Acknowledgement (FBA) messages, where the latter is used for the acknowledgement of the FBI message.

The application and real-life deployment of the HA Initiated Flow Bindings scheme is till a hot research topic. It is still an open question how to map between flows and networks based on cost, QoS, security, user preference, etc. and how to deploy the scheme in a fully/partially distributed networking architecture. Decision algorithms and optimization possibilities are also open topics of actual researches as well as practical evaluations, applicability studies and extensions for other mobility protocols than MIPv6. For example authors of [Tran_Minh_Trung_11] create a network-based Flow Mobility based on Proxy Mobile IPv6 which requires minimum modification at the mobile node and all of its signalling processes are performed on the network side. Evaluation of this scheme proves that network-based flow mobility is a promising technology helping network operators to extend their network capacity at low cost by exploiting benefits of various heterogeneous access networks. In [Wang_07] a hybrid approach is proposed to manage both kinds of flow handovers (i.e., user and network initiated) in a flexible and standardized way by enhancing Mobile IPv6 (MIPv6) or its variants. Authors investigate address management strategies for mobile host multihoming, and present the correspondent architectural and protocol choices, together with a comprehensive handoff management framework designed for the scheme.

### 3.4.2.4 Network-based flow mobility for high granularity traffic management

By relying on the basic concepts introduced by the soon to be RFCd HA initiated flow binding proposal, it is also possible to extend the functionality of the model by defining monitoring points for traffic state and analysis overall the network and adding policy servers to manage the monitoring points and enforce policies based on the processed data. Our proposed scheme tries to implement this theory into practice (Figure 3-6).

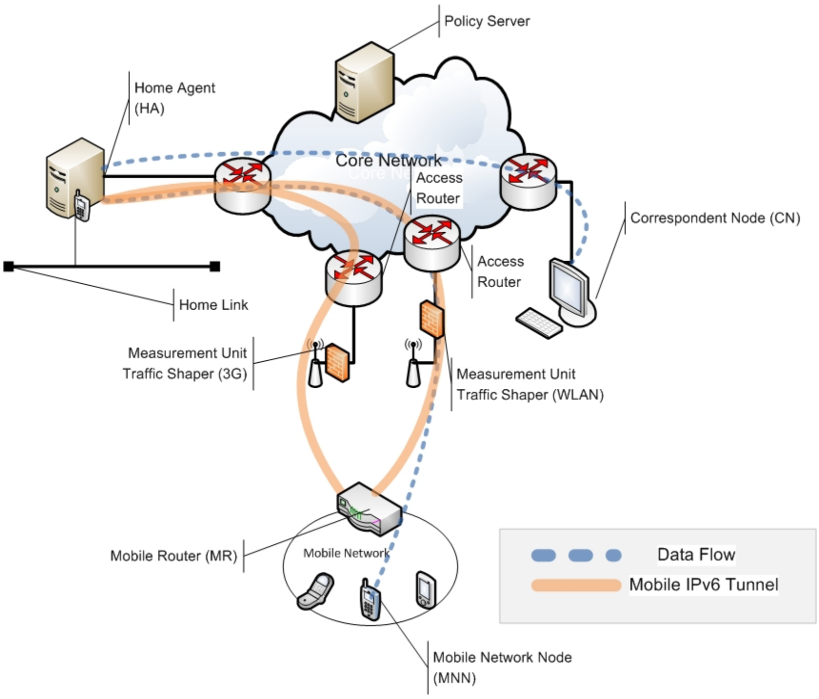


**Figure 3-6:** ***NB-IFOM for high granularity TM – Main concept and architecture***

In order to perform network based and traffic management oriented flow-mobility operations the existing Mobile IPv6 (NEMO/MCoA) architecture must be extended with the following special nodes:

- Measurement Unit: One or more DPI capable (Deep Packet Inspection) devices throughout the core network. They passively monitor the overall and flow based network usage statistics for a given link. When DPI is not available, i.e., when the tunnelled traffic is encrypted, it reports only aggregated statistics on a given link.
- Mobile Node/Router (MN/MR): Mobile IPv6 node with extended functionalities. Performs policy routing and flow binding based on network events. Such policies received from the Mobile IPv6 network management entity (Home Agent) always overrule the local decisions and predefined settings.
- Home Agent (HA): Mobile IPv6 central management entity with extended functionality. Relays and enforces network-based policies received from the Policy Server. Synchronizes its Binding Cache to the Policy Server.
- Policy Server (PS): A single central entity which performs flow binding based on overall network parameters. It receives link and flow usage information from multiple Measurement Units and maintains an aggregated Binding Cache from multiple Home Agents. Knowing the actual flow binding usage on the network it activates policies when trigger conditions are met.

#### *3.4.2.4.1 Policy Server communication*

The following protocol messages are sent from/to the Policy Server (Figure 3-7).

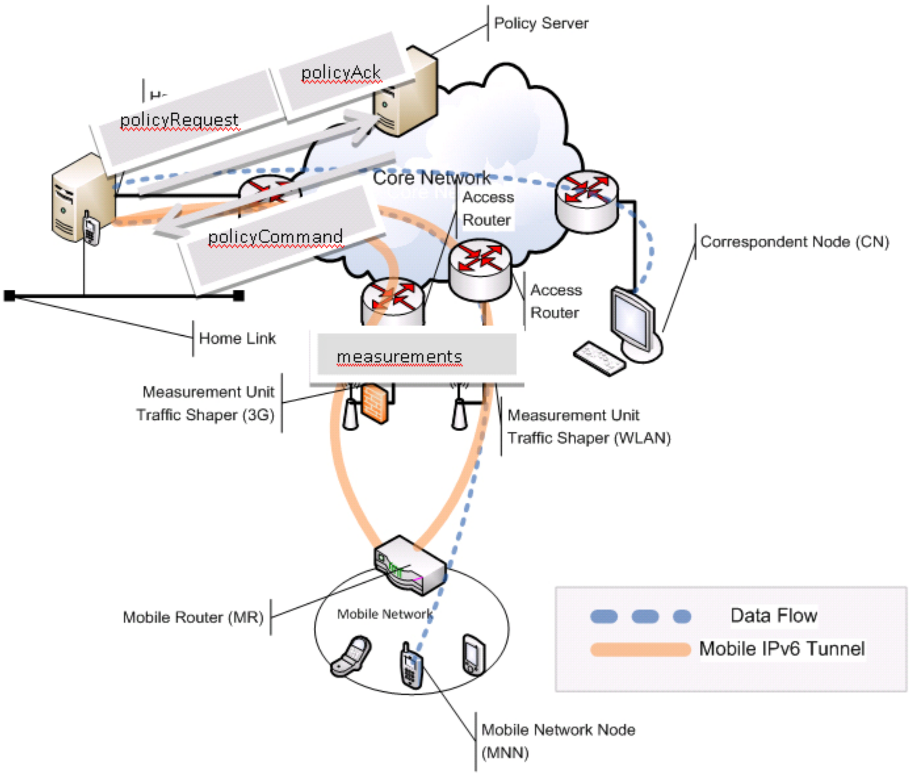


**Figure 3-7:** ***Communication with the Policy Server***

Measurements

Measurement nodes are periodically reporting towards the Policy Server. The messages are not acknowledged; therefore policy management operation must not depend on continuous reports.

The protocol may report overall statistics along with usage statistics of several flows. Flows are predefined and correspond to the flows defined on the Policy Server.

| | |
|---|---|
| Protocol transport | UDP |
| Protocol format | XML or JSON |
| Protocol fields | entity (flow id, link), bandwidth, packet loss, latency |

Policy Request

The message is sent from the Home Agent to the Policy Server. HA requests new policies from the PS for a given Binding entry (HoA, BID, CoA triplet) when the Binding Cache Entry is changed on the HA.

| | |
|---|---|
| Protocol transport | UDP |
| Protocol format | XML or JSON |
| Protocol fields | HoA, BID, CoA, optional flags |

Policy Command

The Policy Server may send a policyCommand message to the Home Agent, indicating the applicable flow policies on the Home Agent.

| | |
|---|---|
| Protocol transport | UDP |
| Protocol format | XML or JSON |
| Protocol fields | Task ID, Flow descriptor, BID, HoA |

Policy Acknowledgement

The Home Agent acknowledges the policyCommand with a policyAcknowledgement message.

| | |
|---|---|
| Protocol transport | UDP |
| Protocol format | XML or JSON |
| Protocol fields | Task ID, status |

*3.4.2.4.2 Home Agent communication*

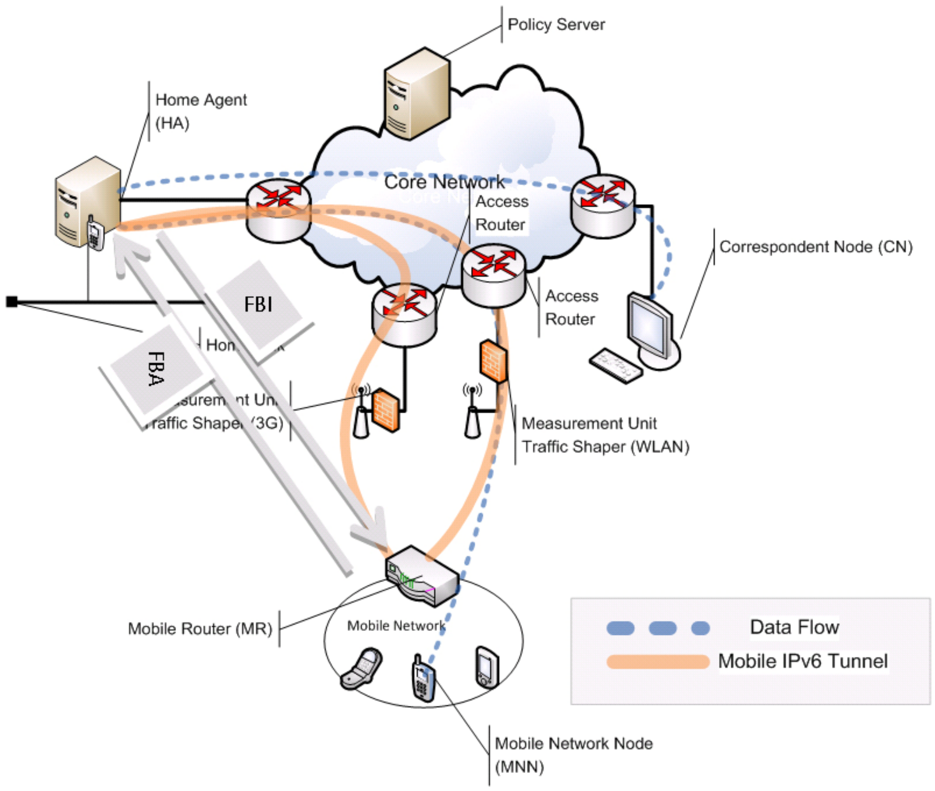


**Figure 3-8:** ***FBA-FBI exchange***

Figure 3-8 introduces the main information exchange between the Home Agent and the Mobile Node. The used messages are presented below.

Flow Binding Indication

HA indicates towards a MN if a status of the policy is changed (updated, revoked) or a new policy should be installed. According to [Yokota_11] the protocol introduces a new Mobility Header type.

| | |
|---|---|
| Protocol transport | ICMPv6 |
| Protocol format | Mobility Header |
| Protocol fields | TBD |

Flow Binding Acknowledgement

MN sets the policy in its routing table and send back and acknowledgement message. According to [Yokota_11] the protocol introduces a new Mobility Header type.

| | |
|---|---|
| Protocol transport | ICMPv6 |
| Protocol format | Mobility Header |
| Protocol fields | TBD |

*3.4.2.4.3 Protocol usage*

The newly introduced protocol messages extend the Mobile IPv6 Binding Update/Acknowledgement process the below way (Figure 3-9). Some of the messages are not dependent on each other, meaning that Policy Server may not send policies for a Binding Cache Entry change. The following messages must be acknowledged/answered:

- Flow Binding Indication (Flow Binding Acknowledgement)
- policyCommand (policyAcknowledgement)
- Binding Update (Binding Acknowledgement)

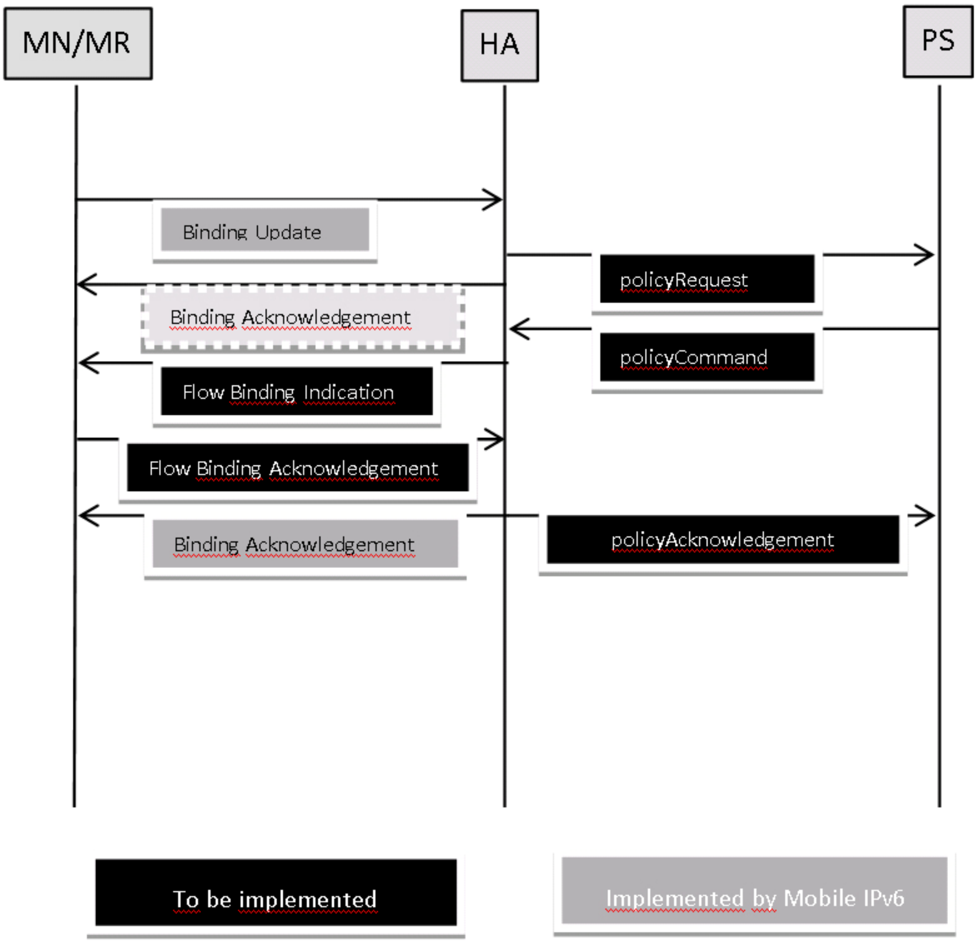


**Figure 3-9*: Protocol usage***

Implementation considerations have been addressed in Appendix A.

### 3.4.3 QoE-aware Traffic Engineered Handovers

The high growth of QoE demanding applications in a context of scarce resources stresses the need to optimize the EPC path w.r.t. QoS & QoE. One way to achieve this is the RAT change in order to balance the traffic or get a better QoS with the hope to improve the QoE. Where as QoS is achievable at the EPS level, there is a need to anticipate the QoE impact of an EPC path change at the e2e level. The key tools to drive traffic switching between 3GPP and non 3GPP technology are ANDSF and MIH; however, their view on a connection restricted to EPS. A transport topology aware protocol to assist the selection of overlay application CN, such as ALTO has an e2e view on a connection, but it has no insight at the EPS level. The solutions exposed in this section address the need to harmonize both perspectives and scopes without decision conflicts.

This section presents decision mechanisms to be involved in RAT changes. They are referred to as Traffic Engineered Handovers (TEHO) in that their goal is not restricted to cope with degradation of signal conditions but covers the improvement of traffic conditions in the IP Connectivity Wireless Access Network (IP-CWAN) together with the improvement and maintenance of the user QoE.

The MEVICO deliverable D4.2.1, on the TE building blocks, in section 3.1.4.2, stresses the need for a TE solution that makes the HO procedure more QoE and network state aware. Indeed nowadays traffic in wireless network is geared towards bursty, QoE demanding and resources greedy applications, and signal quality-based HO decisions, initially tailored for voice traffic are no more sufficient to ensure the satisfaction at both the network and the user level.

A fundamental ingredient to QoE in wireless networks is the QoS achieved in the IP-CWAN. Solutions to improve the IP-CWAN QoS include offloading techniques. The decisions functions driving the offload have a visibility and impact limited to the QoS in the IP-CWAN and it is necessary that the QoE on the impacted flows is also considered.

### 3.4.3.1 Best prior art work on QoE aware TEHO features

The most significant prior art related to the above mentioned necessity has been exposed in IR431, § 3.4.3.1. To improve IP-CAN QoS:

- IFOM: is described in §3.4.2.3 “IP Flow Mobility (IFOM)” of this document. It refers to the capability of using the same APN across two wireless access networks (e.g. LTE and WIFI) and enables seamless roaming of applications across LTE and WIFI technologies.
- Multi-Access PDN Connectivity (MAPCON): is defined in the 3GPP specification, Release-10 and designed for terminals having multiple interfaces. It refers to the capability of simultaneously using two or more Access Point Names (APN) and enables use cases such as using LTE for QoS demanding applications and WIFI for best effort traffic.
- ANDSF: the Access Network Discovery and Selection Function (ANDSF) is a set of policy rules used to drive IP traffic over a non-3GPP technology. The ANDSF transfers to the UE the mobile operator policy to connect through non 3GPP access technologies such as WiFi and WiMAX and enables thus a traffic steering that adapts to the QoS and traffic of the controlled LTE network.
- MIH: another such tool is Media Independent Handover Client-Server protocol that provides information and Handover assistance services to 3GPP access technologies such as WiFi and WiMAX. the three services offered by MIH: the “event notifications” e.g. on links, the “commands” assisting the HO and the “information service” passing information between layers 2 and 3, such as the ISP Name from the IP layer or the rate or delay from the wireless layer.
- ALTO: to improve QoE for applications, the UE may use the IETF Application layer Traffic Optimization (ALTO) protocol, whose design goal is to optimize both the user QoE and the usage of network resources by providing to the UE information helping it to choose the best possible location from which to download the whole or piece of video.

While the ANDSF enables traffic offload and optimization of the network resources, its visibility and decision scope however is limited to the LTE network: it cannot see the end to end path and thus take into account the QoE perceived at the UE, which is more and more challenged by massive use of resources and performances greedy applications. On the other hand, it is not the responsibility of the ALTO protocol to care about the UE mobility. However, the mobility of a UE can impact its path to the PDN and thus the path to the content and thus the related QoE. Therefore, it is necessary to inform the UE, which could take the appropriate decisions, concerning the changes occurred in its path,. Currently, during mobility there is no association between network level information and application level information when a handover occurs.

The proposed solutions address the need to harmonize both perspectives and scopes without decision conflicts.

### 3.4.3.2 Basic idea of the proposed solutions

A set of three solutions is proposed to jointly improve QoS and QoE in HO decisions. Their principle is to proactively or reactively associate QoE continuity to the HO decision. Two solutions introduce the usage of selection mechanisms of correspondent nodes in overlay applications that are triggered after of during an inter-RAT HO. A third solution proposes the possibility of notifying a UE of a change in the S5/S8 bearer upon intra E-UTRAN HO to ensure QoE by checking the e2e path to correspondent nodes upon notification.

### 3.4.3.3 Solution 1 - Adaptive: Inter RAT HO based connection management

Solution1 is applicable to an IFOM-capable UE with one IP address with a scenario illustrated in Figure 3-10.
IFOM causes a change from SGW1 to SGW2 and subsequently a change of the associated path cost from UE to the EPs. Once a HO is performed the CM requests ALTO Client to update its path costs to the current CNs/EPs.

Suppose cost of e2e path from UE to EP defined as MAX[P(EP, PGW), P(PGW, UE)], and is to be minimized. In this example, path cost from UE to PGW evolves from 7.5 to 5. With SGW2, the least cost EP becomes EP2 with C=5, where as the cost with EP1 equals 7.5, so EP2 is preferable. Note that cost of type C = MAX_i(Ci) are frequent when the worst value must be taken over the Ci, for instance to evaluate the cost in terms of bandwidth availability.

On the IP route between UE and PGW lies the Serving Gateway (SGW) to which the UE is attached. Suppose that an IFOM causes a change of SGW: the path between the SEP and the UE is thus changed, in its last hop, that is between the PGW and the SGW.

Although, after a IFOM, the list of candidate EPs remains the same, the associated downloading and routing cost may have changed and needs to be updated. A possible consequence is that the currently used EP to download from is no more optimal and needs to be changed. The UE is notified by the CM of the change in the EPC path in order to re-evaluate the cost of relevant EPs.

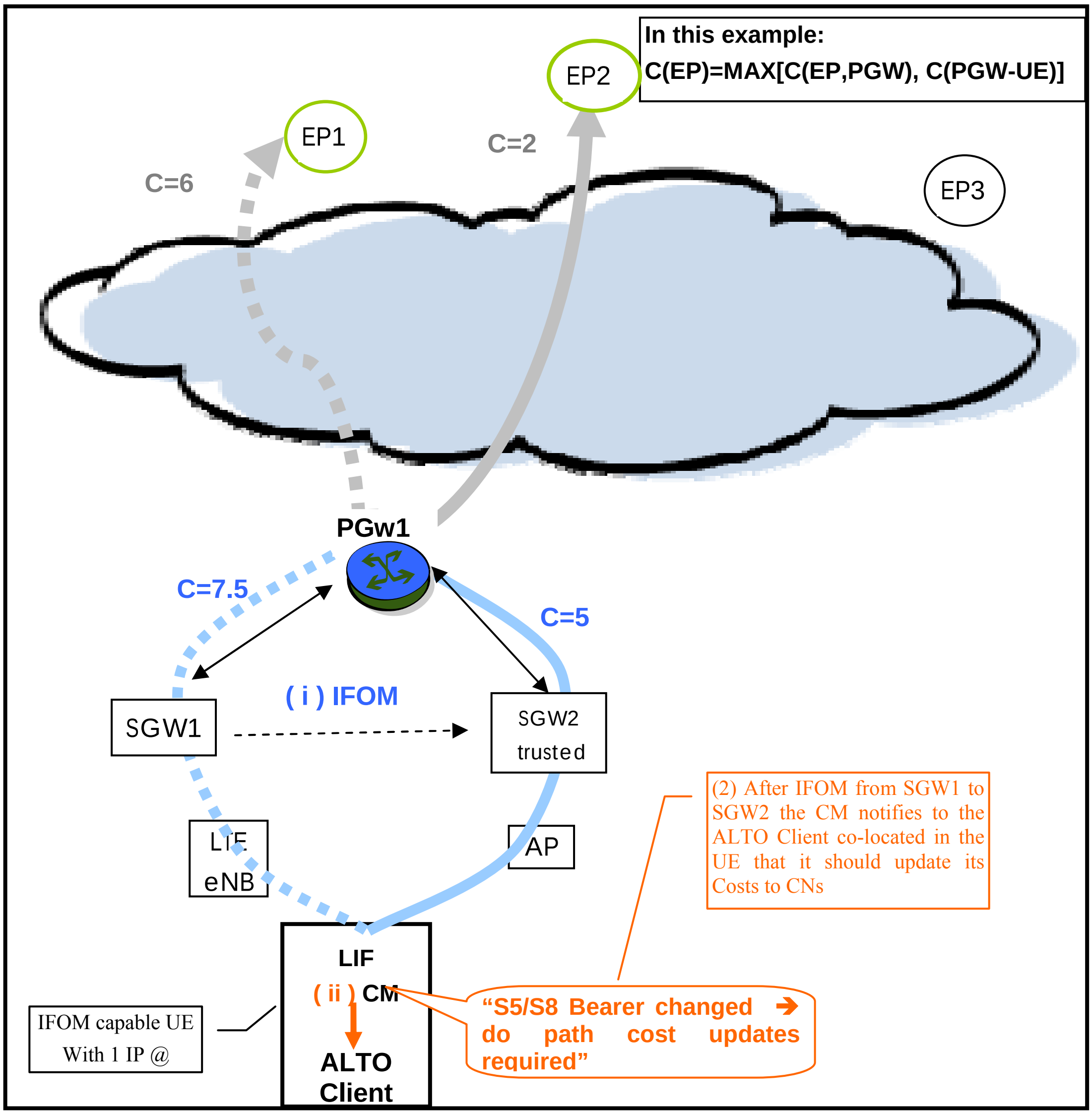

In this example:
C(EP)=MAX[C(EP,PGW), C(PGW-UE)]
EP2
EP1
EP3
C=6
C=2
PGw1
C=7.5
C=5
( i ) IFOM
SGW1
SGW2
trusted
LTE
eNB
AP
(2) After IFOM from SGW1 to SGW2 the CM notifies to the ALTO Client co-located in the UE that it should update its Costs to CNs
LIF
( ii ) CM
ALTO
Client
IFOM capable UE
With 1 IP @
"S5/S8 Bearer changed ➔ do path cost updates required"

**Figure 3-10*: Adaptive ALTO-COMEPS – example operation***

Figure 3-11 depicts the functional blocks required in the UE to get this solution operational. The Connection Manager is implemented according to the following specifications (http://tools.ietf.org/id/draft-seite-mif-connection-manager-02.txt) and it exchanges messages with the ANDSF (or MIH) client always installed in the UE. The CM is also aware of the IP flows sent or received by the UE. When a change in the IP flow routing policies is detected, the CM triggers the ALTO client who can subsequently request an update of the costs to the eligible SEPs. Upon reception of the new cost values, the ALTO client could decide to change its corresponding SEP.

It should be noted that the UE implements the specifications described in http://tools.ietf.org/id/draft-ietf-netext-logical-interface-support-02.txt.

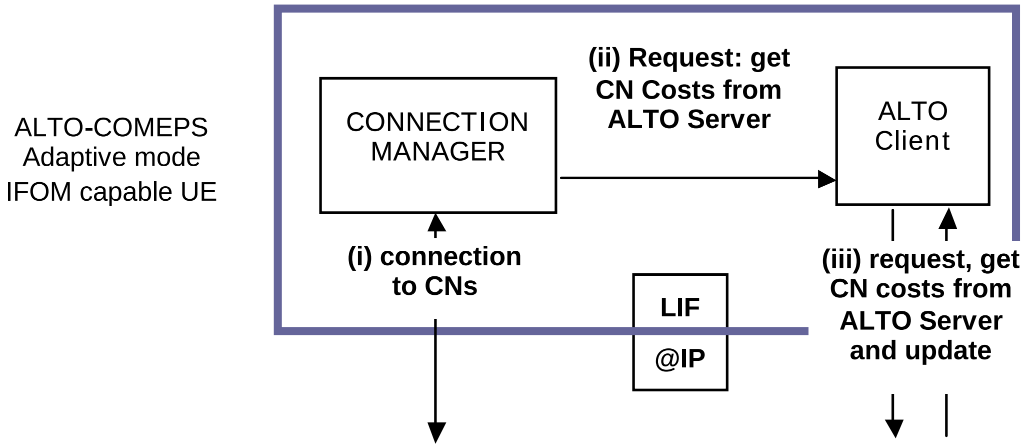

ALTO-COMEPS
Adaptive mode
IFOM capable UE
CONNECTION MANAGER
(ii) Request: get CN Costs from ALTO Server
ALTO Client
(i) connection to CNs
LIF
@IP
(iii) request, get CN costs from ALTO Server and update

**Figure 3-11*: ALTO-COMEPS – architecture on an IFOM capable UE***

### 3.4.3.4 Solution2 – pro-active: Inter RAT HO based connection management

Solution2 is applicable to a MAPCON-capable UE with 2 IP addresses. with a scenario, illustrated in Figure 3-12:

- UE can be connected to both APN1 and APN2. The CM in the UE holds a set of eligible CNs accessible via APN1(PGW1) and APN2(PGW2).
- The traffic is distributed among both APNs based on ANDSF/MIH.
- Before selecting a CN, CM requests the ALTO Client to get the path Costs to the eligible CNs (EPs), from ALTO server.
- The CM selects the CNs upon the information provided by the ALTO Client and sets the appropriate connections

When a new application is started in the UE, an IP address and an interface need to be selected to send and receive data. The selection process can follow standard RFC 3484 procedures, or it can be enhanced with the costs metrics provided via the ALTO client. Figure 3-12 describes the procedures necessary to perform the selection process. First, the CM triggers the ALTO client. Upon reception of the costs to the eligible SEPs from the ALTO client, the CM can select the corresponding SEP *across the APNs*. It then informs the application of the IP address to be used to connect to the selected corresponding SEP. In this sense, the IP address selection procedure becomes both network-aware thanks to the ANDSF or MIH guidance and QoE-aware thanks to the ALTO guidance that provides optimal SEP selection w.r.t. the end to end path costs.

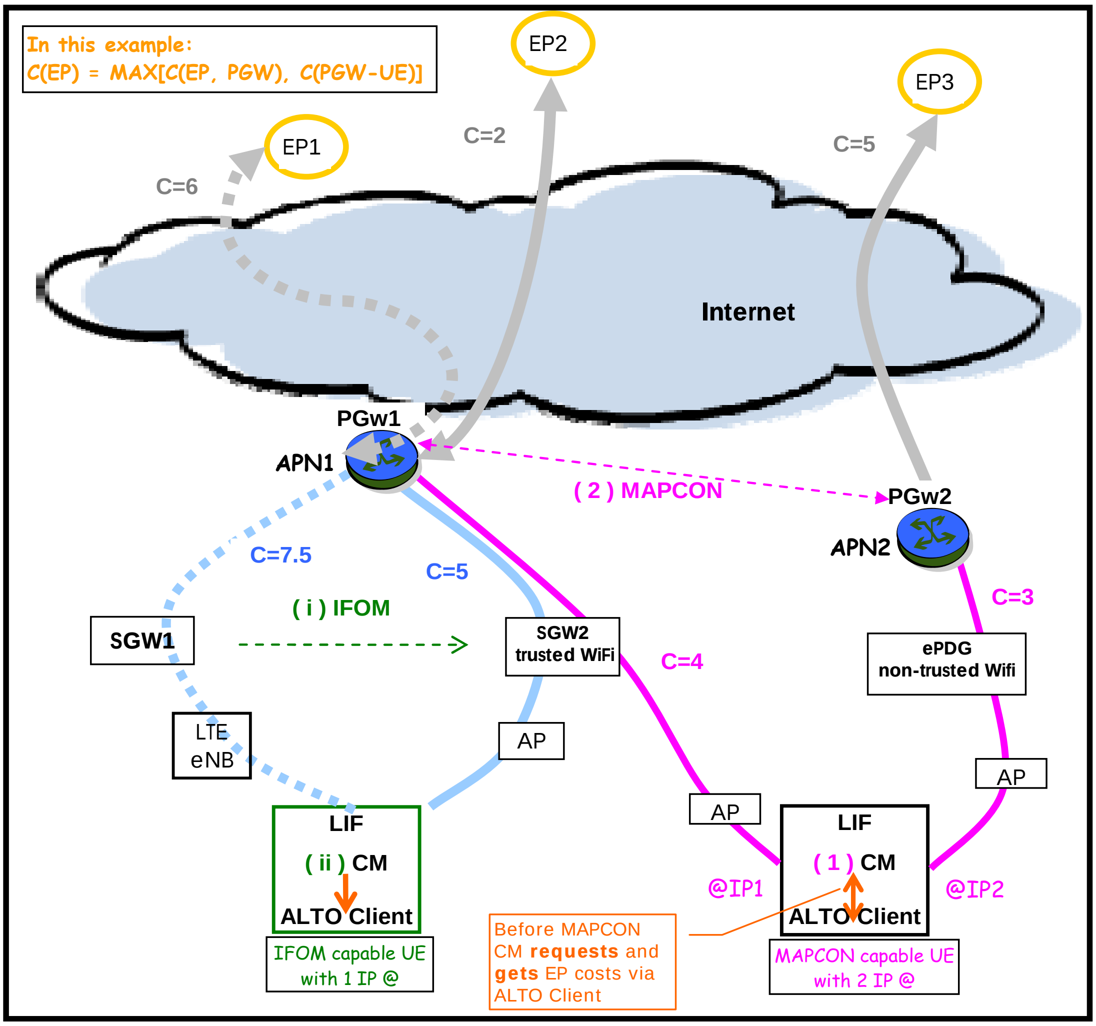


**Figure 3-12:** ***Proactive ALTO-COMEPS – with a MAPCON-capable UE with 2 IP addresses***

### 3.4.3.5 Solution3: S5/S8 bearer change notification to the UE upon E-UTRAN HO

Solution3 is applicable to any LTE capable terminal and does not require inter RAT HO capabilities. Its purpose is the enhancement of E-UTRAN HO in order to maintain QoE on flows between the UE and connection nodes previously selected with the ALTO protocol.

The idea is to: inform the UE of a traffic path change in the EPS due to HO, that may require an action at the application layer level. A particular action to take is that: if the UE is performing an overlay application session with Endpoints selected by ALTO, then having the information of a change in the EPS path, the UE must update its path costs to the Endpoints and possibly reconnect to other Endpoints that have a better cost. What action needs to be taken by the UE is out of scope of this mechanism and the usage of ALTO is one example use cases.

Solution3 performs the following: the MME sends to the UE with the help of Non Access Stratum (NAS) protocol, a notification of an S5/S8 bearer change. Such a change covers the change of SGW with subsequent cost changes in the access and backhaul paths. Changes only in the radio bearer are not considered for the moment for the sake of time scalability.

An example of intra E-UTRAN HO from “source eNB” to “destination eNB” that causes a change is illustrated in Figure 3-13:

- A HO causes a from SGW1 to SGW2, and subsequently a change of the path cost from user equipment (UE) to candidate endpoints (EPs) in the packet data network.
- Suppose the cost of e2e path (UE ➔ EP) is calculated as MAX[P(EP, PGW), P(PGW, UE)], and is to be minimized.
- In this example, the path cost from UE to PGW evolves from 7.5 to 5.
- With SGW2, the least cost EP becomes EP2 with C=5, where as the cost to EP1 equals 6 therefore EP2 is preferable.

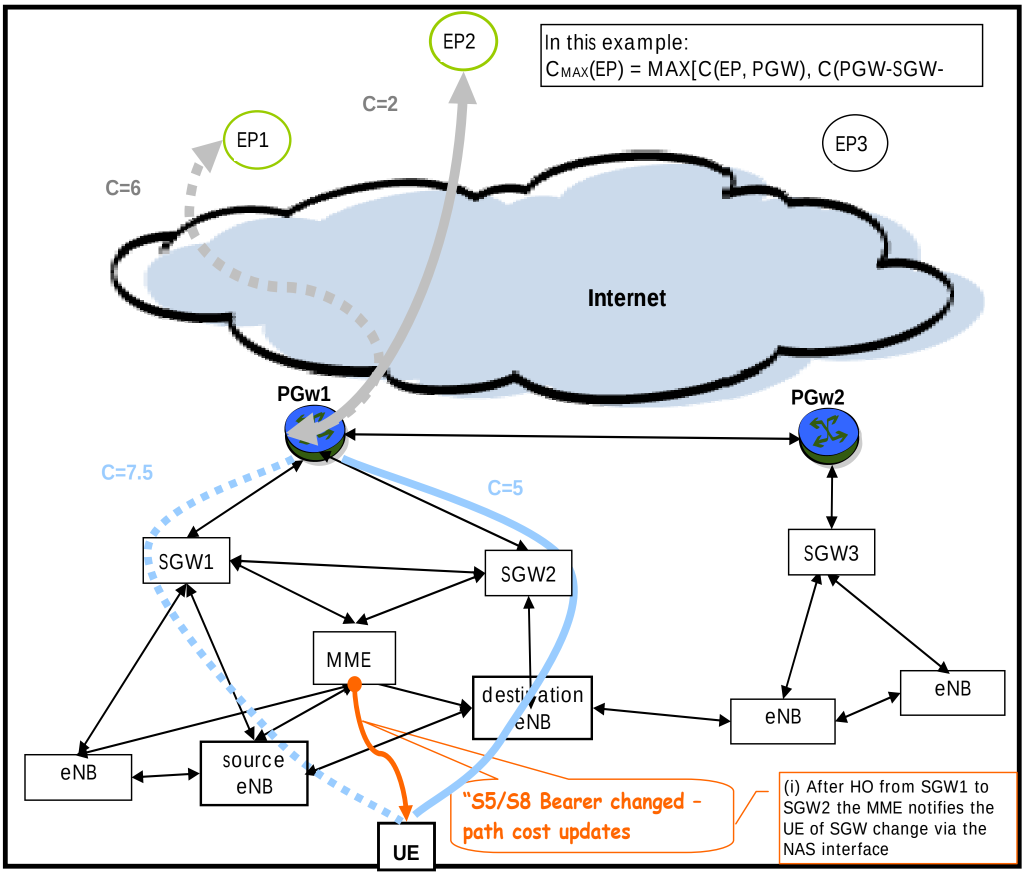

EP2
In this example:
CMAX(EP) = MAX[C(EP, PGW), C(PGW-SGW-
C=2
EP1
EP3
C=6
Internet
PGw1
PGw2
C=7.5
C=5
SGW1
SGW2
SGW3
MME
destination eNB
eNB
eNB
eNB
source eNB
"S5/S8 Bearer changed – path cost updates
(i) After HO from SGW1 to SGW2 the MME notifies the UE of SGW change via the NAS interface
UE

**Figure 3-13:** ***Example of a HO from source eNB to destination eNB***

The HO protocol currently specified at the 3GPP includes the following particular steps:

1. User equipment UE sends a measurement report to source eNB, according to the eNB specifications,
2. if decided by the source eNB, the HO procedure is done. The EPC is not involved in the decision process,
3. The target eNB notifies the MME that the UE has changed eNB,
4. The MME requires the SGW to change its data path to the UE accordingly,
5. if the SGW has also changed, the data path update must be notified until the PGW level.
6. the MME then gets a User Plane Update Response from the new SGW.

In the standard E-UTRAN HO protocol, the UE is never informed when a change of SGW occurs.

The proposed solution introduces an additional message sent by the mobile core network to the user equipment (UE) after step 6, and notifying the UE that a change has occurred in the data path that requires an action from the UE.

The HO impacts the cost associated to the Content Locations from which it downloads/receives content and require the UE to make a Content Location Cost Update (CLCU).

After Step 6, the MME is aware of the change of SGW because it got a PATH SWITCH message from the new SGW via the S11 interface.

1. The MME upon this information notifies the UE of a path change that requests action. Via the 3GPP defined Non Access Stratum EPS Mobility Management protocol (NAS-EMM).
2. It uses an information element called Notification Indicator (NI) and specified by the 3GPP in TS 24.301 §8.3.18A.

Thus, the NI carries a message from the MME to the UE that signifies: "S5/S8 bearer has changed". As an option, the proposed Notification Indicator may only be sent if the UE needs it. The UE may express this needs in several ways, including setting and activating an option that allows the UE to recognize such a NAS message, or notifying the MME that it is willing and ready to receive such a NI.

### 3.4.4 Support of multipath flows

TCP is the prevalent reliable transport layer protocol used in the Internet to carry user data. Usually, a single path between the TCP sender and receiver is used to carry TCP traffic. In that case, TCP throughput is limited to the capacity of the bottleneck link on the path. Instead, if multiple paths are used simultaneously to aggregate the bandwidth the TCP performance may increase. It has been already shown that the potential of the multipath solution lies not only in providing robustness but also, in conjunction with an appropriate congestion controller, in providing means to balance the Internet congestion in a stable way. Despite these benefits several issues concerning multipath transport of TCP still remain to be addressed before it can be successfully deployed. These include (i) reducing the impact of out-of-order delivery, and (ii) relaxing the requirement of support from the end-hosts.

#### 3.4.4.1 Out-of-Order Packet Receptions

Out-of-order delivery is owed to the fact that TCP data segments are carried over multiple paths with different (propagation) delays. TCP performance suffers from out-of-order packets in different ways as described in [Leung07]. Among these, the most important one is that out-of-order arrivals at the receiver cause generation of duplicate acknowledgements (DUPACKs), which in turn unnecessarily triggers the fast retransmission/recovery algorithms [Allman09].

There are several possibilities to minimize the impact of the out-of-order deliveries, namely by (1) using delay estimations of the path, (2) receiver buffering and reordering or (3) tweaking the ACK information and the fast retransmission triggering threshold.

BAG (Bandwidth Aggregation) proxy [Chebrolu05] and Simula Proxy [Evensen09] are two examples of multipath solutions that use delay estimations of the paths to minimize the number of out-of-packet receptions by the TCP receiver. BAG proxy uses PET (Packet Pair based Earliest-Delivery-Path-First) scheduling algorithm to distribute packets to multiple paths: PET selects the path which will deliver the packet earliest to the client. In contrast, Simula Proxy [Evensen09] uses path delay estimations to calculate packet arrival times and buffer the potentially out-of-order packets on the shorter paths, compensating the different path delays.

However, out-of-order packet receptions are inevitable since scheduling algorithms are based on delay estimations which may not be valid because of variations in the network. Thus, buffering and reordering of the out-of-order packets may be necessary before reaching the TCP receiver. A common method used in that case is to buffer and reorder out-of-order packets by means of a component on the receiver host before they are passed to the TCP receiver entity [Chebrolu05], [Radunovic08].

When neither receiver-side buffering nor reordering of out-of-order TCP packets is used, DUPACKs will be generated by the receiver as a response to out-of-order packet receptions. Then, one of the options may be to increase dupthresh value logarithmically based on the number of paths used, as shown in [Lee02]. Alternatively, path utilization information can be used for more accurate scheduling of the packets so that out-of-order deliveries are minimized. PRISM [Kim07] is an example of such a solution using both utilization and delay of the paths, with path utilization having a higher precedence than the path delay: the least utilized path is selected to schedule a packet. PRISM processes TCP selective acknowledgment (SACK) information in DUPACKs to infer sequence number of the out-of-order packets that caused generation of the DUPACKs and modifies packets' cumulative ACK numbers with the inferred sequence number to hide reordered packets from the TCP sender. It buffers and re-orders ACKs based on these cumulative ACK numbers.

#### 3.4.4.2 Deployment and Adoption

From the deployment point of view, the solutions require support from the TCP sender and/or receiver. TCP-PARIS [Karrer05] receiver modifies connect() socket call to pass multiple server addresses. In addition, TCP-PARIS sender uses TCP options to signal its current cwnd value to the receiver. MPTCP [Ford11] also uses TCP options [Ford12] to setup multiple TCP subflows between the TCP end-points and transfer data over them. pTCP [Hsieh05] gets the number of interfaces to use through a socket option.

TCP-PRISM sender side uses additive increase proportional decrease instead of the TCP additive increase multiplicative decrease algorithm to grow the cwnd. TCP-PRISM uses also negative ACKs to learn about the lost packets and decide on the proportion of the cwnd decrease for a path. TCP sender's dupthresh value is modified in [Lee02] based on the number of paths used to transfer TCP data. In addition, TCP receiver uses TCP delayed ACKs with minor modifications in [Lee02]. BAG clients send their interface addresses to the BAG proxy, Simula proxy/clients must be configured to use network address translation (NAT), and TCP-PRISM proxy has to know WWAN interface addresses of the neighbors of the TCP receiver to send TCP data packets via GRE encapsulation.

Kostopoulos et al. [Kostopoulos10] mention three requirements for a solution that requires changes at the both ends of a connection: (1) the solution must be implemented by the OS, (2) at least one of the end-hosts must have simultaneous Internet access across different network interfaces, and (3) both sides must be capable of using the solution. Not surprisingly, solutions that require changes on the end-points have little chance of being adopted and deployed. Therefore, it is necessary to design TCP over multiple paths solutions that are independent of the support from the end-hosts.

#### 3.4.4.3 Transparent TCP over Multiple Paths Proxies

Our proposal aims at addressing both of the mentioned shortcomings allowing therefore an agnostic TCP over multiple paths solution. We aim to design and implement a transparent proxy solution to increase TCP performance over multiple paths (with different RTTs).

We introduced a TCP splitter/combiner architecture (SCA) which enables development of transparent TCP over multiple paths proxies in [Ayar12a]. We documented the architecture and its use cases in an RFC draft [Ayar12b]. Main idea is to place a proxy pair in between TCP sender and receiver without any need for deployment on end-hosts. The proxy pair may split and shape TCP traffic over multiple paths agnostic to TCP sender and receiver. Signaling between proxy pair will ensure that use of multiple paths doesn't harm TCP throughput. The idea has been presented at the 83rd IETF meeting.

Besides working as a pair, proxies may need to work standalone (i.e., without any peer proxy). A transparent standalone SCA proxy that aims at solving the out-of-order packet reception problem is introduced in [Ayar12a]. While scheduling packets to the multiple paths, the proxy estimates out-of-packet arrivals and filters DUPACKs that are resulted from the early packet arrivals on shorter paths. Initial results show the potential in using our architecture to develop transparent TCP over multipath solutions without any end-system modifications.

As future work, we will extend capability of our previously proposed DEF (DUPACK Estimation and Filtering) SCA proxy [Ayar12a] to shape ACK traffic. ACKs will be buffered and processed at the proxy before being sent to the TCP sender. That way ACKs will arrive at TCP sender paced with the RTT of the longest path used. Thus, TCP sender's RTT/RTO estimations will not be inaccurate.

## 3.5 Multi-Criteria cell selection

The growth in next-generation wireless networks is driven by the ever increasing popularity of wireless data applications. In addition to supporting voice services, future mobile networks are envisioned to be truly multimedia networks, which integrate a variety of different applications.

Their success depends on the ability of future network architectures to provide the necessary capacities to support high data rate services, while at the same time dealing with the different mobility patterns of the users. To this effect, hierarchical network architectures with different cell layers are presented. Different layers are distinguished by their respective cell sizes, their maximum throughput, and the number of supportable users.

For simplicity, in such kind of networks, generally two different types of cells are deployed in a hierarchical manner as shown in Figure 3-14. The first type is a wide area cell (macrocell) that provides moderate data rates for users moving at high speed. The second type is a local area cell (microcell) that covers a certain area with high traffic density for static or nomadic users.

One of the advantages of hierarchical architecture is the fact that the distance between a mobile station (MS) and a base station (BS) can be reduced by assigning each user to its closest BS. Thus, both stations are required to use smaller transmit power and experience a lower level of interference. Moreover, the users in the coverage of local area cells have an opportunity for selecting not only the local area cell but also the overlaid wide area cell as a serving cell. Therefore, efficient load balancing can be achieved by applying properly designed cell selection criteria.

The cell selection has an effect on evolved Node B (eNBs) belong to different serving gateways (SGWs). In addition to that, traffic redirection is also requested in the core network.

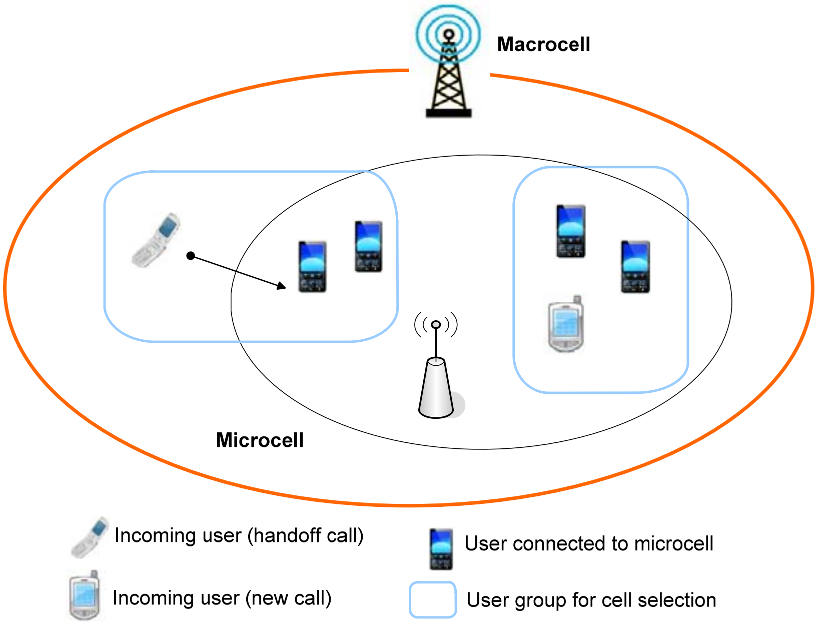

Macrocell
Microcell
Incoming user (handoff call)
User connected to microcell
Incoming user (new call)
User group for cell selection

**Figure 3-14:** ***System model of hierarchical cellular networks [Moon10a]***

Although, using small cells in the network increase the system capacity and number of users to be served, it brings an expense called handoff. It is a kind of switching between cells of the users. Handoff problem is a very crucial and interesting issue to be overcome for hierarchical networks. Many metrics such as signal strength, distance, signal-to-noise ratio (SNR), bit error rate (BER), traffic load, quality indicator and some combination of these indicators can be used in order to decide if the handoff is required or not. Generally, more handoffs imply more overhead, and the risk of call dropping will also increase. Therefore, it is a significant topic to decrease the number of handoff through proper network or cell selection. In these hierarchical systems, cell selection which is the process of determining the cells that provide service to each mobile station is an important issue. Cell selection is responsible for guaranteeing the required Quality of Service (QoS) for MS, keeping MS always camp on a cell with good enough quality, balancing the system load and number of handoffs.

Cell selection has received much attention in recent years, while the existing researches focus mainly on multiple access techniques, power control schemes and handoff protocols. In [Hanly95], a cell selection algorithm has been presented to determine power allocation among different users to satisfy per-user signal-to-interference-noise ratio (SINR) constraints. In [Sang04], a High Speed Packet Access (HSPA) based handoff/cell-site selection technique has been studied to maximize the number of connected mobile stations, and presented the scheduling algorithm to achieve this objective. These techniques take neither variable BSs' capacities nor MSs' QoS requirements into account. An integrated design of optimal cell-site selection and frequency allocation, which maximizes the number of connected MSs and meanwhile maintains quasi-independence of radio based technology, has been provided in [Mathar02].

A method called velocity based cell selection where the cell assignment decision is made based on the mobile's estimated velocity since the mobile speed is an important factor in order to select the proper cell for a mobile user in hierarchical networks. In a real network, cell selection with this criterion can be done with the knowledge of either an accurate velocity estimate or the dwelling time within a cell. A tiered network must have certain velocity thresholds, which is determined by the users' mobility pattern, system QoS, and system capacity, traffic balancing between tiers, handoff rate constraint and etc., to use the velocity as selection criteria. A mobile user with the higher speed than a threshold value is assigned to a larger cell; otherwise it is assigned to a smaller cell. For example, slow mobiles can be assigned to microcells, while fast mobiles are assigned to macrocells for a two-tier network in order to minimize the number of handoffs [Anpalagan99]. In [Chung02], different velocity estimation methods are examined in order to obtain a perfect cell selection. This velocity estimation method based on cell sojourn time offers the advantages of detection of a direction-changing effect, efficient user allocation to cells, estimation of instant velocity, easy implementation, and low power consumption. In [Klein04], a cell selection is performed according to the optimal decision thresholds of the velocity and the amount of data. These thresholds are analytically derived to minimize expected system load. In [Li09], the accurate velocity of the MS is estimated by using the wireless positioning system such as fingerprint technology and back-propagation neural network and the performance results are obtained for multi-frequency band in WiMax systems. Three-tier network is used to classify three kinds of velocity, high speed (like speed of the traffic on the high way), middle speed (like speed of the traffic on the road) and the low speed (like the walking speed of the people). If the velocity of MS is high firstly the macro-cells are considered as the target cells, if the speed is middle type, then the micro-cells are considered and pico-cells are considered as the target cells when the speed is low. To avoid inaccurate estimation of mobile speed, the dwelling time of a mobile user in its original cell is often used to decide if the mobile user should be assigned smaller or larger cell [Lagrange96] [Cha08]. In [Lagrange96], a simple strategy is proposed in a macrocell/microcell network in which a new mobile is assigned to the microcell firstly. After a certain time threshold, if the dwelling time of this mobile user in its first cell is longer than the threshold, the user will be assumed as a slow user and let it to stay in the microcell, otherwise it will be assumed as s fast user and it will be assigned to the macrocell. In [Cha08], the cell selection is obtained using dwelling time for the integrated WLAN (picocell) and WiBro (microcell) networks. Proper dwelling time threshold has been calculated in order to maximize throughput and reduce unnecessary handoffs.

A hierarchical network not only contains macrocell, microcells and picocells but also contains femtocells which have been investigated in the literature. In [Mahmoud10], two cell selection algorithms are compared for the hierarchical networks containing macrocells and femtocells. The first algorithm is based on the Received Signal Strength (RSS) and the cell with the strongest RSS is chosen as a target cell. However, interference and bandwidth are two other critical parameters affecting the capacity of MSs, and should be also taken into account. Especially for femtocell deployment scenarios, there may be significant imbalance in available bandwidth per user when a user is connected to a femtocell compared to when it is connected to the macrocell. Even though the link quality from a certain cell may be better, it may be overloaded with users, allowing only limited capacity for a new MS that requests to join the cell. On the other hand, a different cell with comparably worse link quality may have better spectrum resources (e.g., a femtocell serving to no users, or only few users), and may provide better capacity for a new MS that requests admission. Therefore, a capacity-based cell selection metric may be preferable over a link-quality based cell selection metric [Mahmoud10]. A queue based cell selection algorithm which use the Received Signal Received Power (RSRP) metric is presented for the heterogeneous systems in [Qu10]. In [Kwon11], a new cell selection method is presented for indoor femtocell network where many femtocells are deployed together. This method increases the reliability of femtocell network in indoor environment by minimizing handover failure and achieving a load balancing between neighboring cells.

In current cellular systems, the cell selection process is done by a local procedure initialized by a mobile device according to the best detected SNR. However, this local selection decreases the system capacity since the users only consider the channel quality while selecting their serving cells and ignores the system load information. In order to converge to the optimum solutions while selecting the cells, global cell selection procedure which uses both channel quality information and system's load information is better than local cell selection procedure. The importance of global cell selection procedure is shown in [Amzallag08] by illustrating an example. According to Figure 3-15, let's assume that the best SNR for MS1 is detected from microcell A, and thus MS1 is being served by this cell. When MS2 arrives, its best SNR is also from microcell A, who is the only cell able to cover MS2. However, after serving MS1, microcell A does not have enough capacity to satisfy the demand of MS2 who is a heavy data client. However, if MS1 could be served by picocell B then both MS1 and MS2 could be served. This simple example illustrates the need for a global, rather than a local, cell selection solution that tries to maximize the global utilization of the network, and not just the SNR of a single user. This cell selection optimization problem is defined as all-or-nothing demand maximization problem.

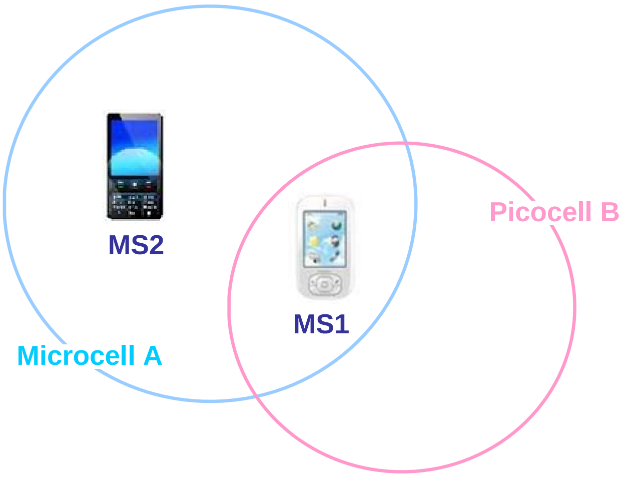


**Figure 3-15:** ***Cell selection scenario for a high loaded cell [Amzallag08]***

One of the most important features of the evolving fourth generation (4G) wireless networks is heterogeneous wireless access networks. Moon et al. also have examined a global cell selection algorithm in [Moon10] [Moon10a] Cells are selected using the coordination of multiple users as a basis; rather than relying on the choice of a single user and uplink transmit power is used as a key parameter. The coordination is designed to reduce the total transmit power of users in both a macrocell and a microcell, i.e., the total power required in each cell. In [Chang10], global cell selection scenario is used and a preference value-based cell selection (PVCS) scheme has been investigated for heterogeneous wireless access environment in which the MS has the ability of connecting to different wireless access networks (for example, WiMAX, LTE and WiFi) as seen in Figure 3-16. The PVCS scheme contains three stages, candidate cells selection, preference value calculation, and target cell determination. The candidate cells selection is used to filter out unsuitable cells by checking three thresholds including the received signal strength constraint, cell loading constraint and dwell time constraint. All these suitable cells constitute a candidate cell set. The preference value calculation of each cell in the candidate cell set is optimized by considering the factors of loading, QoS, and mobility to maximize overall system utilization, maintain the call request's QoS requirements, and minimize handoff occurrence frequency. Finally, the target cell, which has the maximum preference value, could be selected for the call request. A fuzzy multiple objective decision algorithms has been applied in [Guo06] to select the optimal cell from all candidates and an approach based on game theory is used in [Gao11] to solve the cell selection problem.

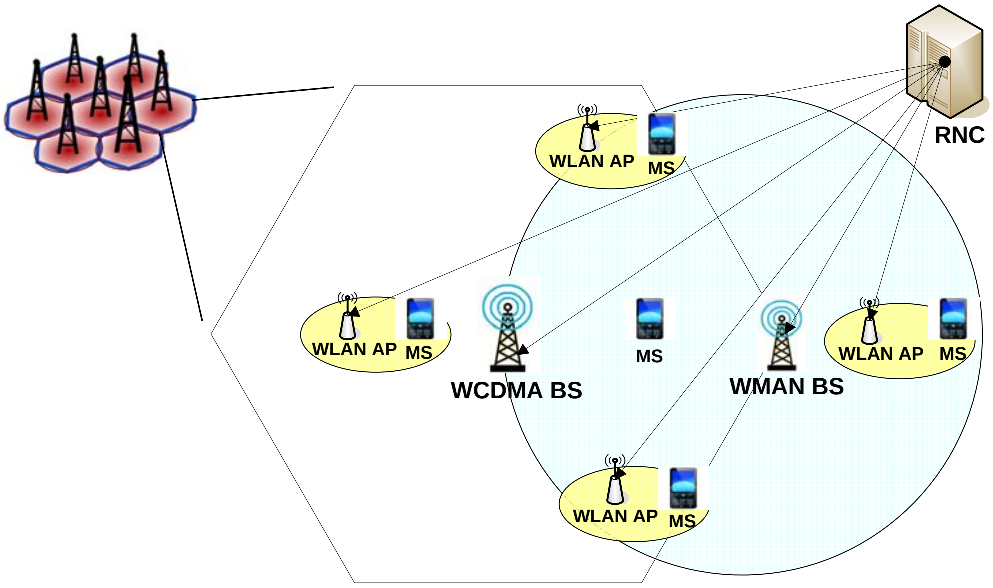


**Figure 3-16:** ***Global cell selection for heterogeneous wireless access environment consisting of a WCDMA cellular system, an IEEE 802.16 WMAN system, and an IEEE 802.11 WLAN system [Chang10]***

Fast Cell Selection (FCS), that is also substantial topic for cell selection procedure, is an important link adaptation technique implemented in 3G WCDMA downlink networks to enable more reliable communications while simultaneously improving data throughput and system capacity. Using FCS, the mobile user can select the best active BS which should serve it on the downlink. It can be treated as alternative to soft handoff. Note that in current WCDMA-HSDPA and CDMA-HDR downlink systems, no soft handoff is used. Various studies have also been performed on FCS in the literature. A network controlled cell selection and its advantage were studied in [Das04] for HSDPA whose network architecture is shown in Figure 3-17. Two types of FCS schemes called inter and intra BS FCS have been presented. The inter BS FCS considers any cell in the network as a candidate for an active set defined as a collection of cells in good condition enough to be considered as a communication candidate for the transmission of next frame. Since the active set comprises links controlled by different BSs, the MS determines a serving cell without taking into account the resources available in the target cell. On the other hand, the intra BS FCS sets up an active set comprising cells covered by a common scheduler that controls the FCS operation considering both the channel quality and load balancing. The common scheduler determines a serving cell for each MS that reports the channel quality information of all links in the active set. In [Gomes09], a network controlled collaborative cell selection scheme is presented in which the RNC, the Node-Bs under the RNC and the UE collaborate in selecting the serving cell for a UE to ensure user satisfaction while maintaining overall network health. Many factors are incorporated in the decision making process such as QoS requirements, device capability, and available network resources, and channel condition. In [Lee06] a FCS scheme is proposed with the use of interference avoidance in the downlink of a packet based Orthogonal Frequency Division Multiplexing (OFDM) cellular system with the use of a universal frequency reuse factor. Most of previous FCS works have been researched in the single-input single-output (SISO) systems. In [Kim09], a FCS scheme is considered for multiple-input multiple-output (MIMO) systems in multi-cell environments.

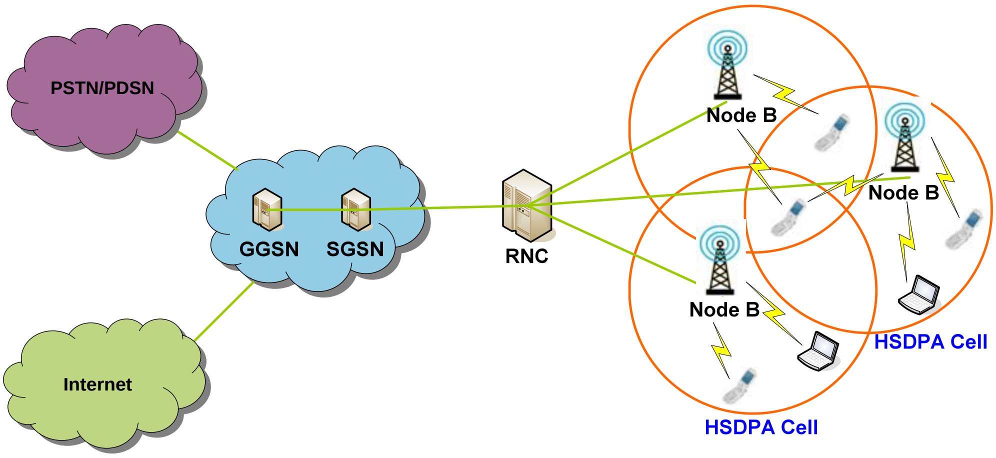


**Figure 3-17:** ***HSDPA Network Architecture***

For LTE (release 8) and LTE-Advanced (release 10) systems, heterogeneous networks where femtocells and picocells overlaid onto macrocells are extensively discussed in addition to traditional well-planned macrocell deployment to improve further the system throughput. There are some differences between the LTE and LTE-Advanced (LTE-A) system. First difference is while LTE Advanced is backward compatible with LTE, LTE is not backward compatible with any 3G networks. The second difference is that the LTE Advanced can be forward and backward compatible with LTE and LTE can also forward and backward compatible with LTE Advanced. In addition LTE can offer 326 Mbps and LTE-Advanced can offer 1200 Mbps. The cell selection methods mentioned in the following paragraphs can be suitable for both LTE and LTE-A systems.

One of the methods of cell selection in LTE and LTE-A is the Reference Signal Received Power based cell selection (RSRP-based cell selection) which compares the downlink (DL) signal transmitted from neighboring cells and the largest selected RSRP as the serving cell by UEs. The cell index can be selected according to the following criteria:

$$i_{RSRP} = \arg\max\left\{c_i^{'}\right\}$$

$$\begin{aligned} c_i^{'} &= c_i && \textit{for macrocell} \\ c_i^{'} &= c_i + \alpha_{RSRP} && \textit{for picocell} \end{aligned}$$

where $i_{RSRP}$ represents the cell index based on the RSRP based criteria, $c_i^{'}$ represents the i$^{th}$ cell, $c_i$ represents the RSRP of the i$^{th}$ cell to be used for cell selection and $\alpha_{RSRP}$ represents the offset value. The RSRP value is defined as the averaged value over the instantaneous variation. The offset value is used to compensate for the difference in interference level between the picocells and macrocells.

Another method for cell selection is Reference Signal Received Quality (RSRQ) which provides additional information when RSRP is not sufficient to make a reliable handover or cell reselection decision. RSRQ is the ratio between the RSRP and the Received Signal Strength Indicator (RSSI), and depending on the measurement bandwidth, means the number of resource blocks. RSSI is the total received wideband power including all interference and thermal noise. As RSRQ combines signal strength as well as interference level, this measurement value provides additional help for mobility decisions. The RSRQ can be defined as :

$$RSRQ = \frac{RSRP}{RSSI} \times N$$

where N is the number of resource blocks. The RSRQ value is defined as the averaged value over the instantaneous variation. Therefore RSRQ is proportional to

$$\frac{SINR}{1+SINR}$$

Here SINR is the signal to interference plus noise ratio. According to the increase in the SINR, the RSRQ is increased. By using the offset value, $\alpha_{RSRQ}$ , the UE selects the cell index based on the following criteria :

$$i_{RSRQ} = \arg\max\left\{\rho_i^{'}\right\}$$

$$\rho_i^{'} = \rho_{Nonprotected,i} \qquad \textit{for macrocell}$$

$$\rho_i^{'} = \rho_{\mathrm{Pr}otected,i} + \alpha_{RSRQ} \qquad \textit{for picocell}$$

where $i_{RSRQ}$ represents the cell index based on the RSRQ based criteria, $\rho_{Nonprotected,i}$ represents the RSRQ in the non-protected resources for the i$^{th}$ cell, $\rho_{Protected,i}$ represents the RSRQ in the protected resources for the i$^{th}$ cell, $\rho_i^{'}$ represents the RSRQ value of the i$^{th}$ cell to be used for cell selection and $\alpha_{RSRQ}$ represents the offset value. By introducing the offset value in RSRQ-based cell selection, different offset values will be required compared due to the different gradients in the RSRQ function.

To further improving the heterogeneous deployment gain, a new cell selection scheme called "Range Expansion (RE)", in which a bias value in RSRP is used for picocell selection to drive more users selecting low power node as serving node [Tongwei10]. The cell index can be selected according to the following criteria:

$$\beta_i = argmax\left\{RSRP + \alpha_{RE}\right\}$$

where $\alpha_i$ is the offset value for cell selection, $\beta_i$ represents the RE value of the i$^{th}$ cell to be used for cell selection.

In heterogeneous network deployment, cell selection as well as intercell interference coordination (ICIC) is very important to improve the system and cell-edge throughput. In [Sangiamwong11], a new technique called signal to interference plus noise power ratio (SINR-based cell selection) is mentioned to consider ICIC effect. In this scheme, the average SINR for pico cells is used for cell selection. In order to obtain the offload effect from macro to picocells an additional offset value is introduced and the UE selects the cell index based on the some specified criteria using this offset value. If ICIC is employed, RSRP-based cell selection is not optimum, since it only reflects the received power from each cell and does not reflect the channel quality of the respective resources. However, simple extension by adding offset values for the picocells can be used to compensate for the difference in channel quality between macro and picocells. Although the offset value in the SINR-based cell selection is mainly used to obtain the offload effect, the offset value in the RSRP-based cell selection is also used to compensate for the difference in interference level. Therefore, a higher offset value is required for the RSRP-based cell selection. The offset value, $\alpha_{SINR}$ can be integrated to the selection of the cell index based on the following criteria:

$$i_{SINR} = \arg\max\left\{\gamma_i^{'}\right\}$$

$$\gamma_i^{'} = \gamma_{Nonprotected,i} \qquad \textit{for macrocell}$$

$$\gamma_i^{'} = \gamma_{\mathrm{Pr}otected,i} + \alpha_{SINR} \qquad \textit{for picocell}$$

where $i_{SINR}$ represents the cell index based on the SINR based criteria, $\gamma_{Nonprotected,i}$ represents the SINR in the nonprotected resources for the i$^{th}$ cell, $\gamma_{Protected,i}$ represents the SINR in the protected resources for the i$^{th}$ cell, $\gamma_i^{'}$ represents the SINR value of the i$^{th}$ cell to be used for cell selection and $\alpha_{SINR}$ represents the offset value.

In case of femtocell deployment in the heterogeneous networks, some additional cases should be taken into account, such as access methods. There are three different access methods in femtocell deployment. These are closed access which requires authentication from its licensed users, open access which is available for every user and hybrid access method which allows the connectivity of nonsubscribers while restricting the amount of OFDMA sub channels that can be shared [Lopez2009]. According to these access methods, different cell selection methods are introduced in [Simsek11] as shown in Figure 3-18. In [Simsek11], it is explained that the best performance can be provided by the upper bound access mode, in which each UE can feel free in selecting a BS no matter if it is a macro or a femto UE. For the average normalized UE throughput the upper bound (case E) supports mainly UEs with low SINR values, whereas for high SINR values the access mode, which gives the femto UEs the permission to get access from a macro BS, shows the best performance.

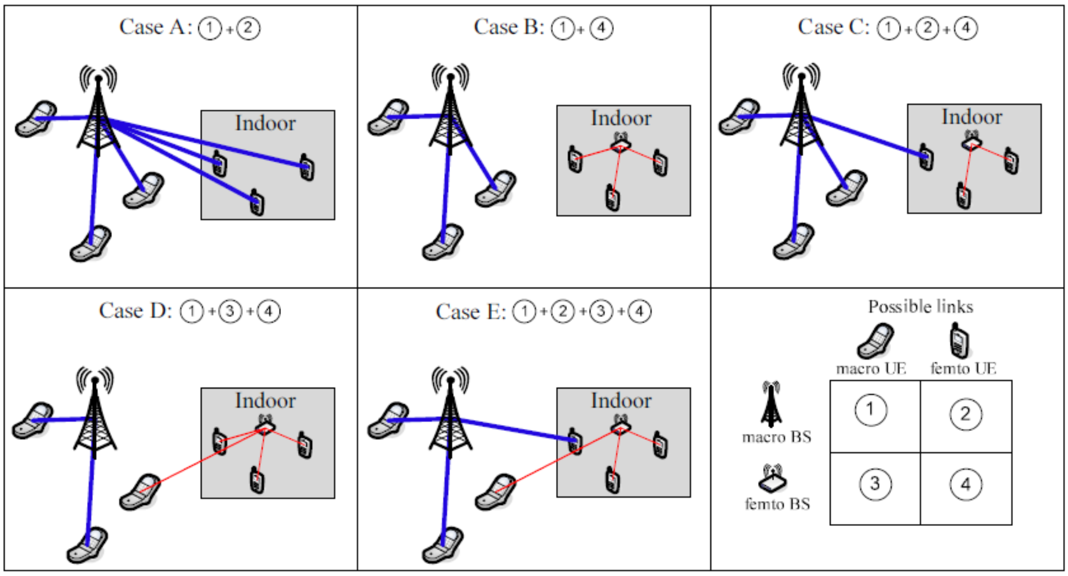


**Figure 3-18:** ***Cell Selection Methods in LTE Heterogeneous network***

We investigate various cell selection algorithms based on different criteria for heterogeneous networks and evaluate their suitability for the future mobile communication networks considering the effect on core network.

The existing various cell selection algorithms are:

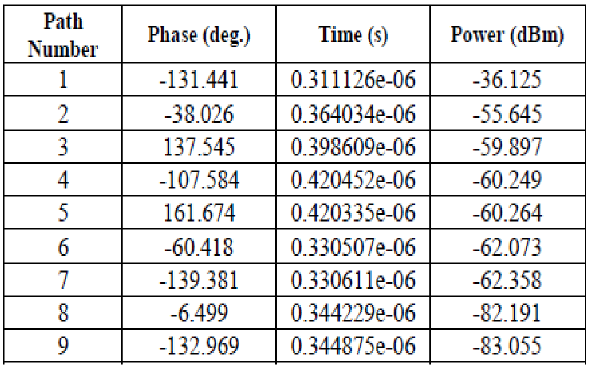
- SINR Based Cell Selection Algorithm
- Distance Based Cell Selection Algorithm
- Velocity Based Cell Selection Algorithm
- Received Signal Strength Based Cell Selection Algorithm
- A Queue Based Cell Selection Algorithm
- Preference Value Based Cell Selection Algorithm
- Fast Cell Selection Algorithm
- Fuzzy Multiple Objective Decision Algorithms

In order to compare the various cell selection algorithms, firstly a realistic simulation environment is generated and channel coefficients are obtained using Wireless Insite simulation tool. Then, the channel coefficients are supplied to MATLAB for performance evaluation.

A table of sample channel coefficients is given in Table 3.1 where the channel impulse response is computed by

$$g_{ij}(t)=\frac{1}{M}\sum_{k=1}^{M}\sqrt{P_k e^{j\theta_k}}\,\delta_k(t-\tau_k)$$

Table 3.1: Sample Channel Coefficients obtained from Wireless Insite



| Path Number | Phase (deg.) | Time (s) | Power (dBm) |
|---|---|---|---|
| 1 | -131.441 | 0.311126e-06 | -36.125 |
| 2 | -38.026 | 0.364034e-06 | -55.645 |
| 3 | 137.545 | 0.398609e-06 | -59.897 |
| 4 | -107.584 | 0.420452e-06 | -60.249 |
| 5 | 161.674 | 0.420335e-06 | -60.264 |
| 6 | -60.418 | 0.330507e-06 | -62.073 |
| 7 | -139.381 | 0.330611e-06 | -62.358 |
| 8 | -6.499 | 0.344229e-06 | -82.191 |
| 9 | -132.969 | 0.344875e-06 | -83.055 |

Initially we use SINR and Distance Based Cell Selection Algorithms.

*SINR Based Cell Selection:*

$$i_{SINR} = \arg\max_{\forall i} \left(\gamma_{ij}\right)$$

where $i_{SINR}$ is the selected cell index based on the SINR-Based criteria, and $\gamma_{ij}$ is the SINR value of the ith cell to be used for cell selection. SINR for each user is calculated as

$$SINR_{ij} = \frac{\left|g_{ij}(t)\right|^2}{N_0 + I_{uj}(t)}, \quad where \quad I_{ij} = \sum_{u=1;u\neq i}^{U} \left|g_{uj}(t)\right|^2$$

where $I_{ij}$ is the total amount of interference which is caused by other cells.

*Distance Based Cell Selection:*

$$i_{\text{Distance}} = \arg\min_{\forall i} \left(d_{ij}\right)$$

where $i_{\text{Distance}}$ is the selected call index based on the distance-based criteria, and $d_{ij}$ is the distance value of the *i*-th cell to be used for cell selection.

The comparisons of the cell selection algorithms in MATLAB are provided in validation results documents. As future work, the network performance of the cell selection algorithms will be compared in NS3 network simulation environment for LTE/EPC.

## 3.6 Cross Layer Interference Detection

Interference is a physical layer phenomena causing data packet corruption at a receiver in wireless networks, caused by simultaneous packet transmissions from multiple senders [Fussen05]. Measuring interference at physical layer is a difficult task: because measurements are usually performed at the receiver prior to packet transmissions, they do not reflect in all cases the present inference situation. Interference at the next packet transmission cannot be predicted [Wu08]; colliding packet transmissions are a stochastically process and are not deterministic. Another reason is that end-systems are moving and changing the interference conditions dependent from their location in time. And clearly, if the number of senders becomes very dense, the complexity of the algorithms for interference detection increases dramatically [Wee11].

Therefore research efforts aimed to identify interference at the MAC and the transport layer protocols. The main problem is to differentiate between different causes of packet losses.

The reaction of a protocol should be different on congestion, path loss errors or fading channels.

On MAC level, the dominating parameter for interference detection is the signal strength indicator (RSSI) or related parameters (e.g. SNR). If an end-system or base station receives some radiation from another device, this device will be identified as a potential source of disturbances. This principle has been widely implemented in wireless systems that operate on a master/slave principle like IEEE 802.15.3 [IEEE802.15.3] or WiMax [IEEE802.16-2004]. But RSSI measurements filter signals that must have a certain energy level (strong senders) and neglect very weak senders [Maheshwari09], which can have an influence on interference, noticed by additional packet losses. The influence of such weak senders can be observed by the hidden terminal problem in IEEE 802.11 wireless LANs [Li08], [Abusubaih08].

To overcome this problem, a lot of MAC layer proposals exist that will implicitly reduce interference, namely access point selection policies [Abusubaih08a], hybrid MACs [Abusubaih09], cooperation of multiple access points or loss differentiations mechanisms. E.g. if end-systems want to communicate outside the wireless network, they use access-point selection policies to choose an access point. The decision to choose an access point can be made as usual dependent on the received signal strength or based on more sophisticated algorithms that take into account the load of an access point or the number of lost packets. The used metrics (packet losses, load etc.) could be a sign of interference.

Only little attention had been made to make TCP aware of different causes of losses. TCP Veno [Cheng03] and TCP Westwood [Gerla01], all extensions of TCP Reno that claim to differentiate between congestion losses and random packet losses. Both protocol cannot explicit distinguish between path losses or interference losses. All random losses are treated equally.

It seems that neither MAC layer, nor transport layer can determine interference correctly. For this reason, cross layer approaches have been made to combine th$f$e advantages of the MAC and transport layer. Authors of [Lohier08] are coupling IEEE 802.11 MAC retransmissions with the operation of TCP. Whenever a packet could not transmitted within the retransmission timeout value of TCP, the IEEE 802.11 MAC will alert TCP. This will give TCP the possibility to faster recover from non-congestion losses. Also TCP Veno has been extended (TCP VenoPlus [Shetty10]) to use the RSSI indicator to differentiate between congestion and random losses.

# 4 Microscopic traffic management

## 4.1 QoS differentiation based on applications and user profiles

### 4.1.1 Application and user classification

Application classification consists of identifying the type or the class of application. The new advanced radio technologies providing real mobile broadband packet data services comparable to the fixed internet, the penetration of smart phones combined together with the flat rate pricing used by the operators, contributed (and continue) to the tremendous growth of the mobile data traffic. There are many reasons behind the need for application classification techniques: These reasons make application classification essential in traffic management in order to prioritize different application traffic in the network.

There are different techniques for application classification: (1) payload based classification that is based on the inspection of the packet content including or not the packet payload, and, (2) statistical based classification that consists on analysing the behavioural and statistical characteristics of the traffic (jitter, session time, inter-arrival, UL/DL distribution, packet size, etc.). Montimage and Ericsson Turkey work on the application classification problem. Montimage is working on their technique based on inspection of the contents of packets, whereas Ericsson Turkey is working on a method which can classify applications into categories Video, VoIP, Instant Messaging, P2P, Web Browsing, File Transfer and Gaming without necessarily knowing which particular application was utilized. Montimage is also supporting Ericsson in this effort.

#### 4.1.1.1 Application Classification Mechanism

Montimage's application classification technique is mainly based on the inspection of the contents of the packets. The inspection is performed by comparing the packet headers and application data to already defined signatures that identify different applications. The signatures which are used describe patterns that identify the nature of applications. It is clear that the accuracy of this method depends on the non-overlapping of the signatures. This property is not always easy to satisfy because of some similarities between the applications. Therefore, in addition to the pattern analysis, we add another layer for state analysis that consists on exploiting the sequence of steps of a protocol when it can be modelled using a state machine (example: an HTTP GET request will be followed by a valid HTTP response).

#### 4.1.1.2 Application Classification initial results

The evaluation results presented in the following are based on sample trace files collected by Ericsson Turkey. The objective of the evaluation was mainly to build a ground truth base for the Bulk Traffic Analysis. The traffic trace files were collected on PC machines connected to the fixed Internet by running the applications of interests (set of P2P applications, Web video, Skype) and recording the traffic activity.
Table 4.1 presents an overview of the application classification results performed on Ericsson's trace files. The results show that the classification accuracy is relatively high reaching around 96% of the traffic data in terms of volume and number of packets. However, the accuracy in terms of number of flows is lower (82%). This result was expected as the trace files were captured on a local network, where the broadcast signalling (using UDP) is relatively high. Only 0.74% of TCP flows were left unclassified. Among the unclassified flows (TCP and UDP flows) only 5% has more than 10 packets. Table 4.2 provides the distribution of unclassified flows based on the number of packets. It shows that the majority of these flows contain few packets. While analysing the unclassified flows, we noticed that some of them were already initiated when the sniffing was done. These cold start flows account for 62% of the data volume of unclassified flows though there number is account for only 0.34%.

**Table 4.1: Overview of the application classification results**

| Application Name | Flows Number | Packets Count | Bytes Count |
|---|---:|---:|---:|
| bittorrent | 15459 | 1371049 | 1198970278 |
| skype | 1370 | 617276 | 251476274 |
| https | 301 | 376751 | 181093125 |
| http | 4646 | 103548 | 82374039 |
| youtube | 44 | 83611 | 79238629 |
| unknown | 227 | 99657 | 75449273 |
| gnutella | 206 | 16073 | 12545929 |

| | | | |
|---|---|---|---|
| dailymotion | 20 | 11429 | 10789797 |
| rtsp | 2 | 9485 | 9243015 |
| winmx | 1 | 9791 | 7136343 |
| edonkey | 12 | 4674 | 3777138 |
| udp | 5205 | 11005 | 1142991 |
| ssdp | 172 | 2698 | 933242 |
| dns | 1969 | 4096 | 368160 |
| icmp | 71 | 6046 | 324024 |
| nbns | 116 | 3278 | 304096 |
| google | 16 | 401 | 210003 |
| facebook | 16 | 306 | 171212 |
| dhcp | 255 | 387 | 133235 |
| google_ads | 9 | 181 | 113998 |
| ymail2 | 2 | 130 | 98068 |
| ares | 247 | 730 | 76568 |
| twitter | 15 | 178 | 74266 |
| wikipedia | 4 | 115 | 71108 |
| smb | 90 | 207 | 46226 |
| jabber | 1 | 56 | 13502 |
| netbios | 2 | 55 | 7750 |
| soap | 3 | 35 | 6081 |
| icmp6 | 21 | 39 | 3430 |
| openvpn | 2 | 32 | 2177 |
| linkedin | 1 | 6 | 1786 |
| igmp | 8 | 22 | 1164 |
| **TOTAL** | **30513** | **2733347** | **1916196927** |
| **Total Classified** | **25081** | **2622685** | **1839604663** |
| **Total Classified (%)** | **82.2%** | **95.95%** | **96.00%** |
| **Total Unkown** | **227** | **99657** | **75449273** |
| **Total Unkown (%)** | **0.74%** | **3.64%** | **3.93%** |
| **Total Unclassified (Unknown + UDP)** | **5432** | **110662** | **76592264** |
| **Total Unclassified (%) (Unkown + UDP)** | **17.8%** | **4.04%** | **3.99%** |

**Table 4.2: Distribution of unclassified flows based on the number of packets**

| Packets Distribution | Flows Number | Bytes Count | Flows Number (Cold Start) | Bytes Count (Cold Start) |
|---|---|---|---|---|
| **2 <= Packets** | 4751 | 0.67 MB | 0 | 0 |
| **[3 : 10] Packets** | 423 | 0.34 MB | 0 | 0 |
| **[11 : 100] Packets** | 213 | 0.92 MB | 2 | 5.2 KB |
| **100 + Packets** | 60 | 235.2 MB | 17 | 147.7 MB |
| | **5447** | **237.15 MB** | **19 (0.34 %)** | **147.7 MB<br>(62.25 %)** |

#### 4.1.1.3 Application classification in the network architecture

Application classification of the network traffic can be used for different objectives.

- Understand the application mix and the usage trends: This can be done by periodically (weekly or monthly basis) recording traffic samples on the interface between the core network and the operator's PDN. The traffic application classification in this case can be performed offline. The objective here is to support the operator with up-to-dated view about the network utilization.
- Policy control and enforcement: requires application classification in order to grant priority levels (based on QoS requirements for instance) to specific applications. In this case, there is a need for live application classification with high performance constraints with respect to the speed and latency (wire speed, low latency). The application classification engine can be integrated, in this case, integrated into the P-GW or installed on a dedicated probe to inspect the S5/S8 or SGi interfaces.

##### *4.1.1.3.1 Bulk Analysis of Network Data and Classification*

Deep Packet Inspection tools in network operators investigate the payload and can determine the exact application type such as Skype, Youtube, Dailymotion, Google+, etc. Please see Figure 4-1.This may be required due to many reasons such as the operator's pricing strategy, campaigns or regulations. However trying to map all the network traffic via DPI would be extremely costly and many times unnecessary.

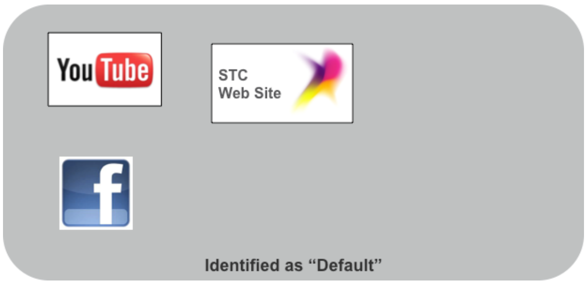


**Figure 4-1:** ***Illustration of how DPI classifies network traffic***

For example, in case there is no specific restriction required, or any pricing strategy, there is no reason to identify whether a VoIP application is Skype or Google talk, however in order to satisfy QoE requirements and manage network efficiently, determining the family of application – that the application is VoIP- would be necessary.

The method which is being developed aims to draw a map of the network traffic, that is, to classify the total usage according to varying time into the following classes:

- Videostreaming
- VoIP
- Instant Messaging
- P2P filesharing
- Web surfing
- Gaming
- M2M

This idea of mapping is illustrated in Figure 4-2.

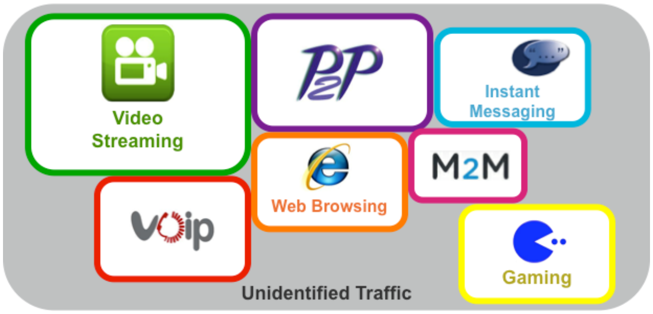


**Figure 4-2:** ***Bulk data analysis aims to bring out the distributions of applications in the MNO.***

The initial aim is to get a mirror copy of data traffic going over the network for a period of time like one hour in a set pattern such as:

- Weekday work hour
- Weekday after work hour
- Weekday night
- Weekend daytime
- Weekend evening
- Weekend night

and get a map of the network as in Figure 4-3. In case there are enough many such intervals, variation of usage during the week can be observed and the distribution of traffic into different applications can be observed.

This can be utilized to set QoS parameters according to distribution and the prioritization of the network operator.

Another aim could be to classify users according to the applications and times they utilize the network and make proper campaigns to customers in order to make better use of the limited bandwidth.

In case this method can be improved so that it can run in real-time, it can be run on the flows where dpi is not running on and help network utilization in real-time and more importantly it may help network neutrality as well.

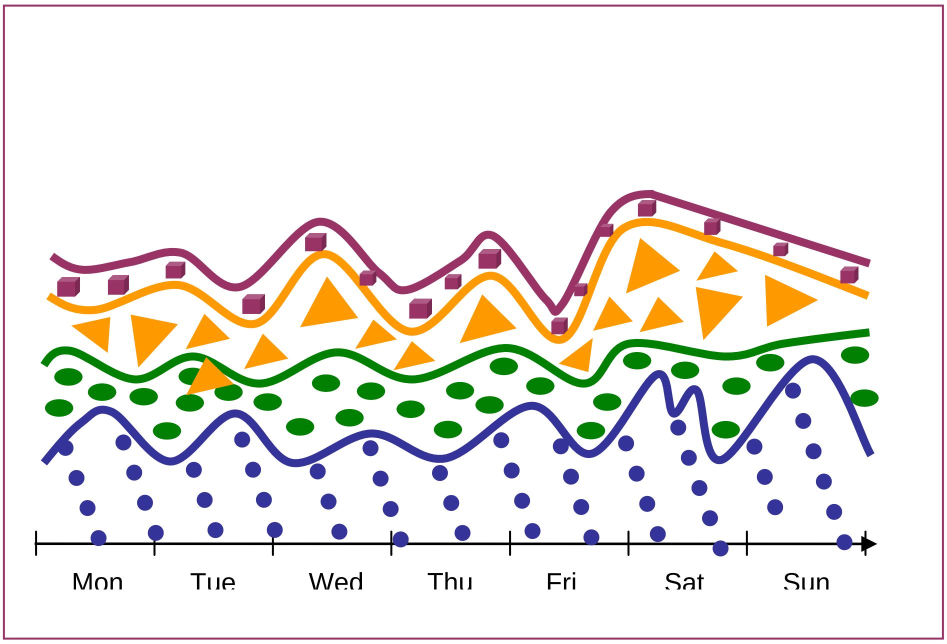


**Figure 4-3:** ***Weekly distribution of network data traffic into applications.***

### 4.1.1.4 Challenges of traffic classification and QoS enforcement

It is a challenging task to classify applications accurately. None of the mentioned methods can provide satisfactory classification of all applications and therefore using together different techniques is typical in modern application classification modules (usually part of DPI systems).

In LTE, policy control is mandatory, meaning that policy enforcement is an essential requirement. This will require DPI functionalities, including application classification. For example, policy decisions can be initiated by Application Functions (AF) that might detect that a particular application is being initiated and notify the PCRF in order to get a decision. To identify the application in use, the AF can include or be linked to a classification function (DPI). We should note that the identification of managed applications (telephony, SMS) is simpler as the operator controls them.

Although, 3GPP standards specified a sophisticated QoS and bearer management model for LTE, it is expected that most Internet traffic will be assigned to the default bearer. In this case, application identification and classification will likely be needed to differentiate and manage internet traffic within the default bearer.

### 4.1.2 Application and user based differentiation

Differentiation of traffic flows for certain applications is increasingly requested and needs to be targeted on a flow or flow class model. This requires the above mentioned classification and detection efforts as well as several means for microscopic traffic management.

Commercial and Linux based routers are in general capable of such traffic manipulation, i.e. traffic shaping, dropping, delay management and bit manipulation.

In MEVICO it is envisioned to develop a microscopic traffic management framework, which derives the required traffic management actions from application QoS profiles associated with specific application behaviour (Skype/YouTube) models.

Starting with the application flow detection, it will be possible to lookup the essential QoS parameters thresholds for satisfying QoE levels and to apply the required actions in a distributed fashion. This concept is shown Figure 4-5.

This however requires a decision on the placement of detection and manipulation nodes within the operator network. The placement task will be solved through simulation as well as optimization efforts. Possible placement options are shown below in Figure 4-4.

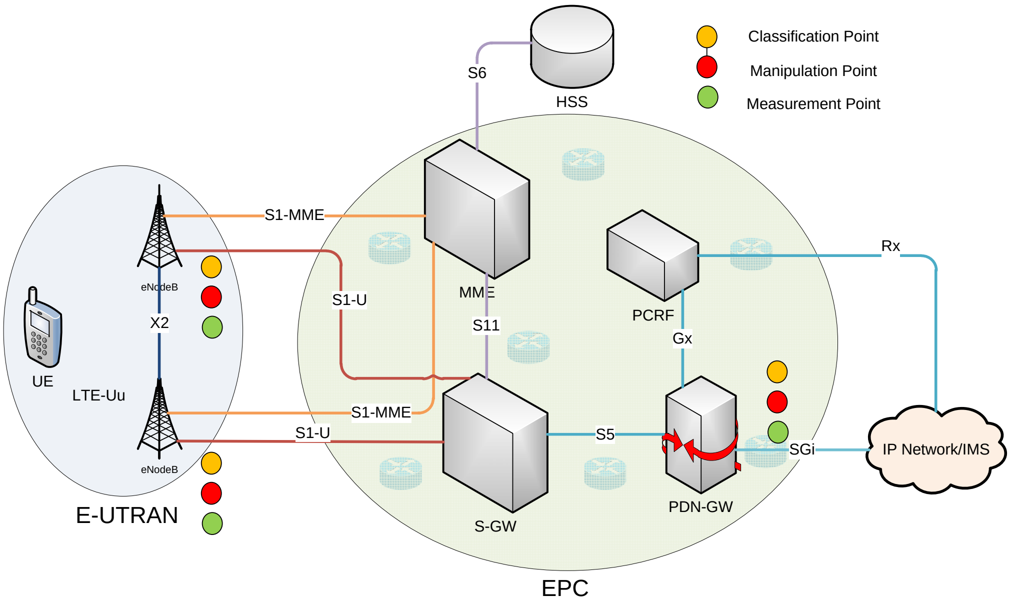


**Figure 4-4:** ***Possible points of presence within the network***

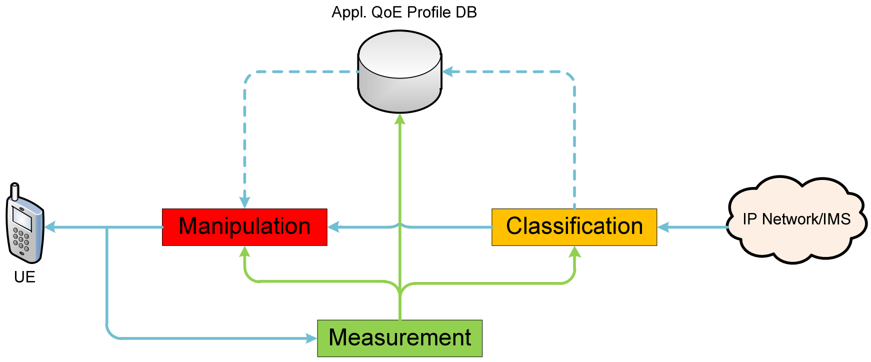


**Figure 4-5:** ***Measurement and control procedure***

One of the major means for differentiation of applications is Quality of Service. Therefore here we focus on Quality of Service and how application and user based differentiation can be made.

## 4.2 End-to-end QoS

The performance of wireless cellular networks is often evaluated in terms of parameters such as the spectral efficiency or the outage probability in academic research and within 3GPP. However, from a network operator point of view, it is important to measure and calculate KPIs in terms of QOS parameters such as latency, jitter, packet loss rate, throughput etc. depending on the application type (such as voice, video, data, control signaling, IT traffic, etc.), from which SLA compliance and end user perceived QoE. This relation is crucial for minimizing the network cost while ensuring QoS as demanded by applications for business and residential users.

### 4.2.1 End-to-end (E2E) QoS in 3G

3GPP defined e2e QoS in Release R99 based on four services and traffic classes:

1. Conversational (e.g. voice),
2. Streaming (e.g. streaming video),
3. Interactive (e.g. web browsing),
4. Background (e.g. background download of emails, files etc.)

In HLR (Home Location Registrar), there are some QoS related parameters such as GBR (Guaranteed Bit Rate), MBR (Maximum Bit Rate) and THP (Traffic Handling Priority), ARP (Allocation and Retention Priority). There are three levels of ARP, 1- high, 2-medium, and 3-low.

The scheduling of packets is done based on SPI (Service Priority Index). There may be a different number of SPI levels. In the Node-B scheduling process, RABs (Radio Access Bearers) are mapped based on SPI with an operator definable weight on each of 15 SPIs (Scheduling Priority Indicator).

Node-B shall use these different SPI levels with appropriate scheduling algorithms (e.g. maximum C/I, Proportional Fair, minimum GBR, etc) to differentiate individual HSDPA flows, taking into account both radio conditions, resources and call priorities.

#### 4.2.1.1 Bearer concept for QoS support

An end-to-end service may have a certain Quality of Service (QoS) which is provided for the user of a network service. A bearer service includes all aspects which comprise control signalling, user plane transport and QoS management functionality, to enable the provision of a contracted QoS. In UMTS bearer service layered architecture, each bearer service on a specific layer offers it's individual services based on the layers below.

A *bearer* uniquely identifies packet flows that receive a common QoS treatment between the terminal and the gateway. A TCP/IP packet flow is defined by a quintuple based on information in the TCP and IP headers:

- the source and destination IP address,
- the source and destination port number,
- a QoS marker (DSCP: DiffServ code point) and/or protocol ID

A bearer is the level of granularity for bearer level QoS control in the EPS. That is, all packet flows mapped to the same bearer receive the same packet-forwarding treatment (e.g., scheduling policy, queue management policy, rate-shaping policy, link-layer configuration, etc.).

One bearer exists per combination of QoS class and IP address of the terminal. The bearer is the basic enabler for traffic separation, that is, it provides differential treatment for traffic with differing QoS requirements. The concept of the bearer and the associated signaling procedures further enable the system to reserve system resources before packet flows that are mapped to that bearer are mapped into the system. Bearer concept is illustrated in Figure 4-6.

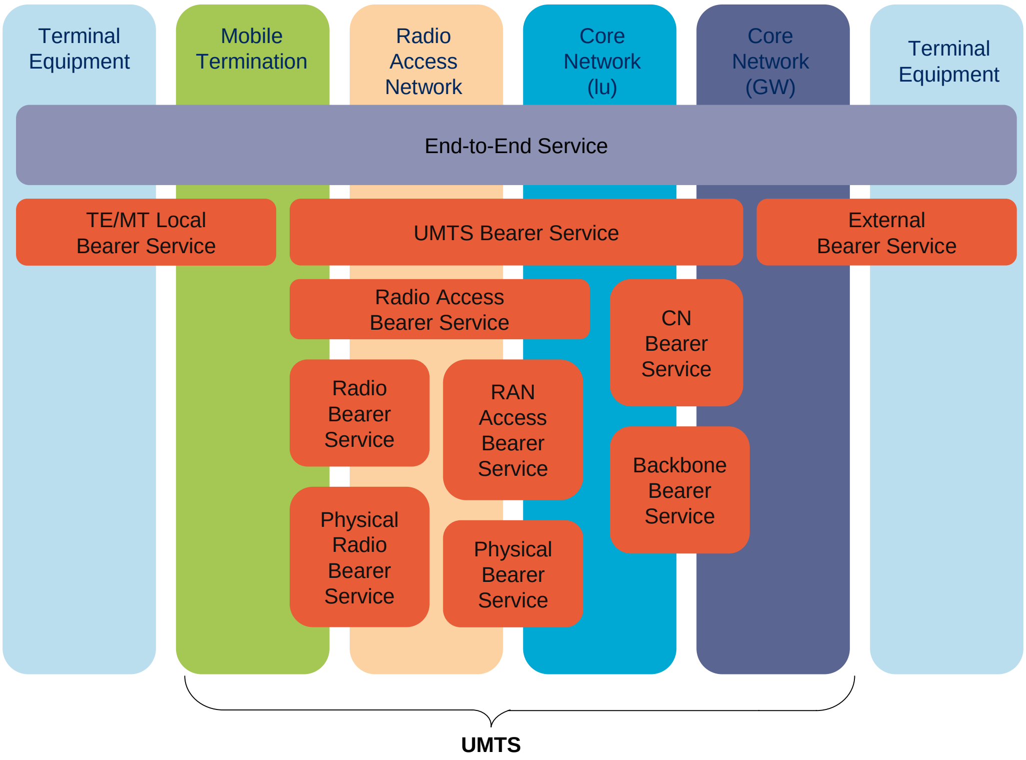


**Figure 4-6:** ***UMTS bearer service layered architecture***

### 4.2.1.2 Types of bearers

**GBR vs. non-GBR bearers**

Two types of bearers exist: guaranteed bit-rate (**GBR**) and non-guaranteed bit-rate (**non-GBR**) bearers.

The traffic carried by a GBR Bearers conforms to the value of the GBR QoS parameter associated with the bearer. On the other hand, the non-GBR bearer does not guarantee such traffic.

**Default vs. dedicated bearers**

Orthogonal to being classified as GBR or non-GBR, a bearer is either a default or a dedicated bearer. The default bearer is the bearer that is set up when the terminal attaches to the network. One default bearer exists per terminal IP address, and it is kept for as long as the terminal retains that IP address. The operator can control which packet flows are mapped onto the dedicated bearer, as well as the QoS level of the dedicated bearer through policies that are provisioned into the network policy and charging resource function (PCRF). A dedicated bearer differs from default bearer based on the QoS parameter values of the PDN connection type. Any additional EPS bearer that is established to the same PDN of the default bearer is referred to dedicated bearer.

And policy controller defines them using IP quintuple.

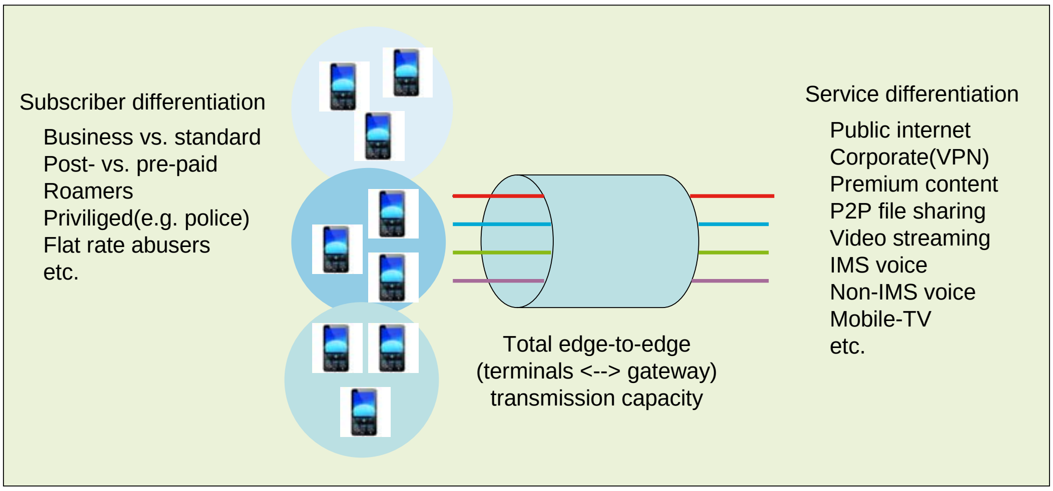


**Figure 4-7:** ***Bearer concept for QoS treatment***

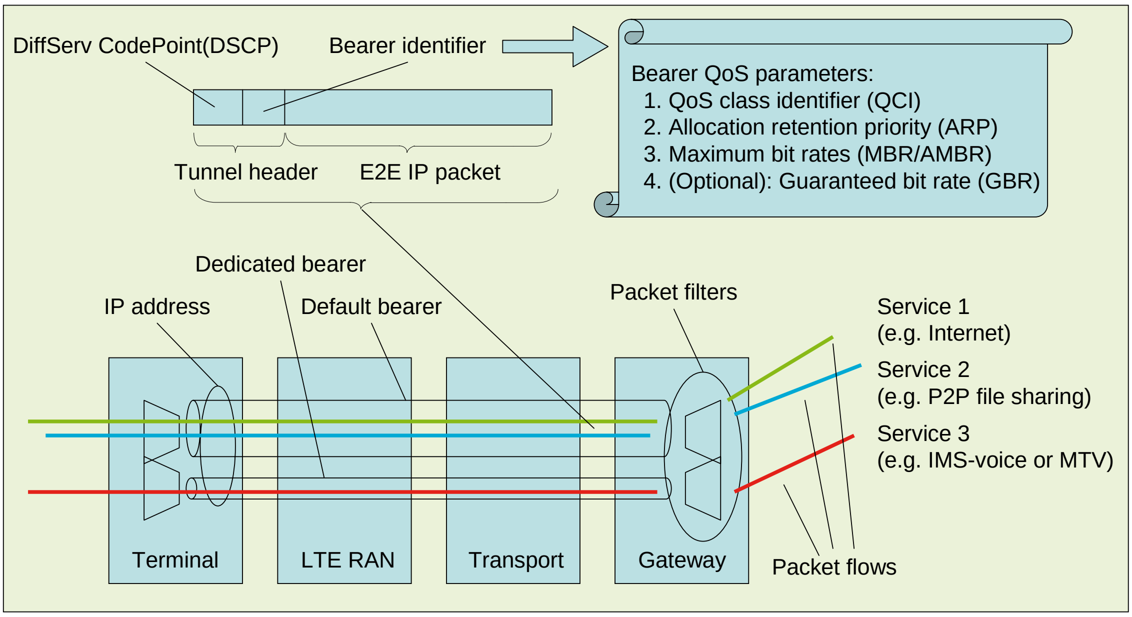


**Figure 4-8:** ***Default and dedicated bearer***

### 4.2.1.3 QoS parameters of the bearer concept

In 3GPP, each EPS bearer is associated with the following bearer-level QoS parameters:

1. QoS class indicators (QCI):

   QCI is a scalar parameter indicating the packet forwarding treatment on the bearer level.

2. Allocation & Retention Priority (ARP):

   ARP is stored in HSS on a APN basis. The ARP gives the control capability to eNodeB for pre-emption when there are insufficient resources to establish a new radio bearer (RB). This depends on many factors such as maximum numbers of UEs and RBs and maximum number of RBs on each GBR (Guaranteed bit rate).

In addition, for each bearer the following bearer level QoS parameters are defined:

1. Guaranteed bit rate (GBR)
2. Max. bit rate (MBR)
3. APN AMBR (Aggregate max. bit rate) for each APN
4. UE AMBR for each UE association

The EPS QoS concept is class-based, where each bearer is assigned one and only one QoS class identifier (QCI) by the network. The QCI is a scalar that is used within the access network as a reference to node-specific parameters that control packet-forwarding treatment (e.g., scheduling weights, admission thresholds, queue management thresholds, link-layer protocol configuration, etc.) and that were preconfigured by the operator owning the node (e.g., the LTE base station).

Whereas the QCI specifies the user-plane treatment that the packets carried on the associated bearer should receive, the **allocation and retention priority (ARP)** specifies the control plane treatment that the bearers receive. More specifically, the ARP enables the EPS system to differentiate the control-plane treatment related to setting up and retaining bearers. That is, the ARP is used to decide whether a bearer establishment or modification request can be accepted or must be rejected due to resource limitations. In addition, the ARP can be used to decide which bearer to release during exceptional resource limitations.

The **maximum bit rate** (**MBR**) and **guaranteed bit rate** (**GBR**) are defined only for GBR bearers. These parameters define the MBR, that is, the bit rate that the traffic on the bearer may not exceed, and the GBR, that is, the bit rate that the network guarantees (e.g., through the use of an admission control function) it can sustain for that bearer.

The main purpose of the **aggregate maximum bit rate** (**AMBR**) is to enable operators to limit the total amount of bit rate consumed by a single subscriber. As such, it is not defined per bearer, but rather per group of non-GBR bearers. This parameter gives operators the tools to offer differentiated subscriptions.

The 3GPP has agreed on defining two different AMBR parameters:

- APN-AMBR: defined per subscriber and APN and known only to the gateway
- Terminal-AMBR: defined per subscriber and know by both the gateway and the radio access network

Both of these AMBR values are defined for an aggregate of non-GBR bearers and are applied separately for uplink (UL) and downlink (DL) direction.

#### 4.2.1.4 Applying QoS parameters and bearers in LTE

3GPP standards in LTE and LTE-Advanced provide an EPS bearer level QoS mechanism with a variety of QoS support and each bearer is associated with QoS parameters such as QCI (QoS Class Indicator), MBR (Maximum Bit Rate), AMBR (Aggregated MBR), GBR (Guaranteed Bit Rate) and ARP (Allocation and Retention Priority). An EPS bearer provides a common QoS or packet forwarding treatment for all packet flows.

Some of the benefits of LTE networks with QoS to network operators are controlled latency and jitter, dedicated bandwidth, improved packet error loss efficiency and priority handling. Guaranteeing QoS requirements at the network level can lead to more efficient resource utilization while maintaining performance for critical applications such as delay-sensitive services (voice, live streaming).

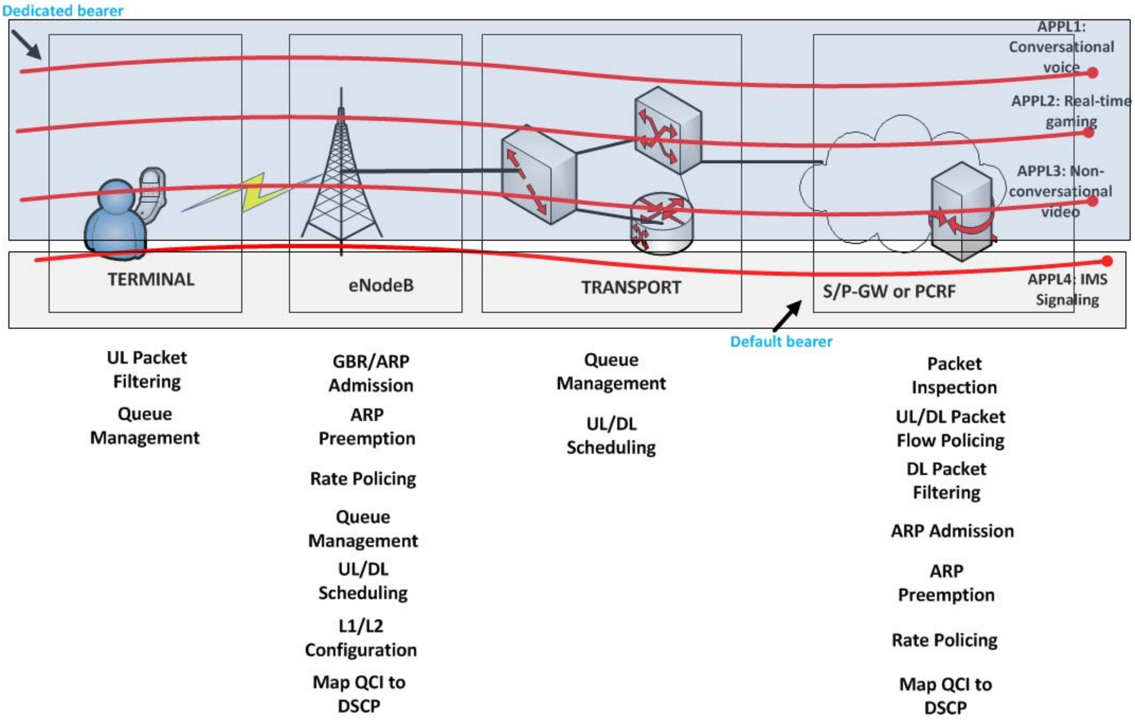


**Figure 4-9: User plane QoS function in EPS**

Figure 4-9 shows the role of each network domain for e2e QoS control in LTE networks. QCIs set the priority and treatment of each traffic type and each type of traffic is associated to specific QCIs. As seen in Figure 4-9, when the packets enter the transport network, each packet will be marked with a specific Diffserv code point (DSCP) value, see also [Hannes] for 3GPP Release 8 version. For DSCP, markings are written to the IP header to implement scheduling, prioritization and classification. Each QCI maps into a unique DSCP value and the mapping should be the same across different LTE equipment such as eNodeB, MME, S-GW and P-GW.

Note also from Figure 4-9 that a dedicated bearer can either be a GBR or a Non-GBR bearer while a default bearer shall be a non-GBR bearer. The LTE QoS "tool box" is very open for use cases, where non-GBR traffic is mapped to dedicated bearers being static or dynamically created, e.g. for application differentiation APPL1 – APPL4 in Figure 4-9.

### 4.2.2 Related QoS standards for IP and other packet networks (MPLS, Ethernet)

For different technologies and transport layers in the aggregation and core EPC network, QoS is supported at various levels being standardized by different bodies, including IETF and IEEE. The following QoS support architectures and mechanisms have been proposed by IETF working groups:

#### 4.2.2.1 QoS Mechanisms in IP networks: IntServ, DiffServ

There are two principal approaches developed by the IETF [IETF] to support QoS in IP networks:

(i) **Integrated services (IntServ)**: In this model, applications use RSVP to request and reserve resources through a network. This is a flow-based mechanism.

- RSVP (Resource Reservation Protocol): This is an end to end bandwidth reservation protocol defined in 1997 by IETF (RFC 2205).
- RSVP-TE (Traffic Engineering Version): This is used to establish traffic engineering via MPLS label switched paths.

(ii) **Differentiated services (DiffServ)**: In this model, DSCP markings are used to implement various queuing techniques on routers and switches to adjust the QoS expectations. This is a class based mechanism. **Per-hop behavior (PHB)** is determined by the DS field of the IPv4 or IPv6 header. The DS field consists of 6 bits DSCP value and 2 bits ECN (Explicit Congestion Notification) parts. The following PHBs have been standardized by the IETF:

- Default PHB: this is best effort (BE) traffic.
- EF (Expedited Forwarding) PHB: This is dedicated to low-latency, low jitter and low error loss traffic. Examples are voice, video and other real-time services. Although there is no guaranteed resource reservation in the DiffServ concept, EF traffic should experience loss and non-neglegible buffering delay only in rare cases.
- AF (Assured Forwarding) PHB: This provides assurance for delivery as long as the traffic is not exceeding the subscribed data rate.
- CS (Class Selector) PHB: This is used for compatibility with non-DiffServ aware routers.

(iii) In addition, **NSIS (Next step in signaling)** as another IETF working group is handling QoS signaling as a development and simplification of RSVP.

### 4.2.3 Challenges of the 3GPP Architecture for ensuring E2E QoS control for LTE

LTE network equipment, such as eNodeB, MME, S-GW and P-GW, will map QCI to DSCP before forwarding traffic in the transport or backhaul network domains. This is important because the Ethernet transport network does not understand the concept of bearers. Instead, DiffServ is commonly used for QoS in transport networks.

In the backhaul or transport network, LTE specifies a DSCP value at the EPC bearer tunnel header so that the network nodes can associate the packet with the correct QoS parameters. However, in the case that QCI to DSCP mapping in the transport domain cannot be perceived accordingly in the packet, the QoS control will be lost in the network. Therefore, it is important to enforce QoS control especially in the backhaul domain and the DSCP mappings are done perfectly at the transport network domain. Therefore, the primary goal in the transport network is to map each QCI into unique DSCP values.

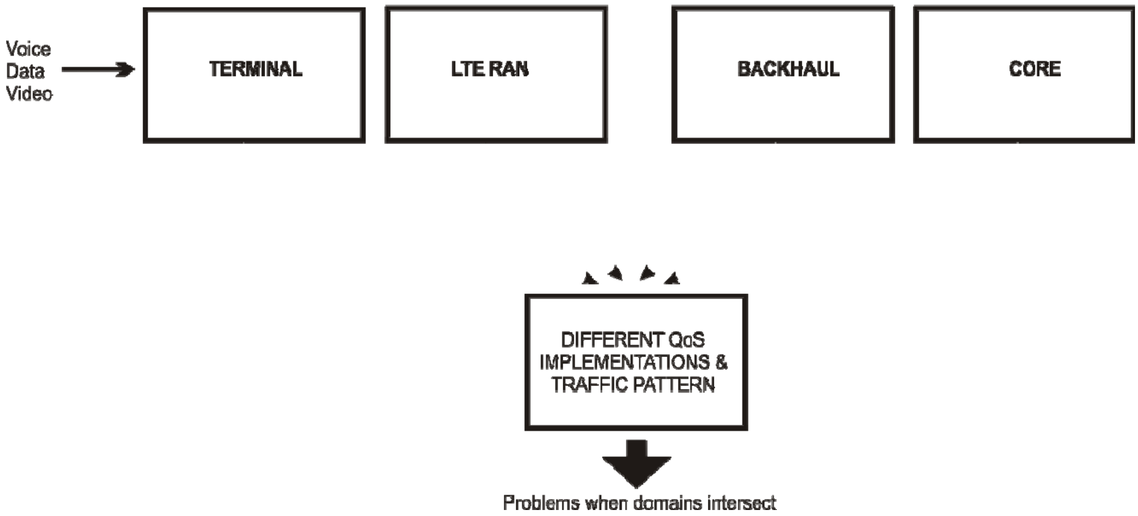

Voice
Data
Video
TERMINAL
LTE RAN
BACKHAUL
CORE
DIFFERENT QoS IMPLEMENTATIONS & TRAFFIC PATTERN
Problems when domains intersect

**Figure 4-10:** ***LTE Network multiple domains***

There are some important challenges for ensuring e2e QoS guarantee in multiple domains in a heterogeneous network environment, see Figure 4-10.

First, there are different service providers and a given IP based QoS marking does not ensure QoS guarantee or a specified service-level agreement (SLA) in the network. LTE defined QCI to DSCP mapping function may fail since the DSCP values and their corresponding scheduling polices are not uniquely defined in all network domains in current IETF standards. Moreover, the VLAN uses 802.1p bit as the QCI whereas MPLS networks use 3 traffic class (TC) bits, see also Table 4.3:.

Second, in a heterogeneous network environment, QoS cannot be implemented on a single protocol layer, or by a particular network component. Moreover, LTE network elements alone cannot guarantee E2E QoS in all network domains. E2E QoS support requires the QoS requests to traverse across different network portions and protocol layers. All network elements including transport network equipment must be taken into account. A bottleneck node due to a particular router or switch can cause network congestion and service degradations, which may cause long delays and outages in LTE networks.

Therefore, the QoS mapping should be done both as cross-layer mapping and cross-domain mapping because the QoS requests traverse across different protocol layers and network portions. The overall performance will depend on the QoS achieved at each layer and domain. Hence, in the whole network, each layer and domain in the network must have its role in QoS control.

The main problem is to align the 3GPP bearer based QoS with the class based QoS defined in transport standardization (IETF, MEF, etc,). A very good example is the QoS use case “subscriber differentiation”. Within 3GPP architecture, the bearer based WFQ (weighted fair queuing) would fulfill required fairness during transport congestion. In this situation, the class based queuing in transport might be unfair especially when e.g., the number of gold users is high and the number of bronze users is low.

Another general problem for the mapping of QCI values to unique DSCP values is that with the EPS convergence of UMTS and other traffic types, a smaller number of DSCP values will be available for this mapping. If the number of QoS classes is not kept low, this will result in the same treatment of different QoS classes in the LTE network. Therefore the mapping has to be done carefully.

### 4.2.4 An example for standardized QCI characteristics

The traffic model in LTE is different from UTMS traffic classes (conversational, streaming, interactive and background) and concrete mappings have not been finalized yet.

Table 4.3: shows an example of the modified version of standardized QCI mapping design from 3GPP Release 11 [3GPP_TS_23.203]. Note that this table shows only one possible example of a mapping. Subscriber and application differentiation could be mapped in a lot of different ways. The last column represents example services specified in 3GPP Release 11. There are 9 QoS classes (QCI 1-9) supported by LTE equipment. Priority, packet delay budget and packet error loss rate are also given in the table.

Bearers with QCI values from 1 to 4 are GBR bearers. GBR bearers are established on demand, because they can block transmission resources. A certain bit rate is guaranteed. Bearers with QCI values from 5 to 8 are non-GBR bearers. For non-GBR bearers, admission control function does not exclusively reserve transmission resources. The LTE concept would allow for dynamically established and released dedicated non-GBR bearers, e.g. for application differentiation on demand, or DPI detection controlled by policies.

In the mobile backhaul network for Table 4.3:, IP/MPLS/Ethernet is used as the transport mechanism to provide reliability and QoS. The 802.1p bit that is used in VLAN is a 3-bit field that is used to define priority (e.g., with 0 the lowest and 7 the highest priority). The TC field in the MPLS network is used for traffic classification and differentiation between traffic types, which allows to have 8 traffic classes which may be assigned from 0 to the lowest, up to 7 to the highest priority.

Two sub-classes can be distinguished from Table 4.3:

- Streaming-RT (Live, Real-Time): e.g., mobile TV, RTP streaming. When congestion occurs, the corresponding portions of some calls are definitely lost, but the call is not dropped.
- Streaming-NRT (Non-Real-Time): e.g., streaming-video (youtube, dailymotion, on demand video) on the web. When congestion occurs, the corresponding portions of calls are delayed.

The QoS level in Table 4.3: represents four service classes namely, platinum, gold, silver and bronze. Note that this Olympic classification is one deployment example. Other operators might reduce to three or only two (business, economy) subscriber classes:

- Platinum class has higher scheduling priority over other classes and the bandwidth is reserved for all control traffic.
- Gold class requires low latency, low jitter and low error real-time traffic.
- Silver and Bronze classes are for low latency traffic. Silver class behaves as close as possible to gold class if the support is available in the network when the congestion due to user applications occurs. Bronze class is a default class for all applications that does not match other classes. The performance is lower than the silver classes.

**Table 4.3:** ***An example for standardized QCI characteristics***

| QCI | Resource Type | Priority | Packet Delay Budget | Packet Error Loss Rate | Bandwidth | QoS Level | ARP | MPLS | Carrier Ethernet | IP | Example Services |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | TC 3 bit | 802.1p 3 bit | DiffServ PHB (DSCP Value) | |
| - - | --- | ---- | ----- | ----- | 5% | Platinum | 1-15 | 6 | 6 | CS7 DSCP56 | Control |
| 1 | GBR | 2 | 100 ms | $10^{-2}$ | 30% | Gold | 1-15 | 6 | 5 | EF DSCP46 | Conversational Voice |
| 2 | | 4 | 150 ms | $10^{-3}$ | | Gold | 1-15 | 4 | 4 | AF43 DSCP38 | Conversational Video LiveStream |
| 3 | | 3 | 50 ms | $10^{-3}$ | | Gold | 1-15 | 5 | 5 | AF42 DSCP36 | RealTime Gaming |
| 4 | | 5 | 300 ms | $10^{-6}$ | | Gold | 1-15 | 3 | 4 | AF41 DSCP34 | Non-Convers. Video (Buffered Streaming) |
| 5 | Non-GBR | 1 | 100 ms | $10^{-6}$ | 10% | Platinum | 1-15 | 7 | 6 | AF21 DSCP18 | IMS Signalling |
| 6 | | 7 | 100 ms | $10^{-3}$ | | Silver | 1-15 | 2 | 3 | AF23 DSCP22 | Voice,Video (Live Stream) Interactive Gaming |
| 7 | | 6 | 300 ms | $10^{-6}$ | 45% | Silver | 1-15 | 3 | 3 | AF22 DSCP20 | Video Buffered Streaming & TCP-based (WWW, E-mail, chat, ftp, p2p file sharing etc.) |
| 8 | | 8 | 300 ms | $10^{-6}$ | 5% | Bronze | 1-15 | 1 | 3 | CS4 DSCP32 | |
| 9 | | 9 | | | 5% | Bronze | 1-15 | 0 | 2 | CS0 BE or Default | |

## 4.2.5 Services, KPIs and implementation steps for E2E QoS parameters in the network

### 4.2.5.1 Exemplary services

**Voice** refers to VoIP bearer traffic only and does not include call-signaling traffic.

**Interactive video** refers to IP video-conferencing and combines delay constraints as for VoIP with high variable bandwidth demand.

**Interactive online gaming** has extreme demands for non-noticable delay.

**Streaming video** is either unicast or multi-cast uni-directional video which can make use of buffering.

**Best effort bulk data** is intended for background, non-interactive traffic flows such as large file transfers, content distribution, database synchronization, back-up operations and email (e.g. file transfers and downloads via ftp, P2P, etc).

**Transactional data** is intended for foreground, user interactive applications such as database access, transaction services, IM and preferred data services.

**Call-signaling class** is intended for voice and video signaling traffic such as Skinny, SIP, H.323, etc.

**Network management** is intended for network management protocols, such as SNMP, Syslog, DNS, etc.

**Routing data exchange** is intended for IP routing protocols such as BGP (Border Gateway Protocol), OSPF (Open Shortest Path First), etc.

#### 4.2.5.2 Common KPI parameters for different applications

**Data applications**: Typically best-effort services, characterized by variable bit rates, being tolerant to some loss and latency before the user perceives poor quality

- Transaction latency (including time-to-first-byte and time-to-last-byte of data)
- UL/DL throughput of data [Mb/s] or measured as transactions per second
- Concurrent transactions
- Loss rate, Re-transmissions - TCP retries
- Connection latencies and failures
- Service availability (as a percentage, e.g. demand for the five nines 99.999%)

**Conversational applications**: Real-time services requiring a constant bit rate. Voice services are sensitive to latency and jitter, but tolerate of some packet loss.

- Maximum and average jitter
- Delay bounds
- Mean opinion score (MOS)

#### 4.2.5.3 Steps for determining QoS parameters for a given network

**Step 1**

Simulate each UE with applications such as example services mentioned above and associate them to specific QCIs (bearers) valued from 1 to 9. Distinguish subscribers (Platinum, Gold, Silver, Bronze or emergency) and data rate values (5Mbps, 20Mbps, 100Mbps, etc.), and map them to QCIs (in HSS subscriber's profile). Use also MBR to create different levels of services (some exemplary services are also given above) in HSS subscriber's profile. In addition, the application differentiation concept might be configured.

**Step 2**

For validation of DiffServ prioritization in the IP transport network, test QCI-to-DSCP mapping. The eNodeB on one side of the LTE network and the PDN-GW on the other side must map QCIs to DSCP. This mapping between the bearer-aware networks and the transport network (DSCP/802.1p/EXP mappings) must be tested for accuracy.

**Step 3**

Create congestion at radio and transport networks with varying traffic profiles and distinguish between them by inserting heavy load into the network. Measure and report behavior and KPIs from IP and Ethernet transport equipment (IETF, MEF) and from 3GPP network elements. The simulation can also be implemented with traffic simulators such as load generators [DevoTeam]. It's only when the network becomes congested and there is competition for resources that we find out if QoS, policy and prioritization are working properly.

**Step 4**

Measure and report KPIs for each QCI (bearer). The network operator defines the KPI expectations associated to each specific bearer. Most common KPIs are also given above for data and voice traffics. The carrier measures and reports KPI's for each QCI over time and with different traffic and subscriber mixes to observe whether they are within the defined threshold limit.

### 4.2.6 Challenges solution approaches for inter-provider end-to-end QoS

As discussed previously in this section, mobile and wireless network operators have to support quality of service within their platform to meet KPI parameters according to SLAs with other parties and service portfolio. This can be achieved by provisioning of enough bandwidth and coverage and by introducing a QoS policy based on differentiation and/or guaranteed resources for the customers and their applications.

Moreover, end-to-end transport paths of most IP applications are traversing several network platforms beyond LTE access and EPC involving different administration regions and heterogeneous technologies. Enforcing end-to-end QoS in a global Internet environment is an important and challenging task, which mainly depends on interconnection agreements and standardization of QoS support between multiple carriers. Service level agreements (SLA) between service, network and content providers as well as users with special QoS demands can help to set up interconnections with specified QoS properties and responsibilities as illustrated in Figure 4-11, but at the current stage have not been deployed.

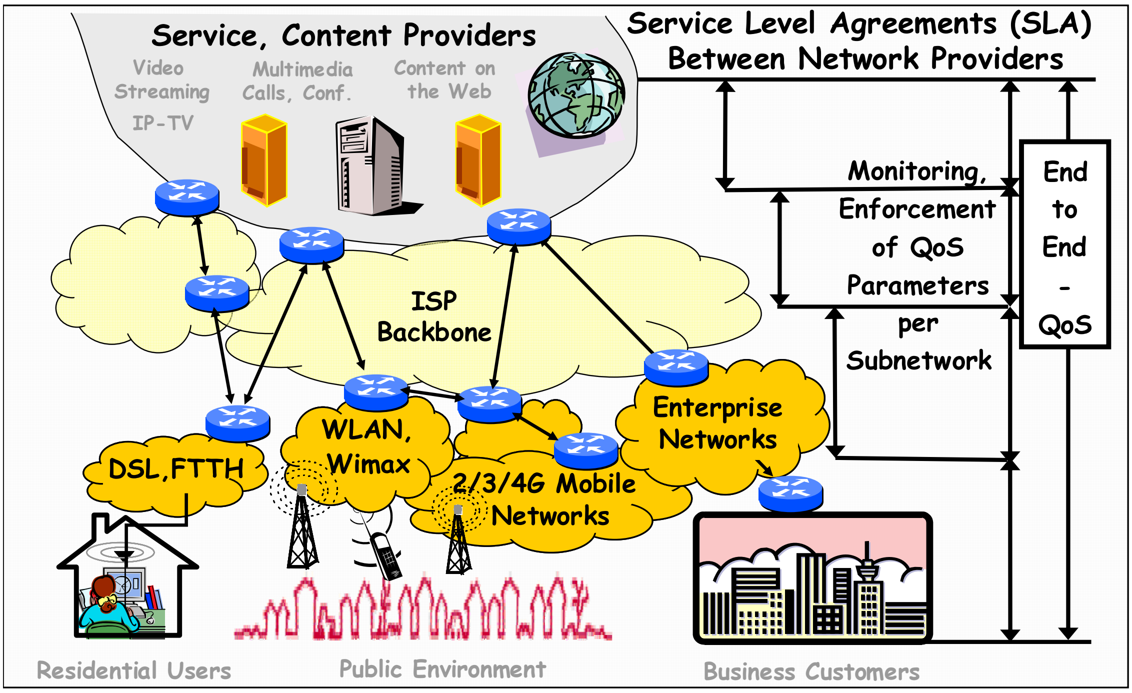


**Figure 4-11:** ***End-to-end QoS for IP service***

Main KPI parameters listed in section 4.2.5.2 have to be negotiated and determined in SLAs of Internet access, user applications, as well as interconnection or transit for network providers.

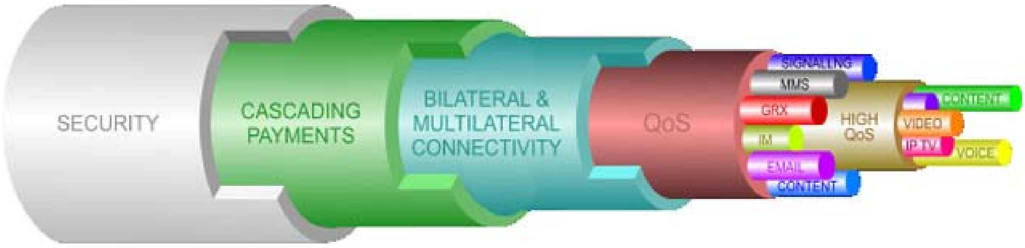
The focus of the involved parties is different, where business customers often demand for extremely high network availability and reliability. Residential users are sensitive to throughput and delay, whereas network providers are concerned about resource efficient service provisioning with regard to the demands of all user and application classes.

Measurement and monitoring of the QoS performance at servers, interconnection points and/or between ingress and egress of a network is crucial for SLA control. On a single transport path, end-to-end QoS parameters usually can only be guaranteed based on worst case assumptions on the resources, i.e. end-to-end throughput is the minimum that is provided on the interfaces and links on the path, and end-to-end delay is the sum of delays on the sections of a path including (de-)coding in terminal equipment etc. Awareness of the current QoS performance is a basis for fast signaling and handling of problems and failure cases, where automated routing and traffic engineering methods involving backup resources are important as a first aid, together with repair and upgrades of resources for long term treatment of bottlenecks.

In addition to single end-to-end paths, content delivery and peer-to-peer networks offer multiple sources for the same content in a distributed architecture and multipath TCP approaches can be used for distributed traffic engineering from the user's devices. Both developments are expected to improve QoS beyond the capabilities of single end-to-end connections, but do not lead to a coordinated or optimized resource usage on interconnected IP core and access network platforms.

QoS support between network and content providers has been addressed by several standardization bodies and research projects. On the IP layer, the differentiated services concept in principle can be applied in different interconnected domains. Therefore the network operators have to agree on similar QoS policies and classification schemes or a set of QoS classes has to be reserved for usage according to a common standardized strategy to be supported by each operator. But neither of both options seems to be commonly used.

Standardization approaches for QoS interworking have been elaborated by the IPsphere forum, which started as a stand alone industry forum and has been integrated in the Tele-Management (TM-)Forum in 2010. In addition, the IPeXchange initiative of the GSM association <www.gsmworld.com/our-work/programmes-and-initiatives/ip-networking/ip_exchange.htm> has specified an interconnect service to be offered by carriers on a competitive basis in a managed network at specific quality levels with traffic engineering. It is based on agreed technical specifications and consistent commercial models. It provides a range of technical and commercial features that enable business models including security features and bi- and multilateral connectivity. Main IPeXchange features are illustrated in **Figure 4-12**.



…

**Figure 4-12:** ***IP eXchange interconnection service framework***

- Secure environment – the IPX is a transparent IP network that is not addressable from the Internet
- Flexible IP Service interconnection – bilateral and multilateral connectivity – one contract, many connections
- Cascading payments – managing the flow of information necessary for settlements to be made between operators – rewarding all players who meet their mutual obligations in the value chain
- Premium quality environment – traffic managed with QoS levels and performance to SLAs that are mutually agreed

Nevertheless, those approaches for QoS support throughout interconnected platforms do not yet seem to find widespread adoption. Other connection oriented approaches e.g. IMS are also under discussion as too complex to be integrated with predominant best effort type data transport in IP platforms. Therefore it remains open how QoS interconnection will be supported in future converged fixed/mobile IP networks. Currently the EU FP7 ETICS project <www.ict-etics.eu> aims at creating a new ecosystem of innovative QoS-enabled interconnection models between Internet service providers allowing for a fair distribution of revenue shares among all the actors of the service delivery value-chain.

Last but not the least, the IETF has established working groups aiming at optimized shortened transport paths for content distribution [IETF]. The ALTO (Application Layer Traffic Optimization) working group is introducing servers with awareness of location of nodes and IP addresses which enable to select sources for downloads in the near of the destination, thus improving long transport paths in peer-to-peer and other globally distributed content delivery schemes. The CDNI (CDN interconnection) working group is focused on improving the interworking between different CDN infrastructures which again aim at serving users from a close by CDN node. Both IETF working group approaches have potential to improve the end-to-end delay as a main QoS property. QoS support in interconnection is relevant on a microscopic as well as a macroscopic level since traffic on network gateways is often exchanged in aggregates and/or VPNs.

End-to-end QoS depends on the performance parameters of the transmission systems including the terminal equipment. Voice and video coding and decoding has a main impact on the perceived quality, e.g. by introducing redundancy to enable for error compensation and thus higher error and loss tolerance on account of higher bandwidth demand. Finally, the user experience (quality of experience, QoE) is the decisive measure for the service quality. It can be determined by subjective estimation e.g. on a mean opinion score (MOS) scale and is partly also assessed by automated methods, e.g. for speech transmission quality as defined by the ITU based on a model for the user sensitivity. User’s sensitivity of quality is often based on logarithmic dependencies, e.g. quality ratings are linear improving for doubling bandwidth or halving delays. A memory effect and other individual factors are also relevant, since users who are accustomed to bad quality are more tolerant and easier to satisfy.

## Streaming optimizations

IMS (IP multimedia subsystem) has been initially defined with 3GPP Rel. 4 and specifies an architecture for managing IP based services within a service domain. Originally defined for voice over IP telephony, other session based multimedia services can be integrated in IMS as well.

In practice many multimedia application for mobile devices are not controlled by IMS. On one side there are operator managed streaming services, which are not using SIP based session control, e.g. RTP-based streaming. On the other hand there are unmanaged over-the top (OTT) applications, which don’t have access to any service control functions within the mobile operator network. The OTT streaming services are primarily HTTP-based. In these cases applications can’t be controlled by the policy and charging rules function - PCRF (via Rx interface) to influence the management of resources associated with these multimedia streams. However excellent user experience for multimedia applications is only possible, if the resources can be managed according to the actual needs of the applications and user profile information. A mobile user may not be aware of resource requirements and also can’t influence the resources to be allocated for a dedicated service if such bearer setup is initiated by an application function, which is situated in the network. Hence PCRF needs to make a best guess about the required resources and usually will apply a default reservation based on user profile information. This approach probably leads to inefficient resource usage and / or suboptimal user experience for the streaming video.

Another important issue is the lack of fast-reactive adaptation for streaming services to changed conditions in the network. One such event is the occurrence of congestion in a 3GPP radio cell. Considering reduced throughput as an indication for congestion is not reliable because many multimedia streams are coded with variable bitrate. Therefore a traffic monitoring entity can't distinguish between congestion and bandwidth fluctuation within short time frame. Once such event has been detected, the associated applications most likely suffer from impaired performance already for some time.

Currently there has been no information exchange specified between media server without IMS control and PLMN network in case of required rate adaptation (especially if a lower GBR could be allocated) and the flow handling in case of handover is undefined. Also for IMS controlled applications there is no possibility to consider load congestion for early adaptation.

Another problem is that currently QoS support for multimedia streaming on mobile devices is insufficient. While this report was created (Feb. 2012) only non-GBR bearer is supported by deployed mobile networks besides GBR for VoIP. Furthermore, due to the variance in the data rate of streaming services in particular, the allocation of GBR bearers with peak-data-rate set to guaranteed data rate would result in a considerably reduced number of supported GBR connections. Since this type of bearer uses resources in shared manner, high load in the network may cause impairment of user experience for delay and loss sensitive applications.

In the following two new components are introduced – a resource aware entity and a media aware entity for streaming services to provide an enhanced QoE for the user and enabling an optimized usage of network resources. Figure 8 shows the new functional components in addition to the 3GPP based core network architecture.

The Streaming Traffic Management Entity (STME), enables traffic control within the 3GPP mobile packet network for multimedia streaming applications, see Figure 4-13. It is responsible for the control and coordination of actions impacting the QoE of one or several sessions. The Media Aware Streaming Entity (MASE), which is located in the data path of the multimedia applications, provides control of the media flow and proxy functionality for the application layer protocols.

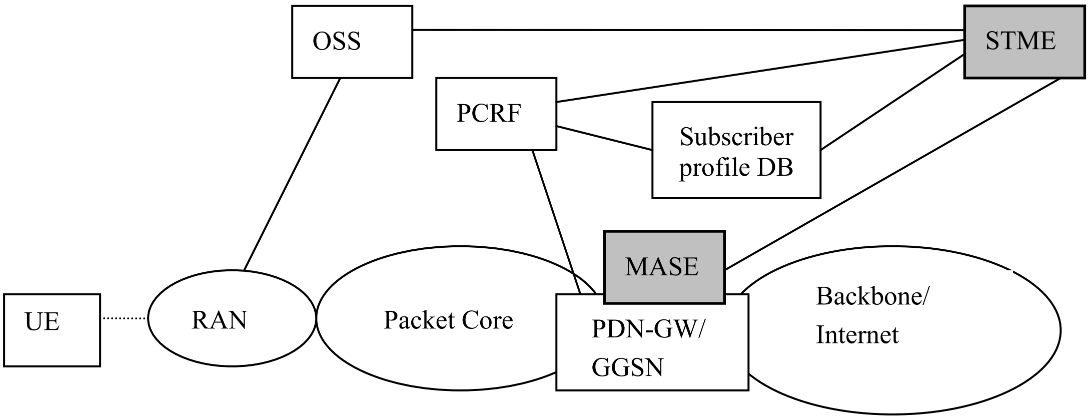


**Figure 4-13:** ***Extended 3GPP architecture with support for streaming optimizations***

### 4.2.7 Streaming Traffic Management Entity (STME)

STME can interact with the PLMN policy control function (e.g. PCRF) in order to provide initial and updated media flow information for the control of the network resources on behalf of the multimedia application. This information is used like described within state of the art to initiate or modify bearer resource allocation. For the information exchange between STME and PCRF, the Rx can be used.

STME has an interface to the operations support system (OSS). The OSS informs the STME about received measurement reports and triggers concerning congestion detection in the radio access and/or changed cell load conditions.

STME may decide to change stream specific properties of a single or multiple flows at a time based on received trigger from OSS by sending a request to the responsible MASE(s). Changing stream specific properties and bearer resource allocation (via PCRF) for a stream has to be jointly and consistently triggered by STME.

### 4.2.8 Media Aware Streaming Entity (MASE)

The MASE implements protocol specific recognition mechanism to extract media flow properties. This information is forwarded to STME in order to initiate or update resource allocation for the bearer via PCRF. For example MASE can extract session description information (such as SDP) from multimedia applications' control messages, thus MASE needs to include proxy support for RTSP protocol to recognize relevant session control messages. Another example of non session based multimedia is HTTP streaming. MASE can identify description for streaming within an HTTP object (manifest file), based on the URL and interpret the syntax of the transferred information.

MASE can analyse performance based on higher layer protocol information. One example is the reception of RTCP report from UE. Another example is the analysis of TCP parameters such as window size for HTTP streaming applications. The MASE informs the STME if the QoS of a streaming session is impacted (e.g. reduced throughput) besides the OSS.

The STME will decide and trigger the appropriate action in MASE. Dependent on the request from STME MASE may forward the information to the streaming server (server performs rate adaptation) or MASE may provide adaptation itself e.g. using transcoding, selecting suitable layer in case of SVC encoded videos, etc..

In case of changed conditions reported by OSS, STME may need to prioritize between media flows in the considered radio cell, therefore STME may request from OSS the identities of all subcribers located in a reported radio cell. Congestion may require to reduce bitrate for one or more media flows, additional resources may allow to increase bitrate for one or more flows. Prioritization can be based on user profile information and application / service specific information for the information received by OSS. Information about user subscription and type of connected user device can be derived from a subscriber profile database by STME. Application / service specific information can be delivered to STME by MASE based on location information (e.g. IP address, URL etc.) and protocol specific information.

## 4.3 QoS support for external content

CDN mechanisms are widely used for the distribution of multimedia content. A commonly used concept is to shift content close to the user for acceptable QoE (quality of experience) – this works fine for many applications and many networks. However some content may have to come from another / remote network, e.g. for the following reasons:

- 3rd party content is not locally available; e.g. a live IP-TV stream injected in a different region. It is assumed that not all network operators have agreements with all the major TV / broadcast companies in order to ensure localized stream injection
- CSP may not have all the infrastructure deployed for local delivery; content might be provided from a central location for different platforms e.g. mobile / fixed access, IP-TV delivery. In that way a multi-national telco may serve content to national subsidiaries

The full path between user and streaming server stretches beyond several domains and the resources therefore can't be controlled by a single entity. However value-added (paid) content needs some reliable delivery, especially for delay sensitive content and applications (e.g. live streaming or video conferencing). Since the delivery may span multiple domains, there are several solution options how to enable QoS support under these circumstances:

- Operator agreements for adjacent domains
- Inter-domain signaling for QoS control
- Resource coordination on application layer

These mechanisms are displayed in Figure 4-14 and the potentials are further discussed below [SEH+10]:

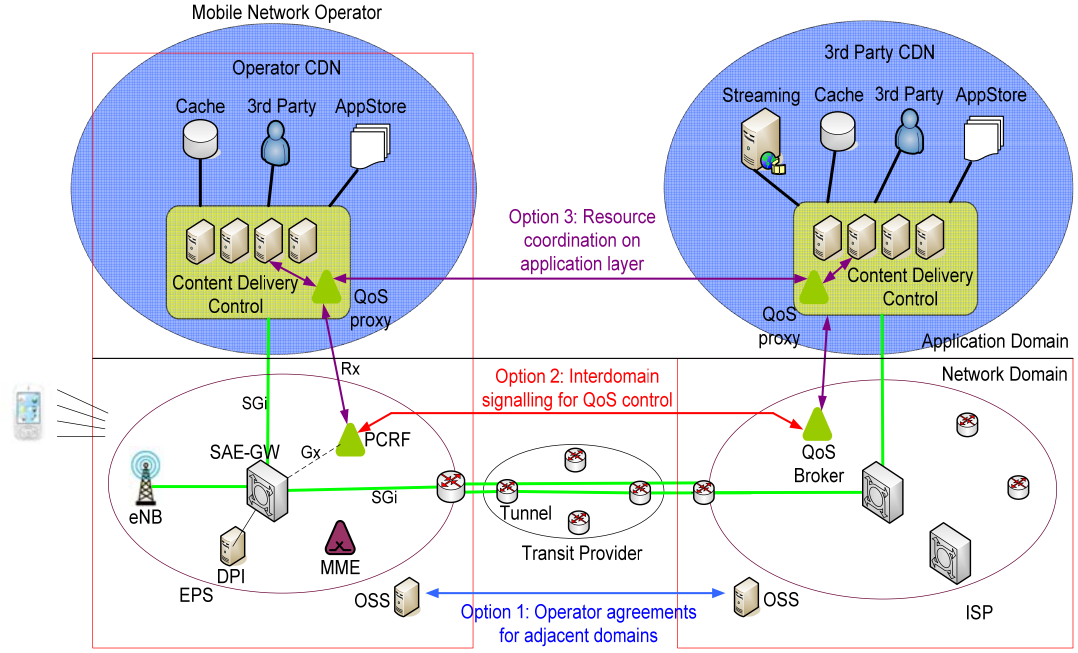


**Figure 4-14:** ***Different solution options for QoS support for external content***

Within a mobile network there is a consuming peer, which requests content from a different site. It is assumed that caching of content is not feasible for the reasons described above.

Option 1 is based on operator agreements between the mobile network and the network domain, which hosts the requested content. These agreements could be used to accomplish QoS reservation. The operations support system (OSS) keeps information about QoS configurations for connections with certain domains. Then OSS e.g. may trigger the policy control function (PCRF) to reserve the necessary resources. However such agreements have the drawback that they are static and therefore the mechanisms can’t react spontaneously to the actual resource situation in the network. In addition this approach is not applicable for most of the OTT content. Usually operator agreements are restricted to a few networks, which are directly connected with the local domain of the mobile operator.

Another possibility (option 2) is based on inter-domain signalling for QoS control: This mechanism implies to exchange control information between different domains in order to control resource allocation along a network path between the domain of the MNO and the network, which hosts the content. One major concern about this approach is that usually operators are not willing to share control information beyond the boundaries of their administrative domains and also transfer part of the control over resources to other entities. This concept has not been accepted also due to the significant impact expected on the EPC architecture. The PCRF would have to be extended with QoS broker functions. Such mechanism has never been considered for PCRF and there is no standardized way for a PCRF to communicate with an external QoS broker and have the required authentication mechanisms in place.

Option 3 includes resource coordination on application layer. Following this approach the impact on the EPC architecture should be limited as much as possible. Bandwidth broker functionality then has to be implemented on application layer, i.e. within the CDN overlay between local and remote CDN components. For the proposed solution it is assumed that transit networks interconnecting the two considered domains offer sufficient resources. The feature set for QoS negotiation would have to be supported on base of a common platform used by all CDN networks in the overlay or based on a standardized interface (horizontal interface of option 3 in Figure 4-14). The lack of open, standardized interfaces for exchange of control information among distributed CDN architecture is now addressed by the IETF working group CDNi. The QoS reservation in the concept proposed here has to be accomplished individually within the local MNO domain and the remote domain, where the requested content is attached to. For this concept a QoS proxy is introduced within each of the considered domains, which implements the QoS broker function and exchanges control information with the policy decision function and the local CDN control infrastructure. Within EPC, the existing interface Rx for exchange of control information between an application function and the PCRF could be used for the proposed concept.

## 4.4 Selective admission control

LTE Rel. 9 introduces the definition of the Home (e)NodeB, allowing the deployment of the femtocells in the indoor environments, e.g., home or small office. The main objectives behind the deployment of femtocells include provision of a high quality broadband connectivity in the indoor environments, fixed mobile convergence, and traffic offloading, to name the few. Apart from all the advantages, the deployment of femtocells will pose some important technical challenges, such as, resource allocation, timing/synchronisation, interference management, mobility management, etc.

The foreseen high QoS requirements together with the fact that deployment of the femtocells will be unmanaged stem the need to look for appropriate resource allocation mechanisms, with one of the crucial aspects being admission control. In such a new landscape the admission control mechanisms traditionally used in the mobile networks seem insufficient to guarantee the desired QoS for the users located in the unmanaged femtocells. Simple capacity overprovisioning may fall short also, taking into account that the increasing diversity of data, voice and high-quality video (e.g., IP-TV services) traffic, e.g., see [Menth08], that may cumulate at the femtocells. This raises the question whether the existing admission control proposals that were originally devised for the Internet might be also applicable in this context.

A good overview and classification of the existing admission control schemes is presented in [Menth10], [Lima07] and by the references therein. Recent trends in admission control demonstrate the need of IP flow analysis to allow dynamic network reconfiguration [Kashihara10], as well as claim the need for distributed admission control solutions, such as the one proposed in [Sakellari10]. On the other hand, the current Internet traffic consists mostly of TCP flows, which are connection-oriented in nature and elastic in their resource requirements (i.e., have strictly concave utility functions). It can be argued that for such flows admission control is barely needed as TCP is adaptive and thus controlling the number of flows sharing the given capacity is not necessary. However, in order to provide the QoS guarantees, e.g., to preserve an acceptable throughput per given flow (that might be the case for the femtocell scenario), it is inevitable to control the admission. The most common approach used so far in context of elastic flows is the measurement-based admission control (MBAC). Usually MBAC approaches estimate the current load, available bandwidth, packet loss probability, etc., and compare that result to the predefined threshold. Due to the large number of TCP flow arrivals and their possible short duration, use of any kind of explicit signalling to control the individual flows seems not to be feasible for the elastic traffic. Without such a per-flow signalling the admission decision on a new flow is implicit. The implicit admission control is then performed by discarding the initial flow packets, as it is enough to inform the source about the rejection decision. An example of such an implicit admission control algorithm has been presented in [Mortier00].

The idea of implicit admission control may be extended for the inelastic traffic, i.e., most of the media traffic, as long as the beginning of the session can be identified and rejected, and thus will form the base for the proposed admission control mechanism. Secondly, the proposed admission control mechanism will aim at fulfilling the paradigm known from the PSTN, admitted user equals satisfied user, and more precisely will provide the QoS support in terms of low packet loss and delay at a guaranteed throughput only for the already admitted users. Such a design allows the network to adjust dynamically to the load without degrading the performance of the already admitted users. That seems to be the appropriate approach to support the increasing volume of heterogeneous traffic that is foreseen for the femtocells.

Foreseen solution (selective admission control algorithm) will provide two modifications to the existent architecture. Firstly, the decision-taking part of the admission algorithm will be located at the end-hosts. There the admission-related measurements will be performed only on the elastic traffic thus requiring modification of the existent TCP stack. Once the measurement is done, packets of the corresponding flow will be marked (e.g., ECN field of the IP header) according to the binary scheme (more states can be introduced, if necessary): provided service is/is not sufficient for the tested flow - new flows may/may not be admitted. This indication (marked packets) will be received at any of the intermediate routers, and there if router is capable of recognizing provided marking (second of the necessary modifications for the provided algorithm) the admission will be executed on any of the incoming new flows (for both elastic and inelastic traffic).

# 5 Improved resource selection & caching

## 5.1 Resource selection

Resource selection is one of the key functional mechanisms to select best suitable resource in P2P and CDN networks. Objectives and some initial concepts are presented in this section.

### 5.1.1 Objectives for resource selection

Requested or selected content could be located on / in

- Operator CDN / cache
- Other localized resource, either end user system or 3rd party provider hosted locally:
- External network (content from CDN, other content provider, end system)

In the following, objectives of resource selection are presented in order to understand why a prioritization between alternative resources could be advantageous.

- React dynamically to changed conditions: Dynamic reaction enables fast adaptation and reduces the period of unfavorable resource usage. A favorable resource may become unfavorable after some time. At first it is required to detect this circumstance.
- Enable configuration of selection policy: Selection of a suitable resource might be rather complex, influenced by many different aspects. For efficient handling it is suggested to have policies in place, which can influence the selection process. Hence it is possible to reduce the operational expenditures (OPEX) for the operator. Dynamic change of selection principles requires configured policies for the consideration of service provider preferences.
- Re-direction to preferred location: A resource should be available at some preferred location for the operator, either within the local domain of within another domain, which is preferred by the operator. A preferred external domain could be characterized by an established service level agreement e.g. with a favorable pricing model (implying lower or no interconnection costs) and / or improved QoS support. If a requesting host (either triggered by the user or the application) sends a request to an arbitrary external resource, the request might be re-directed to a preferred location from service provider perspective. In this sense less data needs to be transferred from an external, un-preferred location. This could save interconnection costs. In addition QoE for the user might be improved.
- Constrain load on network resources – There are several aspects associated with this objective. A resource available within the local network is not used because load in the mobile network should be reduced. Request re-direction should be considered only, if underlying routing mechanism cannot solve the problem. Alternative resource in external network is reached either via alternative access (e.g. non 3GPP access) or local gateway node (e.g. examined within 3GPP SIPTO - Selective IP traffic offload). In the latter case the load is reduced from mobile core and backhaul. In the other situation resources from radio access are saved as well. In that sense also a request from an external consumer to a local resource could be rejected. In addition this would save interconnection costs. If requesting a resource from the home network either an entity in the home or in the visited network may decide to re-direct the request. In a specific case upstream load (for potential bottlenecks such as radio link and mobile backhaul) should be constrained. Content could be hosted on a user device, e.g. a user terminal attached to a cellular network is running a P2P application. Content associated with this application could be frequently requested by other nodes – this puts high load on the radio link. The request could be redirected to other local or external resources. Local or external nodes could initiate the requests. For external requests it should be assumed that re-direction to an external resource is the preferred operator strategy in most cases.
- Constrain load on content resources – A content resource is available within the local domain but it already serves a large number of consumers and needs to be protected (at least temporarily). Alternative candidates might be available within the local domain or otherwise in external networks. It shall be assumed that the general case is a request by a local host. But also an external requestor might be considered, e.g. a roaming user, who wants to get access to a resource in the home network domain.
- Consider impact of events related to user hosted content – e.g. a host attaches, detaches or changes access. This case usually applies to localized requests. If such host suddenly disappears (e.g. without prior notification) before transfer of content has been completed, it

needs to be determined whether only part of the requested content needs to be transferred from a different location. A newly connected resource or better connected resource could also trigger re-connection.

- Consider impact of requesting nodes changing point of attachment (access or network) – At different access or in roaming network the favorable resource might not be the preferred one any more. Selecting a new resource might be useful, if an alternative resource for example is located in proximity of the new access. In one situation the consuming node changes access while connected to a resource. It might be the case that a resource is switched, if this is supported.
- Influence selection in external network: If the selected resource is locally not available and it is not worth or possible to cache this content, it could be selected in an external network (e.g. CDN). By influencing the decision, content delivery via external network might be improved. In an analysis it has been concluded that content in external CDN is not always selected according to shortest RTT but aspects like load balancing may have more relevance [TFR+11].

### 5.1.2 Classification of re-direction mechanisms

There exists a large variety of re-direction mechanisms. In this subsection an overview is provided together with an assessment about potential deployment in a mobile operator network.

#### 5.1.2.1 Transparent re-direction

Transparent selection is based on re-direction mechanisms without involvement of the user or the client node. Re-direction might be implemented by a providing peer / server network, an application proxy or any other entity in the network. These entities control access to resources or provide information about resource location to applications (e.g. a P2P BitTorrent tracker) or entities in the network, which resolve the location of the requested content to the IP address of a specific node. In this situation interaction with the user or the application client is not required. On application layer there is a mechanism based on HTTP, which is called Auto Redirect with HTML. Auto redirect is a technique that automatically forwards user request to another website or webpage. Auto redirects are used for the following reasons:

- If a website or some web pages have been moved to another location
- To show advertisements
- To increase the page rank of a particular page in the website that is highly optimised

The above mentioned mechanisms may be accomplished in a more flexible, powerful way by separating between HTTP frontend and backend servers. HTTP frontends intercept HTTP requests and can dynamically select an appropriate backend server. The URL is changed (‘URL re-writing”) for the internal re-direction. There is no TCP connection between the originator of the request and the nodes, which are serving the content – every exchange of HTTP message is done via the frontend / proxy without the client having any knowledge about backserver infrastructure. One example based on this server load balancing mechanism is the “mod_rewrite” module for Apache web server [APA].

Transparent higher layer mechanisms rely on application proxies to control selection of the resource, otherwise control of re-direction is not possible for application clients connecting to an external application server. For the user client attached to a mobile network, a common principle is to obtain knowledge of application proxies via detection (e.g. WPAD for HTTP proxy) or auto-configuration mechanisms (e.g. P-CSCF via DHCP or PDP context setup).

Network layer mechanisms are based on manipulation of address (locator) information, i.e. the requested destination address is replaced with a different address. There are various mechanisms to accomplish change of IP address for the requested resource, such as,

- Substitution of naming information – e.g. based on DNS
- Change of address translation information (e.g.at NAT)
- Anycast mechanisms: Many hosts serving the same content get the same IP address. A client request is routed to the topologically closest server from the virtual cluster. Usually the mechanism is implemented on content replicas in different IP network domains. Each server advertises a route to the virtual address via routing protocol, such as BGP.
- Policy based routing (PBR): Usually when a router receives a packet, the destination address usually determines forwarding behavior. With PBR other aspects may be used to determine the routing decision based on information that can be derived from the packet in consideration.

- Route maps for CISCO routers: Interact with dynamic routing protocols but also used to circumvent routing tables
- Web cache communication protocol (WCCP): Not as such called a PBR mechanism but it has a lot of similarities. A protocol developed for communication between CISCO routers and Web caches. Based on a redirection hash table in the router, incoming packets matching a specified hash value are re-directed to a Web cache.

#### 5.1.2.2 Non-transparent re-direction

Non-transparent selection enables interaction with the user, which is advantageous for applications where the user should have the possibility to influence the selection process. However in many cases the user may not be aware of the impact when selecting a specific resource. So it is assumed that such interaction should be relevant only where knowledge about user preferences is relevant and is not available by some configurations in the network or in the user device. In principle the interaction requires a specific user-network interface, which may not be feasible in many cases without significant changes for the implementation. If a resource is associated with content specific attributes and this information can be provided to the user or application client, non-transparent selection can be favorable. Otherwise transparent mechanisms shall be considered. Examples of such information can be:

- Change resolution for video streaming (e.g. HD, SD)
- Charging information (e.g. free preview vs. full content version of a video)

In the following, some examples are given for interaction with an application. The interaction may occur with the consuming peer / client node or with the user.

- Navigation hyperlinks [RFC 3040] displayed in a HTML page: The user is requested to select a preferred link.
- HTTP client side re-direct: The HTTP server, which is contacted by a browser, can return status code 3XX to the HTTP client together with the location header (RFC 2616). This provides information for the client to temporarily or permanently send this request to a different server with the new URI specified in the location header. All future requests will be sent by the HTTP browser to the original address. So it can be ensured that the original server is still queried for future requests.
- Stream re-direction: A web server link does not cause immediate play of media file but instead points to a small (type XML) metafile, which is transferred to the web client and passed to the media player on the client. The file includes the address and filename of the content. In such a way a RTSP request could be send to the server specified in the metafile.
- SIP Re-invite sent by the server: the SIP server (where the resource is located) sends an invite message with the address of the new server and possibly changed parameters values for the session (in the SDP body of the message). This implies to transfer a running SIP session.
- Resource information service: IETF WG ALTO proposes a tracker to be used by P2P applications to reduce the chance of unfavorable resource selection, i.e. in this case there is an interface between the UE and a P4P system in the network. It provides information about preferred location of content (from operator perspective).
- Walled garden applications: Software is installed on a client to make use of walled garden applications. The interface as part of the walled garden application could be used to select a specific resource.

#### 5.1.2.3 Assessment of re-direction mechanisms

After having presented various approaches for resource selection, a comparison should elucidate the possibilities and constraints when deploying these mechanisms especially in EPS networks. Therefore some criteria for the assessment of different resource selection mechanisms are explained below.

- Impact on user device: If the mechanism needs to interact with the user device, a software update on the terminal is needed. Such approach usually has limited applicability especially for OTT content, which is mostly accessed via HTTP.
- Time period to adapt selection decision: This aspect decides how fast usually decisions can be changed to satisfy new requests in an improved way.
- Focus applications: The question is whether there is a single application which should be controlled for request re-direction or multiple. Hence an important question for the operator is what applications contribute to the majority of traffic in the mobile access domain.

- Business model of the operator: Re-direction mechanisms may also be impacted by the business, i.e. whether there is a single access domain or international / globally present MNO. From technical perspective the number of name space, IP address spaces which are in the control of the operator could be relevant
- Granularity of resource information: Fine granularity implies that a small number of content resources can be selected with a single information element in a resource data base. For example HTTP re-direction mechanisms can be associated with fine granularity because the URL, which is the location identifier, corresponds to a single or small number of content resources on a Web page. On the other side a DNS domain name may correspond to a large number of content resources hosted within an administrative domain. The resource granularity factor determines number of content resources associated with a single location identifier. Figure 5-1 summarizes the various selection mechanisms.

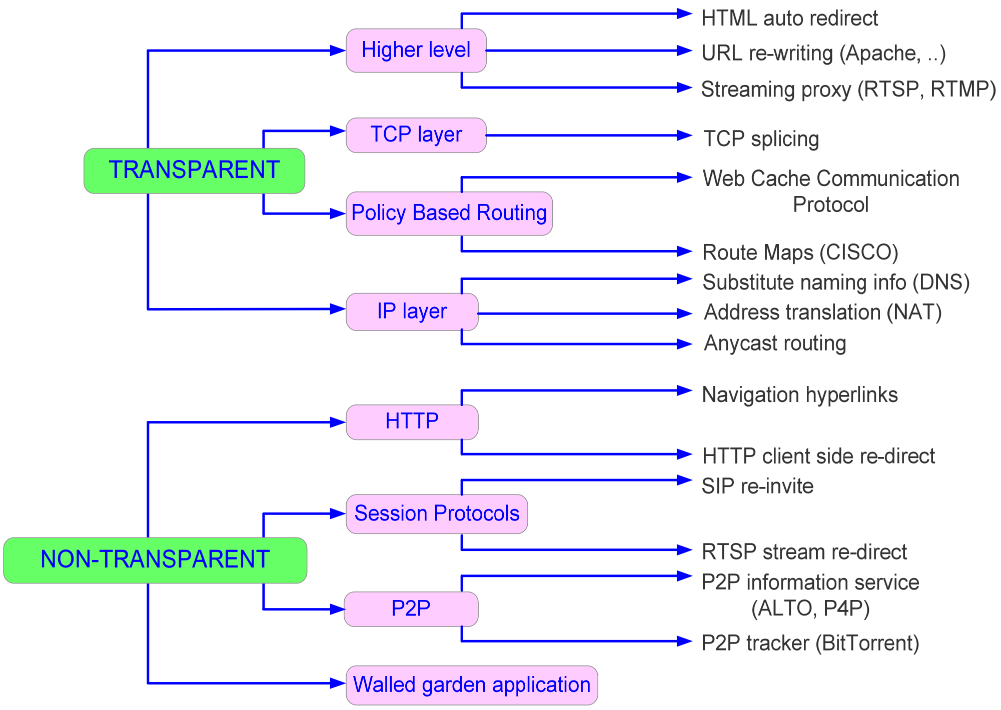


**Figure 5-1: *Classification of re-direction mechanisms***

### 5.1.2.4 Re-direction on application layer

In this section we evaluate re-direction principles in the application layer. Dynamic change of selection policy within short time frame is in principle feasible for all considered higher layer mechanisms. Coding such information within the HTML page is not appropriate, since it does not support dynamic selection. However it is possible to use PHP language for the dynamic creation of HTML code with the relevant re-direction information, which can be retrieved from a database. A PHP interpreter is needed for this purpose. Each request of a Web browser therefore increases load and may reduce Web server performance. Special caches can be used to take pre-processed information in order to accelerate the PHP processing. There is possibly little benefit with this approach, especially if the re-direction information regularly changes and caching information would be useless. Usually more than a single HTTP server would be needed, which makes it difficult to keep re-direction information consistent among different servers. Hence solutions with HTTP proxies are more appropriate for deployment in large-scale networks, like EPS. A proxy represents the front-end server, which receives HTTP requests and propagates to other backend servers. The proxy nodes terminate TCP connections towards the user, and back end servers are invisible to the original requester.

Forward proxies are well known to serve clients within a protected environment for communication with external servers (e.g. in the Internet) through a firewall. It is suggested to make use of this concept also within a mobile access domain in order to control access of the connected user terminals to external or local servers. Reverse proxies are known to delegate connection requests from external host to a localized resource. A front-end server controls access to backend nodes. It is suggested that this concept is also used to control external requests to a resource within the mobile operator network, e.g. if a roaming user requests access to an application in the home network domain.

Access to a wide range of applications, based on HTTP, can be controlled. However, if also other applications (e.g. based on streaming protocols) should be controlled, then either additional higher layer or a lower layer mechanism is required. If more than one higher layer mechanism should be applied in parallel, a higher level of system complexity is implied within the mobile access domain. If resource information is stored in separate databases for each re-direction mechanism, there is the challenge to keep the information synchronized. If a common database is used, specific interfaces are needed to enable application layer re-direction. For a mobile operator it is therefore the question e.g. whether to control also other applications apart from HTTP.

Protocols which are session based (e.g. RTSP and SIP) may have the capability to switch between resources during runtime without significant volume of duplicated data transfer provided that implementations in a client / proxy and the server support this kind of transfer of session state. Since resource switching is more complex than staying with the same resource, it is for further study whether this approach is beneficial.

Some application layer proxies such as HTTP proxies are involved in data plane management. Such kind of proxy is suitable to handle up to several thousand connections, e.g. for a cooperate network. However for a large mobile network domain with possibly up to several million concurrent user connections, a single proxy would presumably become overloaded. Hence a larger number of HTTP proxies are needed. This adds more challenges related to state synchronization for resource management and switching of the serving proxy for connected users.

#### 5.1.2.5 Re-direction with TCP

A TCP proxy has the advantage that not just HTTP but other application flows based on TCP can be re-directed (e.g. FTP). However generic TCP proxies are not common in usage since these proxies add higher layer support for specialized application handling. An example is a HTTP proxy. There is a set of streaming applications, which make use of RTSP and UDP. Those applications cannot be controlled by a TCP proxy for the purpose of resource selection. Other types of TCP proxies can’t be used to control selection of a resource since the passing TCP segments are not changed, e.g. a SOCKS proxy is only used for establishing secure connections. Hence generic TCP proxies should not be further considered for this work.

#### 5.1.2.6 Re-direction on IP layer

All mechanisms, which enable a manipulation of the IP address for a requested resource can in principle influence selection of all types of applications in the Internet (e.g. based on mechanisms like DNS, NAT, BGP, etc.). Hence this mechanism is quite flexible and it is transparent to the user and application client. A further advantage is that re-location can be accomplished on coarse granularity for IP layer re-direction. There are specific issues for different mechanisms of manipulating the IP address. IP layer mechanisms cannot differentiate between application type or content and location. In some cases it could be beneficial to enable such differentiation as explained in the following examples:

- Specific caches for specific content, e.g. some caches may serve only specific type of content
- Content specific processing requires special functions on certain nodes, e.g. transcoding

DNS resolves IP address at connection setup. Afterwards the client is bound to the resource until the end of the lifetime of the TCP / UDP connection. Hence a resource could not be switched during lifetime. Another problem is that a client requesting a resource may not ask a local DNS server to resolve the name of a resource because the information is already available in the local cache of the client. As a solution to this problem the local DNS server could send DNS responses with a short TTL value. However this forces DNS clients to frequently send name resolution requests to the local DNS server. In total this approach could lead to a lot of DNS related signaling traffic and high load on the local DNS server. Therefore it is envisaged to manipulate DNS information only for stable configurations.

Further it shall be assumed that DNS manipulation can be less relevant for resources, which are located on mobile user devices. For pragmatic reasons most (mobile) devices have not allocated permanent IP address/es within their home network domain, but instead they obtain these by configuration mechanisms, such as DHCP. In addition identification of resources on these hosts (in many cases P2P content) is not based on hostnames. Consequently DNS name manipulation is in these cases not a suitable mechanism for resources, which are located on user devices.

For selecting resources located inside the local network, dynamic DNS (DynDNS) can be used to relieve the problem of long lasting mappings between domain-/host-names and IP addresses [DDNS]. This is an additional feature for DNS, which needs to be supported by DNS servers. In order to allow fast mapping (e.g. in the range of 60 s) of a requested domain name with a different IP address for a specified hostname, the TTL (time-to-live) entry needs to be reduced significantly. However with a larger number of DNS records also many lookups are required within the DNS system. This, in general, impairs the performance of the selection process, since DNS was not designed for dynamic mappings between IP address and host name. Another disadvantage is that localization of the closest server or server cluster can be imprecise, if the user's DNS server IP address can't be assigned correctly to the client location, e.g. because the client is far away from the DNS server [ALR+08].

NAT has certain advantages over other IP layer mechanisms because no interaction on protocol level with other network domains has to be considered. Successfully deployed in enterprise networks it is unclear whether NAT can be deployed in EPC serving a number of users, which is several orders of magnitude higher. A NAT device would have to dynamically replace IP address information on the fly. It is for further study whether NAT can be a feasible approach for re-direction.

Re-direction with WCCP requires manual router configuration and therefore this scheme is not suited for spontaneous reaction to changed resource conditions.

Anycast routing can select the most appropriate resource according to given policies from a number of hosts all identified by the same Anycast address. Following problems have been identified with anycast. It is tightly coupled with the IP routing mechanism, thus any change causes a session reset to any session-based protocol such as TCP. Also the IP routing infrastructure can't react to changed conditions [ALR+08]. Once the routing has adapted to new conditions, no additional processing of resource information is necessary. The mechanism is transparent for the user. However synchronization is challenging between Anycast servers in order to keep content identical, especially for regular changing content [WF10].

### 5.1.3 Handling unfavourable resource usage

Previous subsections have described the aspects how a request from a requesting mobile user can be re-directed to a resource, which is more preferable.
The principle idea of the mechanism described here is to detect unfavorable resource usage and perform an appropriate action. A resource may be unfavorable for several reasons:

- Resource itself is unfavorable, e.g. because it can't support requested features (QoE, security, transcoding, etc.) or is not conforming with defined usage conditions (e.g. associated cost)
- Location of the resource is unfavorable, e.g. because it is in an external network or too many hops away
- Path towards a resource is unfavorable, (e.g. because it is temporarily a wireless link or stretched beyond certain gateways, which are not preferred)

There are several situations for unfavorable usage from network perspective. For non-transparent selection, a user or an application client could make an initial unfavorable decision. In addition to a favorable resource, the requesting peer connected with could become unfavorable due to some event. If the resource was selected transparently, the user or application client may have little or no possibilities to change the situation. The following cases should be considered:

- Handover of the requesting node: a user changed the access point and therefore switching to another resource may lead to a more efficient usage of the network
- Un-availability or decreased resource quality: quality of the resource may deteriorate after some time, enforcing a change. One reason could be that an end-user system hosting a resource is switched off
- A new resource, which has higher preference is available / detected after the requesting peer starts to receive content

Detection of the events related to un-preferable resource usage is versatile and are for further study.

In the following there are some possible reactions to the detected situation.

- Continue data transfer / streaming: if the remaining volume (in case it is known) is below a certain threshold volume continuation might be best. For live streaming, the remaining data could be estimated in case there is some indication by the user of the application client
- Stop the transfer of data and connect to a new resource: this is a natural choice for live streaming (RTP based) applications (depending on the time to switch the resource), while for TCP applications such an action might be undesirable.
- Notify user / application client / application server about unfavorable resource usage
- Try to switch user to a different resource. The application needs to support the switch between different nodes. This should happen smoothly, i.e. ideally without a user to notice the switch. Analyze different applications where this is possible and how this can be achieved. For live streaming this should be possible. For downloading it should be possible, if the file is partitioned into multiple parts.

A detailed analysis about cost-gain factors is for further study.

In the following the notification option is further described. The objective is to inform a requesting peer (an end user system), which requests content about unfavorable resource usage and provide additional information about alternative providing peers (an end user system or a node in a service / network domain), which are acceptable from the network provider's perspective and about conditions for resource usage. Information forwarding requires user interaction based on the following principles:

- Pop-up window on the user terminal (e.g. realized by web browser via HTTP or a command line / console window)
- Messaging services (e.g. SMS, MMS, )
- Mail service (e.g. SMTP,)
- Voice message via automated voice call
- Chat services (e.g. MSN,)

Figure 5-2 illustrates the system architecture for the interaction. Unfavorable usage of a providing peer (Providing Peer A) is detected by a traffic monitoring entity in the network provider domain (e.g. a DPI device) based on traffic pattern recognition or payload inspection. Alternatively the detection of the unpreferred peer is already recognized during connection setup (e.g. based on a flow identifier such as the following 'five tuple': source IP address; source (TCP/ UDP) port number; destination IP address; destination (TCP/ UDP) port number, protocol identifier).

The resource selection control keeps the information about preferred resource identities in a database and controls the selection process. It provides also an interface to trigger information forwarding of content from its database, if better alternatives for providing peers are available. This interface is called 'triggering interface' in the following. This interface can be used by trusted entities, which are capable to detect usage of potentially unfavored resources (e.g. a monitoring entity such as a DPI device.), which have been selected by a requesting peer.

Entities reporting potentially unfavored resource usage via triggering interface to the resource selection control can send the following information elements:

- ID of requesting peer
- ID of selected peer
- Type of content
- Protocol ID
- Traffic characteristics

Also a user may trigger notification of candidate peers to obtain the relevant information.

The resource selection control component processes the resource data in the following way: The information obtained via triggering interface is compared with entries from the database. If matching entries are found for the type of content sent within the resource data (e.g. Providing Peer B in Figure 5-2), resource selection control will initiate notification about alternatives either to the user or e.g. the local naming service. For notification to the user, the necessary information is sent to the notification manager.

Purpose of the notification manager is to extract information related to alternative peer selection sent by a resource selection control component, select an appropriate communication channel towards the user (e.g. messaging, pop-up window etc.) based on additional information about user context and trigger notification via the specific communication channel. For the purpose of selecting an appropriate communication channel the notification manager may have access to user profile information and specific session information (e.g. what user equipment is currently connected to the network etc.)

Resource selection control could provide additional information about the conditions of the resource usage to the notification manager. For instance, access to an alternative resource could be associated with connectivity to a specific gateway node, e.g. the APN (Access Point Name) to be used could be specified in the notification to the user.

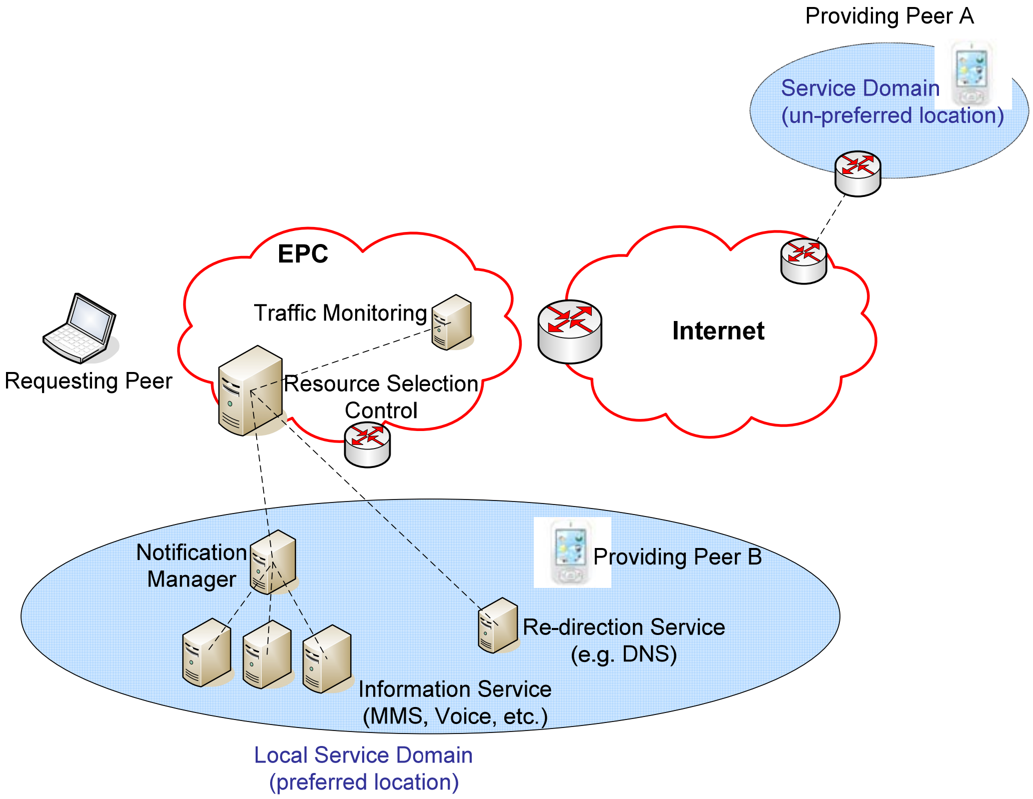


**Figure 5-2:** ***Notification about unpreferred resource selection***

## 5.2 Internet-based Content Distribution via CDN and P2P Overlays

Most of the traffic on the Internet is currently transported via content delivery (CDN) and peer-to-peer (P2P) networks. CDN and P2P systems build overlays of different types, which set up their own communication architecture on top of the underlying IP network. P2P systems interconnect the terminals of the users without or with a minimum of their own network infrastructure and costs, whereas CDNs are based on globally distributed servers, whose connectivity can be supported in a virtual network. Both approaches introduce their own network and traffic management functions within the overlay, which can be organized independently of the network layer. This independent layering facilitates the deployment of new CDN and P2P overlays, but a lack of cross-layer awareness and cooperation leads to inefficient transport. Problems of unnecessarily long paths between peers and possible mismatch between traffic management on CDN and IP network layer are discussed as a main focus of this overview.

From a more general point of view, the need for overlay structures becomes apparent in work on lower layers by the Internet Engineering Task Force (IETF) [IRTF], resulting in several hundred standard documents in the RFC database defining "protocol or network technology x over y" as well as a number of IETF working groups on virtualized networks, pseudo wires, layer 1,2,3 VPNs, etc. Successful overlays, starting with IP over Ethernet, are aware and well adapted to the underlying infrastructure.

### 5.2.1 Content delivery networks (CDN)

Popular client/server-based websites are usually supported by content delivery (or distribution) networks, which are hosted on server clusters on the Internet as depicted in Figure 5-3. CDNs started as enhancements of web browsing and downloads of files but have extended their scope with the broadening spectrum of Internet services to support streaming and IP-TV [DVB09].

A study of transfer paths in Akamai's CDN [AKAM12] [SCK+09] shows how users are redirected from the original website through a hierarchical server farm, consisting of thousands of servers, to a delivery node in the proximity of the client. In this way bottlenecks at a single web server are widened in a distributed network providing higher throughput, while the transport paths are shortened. Reliability and user experience is improving since the CDN connection to a user can be dynamically handed over to another server through prepared backup paths if performance measurement indicates problems on the current path or for load balancing in the CDN server cluster. The study by Su et al. [SCK+09] confirms that CDNs are efficient in shortening transport paths and delays and that they stabilize throughput.

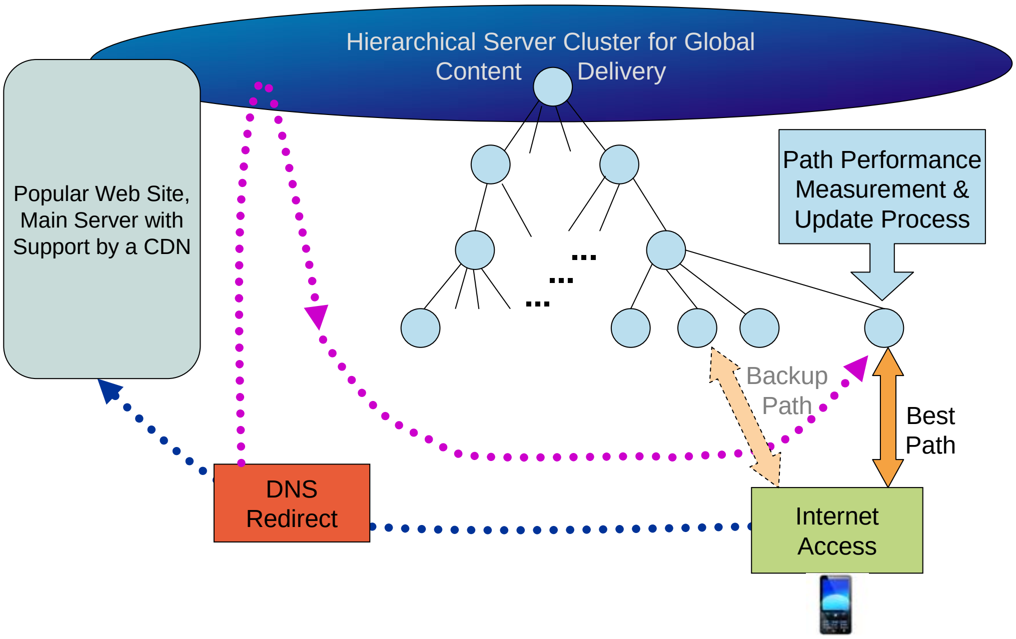

Hierarchical Server Cluster for Global Content Delivery
Popular Web Site, Main Server with Support by a CDN
Path Performance Measurement & Update Process
Backup Path
Best Path
DNS Redirect
Internet Access

**Figure 5-3:** ***Client-server applications supported by global content delivery networks***

Currently, numerous research and engineering activities connected with content delivery networks are being carried out which relate to our main focus on the effect on network provider platforms and in a wider scope:

- In addition to global CDN providers, large network providers are extending or setting up CDNs within their platform, as announced by Level3 and AT&T already in 2007.
- The Internet Engineering Task Force (IETF) has a working group on application layer traffic engineering (ALTO WG) and has recently set up a new working group on CDN interconnection (CDNI) in order to enhance the interworking between CDNs [IRTF].
- Television broadcasters are investigating CDN support for upcoming IP-TV integration [DVB09].
- Community networks also require special content distribution infrastructures [CFM09], which are often based on peer-to-peer networking or combined CDN-P2P architectures [HWL+08].
- Future Internet activities explore new concepts for naming and locating distributed content independent of IP addressing on host nodes to avoid inefficiencies in host-to-host communication, which at the same time pose the challenge of establishing alternative search schemes that are at least as efficient as current IP routing.
- Together with numerous published papers we refer to a book on content delivery networks [BPV08] covering basic properties up to the state of the art as well as to a broad overview of the literature.

### 5.2.2 Peer-to-peer networks (P2P)

Since the beginning of the new millennium, peer-to-peer networking has shown its potential for efficient content distribution and for unleashing idle resources. Popular peer-to-peer protocols have established global file sharing, voice over IP (VoIP), and streaming applications that interconnect millions of users [SYB+09]. Peer-to-peer networks form distributed overlays on users' terminal equipment, offering global services with a minimum of infrastructure. Fast delivery of large volumes of content in globally scattered communities is a main strength of the peer-to-peer principle, with flexible adaptation of network size and resources to demand, even for peers with largely differing access speeds [SHH07]. Peer-to-peer networks are highly efficient for various applications due to their ability

- to exploit vacant resources residing on user equipment (data storage, computation power, access bandwidth),
- to adapt to varying demands with short-time scalability to dynamic flash crowds of huge communities,

- to embed autonomous support for search and replication of data for predefined demands and
- to build and manage self-organizing overlays at low investment in network and server infrastructure.

Measurement studies have identified P2P and user-to-user activity as the main driver of traffic growth in the new millennium [HAS05]. CDNs carry a considerable portion of Internet traffic and have recently become more important because part of user generated content has shifted to client/server-based platforms for video and file up- and downloads, gaming and social networking. Reports by Cisco Systems [CIS10] and Sandvine Inc. [SAN10] indicate that entertainment video and P2P are the most relevant IP traffic components, with a clear growth trend of video streaming and IP-TV but no unique trend for the P2P traffic portion, which is also estimated to largely differ between continents. The broad spectrum of research on P2P overlay architectures and applications can be seen in a handbook [SYB+09] as well as in a number of books on special aspects of P2P networking.

### 5.2.3 Transport paths of CDN and P2P overlays on broadband access infrastructure

Compared to data transfers from a nearby CDN server, P2P downloads are currently subject to much longer transport paths and delays. For network providers, unnecessarily long transfer paths impose higher loads on peering and interconnection routes, including expensive intercontinental links [CIU10] [SCK+09]. Reliability and throughput is also affected when transmissions have to traverse more links than necessary. Figure 5-4 illustrates the different behavior of CDN and P2P networks on the broadband access architecture of the Internet. Tree-shaped access areas are attached to the backbone at points of presence (PoPs), where remote access routers handle the registration of user sessions. Network provider backbones are embedded in the Internet structure via peering links and exchange points to other network providers.

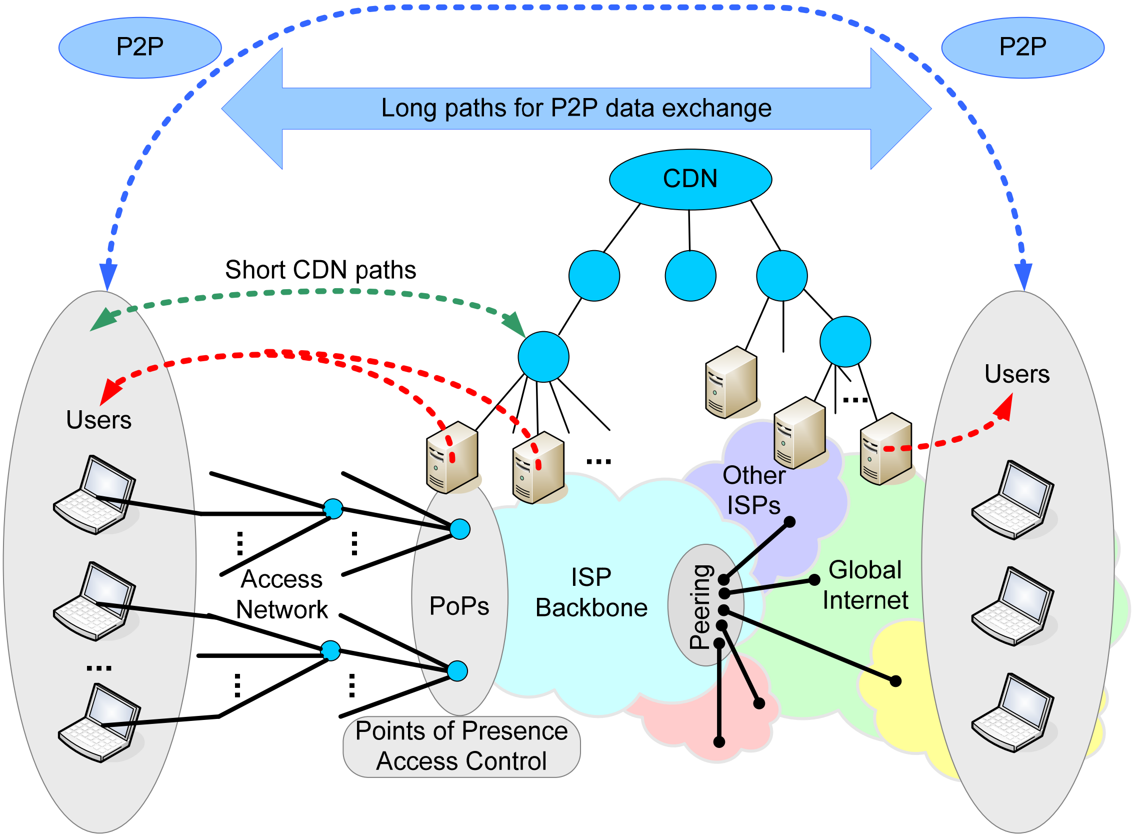


**Figure 5-4:** ***P2P versus CDN overlays on broadband access platforms***

Considering large ISP networks serving millions of subscribers, it can be expected that most of the data of a global file-sharing network is already found to be replicated on the same ISP platform and often in the same access region of a P2P downloader. Due to Zipf law access patterns, most downloads address a small set of currently the most popular files, which strengthens the gain of local downloads [BRE+99] [EUB08] [HHB09].

Some tendency for local P2P exchange may arise due to interaction within social groups. A separation of user communities and content due to different languages is most obvious. When looking at downloads of

German content to a German destination via eDonkey [HAS05], it is not surprising to find a 80 : 20 Zipf distribution rule, such that about 80% of the sources are again located in Germany. By contrast, less than 20% of source locations are found in Germany for downloads of English content. The assignment of peers to the same supernode in hierarchical P2P networks like eDonkey also generates locality within the reach of a supernode, but without correspondence to network layer topology or Autonomous System boundaries.

Nonetheless, the problems of long delivery paths and delays are confirmed in a measurement study for popular P2P live streaming networks (PPLive, SoapCast, TVAnts) [CIU10] with only mild or no preference for peers in the same Autonomous System. Other studies show that locality in the BitTorrent and Coolstreaming P2P networks can essentially be improved by giving preference to peers in the same Autonomous System.

### 5.2.4 Comparison of delays on CDN and P2P transport paths

We have evaluated the delays for traffic via P2P and CDN overlays through packet-based measurement on links in the aggregation of Deutsche Telekom’s broadband access network [HHB09]. As a result, we have not identified the complete traffic of both types but have selected a fraction that can easily be detected via P2P ports for BitTorrent, eDonkey and Gnutella and via IP address ranges indicating Autonomous Systems of Akamai, Limelight, Google and other server sites. The flows classified via P2P ports comprised only a fraction of 2.7% of the traffic volume, while known IP address ranges for CDNs and popular web sites accounted for 10.7% of the total traffic. A one-hour measurement is evaluated at the daily peak rate in mid-2008 in downstream direction on a link with low load, which shows the typical behavior being observed in other cases as well. We used the time stamps of two successive packets sent by the client during the TCP handshake to estimate the round-trip delay, as depicted Figure 5-5, although this may also include delays due to the reaction of the server or peer in response to the TCP connection request.

Figure 5-5 compares the cumulative distribution function of those delays up to 1s in the two continuously increasing curves. There are two additional curves representing the distribution function in a sampling at the 0.01s time scale, i.e. the fraction of delays falling into each 0.01s time slot is shown for both the P2P and CDN flows. As expected, the delays are shorter for CDN delivery. The mean delay for CDN flows is 0.125s, whereas P2P flows have a mean delay of 0.33s. About 10% of the P2P delays exceed 1s and thus would be unacceptable for services with real-time demands.

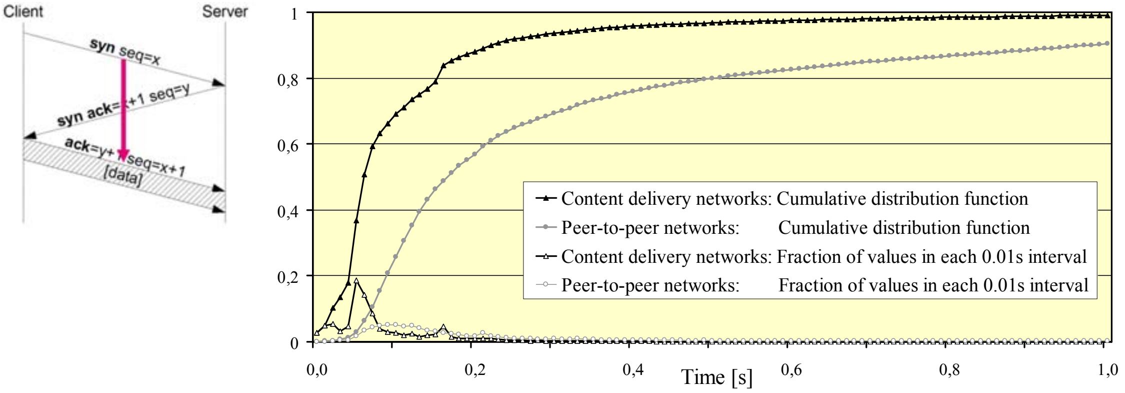


**Figure 5-5:** ***Comparing round trip delays in P2P and CDN overlays using syn packets in the TCP handshake***

### 5.2.5 Summary of characteristics for content delivery in CDN and P2P networks

The main advantage of the peer-to-peer distribution principle is rapid adaptation to new user demands. Large spontaneous flash crowds can trigger an exponential increase in the P2P data exchange volume. In this way, the throughput in P2P networks is rapidly increasing with population until the upstream bandwidths of all involved peers are exploited as the maximum P2P network throughput. The maximum throughput of content delivery networks depends on the bandwidth provided for the network connectivity for all the servers involved, which has a fixed limitation and cannot arbitrarily breathe with user demand.

On the other hand, CDNs prefer servers located near a requesting user. Even if administrative boundaries between global CDNs and network providers prevent fully optimized local content delivery, CDNs achieve essentially shorter transport paths and end-to-end delays, compared to P2P networks. The missing optimization of local transport options is a drawback of P2P networking not least from the network provider’s perspective. On the other hand, the smoothing effects of replicated and distributed content in

P2P networks improve the load balance and the utilization of bandwidth in backbone networks. P2P traffic enters the network from the access and therefore is provisioned inherently through dimensioning access and aggregation links of moderate speed without any danger of sudden shifts in the backbone. CDN traffic enters the backbone via high-speed peering links, such that large CDNs can inject unforeseen traffic shifts if there is no cooperation with a network provider. Table 5.1 compares main CDN versus P2P properties.

**Table 5.1:** ***Properties of content transport on the Internet over CDN versus P2P networks***

| Comparison of characteristics | Hosted on | Scalability | Traffic characteristics | Transport paths; traffic demand | QoS Support |
|---|---|---|---|---|---|
| CDN: Content Delivery Networks | Network nodes, distributed server farms | High, depending on CDN size and bandwidth | Asymmetrical; variability depends on application | Short paths from nearby CDN server; high traffic demands | Low delay; provisioned availability |
| P2P: Peer-to-Peer Networks | User terminals, minimal network infrastructure | Extremely high, growing with user population | Symmetrical; low variability for file sharing | Long transport paths; extremely high traffic demands | High delay; demand driven availability |

### 5.2.5.1 Network and Application Layer Approaches for Short Transport Paths

The challenge of optimizing the transport paths in content delivery for different applications has recently become a main focus in research and standardization. Figure 5-6 illustrates three principle alternatives of

- caching [ASK+10] [BRE+99] [HHB09],
- positioning or coordinate systems [LGS07], and
- information servers [PFA+10].

Caching is an option for network providers and is also useful in the end systems, whereas positioning systems are established on the application layer to estimate distances between clients over the Internet. Last but not least, the introduction of information servers is a cross-layer approach that collects data about client locations on the network layer to be made available for applications. As a result, a cooperative standardization effort is required. Since 2008, the ALTO working group of the Internet Engineering Task Force (IETF) has been exploring ways of improving localized data exchange for peer-to-peer and other types of content distribution [IRTF]. In the following we discuss these alternatives, starting with caching and the properties of Internet access patterns relevant to cache efficiency.

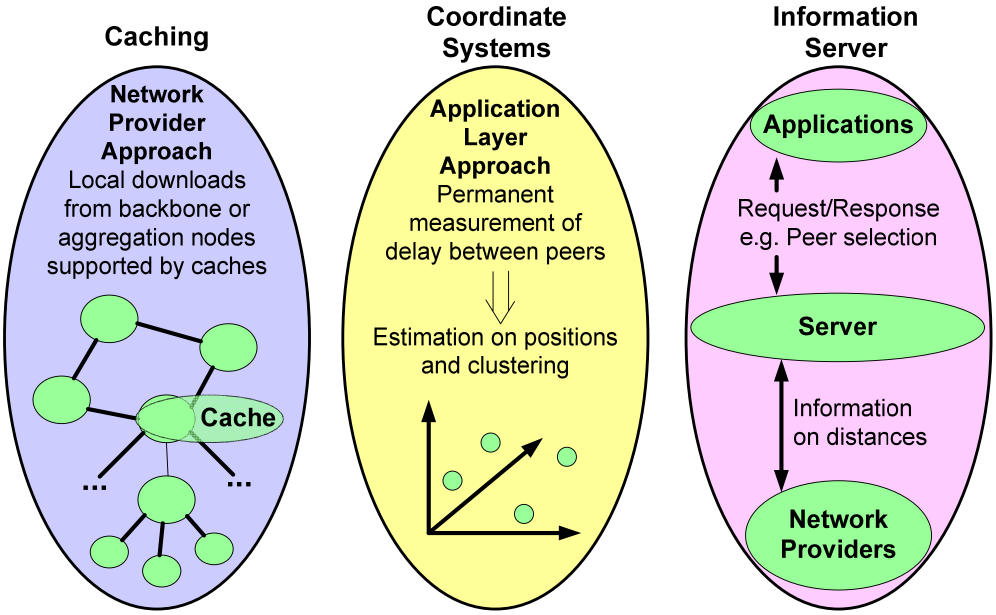

Caching
Network Provider Approach
Local downloads from backbone or aggregation nodes supported by caches
Cache
...
...
Coordinate Systems
Application Layer Approach
Permanent measurement of delay between peers
Estimation on positions and clustering
Information Server
Applications
Request/Response e.g. Peer selection
Server
Information on distances
Network Providers

**Figure 5-6:** ***Approaches for localized content delivery with network and application layer support***

#### 5.2.5.2 Application layer positioning approaches

The measurement of transmission delays between nodes can be evaluated on the application overlay to draw conclusions about the distance between the nodes and about the infrastructure of the underlying transport network. CDNs also make use of latency monitoring to optimize paths as addressed in the previous section. Several projects have investigated the approach for peer-to-peer overlays from lightweight versions to multi-dimensional coordinate systems [LGS07]. In principle, these approaches are most sensitive to changing network conditions, enabling fast response, but they require considerable maintenance effort.

A coordinate system based on probing for delays between BitTorrent peers has been studied and implemented in prototype versions of the Azureus client [LGS07]. The probing is piggybacked on other messages between peers to reduce messaging overhead. Based on transmission delay estimates between pairs of nodes, a coordinate system is set up in a multi-dimensional representation. Among earlier approaches for coordinate positioning, a two-dimensional coordinate system was proposed with an additional height component to account for delays in the access. Considerable effort is needed to maintain coordinate systems due to

- the churn of peers and nodes entering or leaving the population served by the application,
- changes in the underlying network topology or the routing, and
- the variability of measured delays over time due to the changing load on network links.

The study [LGS07] concludes that useful coordinates can be established on the application layer, but that the effort of deploying a coordinate system would seem to be affordable only for large-scale overlays rather than for new evolving services. As a simpler alternative, the lightweight position approach tracks the distances from the view of a node to a set of other nodes, which are classified as belonging to rings according to their transmission latencies, from which close-by nodes in the overlay or an appropriate central node in a cluster can be determined.

On the other hand, it should be noted that successful localized source selection in P2P networks strengthens a tendency to build clusters with strong internal connectivity and a loose coupling between the clusters. With preference for connections between nodes in the same Autonomous System (AS), a cluster or subnet structure is expected to be bounded in an AS. Thus, improved local data exchange may slow down the distribution of content over ISP network boundaries. In the worst case, strict local preference and a high churn in P2P networks may lead to separated subnets for the same content.

An obvious approach for sustaining sufficient inter-domain connectivity is to mix local preference with randomness. A portion of random selection keeps the inter-domain traffic throughput at a low but sufficient level to avoid slowdown of content propagation or separation of a P2P network into local clusters.

#### 5.2.5.3 Traffic engineering support by information servers in current IETF standardization

Since 2008, the Internet Engineering Task Force (IETF) has set up a working group on Application Layer Traffic Optimization (ALTO) [IRTF], which focuses on providing information on node locations to the application protocols in cooperation with network operators and other parties holding IP topology knowledge. Earlier projects and studies already made such information available on websites, e.g. Closest Node <www.closestnode.com> based on the Meridian service or Prefix WhoIs <www.pwhois.org>. Prefix WhoIs queries are used to determine and to prefer peers in the same Autonomous System and district for improved locality in the BitTorrent and Coolstreaming P2P networks.

The main preconditions and questions to be clarified in a standardized framework for location information services are

- Which types of metrics and information are useful for estimating distances and local relationships between different hosts, and which format is appropriate for representing the data?
- Which parties are expected to provide and control such information, based on which own interest or incentives? How can contributions from several sources be integrated into a common total view?
- How should a unique interface of the service to the applications be designed?
- How can availability and scalability be achieved and failure cases be handled?
- How can the benefit of a location information service be monitored, and how can the efficiency be optimized for users and network providers?

In a crude but already beneficial localization scheme, the corresponding Autonomous Systems (AS) can be assigned to the IP address of a potential source as well as the corresponding ISP or corporate network operator. Some information on the number of transmission hops can be extracted from global BGP routing tables, but tunneling mechanisms and other obstacles make it impossible to obtain a clear picture of the hop count from an end-to-end perspective.

For finer granularity, the network providers have the best knowledge to distinguish access regions or other neighborhood relationships within a network under unique administration, but they are usually reluctant to publish their network structure in detail. As a result, opportunities for support without full topology awareness are considered. Download request in P2P networks are often answered by providing the IP addresses of a set of possible sources to be contacted by the peer. A suggestion is to hand over the set of source options to a localization server, which in turn ranks the sources or selects a subset of preferable sources without making the network topology transparent [PFA+10]. Network providers can set up their own policies for dealing with localization servers in order to avoid transport via expensive and overloaded links. Achieving shorter paths is favorable for all parties involved, but a server may partly have no or inappropriate, e.g. outdated, information on different network regions. Applications have to trust or to check if the server information is useful and have to decide what to do otherwise.

On the other hand, applications can collect additional information, e.g. through measurement of inter-node transmission latencies, which can be combined with the server information on application layer or can be made available also for other applications as another input to the server. A clustering of nodes according to the areas supplied via the same server of a global CDN may be considered to check locality relationships [SCK+09]. The overhead and delay introduced by localization servers has to be kept small. Applications may set a threshold on the requested data volume, such that localization servers are only involved for profitable cases of lengthy data transfers.

The scope of the IETF ALTO working group on application layer traffic optimization comprises P2P as well as other applications, including data delivered through CDNs. The information on source locations in distributed platforms can also be utilized by network providers to improve traffic management and quality of service in their platforms. The gain in efficiency based on exploiting provider-aided distance information has been evaluated by [PFA+10].

#### 5.2.5.4 ALTO extension proposal for UE with scarce resources and/or intermittent connection

The Client/Server protocol ALTO provides abstracted information on the transport topology underlying the overlay network of applications such as P2P and CDN. This information includes abstractions of network maps, cost maps and endpoint costs. Abstraction is done by

- Providing topologies of “network locations”, specified by the managing ISP and corresponding to host groups of heterogeneous levels in the Internet hierarchy and
- Associating ISP defined “routing costs” among these network locations.

The current ALTO protocol provides a unique value for the requested cost type, and if up to date values are needed they must be requested as often as the value is changed. However, frequent ALTO transactions for updates are costly. In some cases though the cost value changes are predicted, in particular for the sake of traffic regulation, and having a set of predicted values beforehand would be highly beneficial to the applications and end systems relying on the ALTO service.

In particular, ALTO supported applications such as content delivery spatially shift traffic between network regions in order to lower routing costs for ISPs and regularly move content across their caching nodes to better map to the demand; many non-real time applications have a degree of freedom on *when* to “use a resource”, given that

- a “resource” is a content file or a computation resource stored in a data center and
- ”use” a resource means for example, do a content transfer between caches, access a service, use a physical server for a virtualized application or do time shifted content delivery.

Nowadays, popular applications ran on UEs in wireless network such as content delivery and clouded applications become a challenge for the network and the UEs due to the scarcity of resources and challenged network access. Application clients on end systems with limited access to data centers and/or to the network or using resources scattered around the world need to schedule their access to resources or need to figure out when resource transfer or access costs may change. In some cases of non real time applications, this cost is predictable over a given number of time slots.

An extension to the ALTO protocol called “ALTO Cost Schedule” to support time-shift of traffic has been proposed and presented at the 83rd IETF, see [Randriamasy12]. “ALTO Costs” represented in the “Schedule” mode aim at

- lower traffic peaks and save scarce resources to maintain user QoE and

- providing costs describing resources over a set of time periods.

Particularly suitable ALTO services for this extension are

- ALTO Endpoint Cost Service, which is flexible, "light“ and better suited to endpoint selection based on multivariate optimization
- ALTO Filtered Cost Map Service, which is lighter than providing the entire cost map.

This proposal does not preclude from the need to keep resources information abstract enough to protect confidentiality of network provider information.

**The “Schedule” mode for ALTO costs:**

- Extends Cost Map information in the time horizon
  - defines slots (e.g. hourly) over a period of time (e.g. one day) and
  - has attributes that can be specified for each applicable Cost-Type.
- Adds a new "cost-mode" called : ["schedule"]
- The schedule scope is defined in a new IRD capability with an example as follows:

**"cost-scope": [{"unit": ["hour", 1], "size": 24,**
**"begin": 0, "time zone": "UTC",**
**"lastupdate": mm/hh/dd/mm/yyyy,**
**"nextupdate": mm/hh/dd/mm/yyyy} ]**

#### *5.2.5.4.1 Example of ALTO transaction with costs in the “schedule mode”*

In the following example, as described in [Randriamasy12], an Application Client has the choice to trade content or resources with a set of Endpoints of moderate 'routingcost', and needs to decide with which Endpoint it will trade at what time. For instance, one may assume that the Endpoints are spread on different time zones, or have intermittent access. In this example, the 'routingcost' is assumed constant for the scheduling period and the time-sensitive decision metric is the path bandwidth reflected by a Cost type called 'pathoccupationcost'.

The ALTO Client embedded in the Application Client queries ALTO information on 'pathoccupationcost' for the 24 hours following (implicitely) the date of "lastupdate", as this resource is listed in the IRD.

**REQUEST**:

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: [TODO]
Content-Type: application/alto-endpointcostparams+json
Accept: application/alto-endpointcost+json,application/alto-error+json
{
  "cost-type" : ["pathoccupationcost"],
  "cost-mode" : ["schedule"],
  "endpoints" : {
    "srcs": [ "ipv4:192.0.2.2" ],
    "dsts": [
      "ipv4:192.0.2.89",
      "ipv4:198.51.100.34",
      "ipv4:203.0.113.45"
    ]
  }
}
```

**RESPONSE**:

```
HTTP/1.1 200 OK
Content-Length: [TODO]
Content-Type: application/alto-endpointcost+json
{
  "meta" : {},
  "data" : {
    "cost-type" : ["pathoccupationcost"],
    "cost-mode" : ["schedule"],
    "map" : {
      "ipv4:192.0.2.2": {
        "ipv4:192.0.2.89" : [7, ... 24 values],
        "ipv4:198.51.100.34" : [4, ... 24 values],
        "ipv4:203.0.113.45" : [2, ... 24 values]
        }
    }
  }
}
```

#### 5.2.5.5 Combined CDN, Caching and P2P Overlay Approaches

In principle, P2P traffic has the same properties benefiting cache efficiency as does HTTP-based traffic, i.e. Zipf-like distributions with a high concentration of accesses to popular content and with only few problems of outdated data as in classical web caches, since P2P content is usually permanent and uniquely identified by hash methods. But approaches for including P2P traffic in caches are rare, although P2P networking has dominated Internet traffic for several years and still contributes considerable volume [CIS12] [HAS05] [SAN10].

In 2004, the eDonkey/eMule file-sharing network offered the option of including the web caches of network providers by disguising P2P downloads as normal HTTP requests. When we investigated usage of the cache option in 2006, a small portion of 5-10% of the download volume was found to be supported by caches, although downloading from the cache achieved 5 times higher throughput [HAS05]. Since 2006, neither the caching option nor the eDonkey/eMule network as a whole seem to have been updated or further developed, with the result that the once dominant traffic source in Germany and France [HAS05] is no longer important in current statistics.

The distribution of copyright-infringing or other illegal content through caches is a problem especially if P2P data is included. The content of web caches reflects popular content on the Internet and, therefore, always includes more or less problematic data. Although transparent caches are admissible for enhancing transport performance, any support for copyright-infringing material would lead to counteractions by the affected content owners.

Moreover, a lack of congestion control options is an obvious reason why network providers are reluctant to cache P2P traffic. Due to its symmetrical transfer volume up- and downstream, the throughput bottleneck for P2P traffic is in the upstream bandwidth of the peers. Caching of P2P traffic widens this bottleneck, since a transmission from the cache replaces the upstream limitation by an essentially larger cache bandwidth. Then the downstream speed of a peer becomes the bottleneck, which is up to 10-fold higher in usual asymmetrical broadband access in wireless and DSL lines, as experienced e.g. for the eDonkey/eMule caching option [HAS05]. The throughput and delay performance of the P2P network clearly benefits from cache support, but a network provider could end up with essentially higher P2P traffic volume and increasing transport costs even if transfer paths from the caches are short.

While CDN and P2P distribution schemes can be seen as competing alternatives in current and future Internet architectures, combined CDN-P2P content delivery is a promising way to exploit the advantages of both principles [DVB09] [HWL+08]. A CDN can provide content locally at servers in different regions, while an additional P2P overlay that includes the servers of the CDN can provide more throughput, which scales linearly with user demand and saves CDN connectivity bandwidth. Some streaming services have implemented and evaluated hybrid CDN-P2P content distribution, which is flexible and profitable especially for regions with high bandwidth costs for CDN server connections. More recently, several CDN solution and service providers have integrated hybrid CDN-P2P approaches. The industry forum on digital video broadcasting has also addressed hybrid CDN-P2P delivery technology in reports, with more references to corresponding projects and providers [DVB09].

## 5.3 Content caching

### 5.3.1 Functional analysis

A commonly used mechanism to enhance the performance of CDNs is caching. Caches have been deployed on the Internet for more than a decade in order to shorten transport paths by making a subset of the most popular web content available near users [GO00] [FIE10] [HH10]. In the simplest case, we refer to a cache as storage for content on a network node, but most often several caches are integrated under common management, to form multi-level CDN systems in a network domain [BGW10]. Caches on end systems are also common practice, which completely avoid part of the transmissions arising from repeated user requests for the same web pages. Caches in user equipment can save about 20% of transfer volume [CHA10]. They are most welcome on air interfaces in mobile networks and wherever bandwidth on the last mile is limited and expensive. In addition to shortened round trip time for requested content and load balancing support in the network, the MNO can save interconnect cost due to the reduced amount of data volume received from other network domains. In the following we have listed some criteria for content, which is in principle qualified for caching:

- Size of content is not too small
- Content can be associated with trusted source (no virus, legal and ethical issues)
- Change tendency: content change / update should not happen frequently
- Significant popularity: there is a considerable number of user requests
- Copyright issue: protected content requires special agreements between MNO and content owner
- Application types: caching seems to be less relevant for certain applications, e.g. conversational services
- Generalization: non personalized content may be related with higher number of requests

#### 5.3.1.1 Non-transparent vs. transparent caching

Caching mechanisms can be applied in a way that is non-transparent for an application. Here, the term non-transparent means that the entities in the application overlay know about these caching mechanisms and make deliberately use of them. The second choice for caching mechanisms is the way when they are applied transparently. In this case, the overlay entities are not aware of the existence of a caching entity.

Network level caching relates to a low-level cache of small pieces of information. It is not application-specific, and instead reduces bandwidth for all TCP traffic. Application level caching is protocol specific.

Non-transparent caching was already investigated in the MoPi architecture [HST+09]. This architecture introduces two additional elements, the Cache Peer (CP) and the Index Server (In). Both additional elements can be implemented as stand-alone entities or as functions of another network node e.g. SAE-GWs. They are also visible on application layer.

The CP is considered by the other peers as an ordinary peer and speaks the same protocol as the user peers. A major problem with non-transparent caching is the discovery of caches. This problem is solved in the MoPi architecture by offering a discovery service for content by an operator-controlled entity, the Index Server (In). The In can relay download request to the CP and, in addition, it can instruct the CP to cache popular contents. The major disadvantage of non-transparent caching is how to motivate the users to query the Index Server. This is of particular interest since most of the P2P CDN users are distrustful to operators.

Network level caching mechanisms can be further distinguished by identification mechanisms of cacheable user data. The first identification mechanism for this kind of caching is Deep Packet Inspection (DPI). Here, the payload of the IP packets is investigated and might be eventually changed, e.g. for the redirection of the application data flows to certain cache entity. The advantage of transparent caching using DPI is the more versatile applicability and fine grain decision rules. The disadvantage of this caching mode is high computational power required for DPI, which has to be performed at line speed for a huge number of flows in parallel. In addition, rules have to be specified for every application protocol on which the caching mechanism has to react for. Since these protocols might change frequently, the definition and the updating of the rules might require a significant amount of management effort.

Apart from DPI, the second identification mechanism for transparent caching is hashing. Here, hash values are computed for each IP packet in the user data stream possibly in combination with mechanisms such as deep packet inspection. If a hash value is seen for the second time, then only the hash value is forwarded to egress. The egress will act as a cache and replaces the hash value by the actual data (which has been sent previously and which was stored). In general, this mechanism requires locating caches as near as possible to the UE and the identification elements very close to the ingress.

#### 5.3.1.2 In-line vs. out-of-band caching

In-line caching means that the entity responsible for caching content is located within the data path. Such an entity may listen into the stream of user data and terminate a TCP connection, if requested content is detected. This mechanism does not require redirection of the request to a different location. However the approach is insufficient for user mobility, if the point of convergence between old and new access (anchor point) is further upstream than the cache engine. On the other side, out-of-band caching implies that a cache is located on a node which is out of the original path. A more detailed analysis of the pros and cons of the different approaches is for further investigation especially in the context of mobile networking.

#### 5.3.1.3 Caching metrics

This subsection provides a short overview of the relevant metrics for caching. These might be used in order to determine appropriate key performance indicators (KPI) for the evaluation of a caching solution and therefore is for further specification. The following is considered:

- Cache out capacity: This is the data rate, which can be served by client requests in total and may influence the performance.
- Cache hit rate: This could be measured in content hit rate or byte hit rate. The first determines the successful requests divided by the total amount of requests, while the latter defines the total number of bytes served from caches divided by the total amount of data requested by users.
- Cache response time: This determines how long it takes to get the first data to the user. Cache response should be determined by RTT plus additional value for the connection setup and processing of request.
- Cache size: The size is associated with the amount of data to be stored. Even though price for memory is constantly declining there might be a performance issue with large cache sizes, since access to content increases with a large number of objects to be maintained.
- Number of control message exchanged: This is related to maintain content changes in cache (e.g. to add, modify and delete content) and also to enable cache synchronization.

### 5.3.2 Caching efficiency for web-based content

Before the turn of the millennium, many network providers had installed caches for HTTP-based client server applications, which proved to be efficient despite a non-negligible portion of one-timers, i.e. web pages with only a single request over a long period that rendered caching useless [GO00]. Afterwards, P2P networking comprised the lion’s share of Internet traffic, which bypassed classical HTTP-based web caches and made them inefficient. In recent time, user-generated content has become even more relevant, but server platforms like YouTube for video distribution and community networks have grown faster than data exchange via P2P, with the result that the portion of cacheable HTTP-based content within IP traffic is increasing [ASK+10] [CFE+06] [CIS10].

Caches can be installed on different aggregation levels in broadband access networks or especially at network boundaries to take load from expensive peering or transcontinental links, as illustrated in Figure 5-7.

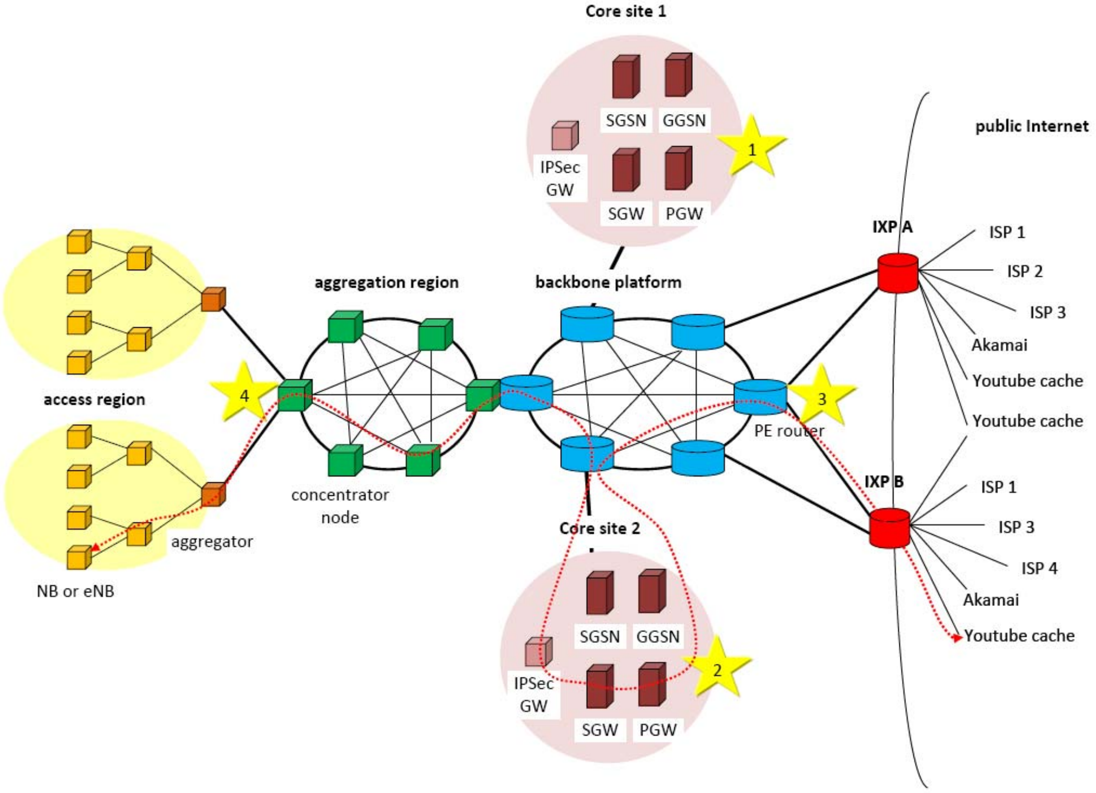


***Figure 5-7: Caches in broadband access networks***

Business case evaluations for caching have to balance the costs of storage and of managing the caching system with the costs saved due to reduced traffic and transmission capacity [HH10]. Improved user experience due to smaller delays also has a positive effect on a service provider's business. When user access patterns are assumed to be random and homogeneous, the access frequency for an item is linear, increasing with the user population. As a result, cache efficiency improves when serving larger populations, while the cache size usually also increases with population. Comparing caches on different aggregation levels, a higher level requires a smaller number of caches, each of which is serving a larger population, but, as a trade-off, the hop distance to the user also becomes greater. Within an administrative domain of a network or CDN provider, caches on different levels can be combined as a distributed and/or hierarchical system under common control and management. The optimization of algorithms and CDN architectures with regard to cache locations has been analyzed in efficient general solutions for tree-shaped structures, whereas heuristics seem to be the only choice for arbitrary network topologies [BGW10] [KAM10].

### 5.3.2.1 Zipf laws to estimate requests for popular content

Zipf distributions are frequently observed in Internet statistics or, more generally, when a large population has access to a large set of items. They indicate the relevance of a small set of the highest-ranked items, such that most requests address a fraction of 10% or less of highly popular items. This property is a basic precondition for caching efficiency. According to the Zipf law, an item of rank $R$ in the order of highest access frequency attracts a number of requests equal to:

$$A(R) = \alpha R^{-\beta} \qquad (\alpha > 0; \beta > 0), \qquad (1)$$

where the parameter $\alpha = A(1)$ is the maximum number of requests observed for an item in the statistics [HHB09]. Alternatively, the same term (1) holds for the probability distribution or the fraction of accesses to the item in rank $R$ if $\alpha$ fulfills a normalization constraint $\Sigma_R A(R) = 1 \Rightarrow \alpha = 1 / \Sigma_R R^{-\beta}$. The exponent $\beta$ determines the decay in access frequencies to the items and thus the variance of the distribution, which increases with $\beta$. A century $0.64 < \beta < 0.85$ ago, Vilfredo Pareto introduced the related form of Pareto distributions for property and income over the population, which also characterizes a major influence of a few extreme outliers on the entire mean value and the variance [REE01].

The relevance of Zipf laws has been confirmed in numerous case studies, e.g.

- for the page requests on popular websites by Breslau et al. [BRE+99] in a series of measurements yielding a range 0.64 < $\beta$ < 0.85 for the exponent,
- for access to content delivery platforms, including YouTube [CHA+07] [GAL+07], where $\beta \rightarrow 1$, America Free TV [EUB08], Amazon and P2P networks,
- for cross-references on the Internet and in literature, for relationships in large social networks, or for the frequency of words in a long text, etc. [EUB08].

Figure 5-8 shows the typical shape of Zipf distributions with regard to the influence of the exponent $\beta$ on the top figure and of the size $N$ of the set of items on the bottom figure.

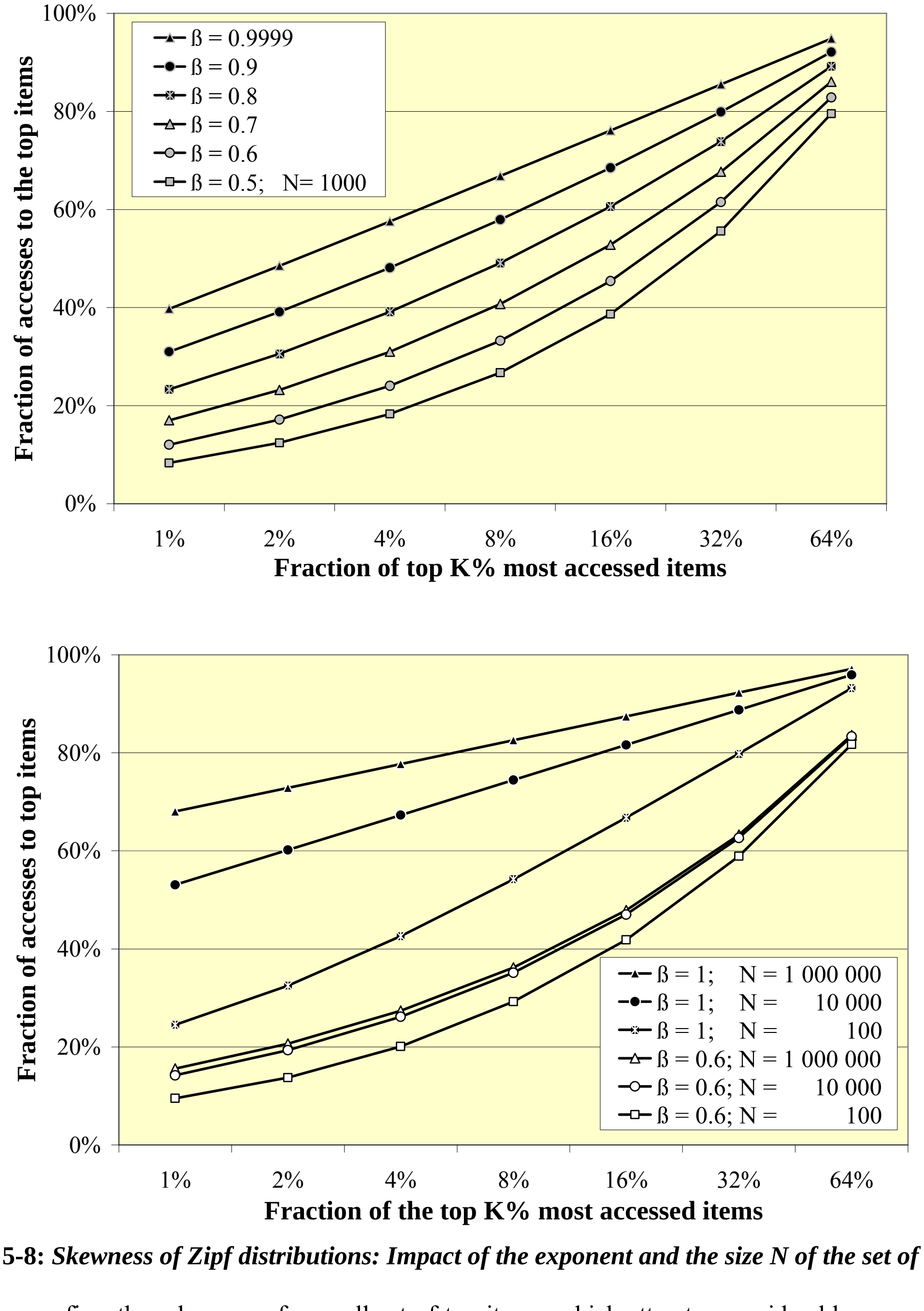


**Figure 5-8:** ***Skewness of Zipf distributions: Impact of the exponent and the size N of the set of items***

The curves confirm the relevance of a small set of top items, which attract a considerable amount of all requests. The fraction of accesses to top elements grows with the exponent $\beta$, such that accesses to the top 1% can be varied in a range of 10% - 40%. Nevertheless, a good adaptation for the top-ranked items goes on account of deviations for the mass of seldom accessed items [BRE+99] [EUB08] [GAL+07] [HHB09]. An alternative distribution form with three parameters for more precise adaptation to less skewed access is proposed by [GUO07]

$$A(R) = (b - a \ln(R))^{c} \text{ with positive, real-valued constants } a,b,c.$$

Developments of the population of items over time have been investigated for BitTorrent files, YouTube videos [CHA+07], and IP-TV channel access [KBB10], which typically start with a steep ramp to maximum popularity, followed by a smooth fade-out. The last study [KBB10] reveals deviations from a Zipf distribution and suggests utility functions of different shapes with more parameters yielding a better fit to include the popularity trend of an item over time.

Another way to improve cache efficiency is to pre-fetch data if information about changing popularity is available, such that new items are expected to become important for future requests. Most relevant is pre-fetching by users who can initiate a download in advance to have it available on time and for delivery in a fixed schedule, e.g. in radio and television programs. Pre-fetching can have different effects on the traffic load: bandwidth and storage may be wasted if pre-fetched content is not subsequently needed, whereas systematic pre-fetching in periods of low network load reduces peak traffic load and improves resource utilization.

#### 5.3.2.2 Measurement of access patterns and evaluation of the effect on caching

Measurement studies of access patterns for YouTube videos have been made available [CHA+07] and analyzed with regard to the achievable hit rate [HH10]. In addition to the static distribution of requests in the long-term trend, the fluctuation in the popularity of items over time is also considered. The popularity of data is observed to slowly vary in a timescale of hours, days or months, even if some new data may become highly popular in the short term [GAL+07].

As a result, static modeling based on the access probabilities over the last hour or a longer timeframe is appropriate for analyzing cache hit rates. For the static case, it is most efficient to keep the videos with highest access probabilities in the cache. The results shown Figure 5-9 have been evaluated from 3.7 billion accesses with references to 1.69 million YouTube videos [CHA+07]. The curve in Figure 5-9 for keeping the most popular items in the cache shows the achievable hit rate depending on the cache size, where a 40% cache hit rate can already be expected with only 1% of the video data volume being cached.

Although access patterns vary only slowly [CFE+06], the cache has to constantly react to changes by introducing a replacement strategy. A simple and usual caching strategy removes the Least Recently Used (LRU) item from the cache whenever a request addresses a new item that does not fit into the cache. In our comparison of LRU with a selection strategy for cache content based on access statistics over a limited timeframe [HH10], essential performance deficits of LRU became visible especially for small caches, as expressed by the difference between both curves in Figure 5-9. Similar gaps between optimum and LRU hit rates are observed for Zipf distributions with parameters in the range $0.5 < \beta < 1$. Thus, it seems worthwhile including request statistics in cache replacement strategies.

Moreover, there are a number of obstacles which detract from cache efficiency for network providers, starting with a lack of standardization regarding the information exchange between original content servers and independent caches. The HTTP draft standard as documented in RFC 2616 [FIE00] and an active Internet draft proposing an update [FIE10] have introduced a number of cache control options, such that the original content server can mark data as cacheable or non-cacheable and can determine policies for expiry, refreshment and allowable delivery of cacheable data. Measurement statistics reveal that a large portion of the HTTP traffic is marked as cacheable with relaxed or no expiry constraints [ASK+10] [CHA10]. Large content volumes in streaming and download applications, which dominate IP traffic, are usually permanent and not subject to temporary changes. On the other hand, notifications from a cache to the content server on successful or failed delivery from the cache do not seem to be supported, although they would be desirable, e.g., to track count statistics for accesses on the original server.

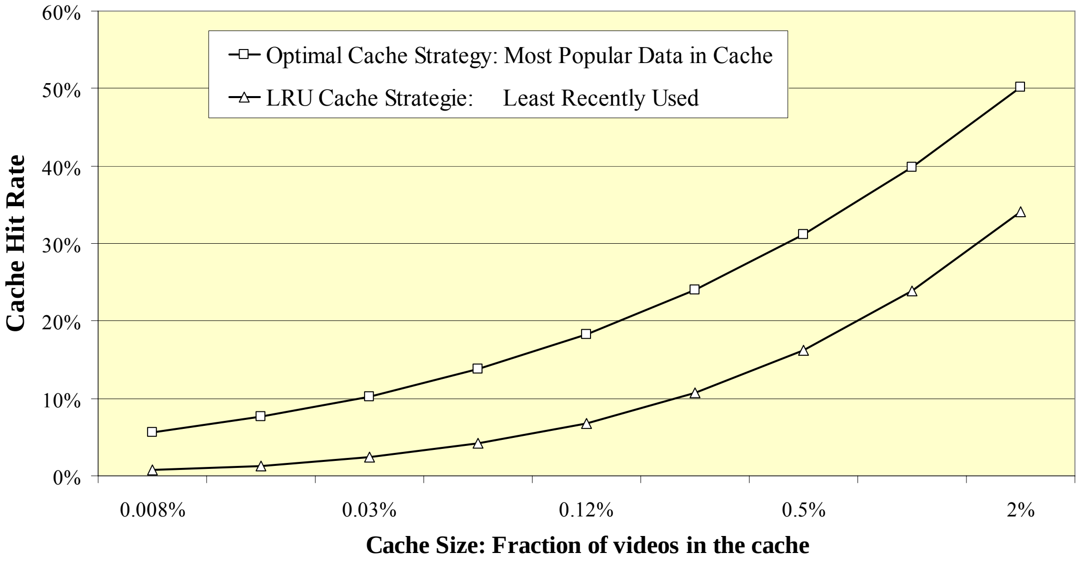


**Figure 5-9:** ***Hit rates depending on the cache size evaluated from a YouTube trace [CHA+07]***

A lack of cooperation beyond the administrative boundaries of content and network providers impedes quality-of-service support in general and optimum end-to-end transport paths in particular. CDN support of large content providers usually ends on peering links at the boundaries of the network platforms performing independent internal traffic engineering policies. In order to make full use of caching, unique hash identifiers are proposed for each content item to detect and avoid duplicate transport over the same path [MCK04], as also successfully applied in P2P networks, but again there is no standardized solution of this kind.

As a result, deep-packet inspection techniques often have to be implemented to identify different requests addressing the same content in the cache. This can impose a considerable additional performance burden on caching systems, even if Web 2.0 provides helpful meta-data to improve the efficiency of caching [ASK+10] [GAL+07]. Content platforms like YouTube often personalize the handling of requests in a way that makes it even more difficult to classify requests addressing the same item based on HTTP data. Nevertheless, a personalized client-server dialog could be separated from included data download phases, such that both the content and the network provider would benefit from keeping large data volumes in caches close to the user. Consequently, it is difficult to predict the efficiency of caching if there is no direct cooperation with the content providers and a continuous update process to changing identification patterns is required. Alternatively, network providers could allow the extension of global CDNs into their broadband access platforms, but there are several drawbacks to such an approach. The network provider would have to cope with different solutions from several major content providers. Unpredictable traffic sources inside the network platform would make traffic management and network planning more challenging. Finally, net neutrality could be violated if enhanced support for some large CDNs is not available for other content.

Moreover, a duplicate detection and suppression method is worth mentioning [MRT10], which is applicable to data segments of a few hundred bytes in IP packets as a link or network layer approach including caches on or even below the network layer. An implementation for traffic reduction on a link puts a compressor device on the input side and a decompressor on the output side. Both the compressor and the decompressor assign and store the same short digests to appropriately segmented data chunks transmitted over the link. The decompressor stores the data chunks for complete IP traffic flows in a cache together with their digests in the same sequence as they were transmitted. When an IP traffic flow in the cache or a part of it is repeated, the decompressor detects duplicates and returns the digests for the next data chunks of the flow to the compressor. If more data chunks of the duplicate flow appear at the compressor which match the returned digests, the corresponding IP packets are suppressed, and a notification is sent to the decompressor to reconstruct the suppressed packets from the cache.

The main advantage of this method is that it does not need to refer to application layer context and does not affect the end-to-end communication except for a small delay introduced at the decompressor to reconstruct IP packets from the cache. Several investigations into the efficiency of the method estimate the saving potential to be in the range of 10-30% of IP traffic volume, which, again, would be most appreciated on expensive links.

## 5.4 Content based measurements

Video content is typically static, and video popularity on platforms like YouTube is assumed to be Zipf distributed [CKR+09], [ZKG10]. Thus video content is seen to be a good candidate for caching. However, several properties of YouTube traffic might have negative impact on cache hit rates and cache performance: One important factor is video popularity. Global popularity of YouTube videos, measured in view counts, must not necessarily match local peculiarities in a specific operator network [GAM07]. Popularity distributions may differ slightly depending on the network that a cache has to serve. Our work aims at quantifying the caching potential of video traffic in end-networks. We picked the YouTube video platform for our study because it is one of the major platforms for user generated video distribution, and has therefore received a lot of attention from researchers, operators and network equipment vendors. For our work, we monitored and analysed all YouTube video traffic from an end-network with more than 120,000 users over a period of a month [BKC+12]. The traffic is examined with respect to relevant parameters for network caches. Building on this traffic evaluation, the caching potential and the reduction of downstream network traffic that can be avoided with caching is estimated.

### 5.4.1 State of the art on measurements

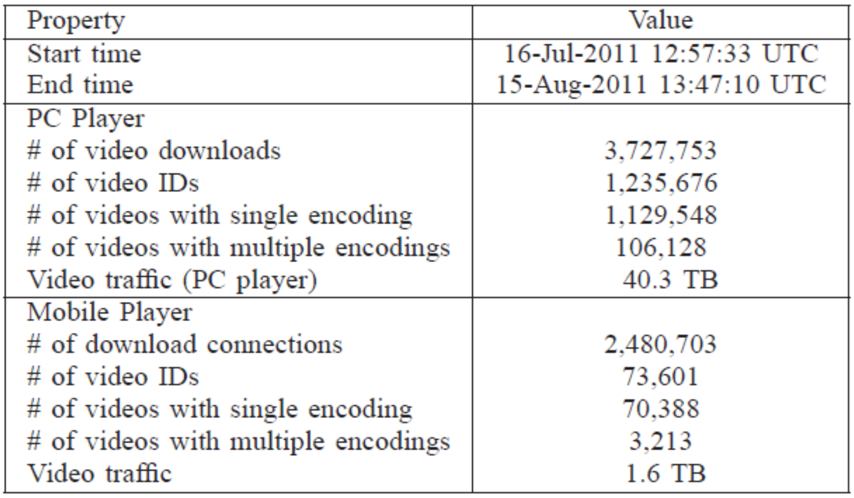
Related work can be grouped into several categories: Some papers discuss on the shares of YouTube traffic in the overall traffic mix. Others focus on YouTube traffic characteristics, YouTube's infrastructure, or caching of video content. Popularity of YouTube videos has been studied from several points of view. One branch of papers describes active crawling of YouTube sites to determine popularity or try to find reasons for popularity of various videos [CSM10]. Figueiredoet al. [FBA11] focus on popularity development over time. Their findings conclude that a lot of videos show a viral popularity growth, indicating potentials for caching. Others find power law patterns with truncated tails in global video popularity distributions and predict good caching potentials [CKR+09], [CKR+07]. Video popularity has also been studied from network local perspectives [ZSG+08], and local and global video popularity have also been compared [GAM07]. These studies show that the global popularity of video content does not have to match the local popularity. For example, Gill et al. [GAM07] find that the Top 100 videos from global YouTube video rankings are viewed in their network but do not have any significant contribution to the overall amount of YouTube traffic observed. Caching strategies must therefore consider local popularity and view counts. Other work tries to provide a better understanding of the YouTube web application or YouTubes' infrastructure Such work includes attempts to inspect YouTubes' policies for picking local data centers for video downloads [TFR+11], or describe load-balancing or traffic asymmetry from the view point of a Tier-1 provider [AJZ10]. Finamore et al. [FMM+11] assess YouTube traffic by evaluating several YouTube traces from different networks. They study traffic and video patterns for YouTube videos from mobile and PC-based devices, and show differences between traffic of these devices. Caching potentials have been considered by Ager et al. for different protocols in [ASK+10] where they outline good potentials for caching HTTP in general. Zink et al. evaluate caching of YouTube videos [ZSG+08]. In their study, the authors collect three one-week traces from a university network and use these as input for cache simulation. The authors consider client, P2P and proxy caches and estimate cache video hit rates. They conclude that high video hit rates can be achieved even with small caches. We extend their work by accounting further important factors such as video encoding formats, aborted video downloads and their impact on caches. Furthermore, we do not only consider video hit rates, but more sophisticated metrics such as content hit rates. Using these metrics, it can be shown that other caching strategies, such as chunk-wise caching strategies, provide better cache performance than the previously proposed and evaluated caching.

### 5.4.2 YouTube traffic data sets

This section builds the base for our YouTube traffic study. We introduce our monitoring setup that we used to observe and analyze YouTube traffic. Our vantage point was deployed in the Munich Scientific Research Network (Münchner Wissenschaftsnetz, MWN) in Munich, Germany. The research network interconnects three major universities and several affiliated research institutions in the area in and around Munich, Germany. Furthermore, the network includes several student dorms that provide housing for the students enrolled in the universities in Munich. In total, the network hosts about 80,000 devices being available to approximately 120,000 users. The Leibniz Supercomputing Center (Leibniz-Rechenzentrum, LRZ), as the operator of this network provides Internet access for all its users via a 10 GBit/s link to its upstream provider the German research network (DFN). Our vantage point was deployed on the border gateway between the MWN and its upstream service provider. Due to this deployment, we were able to observe both office related as well as residential video traffic. Our monitoring setup was built around standard of the shelf PC hardware, operated by a Linux-based operating system. All traffic properties where calculated during an online monitoring run, as we were not able to store the many Terabytes of YouTube traffic that were observed during our monitoring period. We used an optimized capturing setup, including our improvement presented in [BDK+10], based on TNAPI [DER10] to build a multi-core aware monitoring system. The measurement was conducted with the tstat [TSTAT] tool, which has been used for monitoring YouTube traffic [FMM+11].

The monitoring setup was used to log information about the YouTube video downloads for a period of one month. We collected this data in order evaluate long-term statistics of caching relevant parameters. Table 5.2 describes the data set obtained throughout the monitoring process with a distinction between PC player and mobile player traffic.

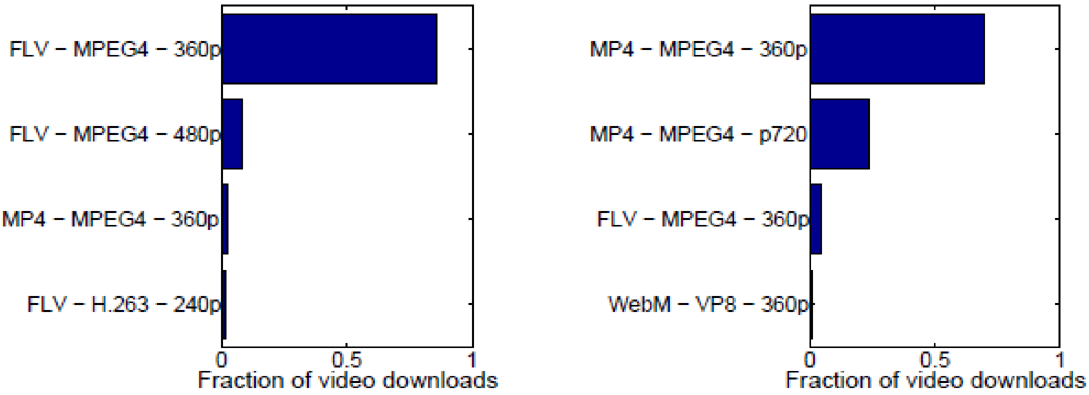
**Table 5.2:** ***Monitoring Data Overview***



| Property | Value |
|---|---|
| Start time | 16-Jul-2011 12:57:33 UTC |
| End time | 15-Aug-2011 13:47:10 UTC |
| PC Player | |
| # of video downloads | 3,727,753 |
| # of video IDs | 1,235,676 |
| # of videos with single encoding | 1,129,548 |
| # of videos with multiple encodings | 106,128 |
| Video traffic (PC player) | 40.3 TB |
| Mobile Player | |
| # of download connections | 2,480,703 |
| # of video IDs | 73,601 |
| # of videos with single encoding | 70,388 |
| # of videos with multiple encodings | 3,213 |
| Video traffic | 1.6 TB |

Similar to [FMM+11], HTTP return codes can be used for this distinction: Video requests from PC players are answered with a HTTP 200 OK return code, while mobile video requests are answered by 206 Partial Content. Mobile downloads are only responsible for a small share of the overall video traffic in the measurements (1.6 TB for mobile downloads vs. 40.3 TB for PC downloads).

However mobile downloads are responsible for quite a large number of connections. This is due to the download procedure, which uses multiple connections for each video.



(a) PC player (b) Mobile Player

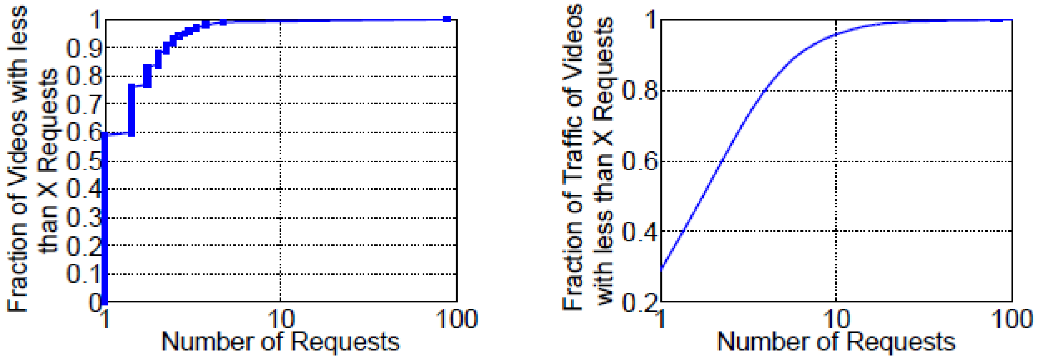
**Figure 5-10:** ***Video encodings for PC and mobile traffic***

One interesting difference between mobile and non-mobile traffic can be found in the video encoding statistics, as shown in Figure 5-10. Most watched videos on a PC platform are transmitted as MPEG-4 AVC (H.264) encoded video with a 360p resolution which is embedded into a flash container. Mobile videos are usually not embedded into flash containers, but are downloaded in a MP4 container. As for video content, the same encoding is used with a 360p resolution. Hence, if the same video is watched with a mobile and a PC-based player, there is a high probability that a cache needs to deliver a completely different video (from a cache's point of view) for the same requested video. Operator networks that provide network access to an equal amount of mobile and PC based devices, might therefore have to cache a lot of videos twice due to different encodings. Networks that have only one dominant type of video requests, such as the investigated network, might thus employ a smaller cache.

The biggest share of the most popular videos were advertisements. Seven of TOP 10 videos can be placed in this category. Possibly these videos were embedded into non-YouTube related sites and automatically downloaded by users who have visited those sites. The TOP 30 videos have a view count of more than 1500 and most of them are clips with a run time of less than two minutes. In the following section, we discuss several parameters that are relevant when caching video traffic.

### 5.4.3 Evaluation

For the sake of brevity and due to the fact that PC player traffic is dominant in the observations, further discussion are constrained to PC player traffic. There are several important factors of video traffic that have large impact on a cache. These properties include video sizes, number of views of a single video or the inter-request times of multiple views. It is distinguished between videos from a caches' point of view: Two videos are considered to be different if they have a different YouTube video id, or if they share the same video id but are encoded in different formats. In the following, we use the term *video* to address unique video content. Furthermore, the term *request* corresponds to a partial or full request of a video, while the term *view* indicates a full download of a video. Figure 5-11a presents the share of videos out of the observed videos that have a particular number of requests. Our data reveals that about 60% of all videos are only requested once in the monitoring interval which results in an average number of 2.7 requests per video. The remaining 40% of the videos can be cached and delivered from the cache to other clients for subsequent requests. The majority of the videos have been requested ten or less times, but some of the videos are watched several hundred or even thousand times. The huge share of videos that are viewed only once or a couple of times on first glance could lead to the assumption of only little potential for caching. However, if traffic volumes are considered, different trends can be observed: Figure 5-11b plots the amount of video content delivered from the YouTube servers for videos that have less than a certain amount of views. The majority of videos that have been requested only once are responsible for only approximately 30% of the video traffic. Videos that are watched more than once account for the biggest part of the traffic, which emphasizes the potential of in-network caches.



(a) Requests per video

(b) Overall traffic depending on the number of requests

**Figure 5-11:** ***Video request counts and their impact on generated traffic***

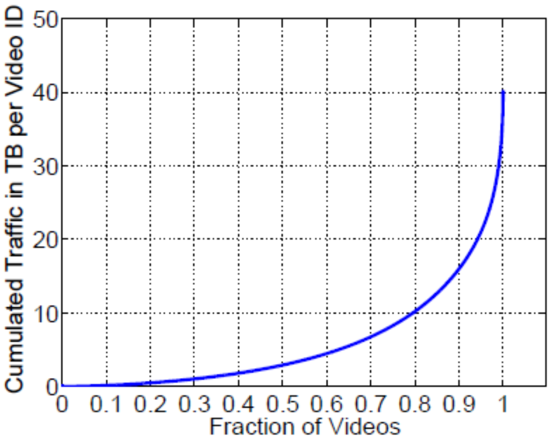


**Figure 5-12:** ***Total requested data of all videos***

Figure 5-12 summarizes the influence of individual videos on the overall amount of download traffic. The graph shows the sum of the requested data per video sorted by traffic size in order to outline the traffic contribution of the individual videos. One can see that 80% of the videos are responsible for about 10 TB of traffic, while the remaining 20% account for 30 TB of all downloaded video data. By identifying and caching such high-profile videos, a cache can significantly decrease the amount of download traffic and achieve good cache hit rates without the need of a large storage capacity, since 20% of the videos generate 75% of the video traffic.

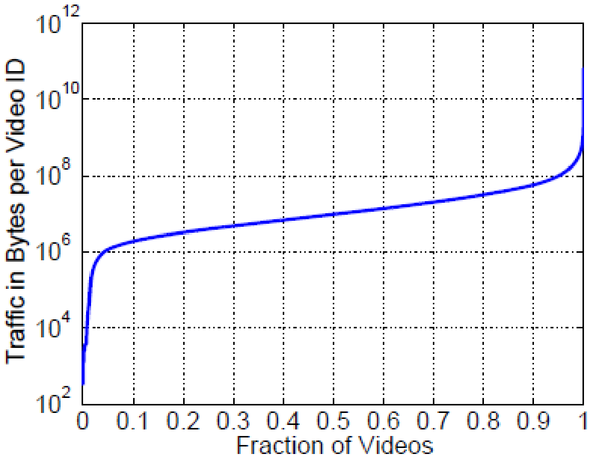
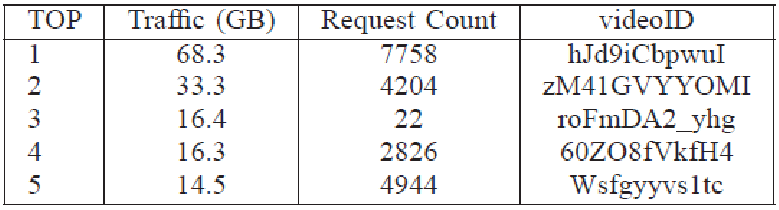


**Figure 5-13:** ***Total requested data per video***

For individual videos, a similar trend can be recognized. Figure 5-13 plots the amount of traffic per video sorted by traffic size. The amount of traffic is summed over all requests of the video. The plot indicates that a very small fraction of videos only contribute a very small share to the overall amount of video data. These videos were probably watched only for a single time and/or were aborted before being downloaded completely. Main observation is that the traffic contribution per video is between 1MB and 100MB for 90% of all observed videos. Request size and number of requests per video are the two factors that contribute to the amount of traffic that is generated by a single video. Very large videos are responsible for a large amount of traffic even if they are watched only a few times. Small videos that are viewed very often can also accumulate a lot of traffic over a larger time interval. From a caches' point of view, videos with high request rates are most beneficial for hit rates and cache efficiency. Table 5.3 lists the amount of traffic and the number of views by the TOP 5 videos which generated the most download traffic.

**Table 5.3:** ***Top 5 video characteristics***



| TOP | Traffic (GB) | Request Count | videoID |
|---|---|---|---|
| 1 | 68.3 | 7758 | hJd9iCbpwuI |
| 2 | 33.3 | 4204 | zM41GVYYOMI |
| 3 | 16.4 | 22 | roFmDA2_yhg |
| 4 | 16.3 | 2826 | 60ZO8fVkfH4 |
| 5 | 14.5 | 4944 | Wsfgyyvs1tc |

As one can see, the characteristics of the Top 5 videos in terms of generated traffic load and number of views differs significantly. The heterogeneity of the videos is indicated by the number of requests, e.g. the video that is ranked third is only requested 22 times but contributes 16.4 GB of traffic whereas the video on fifth place is requested 4944 times. The same characteristics can be recognized for the top 20 videos in terms of requests which we cannot list due to space limitations. However, the number of requests is still a reasonable decision factor whether a video should be cached or not since the average request size of all videos is 12.8MB. Besides the amount of generated traffic per video, request patterns play an important role for the efficiency of a cache. Now, we take a closer look on the average inter-request time of the videos and the time period during which the videos were requested. Due to the heterogeneity of the videos, we decided to evaluate both characteristics for different groups of videos. The videos are grouped according to the number of requests which reflects their popularity. The following intervals were used to group the videos: Low: 2 – 9 views, Medium: 10 – 99 views, High: more than 100 views. These groups are referred to as low, medium and high popular videos, respectively. Figure 5-14 reveals that the average inter -request time of videos differs significantly depending on the number of requests. 95% of the average inter- request time depending on the number of requests of high popular videos are between 1000 and 10000 seconds which is a strong contrast to those with less requests.

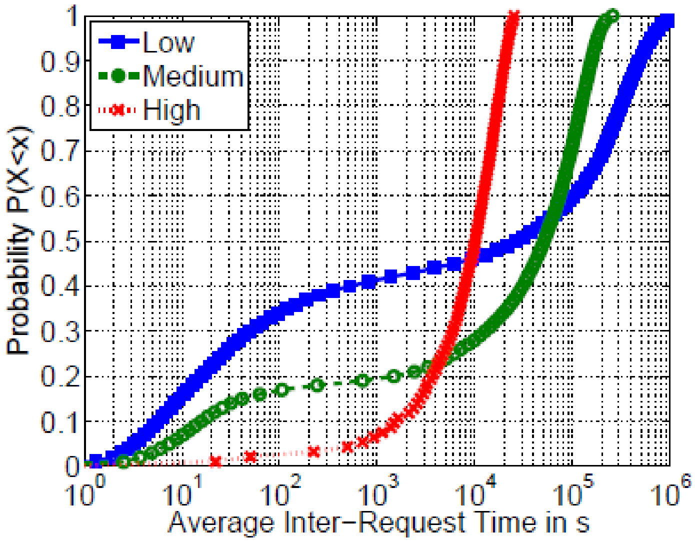


**Figure 5-14:** ***CDF of average inter- request-times***

The videos with low or medium popularity have a much higher variance of their inter-request times. Their results show that the majority of these videos have an average inter-request time of approximately 30 seconds. We assume that this represents a typical value for non-popular videos, which are posted in social networks. Thus, it is likely that friends will request the posted video resulting in a couple of full or partial downloads within a rather short period of time. In addition, a second peak can be recognized for videos with less than 100 requests for an average inter-request time of approximately one day which results in a bimodal distribution. Another important characteristic for caches is represented by the request period which is the time difference between the first and the last request of a video. According to our definition, the maximum of the request period is limited by the monitoring period. The cumulative distribution function of the request period of videos with low, medium and high popularity are shown in Figure 5-15.

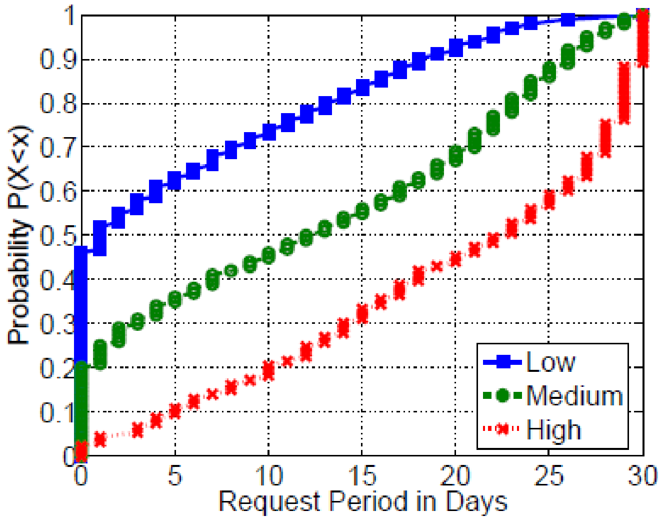


**Figure 5-15:** ***Request period***

The figure points out that depending on the number of requests, videos with a smaller number of views tend to have a shorter request period. 44% of videos with less than 10 requests have a request period of less than a day. This share decreases for videos with medium and high popularity to 20% and 3%, respectively. More than 50% of the high popular videos have a request period of more than two weeks. Almost 12% of all videos were requested over the whole monitoring period which shows that a significant amount of videos are popular over a long time-period. Cache sizes are very important for the estimation of caching benefits. Due to limitations in cache sizes, videos that are no longer watched need to be removed from a cache as soon as its disk is no longer able to store new videos. A video should not be removed if the probability for a subsequent request in the near future is still high. For this reason, we evaluated the probability that a video is requested at least one more time depending on the number of previous requests. Figure 5-16 shows the complementary probability of this event in order to provide a higher readability.

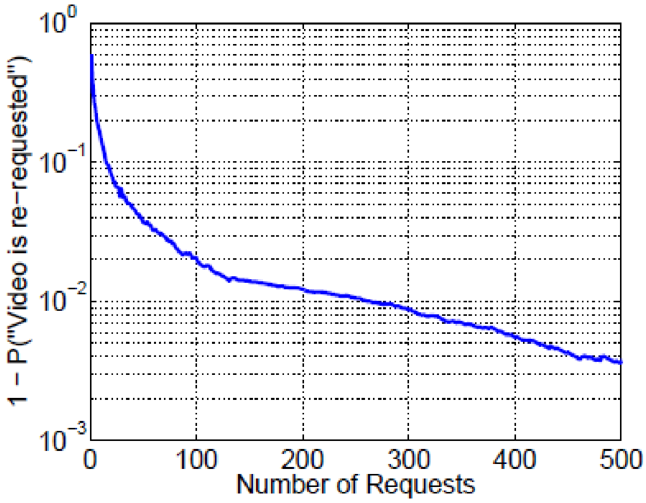
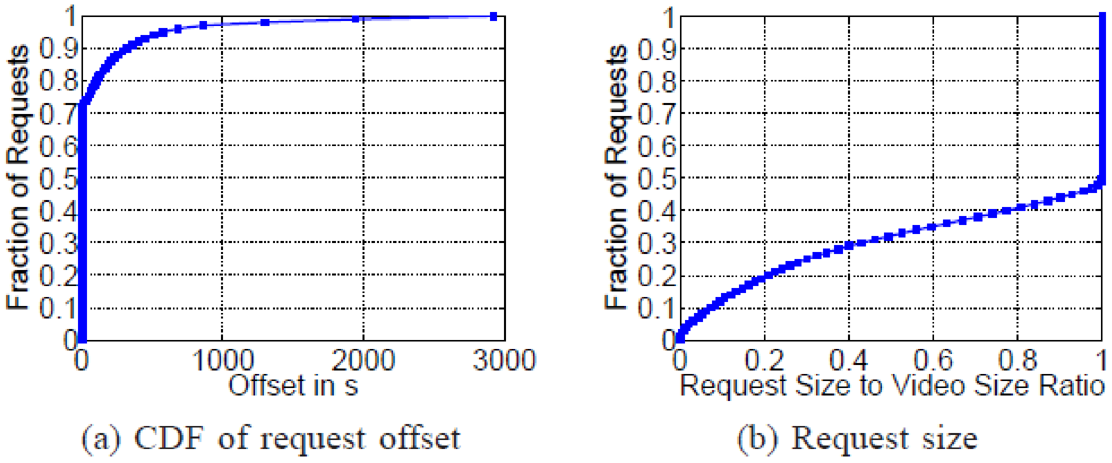


**Figure 5-16:** ***Probability for a video not being requested one more time***

The probability that a video is requested at least one more time increases with the number of requests. The re-request probability of a video that was requested one time in the past is already 60%. This probability increases to 86% for videos that were requested 10 times and exceeds 98% for videos that were requested more than 100 times. The trend suggests that this probability converges against 99.9%. However, the number of videos with such a high number of requests was too low during the monitoring period to support such a statement with a sufficient level of significance. YouTube users do not necessarily watch videos from the beginning since embedded YouTube videos can directly jump into a specific part of video by specifying a starting offset. Furthermore, if a user forwards a video to a not yet downloaded offset, a new TCP connection will be opened which starts a new download beginning from the chosen offset. Figure 5-17a shows the CDF of the offset of all requests. The figure reveals that 72% of all requests have an offset of zero Thus the majority of the users starts at the beginning of a video, which is very beneficial for caches. Only 3.5% of all requests have an offset greater than 1000s due to the fact that the average video duration is 331s. In addition, users can abort a video download before the video has been downloaded completely. This can happen for several reasons, such as the user experiences bad download quality or is not interested in the video [14]. Therefore, we evaluate the behavior of the users by calculating the fraction of request size and video size in order to track how much the user has watched. The results are plotted in Figure 5-17b.



(a) CDF of request offset

(b) Request size

**Figure 5-17:** ***Request characteristic of the video***

The figure shows that more than 50% of the requests are associated with the whole video. Another 20% still request almost half of the video while only a very small fraction requests a small part of the video. Videos that are not watched completely do not need to be fully cached. A cache has to decide whether the complete file is downloaded when a new video is requested, or if it only stores the requested bytes. It is therefore examined whether certain parts of a video are watched with a higher probability, e.g. if the beginning of a video is more likely to be watched than the middle or the end of the video. Thus, we divided each video into chunks and calculated the probability for each chunk to be downloaded. As offsets are defined in milliseconds, we need to calculate byte offsets. For this calculation, we use meta information such as total video length (bytes and duration), which we could only obtain from the flash containers. As a result, we can only evaluate the chunk information for videos which where embedded in flash containers. Figure 5-18 shows the probability for different parts of the video to be watched in a given download.

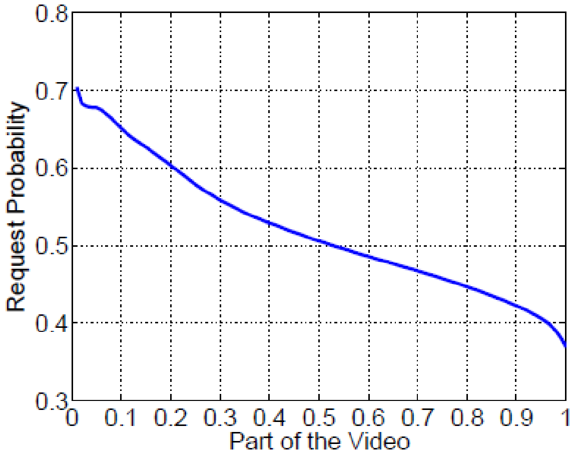
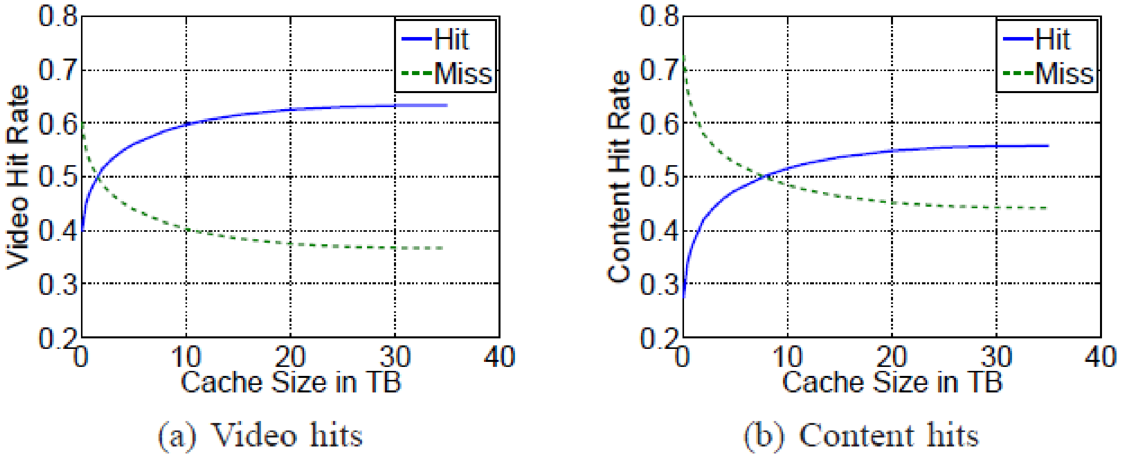


**Figure 5-18:** ***Request probability of different video chunks***

We observe that not all chunks are viewed with the same probability. Video parts from the beginning of the video are more likely to be viewed than the latter parts of the video. The probability for a chunk to be viewed is decreasing with its distance from the start of the video, which is probably due to the fact that users abort a video before it is completely downloaded. We will study the effect of this finding in the following from a caches' point of view. If a cache loads and caches unnecessary chunks, it will on the one hand download too much content. This content will on the other hand fill up disk space which is necessary for caching other relevant chunks. For our evaluation of caching benefits, we use our monitoring data as input for a simulation. The simulation aims at answering the question: ”What if a YouTube video cache had been deployed in the network during our monitoring period?” We calculate benefits that could have been provided by different caches and caching strategies. Caching strategies that define how videos are downloaded and replaced are very important. Another important factor is the disk size, which is a major limitation factor for cache performance. A caching strategy must decide for each user request, whether it will download the complete video or only those parts that have been requested by a user. Zink et al. [ZSG+08] propose to download complete videos upon user request and deliver subsequent requests from this video cache. They also propose a last recently used replacement scheme from the cache: If disk space is exhausted and a new video needs to be stored, the video that has not been requested for the longest time is removed from the cache. We implemented a simulation of this caching strategy and plotted the video and content hit rates for various disk sizes. A video hit is a user request for a video, which can be successfully answered from the cache. Video misses are user requests for videos that need to be fetched from the YouTube video servers. The same is applied to content hits and misses. Here we consider how many bytes of the request needed to be fetched from the YouTube servers and how many bytes could be delivered from the cache



(a) Video hits

(b) Content hits

**Figure 5-19:** ***Request characteristic***

Figure 5-19 shows the cache hit and miss rates for all video requests during our monitoring interval depending on the cache size. We simulated caches with disks sizes between 100 GB and 35 TB, in order to determine hit and miss rates. Similar to Zink et al., we can see good hit rates. About 40% of all requests can be delivered from a cache with very small disk sizes (e.g. 100 GB). A cache with 2 TB disk space, could achieve a hit rate of more than 50%. Our maximum achievable video hit rate is more than 60% for a cache that is able to cache all requested content which corresponds to the video re-request probability for a video as shown in Figure 5-16. However, a hit rate of more than 50% of the videos does not necessarily imply a high content hit rate. Figure 5-19b shows the content hit rate for caches of various sizes. We plot for each requested byte whether it could be delivered from the cache or whether it must be fetched from the YouTube infrastructure. Similar trends can be observed when looking at the hit and miss rates. However, 2 TB of disk space are not sufficient for a 50% hit rate in the cache. At least 8 TB are needed in order to achieve a content hit rate of 50%. The maximum content hit rate is smaller than the video hit rate, but still exceeds 55%. While these figures appear to be amazingly good, this caching strategy requires downloading the complete video. From the previous evaluation, it is known that parts of the videos are not necessarily downloaded. Figure 5-20 shows the number of bytes that have been fetched from the YouTube video servers depending on the cache size.

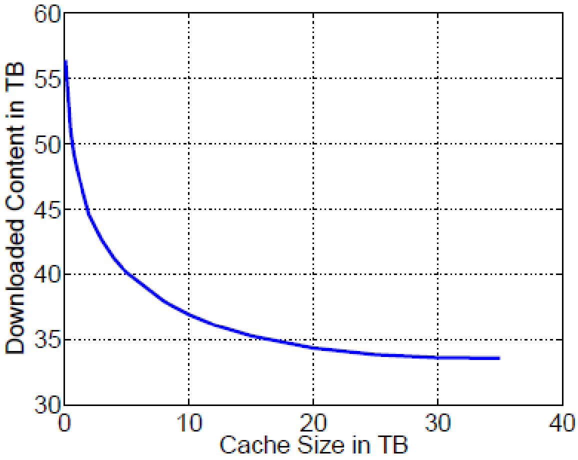


**Figure 5-20:** ***Downloaded content from YouTube video servers***

It can be seen, that this number is very high for small cache sizes and reduces to 33.6 TB with higher cache sizes. The reason for this is that all unique video content, if fully downloaded, results in 33.6 TB of traffic. However, users did not download this unique content completely, but only parts of it. This unnecessarily fetched data must be stored on disk and occupies disk space, which is needed for videos that are requested completely or requested multiple times. For small cache sizes, many important videos are removed from the cache, and need therefore to be downloaded from YouTube several times for subsequent user requests. It is therefore important not only to look at cache hit rates, but also on the number of bytes which have to be fetched from the YouTube video infrastructure. One more important conclusion, according to our monitored data, is that a caching strategy which fetches the complete content instead of the requested content, is not an efficient strategy. Thus, we evaluated a cache, which only stores content chunk-wise (chunk strategy): Videos are separated into 100 chunks, and chunks are only cached on user request. For the reasons outlined before, we can only consider flash content for this evaluation.

Therefore, the numbers for video data and requested content change: The complete size of the flash videos is 29.5 TB (compared to 33.6 TB for all videos). 9.7 TB of this video sizes where not viewed at all, e.g. due to premature download aborts. Storing these parts of the videos in the cache would unnecessarily occupy valuable disk space.

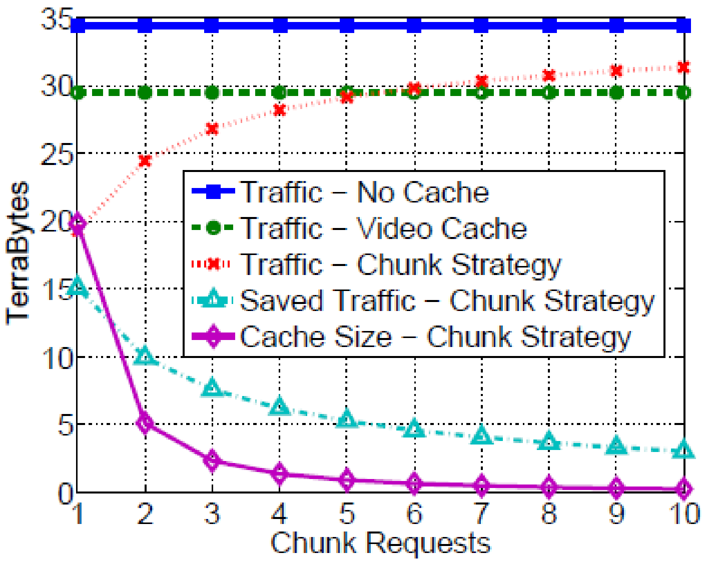


**Figure 5-21:** ***Chunked caching***

User requests to YouTube for flash content sum up to 34.4 TB of video strategies downloads, if no cache is used. A cache which downloads the complete video content if a video is requested (as simulated before), will download 29.5 TB of flash content from the YouTube provided that it is able to cache all requested videos. These two numbers are therefore the base line for our chunked caching strategy. A caching strategy has to provide mechanisms that decide when to store a chunk. Each chunk can be stored when it is requested for the first, the second, or more times. This decision has large impact on storage requirements and traffic reduction of a cache. Popular chunks need to be downloaded twice, three times or more before any cache hit can appear, thus reducing the benefits in download traffic. On the other hand, waiting for a chunk to be requested several times before caching reduces the required cache size. We evaluated the effects and benefits of a chunked caching strategy and plotted the results in Figure 5-21. The figure shows the cache sizes that are required and the traffic to the YouTube infrastructure, depending on the number of requests of a chunk before this chunk is stored. If we store chunks at their first request, a cache needs disk space of 19.5 TB for storing all chunks, and generates the same amount of traffic to the YouTube servers. Hence, when deploying such a cache, the amount of downloads from YouTube can be reduced by 15 TB. If we cache chunks on the second occurrence of a chunk, the required cache size drops to 5 TB (diamond markers), and the amount of traffic to the YouTube servers increases to about 25 TB (cross markers). The amount of reduced download traffic drops by this 5 TB (triangle markers), since popular chunks need to be fetched twice. By comparing the results of the chunked caching strategy with the complete download strategy (triangle markers vs. dashed line), we can see that a properly configured chunked caching strategy performs much better than a properly configured strategy that downloads complete videos. Furthermore, the chunked strategy allows to deploy smaller caches to achieve this high caching benefits.

# 6 Integration of technologies

For the derivation of the network architecture the coexistence of the currently deployed and the newly proposed technologies must be analysed and integration issues must be handled. Besides this the proposed architecture must meet system validation criteria regarding performance, deployment and technology maturity questions.

This chapter describes the integration questions and results for the technologies proposed to handle traffic management related challenges. The analysis considers the coexistence of traffic management technologies and other technologies that influence the successful operation of traffic management without degrading any function in the system. The analysis also deliberates some potential connections with IETF standardization activities on traffic management (see relevant areas in Appendix B).

## 6.1 Process to set the focus of integration work

For well-harmonized integration it was inevitable to work out a process which guarantees that the proposed architecture provides coherent functionalities and tackles most of the challenges raised by usage trends. The difficulty of this task resides in the high number of usage scenarios (~20) , derived challenges (~60), furthermore the high number of proposed technologies (~40) that cover different but probably overlapping sets of challenges and functionalities and that generate multiple architecture options [D1.2].

To tackle this problem, firstly, we have prioritized the usage cases. We analysed the number of challenges raised by the usage scenarios and selected the top three scenarios having the most challenges connected and targeted by the most of the proposed technologies. Section 6.2 describes these usage scenarios. The integration work should focus on the communication phases and important phenomena induced within these scenarios.

The high number of challenges has been addressed by the definition of system validation criteria measuring the system performance, the deployment cost of the technologies and maturity level of the technologies. The performance indicators for system validation were defined based on the expected performance gains and rationales of the technologies and often represent collaborative improvements of technologies [D1.2]. A crosscheck between system validation criteria and challenges have been made, hence the fulfilment of a system validation criterion by a combination of technologies in a given usage scenario also means that a given set of challenges were successfully addressed. Section 6.3 summarizes the system validation KPIs and the related challenges, and also specifies how to evaluate an architecture option under a given system validation KPI.

Technologies have been mapped with challenges based on the rationales of the technologies. Knowing which challenges are covered by which system validation KPIs, we can also enumerate the technologies which are relevant under a given system validation KPI. An initial ranking of the technologies have been made according to the KPIs, basically representing that the technology contributes to a given KPI or not.

The system validation alternatives are architecture options, i.e., combinations of proposed technologies integrated within the EPC. In order to filter out architecture options we made the following steps. Technology rankings under the system validation KPIs have been summed up. The technologies with the highest value were considered as important part of the focused architecture. The reasoning behind this procedure is the intention to cover most of the challenges by fewest technologies. The expectation is that the terminal score should reflect both the number of KPIs addressed by the technology and also favour the technologies that perform very well under a given KPI. Different, better score aggregation method could be used but for a first approach, summing up the rankings seemed good enough. Further assumptions behind these thoughts are that if a technology addresses more KPIs, it also means that it addresses higher number of challenges, and higher number of usage scenarios. Checking the split of challenges of each usage scenario between the system validation KPIs, this assumption holds in our case. Additionally it is also assumed that it is more economical, if fewer technologies cover more challenges. Co-existence of top ranking technologies has been analysed, and technologies covering similar functionality resulted in the creation of several architecture options. Section 6.4 summarizes the architecture options.

## 6.2 Usage scenarios

The prioritization of usage scenarios makes possible to discuss the most relevant parts of signalling and data communication and investigate the dependencies and the collaborative effects of proposed technologies on the system validation KPIs. Three usage scenarios are described in this section:

- 3GPP use case 2: Voice/Multimedia and charging: VoIP/Multimedia QoS and mobility bw. Residential, home Wi-Fi and LTE wide area network, Charging schemas adapted to access type, location

- 3GPP use case 3 Video: premium VoD with guaranteed QoS (LTE-Wi-Fi handover, Wi-Fi provided by Fixed broadband provider), QoS update based on available bandwidth, user-specific policy
- Seamless user experience of mobile Internet over multiple GWs and multiple interfaces

The first two usage scenarios are about the two most important end-user services, i.e., Voice/Multimedia communication and Video on Demand services. The proposed architecture must meet the requirements of fixed-mobile convergence, i.e., operator provides the same services to the user over fixed or mobile networks regardless of the access (home mobile net, roaming mobile net, Wi-Fi) and location. Both usage scenarios describe the mobility of a user between residential Wi-Fi, LTE wide area network and home Wi-Fi. In both scenarios the most important requirement is the support of E2E QoE, with suitable QoS adaptation. Other requirements such as adaptation of charging schemes have also been raised as new challenges.

The third usage scenario is about seamless user experience of mobile Internet over multiple GWs and multiple interfaces, and reflects the mobile network operator aspect and architecture evolution. This scenario can cover any event related to traffic management, network management, transport service, where smart traffic steering actions are made.

The most important challenge is to improve user experience by taking advantages of multiple interfaces and multiple paths to the servers. This provides a better user experience of mobile Internet by managing connectivity towards multiple access of the end device. Upon events, the system may decide to move some flow between accesses in order to increase QoE or for load sharing purpose. There are a lot of possible events, hereafter is a non-exhaustive list of events:

- Detection of new potential accesses: UE detects the coverage of a Wi-Fi AP or a femtocell access
- Loss of access link,
- QoS of current access degrades,
- QoS of preferred access improves,
- New application starts, etc…

## 6.3 System validation KPIs

This section describes the system validation criteria, i.e., specifies the criteria and the related performance metrics (i.e., system validation KPI). Furthermore this section provides a view on how to assess/measure each system validation KPI and what are the expectations related to them, i.e., proposes recommendations for ranking assignment method.

System validation KPIs are different from technology validation KPIs. System validation KPIs described in this section are related to the evaluation of architecture options. An architecture option is a set of new technologies integrated into EPC, using the best topology (or distribution level). On the other hand, KPIs specified in the validation plans are technology specific, measure the efficiency of the technology without considering system integration.

### 6.3.1 Throughput in the access and backhaul

The proposed technology will increase the network throughput and will be measured in terms of increase of number of packets and packets size in the access and backhaul. The requirements in 3GPP defined for classes of traffic should be considered when measuring this KPI.

Related technologies and the reasoning behind the expected gains are the followings

- MCCS: multi-criteria cell selection has a direct effect on access selection and the backhaul path. It directly increases the throughput in the access and backhaul.
- NB-IFOM: network-based IP flow mobility enables the offload of traffic flows to non-3GPP interfaces. Indirectly it increases the available bandwidth on the 3GPP-backhaul and access network.
- Gateway Selection: gateway selection provides load balancing in the transport network. Due to the more efficient usage of access and backhaul resources, an indirect influence of the technology is that the capacity of the 3GPP backhaul and access network increases.

- BTA: bulk traffic analysis applied in real-time is an enabler for the enforcement of QoS and better resource utilization. Those have an indirect effect on the available bandwidth in the 3GPP access and backhaul.
- MPTCP-Pr: this technology enables load balancing within the operator network by directing TCP traffic to multiple gateways via the Internet. Better use of network resources may indirectly increase the available access and backhaul capacity in a distributed or flat mobile operator network topology.
- Caching: content caches in the operator network, depending on the distribution level, may provide better load distribution, hence, indirectly, may increase the available bandwidth in the 3GPP access and backhaul part.
- ALTO: selecting the best path to the available content servers by the use of ALTO may either lead to offload of some traffic from 3GPP access and backhaul, or lead to a more efficient load distribution in the access and backhaul. Hence ALTO indirectly increases the available bandwidth of the access and backhaul networks.
- Resource redirector: path optimization may lead to better utilization of transport networks, hence indirectly; it may increase the available bandwidth in the 3GPP access and backhaul.
- DPI: Deep packet inspection is an enabler for application or service level flow monitoring, and provides triggers to offloading or load-balancing mechanisms. Indirectly, it may lead to better access and backhaul utilization, increasing the available bandwidth.
- QoE estimation & traffic manipulation: QoE estimation and traffic manipulation enables better adaptation of flows to available resources in access networks and backhaul. Under a certain threshold the application of the technology enables higher number of served demands with sustained QoE. Hence the available access and backhaul capacity might indirectly increase due to the usage of the technology.

### 6.3.2 Backhaul and RAN influence on E-E delay

The proposed technology decreases the E-E delay due to manipulation in the RAN and backhaul. This can be measured with delays when applying the technology compared to not applying the technology.

Related technologies and the reasoning behind the expected gains are the followings

- MCCS: multi-criteria cell selection influences the selection of radio access (macro, micro, femto-cell,). Hence it has influence on the E-E delay of the traffic.
- Caching, mP4P, ALTO, resource redirection: these technologies have direct influence on E-E path between UE and content, and may lead to the decrease of E-E delay. They don’t change access and backhaul transport functionalities.

### 6.3.3 Efficient load distribution (in backhaul and in the core)

This KPI should show that the application of load balancing mechanism contributes to the non-congested states of the network in case of high traffic demands. The traffic load, the inter-arrival time and transmission delay should be measured either on end points or in the routers/switches if possible. The KPI could also use global packet loss ratio to measure the congestion of the end-to-end network.

Related technologies and reasoning behind the expected gains:

- NB-IFOM, Gateway Selection, MPTCP-PR, ALTO. These technologies contribute to better load distribution.

### 6.3.4 Capacity aggregation and E2E QoE provision

This KPI should measure the throughput gain including goodput. The KPI will also measure QoS packet delay jitter packet loss plus any additional QoE measurements.

Related technologies and reasoning behind the expected gains:

- NB-IFOM, Gateway Selection, BTA, ALTO, DPI, QoE estimation & traffic manipulation: these technologies consider the provision of E-E QoE; hence they are considered under this KPI.
- MPTCP-Pr: this technology is considered due to capacity aggregation and better QoE by e.g., resistance to network link failures.

### 6.3.5 Service Interruption and Handover delay

This KPI will measure the packet transmission additional delay due to flow mobility. The KPI can measure the service interruption delay and it can measure jitter due to HO. Packet delay budgets for guaranteed bitrate real-time services can be considered as hard constraints for induced E-E service interruption delay.

Related technologies and reasoning:

- NB-IFOM: it uses HA-initiated Binding updates, an extension of Mobile IPv6 to remap traffic flows. The service interruption time of the technology should be measured.

### 6.3.6 Handover related signalling load on the network

This KPI can measure transmitted data overhead for HO process. This KPI can measure the number of HO messages and its size.

Related technologies and reasoning:

- NB-IFOM: the technology induces control messages to hand off traffic flows; hence its signalling overhead should be measured.
- MCCS: the technology induces control messages during handovers; its signalling overhead should be measured.

### 6.3.7 E-E delay between UE and content

This KPI can measure RTT (on UE or server). The 3GPP Requirements by application type in terms of e2e delay budget should be considered. This KPI could also measure the path lengths in terms of number of L2/L3 hops or also the air distance.

Related technologies and reasoning behind expected gains:

- NB-IFOM: one possible trigger to redirect flows from one path to another path is the latency of the path, between the UE and the Home Agent. Traffic types that are sensible to latency are taken into consideration by one-way latency measurements and if a threshold is reached then the flow is moved to a better path (if available).
- Gateway Selection: gateway selection has effect on path length; hence E-E delay is influenced by the technology. It depends on the optimization goal whether path length (latency or number of hops) is considered in the decision,
- Caching: if content is available in operator provided caches, then it definitely will shorten the E-E delay
- mP4P: this technology prefers the selection of local seeds and peers in the same domain for P2P content. Hence on average, paths will decrease between UE and content.
- ALTO: ALTO optimizes content server selection for the UEs (ALTO clients). The content is downloaded from the operator provided content servers on the cheapest path. The cost of path may include distance or path length. In that case the technology has direct influence on E-E delay.
- MASE: in streaming optimization the objective is to sustain QoE for the user and efficiently manage the network resources. One constraint the packet delay budget that must be considered for different real-time services, such as live broadcast.
- Resource redirector: resource redirection has influence on the path between the UE and content. The optimization objective can include minimizing the path length,. In that case E-E delay is decreased by the technology.

### 6.3.8 Offload gains for core network elements

This KPI can measure the throughput (i.e. number of data flows) on network elements such as S/P-GW, furthermore, user-signalling reduction (i.e. number of signalling messages, the number of active user contexts per network equipment) on network elements such MME or S/P-GW. It can also measure goodput values on the specific network element

The related technologies and reasoning behind the expected gains:

- Gateway Selection: distributed GWs decrease the number of user contexts, data throughput per GW.

## 6.4 Architecture options

### 6.4.1 Selection process of top ranking technologies that should form a fundamental part of the proposed MEVICO architecture

Technologies have been ranked based on

- Whether they address system validation criteria,
- How mature the technology is, and
- Whether its deployment needs changes in the UE or network or in terms of new network elements.

This evaluation lead to three basic architecture options that must be further analysed from mobility management, traffic management, network management aspect, and co-existence issues, relevant scenarios for the usage of the proposed technologies must be derived. This document deals with the relevant questions related to new traffic management technologies.

The next three subsections present three architecture options having different distribution levels for the EPC functions. The figures for the architecture options include the top ranking technologies. An additional section will discuss technologies that are relevant from the aspect of traffic management, but that were not included in the top ranking technologies due to addressing only a few specific system validation criteria. The description of the technologies can be found in the preceding part of the document.

Two main questions will be addressed regarding the architecture options

- Validation results will prove in the next release of the document that the proposed traffic management technologies bring the claimed performance gains.
- This document addresses the question of co-existence issues of traffic management technologies with other technologies. This means that it describes how the technologies depend on each other, are there any co-existence issues in one network element or on the UE and if technologies are preferred in the same or in different scenarios. The document concludes what are the relevant options for the usage and it shows possible combinations of the proposed technologies in traffic management.

### 6.4.2 Initial proposal for centralized architecture

Figure 6-1 presents the initial proposal of MEVICO for a centralized EPC. The figure includes technologies that address the most of the system validation criteria. Traffic management related technologies are highlighted. The technologies are: NB-IFOM, DPI, GW selection, ALTO and MPTCP-Pr.

NB-IFOM provides load balancing or offloading for the 3GPP access and backhaul by mapping flows to different interfaces of the UE towards the DSMIPv6 Home Agent located in the P-GW.

DPI monitors traffic and is able to provide application/transport/network flow level measurements in a non-intrusive way for technologies enforcing flow mapping, by probing the transport links in the backhaul and core (i.e., the S1-U, S2, S5/S8, SGi interfaces)

GW selection influences PDN- and S-GW selection for the UE based on different characteristics of the network, e.g. measured by DPI, or information from network management.

ALTO enables the selection of the best path between the UE and multiple content server options where the content can be found. It provides better QoE for users and enhances the load distribution of the network.

MPTCP-Pr provides more reliable, less congested, higher capacity transport for TCP traffic. The typical scenario for its deployment is to provide multipath transport for TCP sessions on the Internet, between mobile operator domains. Depending on the location of the traffic splitter, e.g. in S-GW, the load of different P-GWs can be balanced in the core network of the mobile operator on S5/S8 interfaces.

All these traffic management technologies contribute to an efficient resource usage of the mobile operator network, and are relevant in the “seamless user experience of mobile Internet over multiple GWs and multiple interfaces” usage scenario.

Relevant questions regarding co-existence of technologies are described in Section 6.5.

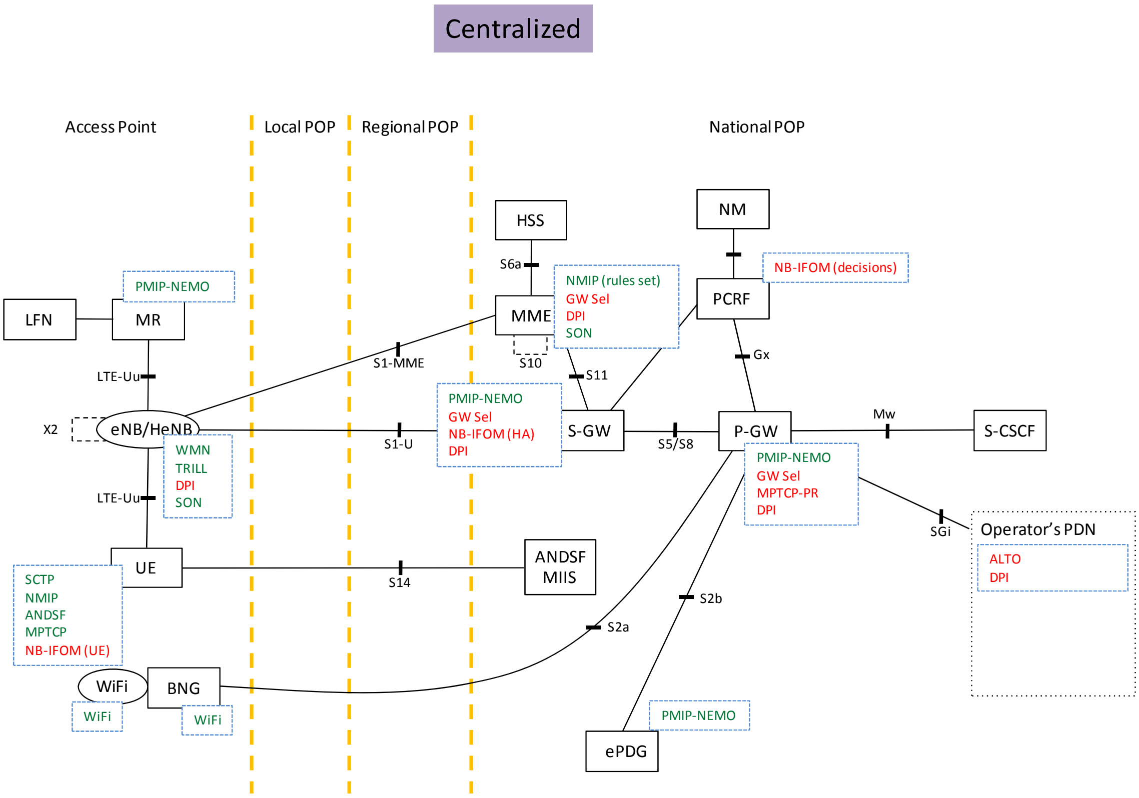


**Figure 6-1:** ***Initial proposal for centralized EPC architecture***

### 6.4.3 Initial proposal for distributed architecture

Figure 6-2 presents the initial MEVICO architecture option for a distributed core. The highlighted technologies that provide traffic management functionalities are the same as in case of centralized architecture option, i.e., NB-IFOM, DPI, GW selection, ALTO and MPCTP-Pr.

Relevant co-existence questions are discussed in Section 6.5.

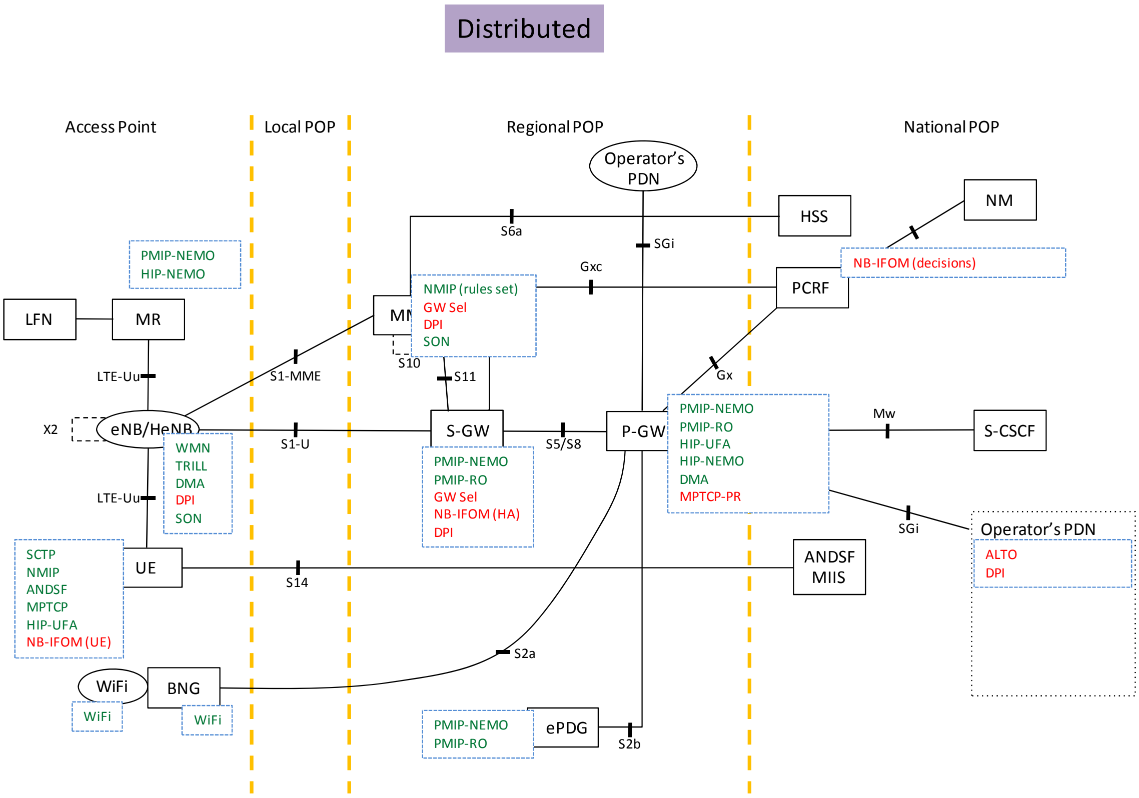


**Figure 6-2:** ***Initial proposal for distributed EPC architecture***

### 6.4.4 Initial proposal for flat architecture

Figure 6-3 presents the initial MEVICO proposal for a flat EPC. The traffic management related technologies highlighted and supporting flat architecture are: NB-IFOM, DPI, ALTO and MPTCP-Pr.

Relevant co-existence questions are discussed in Section 6.5.

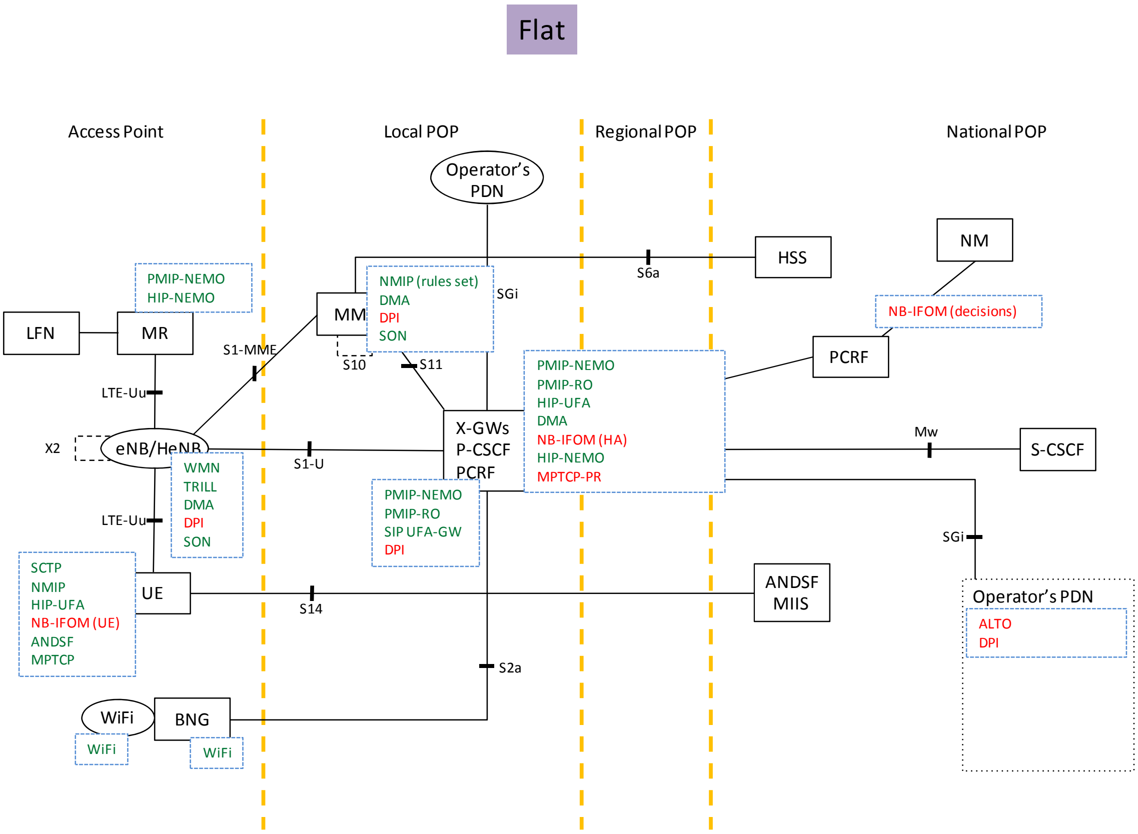


**Figure 6-3:** ***Initial proposal for flat EPC architecture***

### 6.4.5 Additional technologies

Besides NB-IFOM, DPI, Gateway selection and ALTO there are additional technologies relevant for traffic management. Those should also be considered as part of the traffic management of MEVICO proposals for the EPC architecture, even if those technologies address fewer system validation criteria.

- BTA: Bulk traffic analysis applied in real-time provides similar triggers as DPI to load balancing and offloading technologies, QoE or QoS monitoring of application flows is the first main step in quality assurance, and these technologies contribute to that step. Load-balancing and offloading are part of the important challenges for top ranking usage scenarios.
- Caching: This technology optimizes the availability and location of cached content in the operator network. Caching may enhance user experience by decreasing latency, increasing throughput of the content to the UE, and also decrease the volume of demands transported through the mobile operator core network, between the UE and content providers. This exactly fits to usage scenario 3, which targets better network resource utilization, seamlessly for the user.
- mP4P: This technology decreases P2P traffic demands between different operator domains.
- Streaming optimization (MASE, STME): Streaming optimization may enhance QoE of users. That fits exactly to the main challenges of top ranking usage scenario 2.
- Resource redirector: Similarly to ALTO, caching, the resource redirection enables better path allocation between the UE and content, and leads to better QoE and more efficient usage of network resources. Challenges raised by top ranking use cases, especially 2 and 3, are addressed by this technology.
- MCCS: Multi-criteria cell selection minimizes number of handoffs, and provides efficient load balancing in the RAN. Seamless mobile internet usage over multiple interfaces in a hierarchical cellular network requires optimization on the access selection level.

- QoE estimation & traffic manipulation: the technology provides knowledge on the actual QoE of users hence provides triggers for traffic management.

## 6.5 Integration Issues

This section describes co-existence issues and some basic statements on the integration, related to traffic management technologies.

The most important co-existence issues are originated from the fact that different network operations initiate similar operations based on their particular decision scheme. There are for example several technologies related to traffic, network and mobility management that influence flow mapping based on network state information. An outcome of a decision could be the selection of a new access network type, a new cell type (macro, micro, pico, femto) or a new GW for a specific service data flow.

One decision outcome might trigger another decision-making process. E.g., if a new radio access network is selected for a service data flow, and that access network is attached to a different set of GWs, the tunnel that provides IP connectivity to the UE must also be updated/re-established, hence a GW re-selection process might be started. On the other hand, if a GW change is triggered by the GW selection protocol, and the new GW has no connectivity to the current access network of the UE, then an access network reselection might be initiated. To deal with such phenomena we must analyse all technologies that influence service data flow mappings to different paths in the network.

In general, the objective is to avoid traffic fluctuations caused by optimizations on different layers or planes on different time scales. Different scenarios/options must be defined when to use a given technology in combination with other technologies.

### 6.5.1 Macroscopic traffic management technologies

#### 6.5.1.1 Gateway Selection

The main goal of GW selection is to balance the load between GWs, i.e., keep the load of a GW below a given threshold. The network management provides information on the load of the links and GWs. The GW utilization thresholds are evaluated on the order of seconds or minutes. GW selection is relevant during session establishment, and, if the tunnelling option supports inter-GW handover, then during session modification procedures. E.g., WP2's HIP-UFA scheme applies HIP/IPsec based tunnelling, and inter-GW handovers might not only be executed due to radio link degradation (monitored by 802.21 MIH), but also by GW selection. WP2's PMIP-RO basically executes route optimization for any session where UEs attach to the same LMA. Gateway selection could in this case be applied for the decision on enabling route optimization procedure.

GW selection is a higher layer decision system and it could be mapped to any tunnelling solution (e.g., GTP, PMIP, DSMIPv6). GW selection provides triggers to tunnelling protocols about where to map a given traffic flow. Since GW selection is located in the MME, a network-initiated IP-connectivity tunnel update procedures fit better to it than UE-initiated procedures.

In case of PMIP or DSMIPv6-based NB-IFOM, the GW selection could take effect only during the initial session establishment procedure, e.g., selection of LMA/HA for the UE, because these NB-IFOM technologies don't support GW change for on-going sessions. E.g., the NB-IFOM provides network-controlled IP flow mobility using DSMIPv6 based tunnelling (which is orthogonal to the other tunnelling solutions used in MEVICO), GW Selection comes into picture during HA selection for a UE (based on GW load, and probably other parameters of the network). After selecting the HA, NB-IFOM can make more fine grained distribution of service data flows, and may not change then the GW. For more dynamic GW change a HA-to-HA protocol is needed (like Global Home Agent to Home Agent protocol [GHAHA]).

#### 6.5.1.2 NB-IFOM

The main goal of NB-IFOM technology is to provide load balancing and to improve the congestion less state of the network. The decision is based on information, such as bandwidth, packet loss and latency of service data flows or network links. Frequency of decisions can vary from a couple of seconds up to several minutes. The operator or the manufacturer might define this. The PCRF makes the decision based on some PCC rules defined for service data flows, or based on subscription data. The Home Agent in the P-GW executes the flow mapping decision.

NB-IFOM is applicable to DSMIPv6 based tunnelling option, i.e. when UEs get IP connectivity through the S2c interface. Hence it cannot be applied when other tunnelling options are used by the UE, such as GTP on S1 and S5/S8, GTP on S1 extended with PMIP/IP GRE on S5/S8, GTP on S2a/S2b, PMIP/IP GRE on S2a/S2b.

The information for the decision making on the mapping of service data flow to a new UE interface, hence to a new path towards the Home Agent (HA), is collected from the following services. The network management system shall provide performance measures with network management granularity, such as transport link utilization. DPI or BTA shall provide performance measures with service data flow granularity. Note that DSMIPv6 establishes IPsec tunnel between the UE and the HA on the S2c interface. Hence a dependency issue is that DPI cannot get service data flow level information, expect the rare case when a null-encryption algorithm is selected for data protection. Possible issues with GW selection are described in Section 6.5.1.1.

#### 6.5.1.3 ALTO

The main goal of ALTO technologies is to optimize the path between the UE and the CN. That is why it can be considered either as a macroscopic traffic management technology, or an improved resource redirection scheme, or even an application layer traffic management extension. In this case ALTO is combined with 3GPP tunnelling options (such as IFOM and MAPCON), therefore it is an integrated, macroscopic TM solution.

In case of ALTO, ANDSF and UE-controlled MIH technologies, it is the UE that makes decision based on information from network about available CNs or alternative radio accesses.

There is a co-existence issue between ALTO and PCRF-based decisions related to flow mapping to bearers.

Note that the input parameters for decision-making are collected, and the final decision is made either on the UE-side, e.g., in case of ALTO, or at the GW-side, e.g., in case of NB-IFOM, or at the eNodeB, e.g., MCCS.

#### 6.5.1.4 MPTCP-PR

MPTCP-PR aims to allow an agnostic TCP over multiple paths solution with a transparent proxy solution to increase TCP performance over multiple paths (with different RTTs).

WP2's NMIP technology provides IP mobility management on TCP-level for TCP-based service data flows. MPTCP-PR is transparent for NMIP. Note that standard MPTCP (i.e., not the proxy version) and NMIP can not coexist on the same UE because of conflicting modifications of the TCP functionalities.

#### 6.5.1.5 MCCS

The main goal of MCCS technology is, similarly to SON technology in network-management, to optimize the radio usage/resources. eNodeB makes the decision. Both MCCS and SON use information about traffic load in the cell, user's quality of services, user's channel quality etc.

### 6.5.2 Microscopic traffic management

#### 6.5.2.1 DPI/BTA

Deep packet inspection (DPI) mechanisms are being deployed at the operators, however the amount of investment required inspecting all the traffic in detail is huge. Usually only the traffic types which may be important for the operator's business is detected and the rest is not identified. This shows that some kind of bulk traffic analysis (BTA) may prove to be very useful to classify the total data into applications so that the network operator is able to see the big picture. This does not need to be done in real time and not the whole traffic needs to be analysed, but some time periods can be selected to reflect to the whole week.

As their operation mainly intents to provide network monitoring and traffic analysis, DPI/BTA solutions can coexist with any other technologies in the architecture. Possible limitations in case of NB-IFOM are described in Section 6.5.1.2.

#### 6.5.2.2 MSO

The main goal of Multimedia streaming optimization (MSO) techniques is to enable the coordination of streaming application requirements and conditions with the bearer resource management in EPS, and also to communicate with the application higher in order to enable adaptation of the video playback to the changed conditions. Main coexistence considerations:

- Distributed architectures will require distribution of MASE functions
- Gateway Selection could result in changes of serving MASE
- Caching could impact the operation of MSO (e.g., streaming server roles can be played by the cache server), therefore information exchange between the caching mechanisms and the MSO solution could be required
- MSO is transparent to different tunneling options providing IP connectivity to UEs

### 6.5.3 Interactions between MicTM and MacTM

This section provides a first insight about how the traffic management building blocks, MicTM and MacTM, could interact in a layered fashion, as depicted in Figure 2.3.1.1 1.

One general service that MicTM mechanisms could offer to MacTM mechanisms is that it passes the traffic measurements, conducted for their own operation. The MacTM mechanisms could aggregate these measurements from many MicTM mechanisms to a coarser resolution and use this as an input parameter. This has the advantage, that fewer measurement functions are needed within the network.

A more specific use case would be, if the MicTM mechanism “Cross Layer Interference Detection” could periodically signal to the MacTM mechanisms “Traffic engineered handovers and network-based IP Flow Mobility” and “Multi-Criteria Cell Selection” the interference situation. These MacTM mechanisms could use this information as an input parameter for their own operation.

### 6.5.4 Improved resource selection and caching

#### 6.5.4.1 Caching

Another possible issue is related to the candidate locations of caching servers inside the MNO domain and is strongly connected to the problem space of different MEVICO architecture proposals (i.e., Centralized, Distributed and Flat mobile architectures). The most apparent location of caching nodes is beyond the GTP tunnel endpoint at the SGi interface (e.g., co-located with PDN-GW) or at the S5 interface. The problem here is that placing cache servers there could result in loss of connectivity to the cached material after handovers between 3GPP and non 3GPP accesses. Applying cache nodes at the S1 interface (e.g., co-located with eNodeB) might also be considered, but the benefit of accessing content closer to the users comes along with potential security issues (operators commonly use IPSec) or other limitations (e.g., deployment costs, problems of outdoor deployment). The above motivations make co-location of cache nodes within gateways to be also a promising option.

Besides the possible architectural impacts, the co-location of caching and gateway nodes may also affect gateway re-selection based traffic engineering and/or mobility management mechanisms. If a cache server can be accessed only via a specific gateway node, gateway re-selection and/or mobility management would break the connectivity to the serving cache. Solving this problem would require a complex and presumably costly cache node session transfer protocol. Therefore these mechanisms can co-exist and work simultaneously, if gateway re-selection and mobility management does not force cache server re-selection of a running session.

Microscopic traffic management schemes and techniques (such as support of multipath communication based on IFOM and/or MPTCP) address operator-centric handling of individual flows. Even though content accessed from a cache node normally doesn’t have strict real-time requirements (in terms of latency) there still might be certain quality of experience requirements for the traffic flow associated with the requested cached content. It is therefore a challenge to synchronize the access to the cached content with EPS bearer management and possibly consider additional aspects such as user device characteristics, radio access conditions and other network parameters. Therefore a tighter co-operation might be required between caching and microscopic traffic management solutions at least for some type of cached content.

#### 6.5.4.2 mP4P

The goal of mP4P is to optimize routing of P2P based content sharing within the LTE/EPC. As the scheme provides a specific type of resource selection and redirection, works with a specific type of traffic (i.e., P2P), it is completely transparent to other technologies but the following issues must be noted:

- Coexistence with other resource redirection technologies such as ALTO could be a subject for further investigations
- Currently it seems that P2P and P4P applications shall not be ALTO-aware
- mP4P solution does not consider caching issues which could lead to coexistence problems with caching and CDN schemes.

### 6.5.5 Considerations of a preliminary TM architecture

Traffic management functions usually require access to higher layer user plane data, i.e. IP packets, TCP segments and application layer protocols. Placements of such functions at SGi or S5 interfaces within the EPC architecture are possible options, since GTP tunnels are terminated at these locations. In the following, a brief analysis is done on possible impacts. Positioning of TM functions at S-GW implies that all user plane data can be managed by the considered function unless there is a handover between 3GPP and non-3GPP access network. In such a case user plane traffic could not be processed or handled by the same node, hosting the TM function. If this can't be avoided, possibly different instances of the TM function have to coordinate in order to ensure continuous TM operation, in case such feature is supported by the TM function in consideration. Positioning the TM function at the SGi interface – e.g. co-located with PDN-GW – may cause problems, if user data is transferred using different access point names (APN). This usually implies that data paths stretch along different SGi interfaces. It is common practice in currently deployed networks to allocate the same APN to a user for all over-the-top (OTT) services. However managed operator services may use different APNs. As a consequence the same TM function may not be used for connection via different APNs. This situation would increase equipment cost (CAPEX) as well as operational cost (OPEX). As a consequence, the suitable location of TM functions depends on mobility aspects (whether a TM function needs to be supported after 3GPP-non-3GPP handover) or connectivity aspects (whether the same TM function shall be in usage for services using different APN).

## 6.6 Validation of traffic management architecture options

The system validation results gained in the MEVICO project will be described in the next release of this document, that is, D 4.3.2. Architecture options should be validated under specific scenarios that are relevant for traffic management, and the previously described system validation KPIs should show the benefits of the technologies. The results should indicate what is the best technology combination for the traffic management technologies. The investigation will show whether there are collaborative improvements of technologies, considering the top ranking usage scenarios, and will consider the selection of the best architecture distribution level.

Besides this, a continuation of the MEVICO project should also focus on testbed implementation of the proposed logical structural solutions represented by the architecture options. Such a future work is essential to confirm the real-life relevance of the introduced architecture alternatives.

# 7 Conclusions

In this document we have described the state of the art in traffic engineering and have also provided a preliminary design for the MEVICO traffic engineering architecture.

A set of traffic engineering methods has been classified as macroscopic traffic management. Some topics that fall into this set and need enhancement are summarized below.

In the mobile backhaul network, alternative pre-computed routing paths have been mainly used for supporting connection survivability, coarse-grained QoS routing, or simple load balancing. To optimize resource usage, load and capacity aware adaptive routing would be desirable in the backhaul, similar to 3GPP-defined self-organized load balancing on the radio network layer. Such traffic steering solutions on the radio and/or transport network layers need to be coordinated across the layers.

Traffic offloading techniques are topics of high interest to mobile operators nowadays. The reason of the accentuated attention is that these advanced schemes enable user traffic offload at specific points in the network in a really cost-efficient manner. Offloading solutions also have the capabilities to enormously increase architecture scalability and offer mobile operators the required flexibility to efficiently deal with the ever-growing scale of smart phone applications, services, and the associated growth of mobile data traffic.

The power consumption of the Radio Access Network is a significant part of the overall network consumption. Therefore the on-off switching is applied as a power consumption method. On the other hand, the traffic redirections (as a part of an optimization method) can also be resulted in switching the network elements. These processes can have significant effect on the Core Network and on the traffic engineering solutions that are applied in it. Therefore it is necessary examining the link capacities and the remaining free capacities. It is also required to examine how the transmitted traffic should be distributed and how the free resources should be handled.

Yet another set of techniques have been classified as microscopic traffic management are more involved with QoS requirements of individual applications.

As far as the techniques for QoS support for external content is concerned, resource coordination on application layer is the preferred method. It is expected that the required changes within the EPC is relatively low (limited to the Rx interface) and is therefore a viable solution. For further study is the synchronization of resource control between the different network domains and details of the flow of control information.

Improved resource selection & caching is the third main building block which has been handled in this document. Popular content on the Internet is mainly delivered via global CDNs, which shorten the transport paths through distributed server architectures, and via peer-to-peer networks, which are currently subject to longer transport paths, resulting in unnecessarily high network load and delays. Alternative ways of optimizing traffic paths on CDN and P2P overlays with support from location servers, caches or traffic engineering have been addressed, based on delay measurement. The broadly confirmed relevance of Zipf laws in access patterns is favorable for caching popular content close to users. Avoiding unnecessary traffic load and minimizing delays is a decisive factor for remaining competitive for an upcoming wave of broadband video and IP-TV streaming on the Internet, where hybrid CDN-P2P solutions are the most promising approach for maximizing throughput and exploiting the bandwidth provided in data centers as well as in the broadband access.

Optimization tools for traffic engineering and load balancing improve the throughput on a network platform if a fine granular control of resource usage is enabled by directing explicit traffic paths through the network. The flexibility to redistribute imbalanced load in parts of a network is most valuable after shifts in the traffic matrix, topology changes and failures as well as for planning processes to evaluate and optimize topology design alternatives. Dynamic environments in mobile and wireless networks make traffic management more challenging, where support functions are often missing especially in new networking technologies.

Today, optimization for short delivery paths and for load balancing is partly performed on global CDN and P2P overlays and partly by the network providers, but there is a lack of coordination between different administrative domains. Since CDN, P2P and network providers have the common goal of offering high service quality in a cost-efficient way, more cooperation and standardization effort is expected to overcome current cross-layer inefficiencies.

MEVICO partners are conducting research on these technologies and suggested improvements – The final traffic engineering architecture- will be documented in the D 4.3.2. An important area of work is seen as merging suggested technologies, considering the possibilities of using them together.

# Acknowledgement

The work has been carried out in the framework of the CELTIC project MEVICO. The views expressed in this document are solely those of the authors and do not represent the views of their employers.

Our thanks go to all colleagues in the CELTIC project MEVICO, who have contributed to relevant documents in the project and in addition to Lothar Braun, Alexander Klein, Georg Carle from Chair for Network Architectures and Services, Technical University of Munich for the work presented in Sec. 5.4.

# References

[3GPP_23.401] 3GPP TS 23.401 V10.4.0, Technical Specification, “General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access”, Release 10, June 2011.

[3GPP_29.303] 3GPP TS 29.303 V10.2.1, Technical Specification, “Domain Name System Procedures, Stage 3”, Release 10, July 2011.

[3GPP_TR_23.829] 3GPP TR 23.829: Local IP Access and Selected IP Traffic Offload, Release 10, V1.3.0, Sept. 2010.

[3GPP_TR_23.861] 3rd Generation Partnership Project, Multi access PDN connectivity and IP flow mobility (Release 9), 3GPP TR 23.861 V1.3.0, February 2010.

[3GPP_TR_23.919] 3GPP TR 23.919 V10.0.0, Technical Report, “Direct tunnel deployment guideline”, March, 2011.

[3GPP_TR_23.919] 3GPP TR 23.919: Direct Tunnel Deployment Guideline, Release 7, V1.0.0, May 2007.

[3GPP_TS_22.220] 3GPP TS 22.220 V10.7.0, “Service requirements for Home Node B (HNB) and Home eNode B (HeNB)”, Release 10, June 2011.

[3GPP_TS_23.002] 3GPP TS 23.002: Network architecture, V10.1.1, Release 10, Jan. 2011.

[3GPP_TS_23.060] 3GPP TS 23.060 V10.4.0, Technical Specification, “General Packet Radio Service (GPRS); Service description”, Stage 2, Release 10, June 2011.

[3GPP_TS_23.261] 3GPP TS 23.261 V10.1.0, Technical Specification, “IP flow mobility and seamless Wireless Local Area Network (WLAN) offload, Stage 2”, Release 10, September 2010.

[3GPP_TS_23.401] 3GPP TS 23.401: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, Release 8, V8.12.0, Dec. 2010.

[3GPP_TS_23.402] 3GPP TS 23.402 V10.4.0, “Architecture enhancements for non-3GPP accesses”, Release 10, June, 2011.

[Abusubaih08] M. Abusubaih, B.Rathke, and A.Wolisz, "Collaborative Setting of RTS/CTS in Multi-Rate Multi-BSS IEEE 802.11Wireless LANs", In Proc. of the 16'th IEEE Workshop on Local and Metropolitan Area Networks, IEEE LANMAN'08, Cluj-Napoca, Romania, September 2008

[Abusubaih08a] M. Abusubaih and A. Wolisz, "Interference-Aware Decentralized Access Point Selection Policy for Multi-Rate IEEE 802.11 Wireless LANs", In Proc. of the 19'th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008., Cannes, France, September 2008

[Abusubaih09] M. Abusubaih, B. Rathke, and A.Wolisz, "A framework for Interference Mitigation in Multi-BSS 802.11 Wireless LANs", In Proc. of 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (IEEE WoWMoM 2009), Kos, Greece, June 2009

[AJZ10] V. K. Adhikari, S. Jain, and Z.-L. Zhang, “YouTube Traffic Dynamics and Its Interplay with a Tier-1 ISP: An ISP Perspective,” IMC ’10: Proceedings of the 10th annual conference on Internet measurement, 2010.

[AKAM10] Akamai, State of the Internet, Quarterly Report Series (2011) <www.akamai.com>

[ALI11] “ALTO Protocol”, R. Alimi et al., June 27th 2011, http://tools.ietf.org/html/draft-ietf-alto-protocol-09

[Allman09] M. Allman, V. Paxson, and E. Blanton, ”TCP Congestion Control,” RFC 5681, September 2009.

[Almeida_99] S. Almeida, J. Queijo, L.M. Correia: “Spatial and temporal traffic distribution models for GSM”, Vehicular Technology Conference, 1999. VTC 1999 - Fall. IEEE VTS $50^{th}$, Vol. 1. , pp. 131-135

[ALR+08] H. A. Alzoubi, S. Lee, M. Rabinovich, O. Spatscheck, J. E. van der Merwe “Anycast CDNS revisited” WWW 2008 / Refereed Track: Performance and Scalability April 21-25, 2008. Beijing, China, pp277-286

[Al-Sanabani08] M. Al-Sanabani & al., “Mobility Prediction Based Resource Reservation for Handoff in Multimedia Wireless Cellular Networks”, the International Arab Journal of Information Technology, Vol. 5, N°2, April 2008.

[Amzallag08] D. Amzallag et al., “Cell selection in 4G cellular networks,” in Proc. *IEEE Conference on Computer Communications* 2008, pp. 700-708, Apr. 2008.

[Anpalagan99] A.S. Anpalagan and I. Katzela, "Overlaid Cellular System Design with Cell Selection Criteria for Mobile Wireless Users", *IEEE Canadian Conf. on Electrical and Computer Eng.*, vol. 1, 1999, pp. 24-28.

[APA] Apache web server module “mod_rewrite”
http://httpd.apache.org/docs/2.2/mod/mod_rewrite.html

[ASK+10] B. Ager, F. Schneider, J. Kim and A. Feldmann, Revisiting cacheability in times of user generated content, Proceedings of the 13th IEEE Global Internet Symposium, San Diego, CA, USA (2010)

[Ayar12a] T. Ayar, B. Rathke, Ł. Budzisz, and A. Wolisz, "A Splitter/Combiner Architecture for TCP over Multiple Paths", TKN Technical Report Series TKN-12-001, February 2012, available at: http://www.tkn.tu-berlin.de/menue/publications/tkn_technical_report_series/.

[Ayar12b] T. Ayar, B. Rathke, Ł. Budzisz, and A. Wolisz, “A Transparent Performance Enhancing Proxy Architecture To Enable TCP over Multiple Paths for Single-Homed Hosts,” IETF Internet Draft, draft-ayar-transparent-sca-proxy-00, Work in Progress, February 2012.

[BDK+10] L. Braun, A. Didebulidze, N. Kammenhuber, and G. Carle, “Comparing and improving current packet capturing solutions based on commodity hardware,” IMC ’10: Proceedings of the 10th annual conference on Internet measurement, Nov. 2010.

[BGW10] S. Borst, V.Gupta and A. Walid, Distributed caching algorithms for content distribution networks, IEEE Infocom (2010).

[BMV10] A. Balasubramanian, R. Mahajan, and A. Venkataramani, “Augmenting mobile 3G using WiFi,” in *Proc. of MobiSys.* ACM, 2010, pp. 209–222.

[BKC+12 “Analyzing Caching Benefits for YouTube Traffic in Edge Networks - A Measurement-Based Evaluation” L. Braun, A. Klein, G. Carle, H. Reiser, J. Eisl; NOMS 2012 IEEE Network Operations and Management Symposium Maui Hawaii USA Apr. 16th-20th 2012]

[BPV08] R.Buyya, M. Pathan and A. Vakali (Eds.), Content delivery networks, Lecture Notes in Electrical Engineering 9, Springer (2008)

[BRE+99] L. Breslau et al., Web caching and Zipf-like distributions: Evidence and implications, Proc. IEEE Infocom (1999)

[CFE+06] K. Cho, K. Fukuda, H. Esaki, A. Kato, The impact and implications of the growth in residential user-to-user traffic, ACM Sigcomm Conf., Pisa (2006) <www.acm.org/sigs/sigcomm/sigcomm2006>

[CFM09] R. Canonico, C. Fuererro and A. Mauthe (Eds.), Content distribution infrastructures for community networks, Computer Networks Special Issue Vol. 53/4 (2009) 431-568

[CHA+07] M. Cha et al., I tube, you tube, everybody tubes: Analyzing the world’s largest user generated content video system, Internet measurement conference IMC’07, San Diego, USA (2007)

[Cha08] B. Cha, S. Seo, Y. Choi, and J. Song, “Mobile-Velocity adaptive Vertical Handoff in Integrated WLAN and WiBro Networks,” *ICIAFS 2008*, pp. 384-389. [CHA10] J. Charzinski, Traffic properties, client side cachability and CDN usage of popular web sites, Proc. 15th MMB conference, Essen, Germany, Springer LNCS 5987 (2010) 182-194

[Chandrasekhar_08] V. Chandrasekhar, J. G. Andrews, A. Gatherer: Femtocell Networks: A Survey, IEEE Communications Magazine, 46 (9), pp. 59-67, Sept. 2008.

[Chang10] C.-J. Chang, C.-Y. Hsieh and Y.-H. Chen, “A Preference Value-Based Cell Selection Scheme in Heterogeneous Wireless Networks,”, *IEEE WCNC*, 2010, pp. 1–6.

[Chebrolu05] K. Chebrolu, B. Raman, and R. R. Rao, ”A Network Layer Approach to Enable TCP over Multiple Interfaces”, ACM/Kluwer Journal of Wireless networks (WINET), Vol. 11, No. 5, pp. 637-650, September 2005.

[Chen_08] C. Chen, Y Xu, L. Zhang: "Some Remarks on ON/OFF Network Traffic", Power Electronics and Intelligent Transportation System (PEITS), 4-5 August 2008, Guangzhou, China, pp. 515-519

[Cheng03] P.F. Cheng and S.C. Liew, “TCP Veno: TCP enhancement for transmission over wireless access networks”, IEEE Selected Areas in Communications, vol. 21, February 2003.

[Chung02] Y. Chung, D.-J. Lee, D.-H. Cho, and B.-C. Shin, “Macrocell/Microcell Selection Schemes Based on a New Velocity Estimation in Multitier Cellular System,” IEEE Trans. Vehicular Technology, vol. 51, no. 5, pp. 893-903, Sept. 2002.

[CIS10] Cisco Systems, Visual networking index, forecast and methodology, White paper series (2011) <www.cisco.com

[CIS11] ——, "The Future of Hotspots: Making Wi-Fi as Secure and Easy to Use as Cellular", White paper series (2011) <www.cisco.com>

[CIU10] D. Ciullo, Network awareness of P2P live streaming applications: A measurement study, IEEE Trans. on Multimedia 12 (2010) 54-63

[CKR+07] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," in Proceedings of the 7th Conference on Internet Measurements (IMC) 2007, 2007.

[CKR+09] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "Analyzing the Video Popularity Characteristics of large-scale User Generated Content Systems," IEEE/ACM Transactions on Networking (TON), vol. 17, no. 5, pp. 1357–1370, Oct. 2009.

[CPZ11] "ALTO in Mobile and Wireless Network" W. Chen, J. Peng, Y. Zhang, China Mobile, July 3 2011, http://tools.ietf.org/html/draft-chen-alto-in-mobile-wireless-network-00

[CSM10] G. Chatzopoulou, C. Sheng, and M. Faloutsos, "A First Step Towards Understanding Popularity in YouTube," in INFOCOM IEEE Conference on Computer Communications Workshops , 2010, 2010, pp. 1–6.

[D1.2] Ivan Froger, Didier Becam (ed.) Architecture Design Release 2 Documentation, MEVICO Project Deliverable, 21. December 2011.

[D. Y. Cl_10] D. Y. Cl.: "RF IC design of highly-efficient broadband polar transmitters for WiMAX and 3GPP LTE applications", IEEE Solid-State and Integrated Circuit Technology (ICSICT), 1-4 November 2010, Shanghai, China, pp. 150-153

[Das04] A. Das, K. Balachandran, F. Khan, A. Sampath, and H.-J. Su, "Network Controlled Cell Selection for the High Speed Downlink Packet Access in UMTS," in *Proc. IEEE WCNC*, vol. 4, pp 1975-1979, Mar. 2004.

[DDNS] Dynamic DNS - http://www.webopedia.com/TERM/D/dynamic_DNS.html

[DER10] L. Deri, "High Speed Network Traffic Analysis with Commodity Multi-Core Systems," IMC '10: Proceedings of the 10th annual conference on Internet measurement, Nov. 2010.

[DevoTeam] DevoTeam. Traffic Load Generator. [Online]
http://www.devoteam.co.uk/index.php?option=com_content&task=view&id=445&pays=uk&Itemid=744&lang=2.

[DVB09] Digital Video Broadcasting Project (DVB), Internet TV Content Delivery study mission report, DVB Document A 145 (2009)
<www.dvb.org/technology/standards/A145_Internet_TV_Content_Delivery_Study.pdf>

[E._Oh_11] E. Oh et al.:" Toward Dynamic Energy-Efficient Operation of Cellular Network Infrastructure", IEEE Communications Magazine, June 2011, pp 56-61

[ETSI_96] ETSI GTS GSM 03.02-v5.1.0: Digital cellular telecommunications system (Phase 2+) - Network architecture (GSM 03.02), 1996.

[EUB08] M. Eubanks, The video tsunami: Internet television, IPTV and the coming wave of video on the Internet, Plenary talk, 71. IETF meeting (2008) <www.ietf.org/proceedings/08mar/slides/plenaryt-3.pdf>

[Evensen09] K. Evensen, D. Kaspar, P. Engelstad, A. F. Hansen, C. Griwodz, and P. Halvorsen, "A Network-Layer Proxy for Bandwidth Aggregation and Reduction of IP Packet Reordering," IEEE 34th Conference on Local Computer Networks (LCN 2009), pp. 585-592, 20-23 October 2009.

[FBA11] F. Figueiredo, F. Benevenuto, and J. M. Almeida, "The Tube over Time: Characterizing Popularity Growth of Youtube Videos," in In Proceedings of the 4th ACM Conference on Web Search and Data Mining, 2011.

[FemtoForum_10] FemtoForum: Femtocells – Natural Solution for Offload – a Femto Forum topic brief, June 2010.

[FIE00] R. Fielding et al., Hypertext transfer protocol - HTTP/1.1, Request for Comments 2616 <www.rfc-editor.org/rfc/rfc2616.txt> (2000)

[FIE10] R. Fielding et al., HTTP/1.1, part 6: Caching, Internet-Draft
<https://datatracker.ietf.org/doc/draft-ietf-httpbis-p6-cache/> (2010)

[Fiedler_Hoßfeld_Tran-Gia_10]M. Fiedler, T. Hoßfeld, P. Tran-Gia: A Generic Quantitative Relationship between Quality of Experience and Quality of Service. IEEE Network, Special Issue on Improving QoE for Network Services, Vol. 24 Issue 2, March-April 2010.

[Fiercewireless_10] http://www.fiercewireless.com/europe/story/femtocell-deployment-update/2010-09-17

[FMM+11] A. Finamore, M. Mellia, M. Munafo, R. Torres, and S. Rao, "YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience," Technical Report, 2011.

[Fodor04] G. Fodor, A. Furuskar, and J. Lundsjo. On access selection techniques in always best connected networks. In *ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, August 2004.

[Ford11] A. Ford, C. Raiciu, S. Barre, and J. Iyengar, "Architectural Guidelines for Multipath TCP Development," RFC 6182 , March 2011.

[Ford12] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses," IETF Internet Draft, draft-ietf-mptcpmultiaddressed-07, Work in Progress, March 2012.

[Fussen05] M. Fussen, R. Wattenhofer, and A.Zollinger, "Interference arises at the receiver", In Proc. of Wireless Networks, Communications and Mobile Computing, 2005

[GAL+07] P. Gill, M. Arlitt, Z. Li and A. Mahanti, YouTube traffic characterization: A view from the edge, Internet measurement conference IMC'07, San Diego, USA (2007)

[GAM07] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "Youtube traffic characterization: a view from the edge," in IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. ACM Request Permissions, Oct. 2007.

[Gao11] L. Gao, X. Wang, G. Sun, Y. Xu, "A Game Approach for Cell Selection and Resource Allocation in Heterogeneous Wireless Networks," in Proc. of *IEEE SECON 2011*, Salt Lake City, Utah, USA, 2011.

[Gazis05] V. Gazis, N. Alonistioti, and L. Merakos. Toward a generic "always best connected" capability in integrated WLAN/UMTS cellular mobile networks (and beyond). *Wireless Communications, IEEE*, 12(3):20–29, June 2005.

[Gerla01] M. Gerla, M.Y. Sanadidi, W. Ren; A. Zanella, C. Casetti and S. Mascolo, "TCP Westwood: congestion window control using bandwidth estimation", in proc. of Global Telecommunications Conference, 2001 (GLOBECOM '01), 2001

[GHAHA] R. Wakikawa, R. Kuntz, Z. Zhu, L. Zhang, Global HA to HA Protocol Specification, IETF Internet Draft, draft-wakikawa-mext-global-haha-spec-02, September, 2011.

[GO00] G. Barish and K. Obrazcka, World wide web caching: Trends and techniques, IEEE Communications Magazine (May 2000) 178-185

[Gomes09] J. S. Gomes, "A Rule based Co-operative Approach for Cell Selection in High Speed Cellular Networks", *IEEE International Symposium on Network Computing and Applications*, pp.74-81, 2009.

[Guo06] Q. Guo, X H Xu, J Zhu, et al. "A QoS-guaranteed cell selection strategy for heterogeneous cellular systems". ETRI Journal, 2006, 28(1): 77–83.

[GUO07] L. Guo et al., Does Internet media traffic really follow Zipf-like distributions? ACM SIGMETRICS (2007)

[Hanly95] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," IEEE Journal on Selected Areas in Communications, vol. 13, no. 7, pp. 1332–1340, 1995.

[Hannes] H. Ekstrom, *QoS Control in the 3GPP Evolved Packet System*, IEEE Communications Magazine (2009) 76-83

[HAS05] G. Haßlinger, ISP platforms under a heavy peer-to-peer workload, Proc. Peer-to-Peer Systems and Applications, Eds.: R. Steinmetz and K. Wehrle, Springer LNCS 3485 (2005) 369-382

[Haßlinger05] G. Haßlinger, S. Schnitter and M. Franzke, The efficiency of traffic engineering with regard to failure resilience, Telecommunication Systems Vol. 29/2, Springer (2005) 109-130.

[Haßlinger11] G. Haßlinger, G. Nunzi, C. Meirosu, C. Fan and F.-U. Andersen, Traffic engineering supported by inherent network management: Analysis of resource efficiency and cost saving potential, Internat. Journal on Network Management (IJNM), Special Issue on Economic Traffic Mgnt., Vol. 21 (2011) 45-64.

[HH10]G. Haßlinger and O. Hohlfeld, Efficiency of caches for content distribution on the Internet, Proc. 22. Internat. Teletraffic Congress, Amsterdam, The Netherlands (2010)

[HHB09] G. Haßlinger, F. Hartleb and T. Beckhaus, User access to popular data on the Internet and approaches for IP traffic flow optimization, Proc. ASMTA Conf., Madrid, Spain, Springer LNCS 5513 (2009) 42-55

[HMG+07] G. Haßlinger, J. Mende, R. Geib, T. Beckhaus and F. Hartleb, Measurement and characteristics of aggregated traffic in broadband access networks, Proc. 20. Internat. Teletraffic Congress, Ottawa, Canada, Springer, LNCS 4516 (2007) 998-1010

[Holma_07] H, Holma, A. Toskala, K. Ranta-aho, J. Pirskanen: High-Speed Packet Access Evolution in 3GPP Release 7, IEEE Communications Magazine, 45 (12), pp. 29-36, Dec. 2007.

[Hoßfeld_ Schatz_ Biedermann_11] T. Hoßfeld, R. Schatz, S. Biedermann, A. Platzer, S. Egger, M. Fiedler: The Memory Effect and Its Implications on Web QoE Modeling. Proc. ITC 2011, San Francisco, USA, September 2011

[Hsieh05] H.-Y. Hsieh and R. Sivakumar, ”A Transport Layer Approach for Achieving Aggregate Bandwidths on Multihomed Mobile Hosts,” ACM/Springer Wireless Networks Journal, Volume 11, Number 1-2, pp. 99-114, January 2005.

[HST+09] Hoßfeld, Tobias, Schlosser, Daniel, Tutschku, Kurt, Tran-Gia, Phuoc. Coopereation Strategies for P2P Content Distribution in Cellular Mobile Networks: Considering Selfishness and Heteo-geneity In: Mobile Peer-to-Peer Computing for Next Generation Distributed Environments: Advancing Conceptual and Algorithmic Applications.Editor: B.-C. Seet. IGI Global, Hershey, PA, USA, May, 2009

[HWL+08] C. Huang, A. Wang, J. Li, K. Ross, Understanding hybrid CDN-P2P, Proc. NOSSDAV Conf., Braunschweig, Germany (2008) 75-80

[ICT_ EARTH_11] Project INFSO-ICT-247733 EARTH (EU FP7), WP2, Deliverable 2.2: ”Definition and Parameterization of Reference Systems and Scenarios”, June 30, 2011, pp 28-40

[IEEE802.15.3] Draft Standard for Telecommunications and Information Exchange Between Systems -- LAN/MAN Specific Requirements -- Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPAN) Replaced by IEEE 802.15.3-2003, 2007

[IEEE802.16-2004] Draft IEEE Standard for Local and metropolitan area networks Corrigendum to IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems Corrigendum to IEEE Std 802.16-2004, 2004

[Informa_08] Informa Telecoms & Media: Mobile Broadband Access at Home – Aug. 2008.

[IRTF] Internet Engineering Task Force (IETF) <www.ietf.org>, Internet Research Task Force (IRTF) <irtf.org>,

- working group on Application Layer Traffic Optimization (ALTO) <tools.ietf.org/wg/alto/charters>,
- working group on MultiProtocol Label Switching (MPLS), <tools.ietf.org/wg/mpls/charters>,
- working group on Peer-to-Peer Streaming Protocol (PPSP) <tools.ietf.org/wg/ppsp/charters>,
- working group on CDN interconnection (CDNI) <tools.ietf.org/wg/cdni/charters>,
- peer-to-peer research group <www.irtf.org/charter?gtype=rg&group=p2prg>

[ITU] International Telecommunication Union (ITU) <www.itu.int>

[Joel_72] A. E. Joel, “Mobile Communication System,” U.S. Patent 3,663,762 (May 1972)

[KAM10] N. Kamiyama et al., ISP-operated CDN, 14th NETWORKS Telecom. Network Strategy & Planning Symposium, Warszawa, Poland (2010).

[Karrer05] R. Karrer and E. Knightly, "TCP-PARIS: A Parallel Download Protocol for Replicas," In Proceedings of IEEE International Workshop on Web Content Caching and Distribution (WCW 2005), Sophia Antipolis, France, pp. 15-25, September 12-13 2005.

[Kashihara10] S. Kashihara and M. Tsurusawa, ”Dynamic bandwidth management system using IP flow analysis for the QoS-assured network“, In Proc. of the IEEE GLOBECOM 2010, December 2010.

[KBB10] G. Kandavanam, D. Botvich and S. Balasubranmaniam, PaCRA: A path-aware content replication approach to support QoS guaranteed video on demand service in metropolitan IPTV networks, IEEE/IFIP Network Operations & Mgnt. Symp. NOMS (2010) 591-598

[Khandekar_10] A. Khandekar, N. Bhushan, Ji Tingfang, V. Vanghi: LTE-Advanced: Heterogeneous Networks, In Proceedings of the European Wireless Conference, pp. 978 – 982, April 2010.

[Kim07] K.-H. Kim and K. G. Shin, ”PRISM: Improving the Performance of Inverse-Multiplexed TCP in Wireless Networks,” IEEE Transactions on Mobile Computing, Vol. 6, Issue 12, pp. 1297-1312, December 2007.

[Kim09] S.-W. Kim, Y.-H. Lee, “Adaptive MIMO Mode and Fast Cell Selection with Interference Avoidance in Multi-cell Environments,” *5th International Conference on Wireless and Mobile Communications (ICWMC)*, pp.163-167, 2009.

[Kineto10] Kineto Wireless, “Smart Offload for Smartphones,” Whitepaper, 2010.

[Klein04] T. Klein et al., “Assignment strategies for mobile data users in hierarchical overlay networks: performance of optimal and adaptive strategies,” IEEE J. Sel. Areas Commun., vol. 22, no. 5, June 2004.

[Kostopoulos10] A. Kostopoulos, H. Warma, T. Leva, B. Heinrich, A. Ford, and L. Eggert, ”Towards Multipath TCP Adoption: Challenges and Opportunities,” 6th EURO-NF Conference on Next Generation Internet (NGI), pp. 1-8, Paris, 2-4 June 2010.

[Kwan_Cong_10] R. Kwan, R. Arnott et al., “On Pre-emption and Congestion Control for LTE Systems”, Proc. of IEEE Vehicular Technology Conference (VTC), Fall, 2010

[Kwan_Mob_10] R. Kwan, , R. Arnott et al.: ”On Mobility Load Balancing for LTE Systems”, Vehicular Technology Conference Fall (VTC 2010-Fall), 2010

[Kwan_Radio_10] R. Kwan, R. Arnott et al., “On Radio Admission Control for LTE Systems”, Proc. of IEEE Vehicular Technology Conference (VTC), Fall, 2010

[Kwon11] Y. J. Kwon, D.-H. Cho, “Load Based Cell Selection Algorithm for Faulted Handover in Indoor Femtocell Network”, *IEEE VTC* Spring, pp.1-5, 2011.

[Lagrange96] X. Lagrange, P. Godlewski, “Performance of a hierarchical cellular network with mobility-dependent handover strategies,” *in Proceedings of the Vehicular Technology Conference*, 1996, pp. 1868-1872.

[Lee02] Y. Lee, I. Park, and Y. Choi, ”Improving TCP Performance in Multipath Packet Forwarding Networks,” Journal of Communication and Networks (JCN), Vol. 4, No. 2, pp.148-157, June 2002.

[Lee06] K.-W. Lee, J.-Y. Ko, Yong-Hwan Lee, “Fast Cell Site Selection with Interference Avoidance in Packet Based OFDM Cellular Systems”, *Global Telecommunications Conference*, 2006.

[Leung07] K.-C. Leung, V. O. K. Li, and D. Yang, ”An Overview of Packet Reordering in Transmission Control Protocol (TCP): Problems, Solutions, and Challenges,” IEEE Transactions on Parallel and Distributed Systems, Vol. 18, Issue 4, pp. 522-535, April 2007.

[LGS07] J. Ledlie, P. Gardner and M. Seltzer, Network coordinates in the wild, Proc. USENIX Conf. (2007) 299-311

[Li08] B. J. Li and C. L. Soung, Improving Throughput and Fairness by Reducing Exposed and Hidden Nodes in 802.11 Networks. IEEE TRANSACTIONS ON MOBILE COMPUTING, vol. 7, no. 1, January 2008

[Li09] Hua Li, Md. Humayun Kabir, Takuro Sato. “Velocity adaptive vertical handoff on multi-frequency system”. *In Proceedings of PIMRC'2009*, pp.773-777.

[Lima07] S. R. Lima, P. Carvalho and V. Freitas, ”Admission Control in Multiservice IP Networks: Architectural Issues and Trends“, IEEE Communications Magazine, vol.45, no.4, pp.114-121, April 2007

[Lohier08] S. Lohier, Y. Ghamri Doudane, and G. Pujolle, “Cross-layer design to improve elastic traffic performance in WLANs”, ACM International Journal of Network Management, Volume 18 Issue 3, July 2008.

[Lopez09] D. López-Pérez et al., “OFDMA Femtocells: A Roadmap on Interference Avoidance,” IEEE Commun. Mag., vol. 47, no. 9, Sept. 2009, pp. 41–48.

[LRL+10] K. Lee, I. Rhee, J. Lee, S. Chong, and Y. Yi, “Mobile data offloading: how much can WiFi deliver?” in *Proc. of Co-NEXT*. ACM, 2010, pp. 26:1–26:12.

[Maheshwari09] R. Maheshwari, C. Jing and S.R. Das, “Physical Interference Modeling for Transmission Scheduling on Commodity WiFi Hardware”, In Proc. of IEEE INFOCOM '09, 2009

[Mahmoud10] H. Mahmoud, I. Güvenc, and F. Watanabe, “Performance of Open Access Femtocell Networks with Different Cell-Selection Methods,” in Proc. of *the 71st IEEE Vehicular Technology Conference (VTC)*, May 16–19 2010, pp. 1–5.

[Mathar02] R. Mathar and M. Schmeink, “Integrated optimal cell site selection and frequency allocation for cellular radio networks,” Telecommunication Systems, vol. 21, pp. 339–347, 2002.

[MCK04] J. Mogul, Y. Chan and T. Kelly, Design, Implementation and evaluation of duplicate transfer detection in HTTP, Proceedings 1. Symposium on Network Systems Design and Implementation (2004) 43-56

[McNair04] Janise McNair and Fang Zhu. Vertical handoffs in fourth-generation multinetwork environments. *Wireless Communications, IEEE*, 11(3):8–15, June 2004.

[Menth08] M. Menth, S. Kopf, J. Charzinski and K. Schrodi, ”Resilient network admission control“, Comput. Netw., vol.52, no.14, pp. 2805-2815, October 2008.

[Menth10] M. Menth, F. Lehrieder, B. Briscoe, P. Eardley, T. Moncaster, J. Babiarz, A. Charny, X. Zhang, T. Taylor, K.-H. Chan, D. Satoh, R. Geib and G. Karagiannis, ”A survey of PCN-based admission control and flow termination“, IEEE Communications Surveys & Tutorials, vol.12, no.3, pp. 357-375, 2010.

[Moon10] J.-M. Moon and D.-H. Cho, “Efficient Cell Selection Algorithm in Hierarchical Cellular Networks: Multi-User Coordination,” IEEE Communication Letters, Vol. 14, No. 2, February 2010.

[Moon10a] J.-M. Moon and D.-H. Cho, “Cell Selection Algorithm Based on Competition of Users in Hierarchical Cellular Networks,” *IEEE WCNC*, 2010, pp. 1–6.

[Morimoto_09] A. Morimoto et al.: "Investigation on Optimum Radio Link Connection Using Remote Radio Equipment in Heterogeneous Network for LTE-Advanced", IEEE Vehicular Technology Conference, 26-29 April 2009, Barcelona, Spain, pp. 1-5

[Mortier00] R. Mortier, I. Pratt, C. Clark, and S. Crosby, ”Implicit admission control“, IEEE Journal on Selected Areas in Communications, vol.18, no.12, pp.2629-2639, Dec 2000.

[MRT10] “ALTO in Mobile Core”, Y. El Mghazli, S. Randriamasy, F. Taburet, Alcatel-Lucent Bell Labs France, October 23 2010, http://tools.ietf.org/html/draft-randriamasy-alto-mobile-core-01

[Neu_Mobile_09] Neu Mobile Ltd, Mobile Traffic Growth + Cost Pressures = New Solutions? – Aug. 2009. [Online:] http://www.neu-mobile.com/report_finalv2.pdf (Accessed: July. 12, 2011)

[NokiaSiemensNetworks_1] Mobile broadband with HSPA and LTE –capacity and cost aspects (whitepaper) <http://www.nokiasiemensnetworks.com/>.

[NokiaSiemensNetworks_2] Nokia Siemens Networks, “Improving 4G coverage and capacity indoors and at hotspots with LTE femtocells”, White paper (2011) <http://www.nokiasiemensnetworks.com/>.

[PFA+10] I. Poese, B. Frank, B. Ager, G. Smaragdakis, A. Feldmann, Improving content delivery using provider-aided distance information, Proc. Internet Measurement Conference IMC’10, Melbourne, Australia (2010) 22-34

[Qu10] T. Qu, D. Xiao ,D. Yang, “A novel cell selection method in heterogeneous LTE-advanced systems”, *Proceedings of IEEE IC-BNMT2010*, pp.510-513, 2010.

[Radunovic08] B. Radunovic, C. Gkantsidis, D. Gunawardena, and P. Key, ”Horizon: Balancing TCP over Multiple Paths in Wireless Mesh Network,” 14th Annual International Conference on Mobile Computing and Networking (MobiCom 2008), 14-19 September 2008.

[Randriamasy12] S. Randriamasy (editor) and N. Schwan, “ALTO Cost Schedule”, IETF draft draft-randriamasy-alto-cost-schedule-00, March 5 2012, Presented at the 83rd IETF, Paris (March 2012), http://tools.ietf.org/id/draft-randriamasy-alto-cost-schedule-00.txt

[RAN11] “Multi-Cost ALTO”, S. Randriamasy, Alcatel-Lucent Bell Labs France, July 11 2011, http://tools.ietf.org/id/draft-randriamasy-alto-multi-cost-03.txt

[REE01] W.J. Reed, The Pareto, Zipf and other power laws, Economics Letters 74 / 1 (2001) 15-19

[RFC4655] A. Farrel, J.-P. Vasseur, and J. Ash, A Path Computation Element (PCE)-Based Architecture, IETF RFC 4655, August 2006. http://tools.ietf.org/html/rfc4655

[RFC4960] R. Stewart (Ed.), Stream Control Transmission Protocol, IETF RFC 4960, September 2007. http://tools.ietf.org/html/rfc4960

[RFC5555] H. Soliman (editor), “Mobile IPv6 Support for Dual Stack Hosts and Routers”, IETF RFC 5555, June 2009.

[RFC6089] G. Tsirtsis, H. Soliman, N. Montavont, G. Giaretta, K. Kuladinithi, “Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support”, IETF RFC 6089, January 2011.

[RFC6372] N. Sprecher (Ed.) and A. Farrel (Ed.), MPLS Transport Profile (MPLS-TP) Survivability Framework, IETF RFC 6372, September 2011. http://tools.ietf.org/html/rfc6372

[Sakellari10] G. Sakellari and E. Gelenbe, ”A distributed admission control mechanism for multi-criteria QoS“, In Proc. of IEEE GLOBECOM 2010, Workshop on Advances in Communications and Networks, pp. 1195-1999, December 2010.

[SAN10] Sandvine Inc., Fall 2010 global Internet phenomena report <www.sandvine.com> (2010)

[Sang04] A. Sang, X. Wang, M. Madihian, and R. D. Gitlin, “A Load-aware handoff and cell-site selection scheme in multi-cell packet data systems,” in Proceedings of the *IEEE 47th Global Telecommunications Conference (GLOBECOM).*, vol. 6, 2004, pp. 3931–3936.

[Sangiamwong11] J. Sangiamwong, et al., “Investigation on Cell Selection Methods Associated with Inter-cell Interference Coordination in Heterogeneous Networks for LTE-Advanced Downlink*”, European Wireless 2011*, April 27-29, 2011.

[SCK+09] A.-J. Su, D.R. Choffnes, A. Kuzmanovic and F.E. Bustamante, Drafting behind Akamai, IEEE/ACM Trans. on Networking 17 (2009) 1752–1765

[SEH10+] "Interconnected Content Distribution in LTE Networks" C. Schwartz, J. Eisl, A. Halimiz, A. Rafetseder, and K. Tutschku, IEEE GlobeCom 2010 Workshop on Advances in Communications and Networks (ACN 2010), Miami USA Dec. 6th – 10th

[Shetty10] S. Shetty, T. Ying and W. Collani, "TCP Venoplus — A cross-layer approach to improve TCP performance in wired-cum-wireless networks using signal strength", in proc. of International Conference of Networking, Sensing and Control (ICNSC), 2010

[SHH07] H.M. Sigurdsson, U.R. Halldorsson and G. Haßlinger: Potentials and challenges of peer-to-peer based content distribution, Telematics and Informatics, Elsevier, Vol. 24 (2007) 348-365

[Simsek11] M. Simsek, et al. , "Performance of different cell selection modes in 3GPP-LTE macro-/femtocell scenarios" , *IEEE Wireless Advanced (WiAd) 2011*, pp.126-131, June 2011.

[Stevens-Navarro06] Enrique Stevens-Navarro and Vincent W.S. Wong. Comparison between vertical handoff decision algorithms for heterogeneous wireless networks. In *Proc of 63rd Vehicular Technology Conference, VTC 2006-Spring.*, volume 2, pages 947–951. IEEE, 2006.

[Susitaival09] R. Susitaival and S. Aalto, Adaptive load balancing with OSPF, in Traffic and Performance Engineering for Heterogeneous Networks, ed. D.D. Kouvatsos, pp. 85 - 107, 2009, River Publishers, Gistrup, Denmark.

[SYB+09] X. Shen, H. Yu, J. Buford and M. Akon (Eds.), Handbook of peer-to-peer networking, Springer (2009)

[Taleb_11] T. Taleb, K. Samdanis, S. Schmid, "DNS-Based Solution for Operator Control of Selected IP Traffic Offload", IEEE International Conference on Communications (ICC), ISBN: 978-1-61284-232-5, pp.1-5, 2011.

[TFR+11] R. Torres, A. Finamore, Jin Ryong, Kim; M. Mellia, M. M. Munafo; Sanjay Rao "Dissecting Video Server Selection Strategies in the YouTube CDN" 2011 31st International Conference on Distributed Computing Systems (June 2011), pg. 248-257

[Tongwei10] Tongwei Qu, Dengkun Xiao, Dongkai Yang, "A novel cell selection method in heterogeneous LTE-advanced systems", Broadband Network and Multimedia Technology (IC-BNMT), 3rd IEEE International Conference on, 2010

[Tran_Minh_Trung_11] Tran Minh Trung, Youn-Hee Han, Hyon-Young Choi, Hong Yong Geun, "A Design of Network-based Flow Mobility based on Proxy Mobile IPv6", In proc. of IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 373–378, ISBN: 978-1-4577-0249-5, April 2011.

[TSTAT] Tstat Homepage, http://tstat.polito.it.

[Tuffery_11] A. Tuffery et al.: "A 27.5-dBm linear reconfigurable CMOS power amplifier for 3GPP LTE applications", IEEE New Circuits and Systems Conference (NEWCAS), 26-29 June 2011, Bordeaux, France, pp. 221-224

[Wang_07] Q. Wang, R. Atkinson, C. Cromar, J. Dunlop, "Hybrid User- and Network-Initiated Flow Handoff Support for Multihomed Mobile Hosts", In proc. of IEEE 65th Vehicular Technology Conference, VTC2007-Spring, pp.748–752, ISBN: 1-4244-0266-2, April 2007.

[Wang99] H. J. Wang, R. H. Katz, and J. Giese. Policy-enabled handoffs across heterogeneous wireless networks. In *WMCSA '99: Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications*, page 51, Washington, DC, USA, 1999. IEEE Computer Society.

[Wee11] L. T. Wee, M. Portmann, and H. Peizhao, "A Systematic Evaluation of Interference Characteristics in 802.11-Based Wireless Networks" In Proc. IEEE International Conference on Advanced Information Networking and Applications (AINA), 2011

[WF10] F. Weiden , P. Frost "Anycast as a load balancing feature" LISA'10 Proceedings of the 24th international conference on Large installation system administration

[wikithc] http://wiki.thc.org/vodafone

[Wu08] D. Wu, P. Djukic, and P. Mohapatra, "Determining 802.11 link quality with passive measurements", in Proc. Wireless Communication Systems (ISWCS '08), 2008.

[WW09] S. Wiethölter and A. Wolisz, "Selecting vertical handover candidates in IEEE 802.11 mesh networks," in *Proc. of WoWMoM Workshops*. IEEE, June 2009, pp. 1 – 7.

[Yilmaz05] O. Yilmaz, A. Furuskar, J. Pettersson, and A. Simonsson. Access selection in WCDMA and WLAN multi-access networks. In *in Proc. of Vehicular Technology Conference, VTC Spring, IEEE*, volume 4, pages 2220–2224, 2005.

[Yi-Neng_Lin_10] Yi-Neng Lin, Wen Chen, Shan-Chi Tsai, Yi-Bing Lin: Design and Implementation of An Offloading Technology for 3.5G Networks, IEEE 71st Vehicular Technology Conference (VTC 2010-Spring), pp. 1 – 5, May 2010.

[Yokota_11] H. Yokota, D. Kim, B. Sarikaya, F. Xia, "Home Agent Initiated Flow Binding for Mobile IPv6", IETF Internet Draft, December 22, 2011.

[YSN10] X. Yan, Y. A. Sekercioglu, and S. Narayanan, "A survey of vertical handover decision algorithms in Fourth Generation heterogeneous wireless networks," *Computer Networks*, vol. 54, no. 11, pp. 1848 – 1863, 2010.

[Zhao11] L. Zhao, X. Li, T. Weerawardane, A. Timm-Giel and C. Görg, Joint Load Balancing of Radio and Transport Networks in LTE System, in Third International Conference on Ubiquitous and Future Networks (ICUFN 2011), Dallian, China, June 15-17, 2011.

[ZKG10] R. Zhou, S. Khemmarat, and L. Gao, "The Impact of YouTube Recommendation System on Video Views," IMC '10: Proceedings of the 10th annual conference on Internet measurement, 2010.

[ZSG+08] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Watch Global, Cache Local: YouTube Network Traffic at a Campus Network-Measurements and Implications," Proceedings of the 15th SPIE/ACM Annual Multimedia Computing and Networking Conference (MMCN), 2008.

# Appendix A: Implementation considerations of NB-IFOM prototype

Policy Server

The Policy Server is to be implemented as a multi-threaded background service running on UNIX/Linux platform. It uses the native socket and TCP/IP interface provided by the operating system.

Database

As the Policy Server node has to store quite extensive set of data structures, an external RDBMS is used to better visualize and organize data. The chosen solution must support foreign keys in order to maintain coherency throughout the data tables.

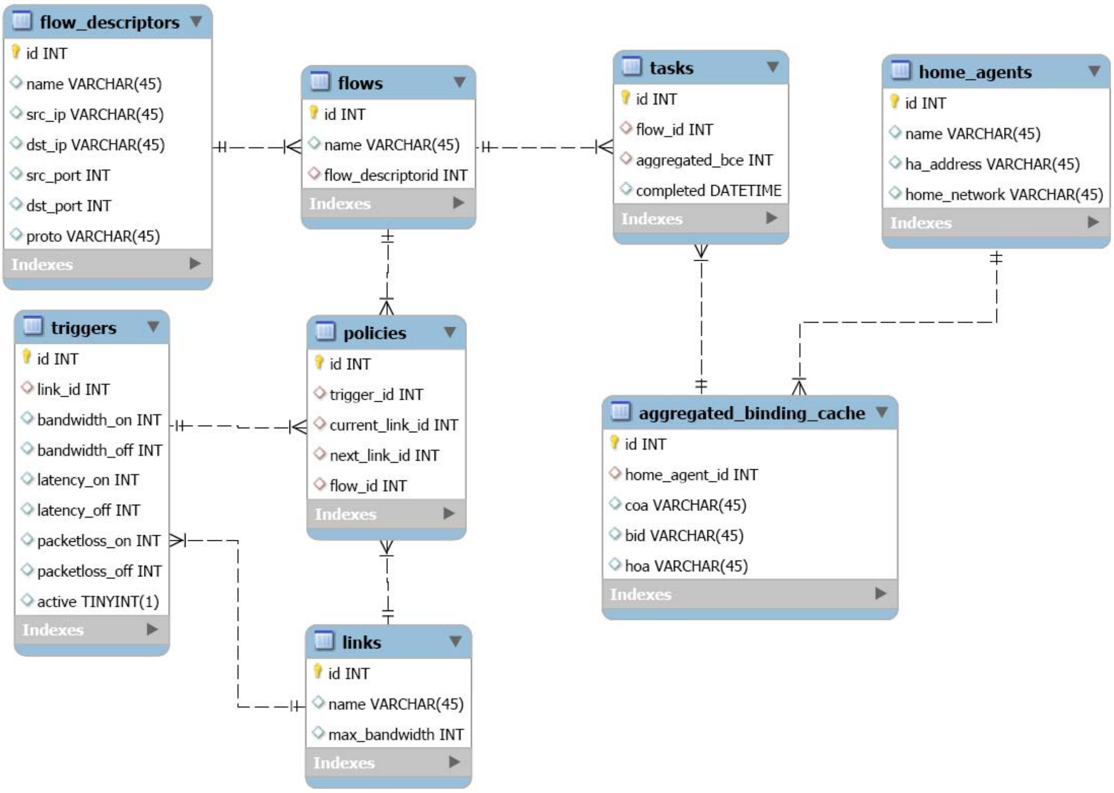


**Figure A.1:** ***Policy Server database layout***

Flow Descriptor Table

Used to identify a single flow based on 5-tuple information.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| name | string | Name of the stream |
| src_ip | IPv6 address | Source IPv6 address |
| dst_ip | IPv6 address | Destionation IPv6 address |
| src_port | integer | Source TCP/UDP port |
| dst_port | integer | Destination TCP/UDP port |
| proto | integer | Transport protocol id. |

Flows Table

As flow descriptors only identify a single flow, to identify a protocol consisting of several different flows, another abstraction layer is used.

| Field Name | Data type | Comment |
|---|---|---|

| | | |
|---|---|---|
| id | integer | Unique key |
| name | string | Protocol name |
| flow_descriptor_id | integer | Flow descriptor id key |

Home Agents Table

This table stores administrative information for each Home Agent.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| name | string | Name of the Home Agent |
| ha_address | IPv6 address | Home Agent address |
| home_network | IPv6 prefix | Home Network |

Aggregated Binding Cache Table

BCE entries received in the policyRequest messages are stored here.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| home_agent_id | integer | Home Agent id |
| hoa | IPv6 address | Home Address |
| coa | IPv6 address | Care-Of Address |
| bid | integer | Binding Identifier |

Tasks Table

Policy decisions are stored here. New entries are sent to the HA with the policyCommand message. Acknowledgement sets the completed field.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| flow_id | integer | Protocol flow id |
| aggregated_bce_id | integer | Aggr. Binding Cache Entry id. |

Triggers Table

Trigger statements are stored here.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| link_id | integer | Link id |
| bandwidth_on | integer | Policy valid if exceeded. |
| bandwidth_off | integer | Policy invalid if exceeded. |
| packetloss_on | integer | Policy valid if exceeded. |
| packetloss_off | integer | Policy invalid if exceeded. |
| delay_on | integer | Policy valid if exceeded. |
| delay_off | integer | Policy invalid if exceeded. |
| active | integer | Turn on/off trigger. |

Policies Table

Applicable policies are stored in this table, filled by trigger. Moves flow from one link to another.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| trigger_id | integer | Trigger id. |
| current_link_id | integer | Current link id. |
| next_link_id | integer | Next link id. |

| | | |
|---|---|---|
| flow_id | integer | Flow id. |

Links Table

Stores link information.

| Field Name | Data type | Comment |
|---|---|---|
| id | integer | Unique key |
| name | integer | Name (BID) |
| max_bandwidth | integer | Link capacity |

Threads

Measurement thread (decision module): Listens for measurements and triggers policies if certain conditions are met.

PolicyRequest thread: Listens for policyRequest messages and updates the aggregated_binding_cache table

PolicyCommand thread: Started by the decision module. Sends new tasks to the Home Agent, and waits for acknowledgement.

# Appendix B: IETF standardization on traffic management

Many IETF working groups are dealing with traffic management aspects which are relevant for Mevico WP4 work, especially those in the transport area. Basic protocols for TCP congestion control mechanisms and QoS support by differentiated services on the Internet have been introduced and developed by the IETF over the past decades.

In this section we give an overview on current activities for enhanced traffic management features. This includes three working groups looking at shortened and optimized traffic paths in overlay networks (ALTO: Application-Layer Traffic Optimization) and via storage on network elements (CDNI: CDN Interconnection; DECADE: Decoupled Application Data Enroute), as well as an improved information base for BGP discussed by the Inter-domain Routing (IDR) working group. Another main focus is on advanced TCP and congestion control functions, which is developed in working groups on Multipath TCP (MPTCP), Congestion Exposure (CONEX), Low Extra Delay Background Transport (LEDBAT) and Datagram Congestion Control Protocol (DCCP). Moreover, recently started working groups on Network Virtualization Overlays (NVO3) and Distributed Mobility Management (DMM) are also related to work on traffic management in MEVICO.

## B.1 Application-Layer Traffic Optimization (ALTO)

The IETF ALTO working group started in Nov. 2008.

### B.1.1 Main goals of IETF ALTO <datatracker.ietf.org/wg/alto/charter/>

The working group is designing and specifying an Application-Layer Traffic Optimization (ALTO) service that will provide applications with information to perform better-than-random initial peer selection. ALTO services may take different approaches at balancing factors such as maximum bandwidth, minimum cross-domain traffic, lowest cost to the user, etc. The WG will consider the needs of BitTorrent, tracker-less P2P, and other applications, such as content delivery networks (CDN) and mirror selection.

### B.1.2 Relevance of the IETF ALTO work in MEVICO WP4, results and current status

The relationship and relevance of the ALTO approach for MEVICO WP4 is described in D4.3.1 sections 5.2.5 and 6.5.1.3. Alcatel-Lucent is active as a MEVICO partner in this working group with two proposed drafts on

- ALTO cost schedule <draft-randriamasy-alto-cost-schedule-01>,
- and multi-cost ALTO <draft-randriamasy-alto-multi-cost-06>.

The working group has issued an RFC 5693 on the problem statement and there are four working group drafts in an advanced state on

- ALTO requirements,
- ALTO deployments,
- the ALTO protocol and
- ALTO server discovery.

Furthermore, about a dozen individual drafts are submitted and considered by the ALTO working group.

## B.2 Content Delivery Network Interconnection (CDNI)

The IETF CDNI working group started in June 2011.

### B.2.1 Main goals of IETF CDNI <datatracker.ietf.org/wg/cdni/charter/>

A Content Delivery Network (CDN) is an infrastructure of network elements operating at layer 4 through layer 7, arranged for the efficient distribution and delivery of digital content. Such content includes, but is not limited to,

- web pages and images delivered via HTTP,
- streaming of continuous media delivered via HTTP, RTSP, RTMP, etc.

CDNs typically provide services to multiple Content Service Providers (CSPs).

The goal of the CDNI Working Group is to allow the interconnection of separately administered CDNs in support of the end-to-end delivery of content from CSPs through multiple CDNs and ultimately to end users via their respective User Agents. The CDNI WG aims at delivering a targeted, deployable solution in a short timeframe (18-24 months) as needed by the industry. It is expected that the CDNI interfaces will be realized using existing IETF protocols for transport and message exchange, and using existing object notation grammars/languages for the definition of CDNI objects and semantics. In the event that protocol extensions or new protocols are deemed necessary by the WG, the WG will recharter.

### B.2.2 Relevance of the IETF CDNI work in MEVICO WP4

The CDN interconnection working group is addressing the problem of a lack of interaction between content delivery architectures under different administration in the Internet. Distribution of popular content on the Internet is often supported by several CDNs, including global CDNs, CDNs in the fixed network core of large network providers, and/or CDNs of mobile network operators. A large portion of Internet traffic is currently delivered via global CDNs. YouTube video streaming is a popular example generating more than 20% of Internet traffic according to measurement statistics [Sandvine, Cisco reports]. The CDN provider Akamai claims to have around 20% of Internet traffic for other application running over his global infrastructure.

### B.2.3 Results and Current status

After about one year of activity, the CDNI WG has adopted four documents as IETF CDNI working group drafts, which are worked out towards RFC status as the basis of the CDNI standardization: The *problem statement*, the *framework*, the *requirements* and *use cases*.

The main *problem statement* is that no standards or open specifications currently exist to facilitate CDN interconnection for support of content delivery in between the content provider and the requesting user. Several types of interfaces are required for control, logging (eventually including accounting data), routing requests, metadata and acquisition of content from surrogates, such as caches.

```
   --------
  /        \
  |   CSP  |
  \        /
   --------
       *
       *
       *                         /\
       *                        /  \
   ----------------------      |CDNI|        ----------------------
  /     Upstream CDN     \     |    |       /    Downstream CDN    \
  |      +-------------+ | Control Interface| +-------------+      |
  |*******   Control   |<======|====|========>|   Control   *******|
  |*     +------*----*-+ |     |    |       | +-*----*------+     *|
  |*            *    *   |     |    |       |   *    *            *|
  |*     +------*------+ | Logging Interface| +------*------+     *|
  |* *****   Logging   |<======|====|========>|   Logging   ***** *|
  |* *   +-*-----------+ |     |    |       | +-----------*-+   * *|
  |* *     *           *   | Request Routing  |   *           *     * *|
.....*...+-*---------*-+ |    Interface     | +-*---------*-+...*.*...
. |* * *** Req-Routing |<======|====|========>| Req-Routing *** * *| .
. |* * * +-------------+.|     |    |       | +-------------+ * * *| .
. |* * *                 .  CDNI Metadata   |                 * * *| .
. |* * * +-------------+ |.   Interface     | +-------------+ * * *| .
. |* * * | Distribution|<==.===|====|========>| Distribution| * * *| .
. |* * * |             | |  .   \  /        | |             | * * *| .
. |* * * |+---------+  | |   .   \/         | |  +---------+| * * *| .
. |* * ***| +---------+| |    ....Request......+---------+ |*** * *| .
. |* *****+-|Surrogate|************************|Surrogate|-+***** *| .
. |*******  +---------+| |   Acquisition    | |+----------+ *******| .
. |      +-------------+ |                  | +-------*-----+      | .
. \                      /                  \         *            / .
.  ----------------------                    ---------*------------  .
.                                                     *              .
.                                                     * Delivery     .
.                                                     *              .
.                                                  +--*---+          .
...............Request.............................| User |..Request..
                                                   | Agent|
                                                   +------+

<==>  interfaces inside the scope of CDNI
****  interfaces outside the scope of CDNI
....  interfaces outside the scope of CDNI
```

Figure 4: Interfaces for CDNI <IETF draft  CDNI problem statement>

Among the main *requirements*, CDN interconnection shall be done without changes to the user agents and the servers of content provisioning systems and shall reuse existing protocols rather than introduce new ones. Inefficient and inappropriate operation modes have to be avoided, e.g. loops in control, routing and delivery procedures or delivery of small data units with high overhead.

The *framework* draft includes a reference model, gives an overview on operations, including DNS and HTTP redirection modes, storing and removal of content etc. Those topics are closely related to the CDN chapters of the D.4.3.1 deliverable. Main interfaces are described. Deployment, trust and security aspects are considered.

The *use cases* include

- delivery of content from popular web sites,
- delivery of web-TV services,
- delivery of IPTV services,
- support for mobile terminals, nomadic users,
- shorting traffic paths and delays for reducing load on interconnection links,
- improving content availability and acquisition resilience in distributed systems,
- robustness against denial-of-service attacks,
- offload for overload handling between different CDNs,
- enhancement of CDN capabilities by support of another CDN.

```
           -----------                              ----------
          /    CDN A  \                            /   CDN B   \
          | +-----+   |                           |  +-----+   |
//=========>| C   |<===============CDNI=============>| C   |<===========\\
||        | +-----+   |             C             |  +-----+   |        ||
||        | +-----+   |                           |  +-----+   |        ||
||  //=====>| D   |<===============CDNI=============>| D   |<=======\\  ||
||  ||    | +-----+   |             M             |  +-----+   |    ||  ||
||  ||    |           |     /-------------\       |            |    ||  ||
||  ||    | +-----+   |     | +--+  CDN Ex|       |  +-----+   |    ||  ||
||  || //==>| RR  |<===CDNI==>|RR|<========CDNI=====>| RR  |<====\\ ||  ||
||  || || | +-----+   | RR  | +--+        | RR    |  +-----+   | || ||  ||
||  || || |           |     |  /\         |       |            | || ||  ||
||  || || | +-----+   |     |  ||   +---+ |       |  +-----+   | || ||  ||
||  || || | | L   |<===CDNI========>| L |<=CDNI=====>| L   |   | || ||  ||
||  || || | +-----+   |  L  |  ||   +---+ |   L   |  +-----+   | || ||  ||
||  || || \           /     \  ||     /\  /       \            / || ||  ||
||  || || ------------       --||-----||--         ------------  || ||  ||
||  || ||                     ||     ||                         || ||  ||
||  || ||                   CDNI RR   ||                         || ||  ||
||  || ||                     ||    CDNI L                      || ||  ||
||  || ||                     ||     ||                         || ||  ||
||  || ||                  ---||-----||----                     || ||  ||
||  || ||                 /   \/     ||     \                   || ||  ||
||  || ||                 |  +-----+ ||     |                   || ||  ||
||  || \\=====CDNI===========>| RR  |<==============CDNI=========// ||  ||
||  ||          RR          |  +-----+ \/     |       RR             ||  ||
||  ||                     |         +-----+ |                      ||  ||
||  ||                     |         | L   | |                      ||  ||
||  ||                     |         +-----+ |                      ||  ||
||  ||                     |  +-----+       |                      ||  ||
||  \\=======CDNI============>| D   |<==============CDNI============//  ||
||            M             |  +-----+       |       M                  ||
||                         |  +-----+       |                          ||
\\===========CDNI============>| C   |<==============CDNI================//
              C             |  +-----+       |       C
                           \         CDN C   /
                            ---------------
```

Figure 5: CDNI exchange in the deployment model <IETF draft CDNI framework>

There are about two dozen *IETF drafts* proposed for inclusion in the CDNI work, whose adoption is open and to be discussed in the next steps.

### B.2.4 Conclusion

On the whole, the IETF CDNI working group provides an overview of the main aspects of CDNs and their interworking with many contributions on details. The main focus is on interfaces, which are currently missing between global CDNs and caching and CDN systems of network providers. The impact on routing is addressed and the handling of content as well as the information exchange between CDNs, which are also relevant for traffic management.

After one year, a broad approach of the CDNI working group is visible and the basic documents are advanced and developing towards RFC status. The working group is expected to proceed in the next year with details addressed in many individual drafts. The CDNI work provides background information for Mevico WP4 and on the other hand still is open for input.

## B.3 Decoupled Application Data Enroute (DECADE)

The IETF DECADE working group started in April 2010.

### B.3.1 Main goals of IETF DECADE <datatracker.ietf.org/wg/decade/charter/>

The main focus of DECADE is on support of P2P applications by in-network storage in order to avoid upload traffic from peers on the last mile which is especially important on air interfaces in mobile networks. An open protocol should allow all P2P applications to make use of in-network storage and usage scenarios by other applications are also addressed.

Besides documents on the problem statement and requirements, a DECADE architecture is worked out and the integration with main P2P protocols is considered.

### B.3.2 Relevance of the IETF DECADE work in MEVICO WP4

In principle, unloading the air interface is a main goal of traffic management in mobile networks. DECADE proposes a special type of caches mainly for P2P networking. P2P applications cannot utilize HTTP-based web caches, with the exception of an approach by the eDonkey file sharing protocol to disguise P2P traffic as usual HTTP traffic. However, P2P networking is less relevant in mobile than in fixed networks and it is questionable whether network providers are motivated to support especially P2P applications by providing in-network storage. The alternative to provide in-network storage on the application layer or by the P2P network providers would require a global infrastructure or would be less efficient.

### B.3.3 Results and current status

The working group has finalized two informational RFC documents

- RFC 6646 for the problem statement and
- RFC 6392 for a survey of in-network storage systems.

The latter gives a framework listing main types, functions, protocols, and deployments of in-network storage for various purpose, thus extending MEVICO WP4 documents on caching and content delivery networks. In addition, three working group drafts are in an advanced state on

- the DECADE architecture,
- DECADE requirements and
- integration examples of the DECADE system, which include P2P file sharing and streaming and ALTO mechanisms.

Figure 6 shows the DECADE client/server interworking as a basis of the architecture, giving the application end-points more influence on the in-network storage mechanisms than in transparent CDN and caching systems.

```
                     Native Application
     .-------------.      Protocol(s)     .-------------.
     | Application | <------------------> | Application |
     |  End-Point  |                      |  End-Point  |
     |             |                      |             |
     | .--------.  |                      | .--------.  |
     | | DECADE |  |                      | | DECADE |  |
     | | Client |  |                      | | Client |  |
     | `--------'  |                      | `--------'  |
     `-------------'                      `-------------'
         |     ^                              |     ^
 DECADE  |     | Standard                     |     |
Resource |     |   Data                   DRP |     | SDT
Protocol |     | Transfer                     |     |
 (DRP)   |     |   (SDT)                      |     |
         |     |                              |     |
         |     |                              |     |
         |     |                              |     |
         |     |                              |     |
         |     |                              |     |
         |     |                              |     |
         v     V                              v     V
     .=============.         DRP          .=============.
     |   DECADE    | <------------------> |   DECADE    |
     |   Server    | <------------------> |   Server    |
     `============='         SDT          `============='
```

Figure 6: DECADE generic protocol flow

### B.3.4 Conclusions

The DECADE work has some relevance to support P2P applications with in-network storage. In this way, P2P traffic can be supported by caches using an alternative standard besides HTTP caching. The most important benefit can be seen in offloading upload traffic from mobile peers via air interfaces to nodes in the EPC. It is open whether this additional caching option for P2P networking will be supported by network providers.

## B.4 BGP link state information distribution (draft to Interdomain Routing IDR)

The inter-domain routing working group (IDR) within the IETF is constantly developing and maintaining the Border Gateway Routing Protocol (BGP) as the de-facto standard for inter-domain information exchange.

A current activity addresses the North-bound interface type information exchange, where end-to-end traffic engineering or other overlay service types might benefit from interior gateway routing information, which is not normally visible outside the Autonomous System's routing domain.

Current activities within IETF that make such intrinsic information partially visible to outside or over-the-top services include e.g. Path Computation Element (PCE) [RFC4655] and the ALTO Server [RFC5693], which have been developed in their own working group.

The activity within IDR now suggests standardizing a common information distribution format for such services based on the Network Layer Reachability Information (NLRI) conveyed within BGP protocol messages. The respective draft can be found at <datatracker.ietf.org/doc/draft-gredler-idr-ls-distribution>.

## B.5 Multipath TCP (MPTCP)

The IETF MPTCP working group started in Oct. 2009.

### B.5.1 Main goals of IETF MPTCP <datatracker.ietf.org/wg/MPTCP/charter/>

The Multipath TCP (MPTCP) working group develops mechanisms that add the capability of simultaneously using multiple paths to a regular TCP session. Key goals for MPTCP are:

- to be deployable and usable without significant changes to existing Internet infrastructure;
- to be usable by unmodified applications;
- and to be stable and congestion-safe over the wide range of existing Internet paths, including NAT interactions.

MPTCP assumes that both peers are modified and that one or both peers have multiple addresses, which often results in different network paths that are at least partially divergent (however, note there is no guarantee that the paths are divergent at all).

### B.5.2 Relevance of the IETF ALTO work in MEVICO WP4, results and current status

The relationship and relevance of the ALTO approach for MEVICO WP4 is described in D4.3.1 sections 3.4.4 and 6.5.1.4. Technische University of Berlin is active as a MEVICO partner in this working group with a draft on

"A Transparent Performance Enhancing Proxy Architecture To Enable TCP over Multiple Paths for Single-Homed Hosts" <draft-ayar-transparent-sca-proxy-00>.

The working group has issued three RFCs on

- Architectural Guidelines for Multipath TCP Development (RFC 6182),
- Threat Analysis for TCP Extensions for Multipath Operation with Multiple Addresses (RFC 6181),
- and Coupled Congestion Control for Multipath Transport Protocols (RFC 6356).

In addition, there are two advanced working group drafts on

- MPTCP application interface considerations < draft-ietf-mptcp-api-05> and
- TCP extensions for multipath operation with multiple addresses <draft-ietf-mptcp-multiaddressed-09>

## B.6 Congestion Exposure (CONEX)

The IETF CONEX working group started in June 2010.

### B.6.1 Main goals of IETF CONEX <datatracker.ietf.org/wg/conex/charter/>

The purpose of the CONEX working group is to develop a mechanism by which senders inform the network about the congestion encountered by previous packets on the same flow. Today, the network may signal congestion by ECN markings or by dropping packets, and the receiver passes this information back to the sender in transport-layer acknowledgements. The mechanism to be developed by the CONEX WG will enable the sender to also relay the congestion information back into the network in-band at the IP layer, such that the total level of congestion is visible to all IP devices along the path, from where it could, for example, be provided as input to traffic management.

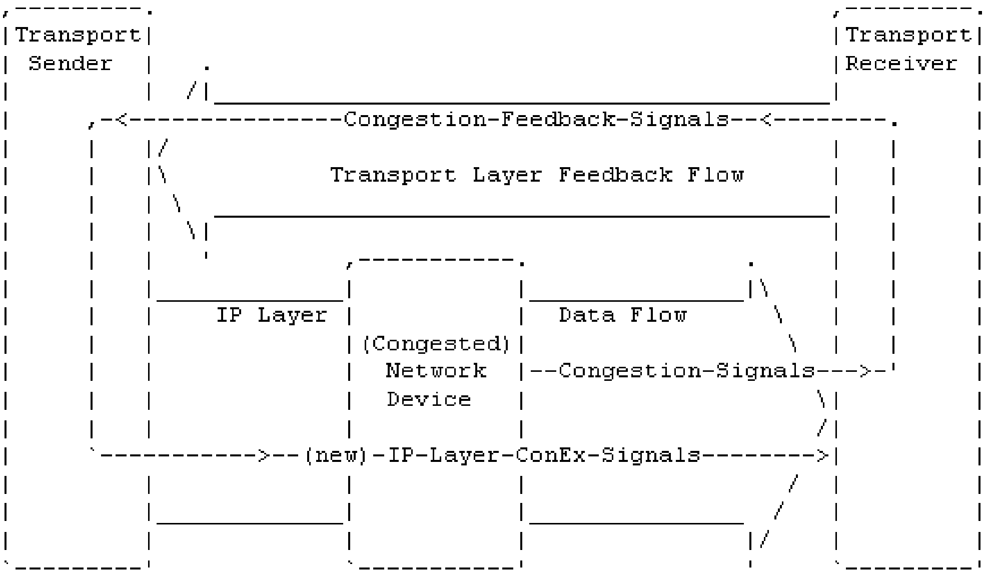

Transport Sender
Transport Receiver
Congestion-Feedback-Signals
Transport Layer Feedback Flow
IP Layer
(Congested) Network Device
Data Flow
Congestion-Signals
(new)-IP-Layer-ConEx-Signals

Figure 7: CONEX information exchange scheme <datatracker.ietf.org/doc/draft-ietf-conex-abstract-mech>

### B.6.2 Relevance of the IETF CONEX work in MEVICO WP4

Information exchange on congestion in the network is essential for traffic management. CONEX aims at making the awareness of congestion at the sender also available to the network. The sender can get such information via feedback of acknowledgements in TCP flows. Better knowledge of the network status enables more timely and precise reactions of the traffic management with congestion control loops involving network elements on the time scale of one or several round trip delays.

### B.6.3 Results and current status

Currently there are five main drafts adopted as CONEX working group drafts. In addition to describing concepts and use cases, they especially focus on

- mobile communication <draft-ietf-conex-mobile-00>
- TCP modifications <draft-ietf-conex-tcp-modifications-02> and
- IPv6 options <draft-ietf-conex-destopt-02>.

In particular, the draft on mobile communication is closely related to MEVICO WP4 work including sections on

- Overview of 3GPP’s Evolved Packet System (EPS)
- CONEX Use Cases in the Mobile Communication Scenario
- CONEX as a Basis for Traffic Management
- CONEX to Incentivize Scavenger Transports
    - Accounting for Congestion Volume
    - CONEX as a Form of Differential QoS
- CONEX in the EPS
    - Deployment Scenarios
    - Implementing CONEX Functions in the EPS

- CONEX Protocol Mechanisms
- CONEX Functions in the Mobile Network

### B.6.4 Conclusions

The CONEX working group of the IETF is considering relevant contributions to improve congestion status awareness in the network including the sender and information available at the end points of a TCP/IP communication. The draft on mobile communication gives a high level overview also on traffic management in the EPC, which has been worked out in detail in the MEVICO WP4 deliverables.

### B.6.5 Remark on Congestion and Pre-Congestion Notification (PCN)

The PCN working group has set up a scheme for data exchange about congestion between different network domains and suggested control mechanisms by not admitting new flows or even by terminating flows in high and overload conditions. The PCN working group has been concluded in 2012.

## B.7 Low Extra Delay Background Transport (LEDBAT)

The IETF LETBAT working group started in November 2008.

### B.7.1 Main goals of IETF LEDBAT <datatracker.ietf.org/wg/ledbat/charter/>

The main focus of LEDBAT is on delay problems of the transport protocol TCP in case of congestion, when queues of an IP router on the path are filling and add to the end-to-end delay. The goal is to utilize the bottleneck and at the same time to keep buffering delays low.

### B.7.2 Relevance of the IETF LEDBAT work in MEVICO WP4, results and current status

In principle the performance of transport protocols and their reaction in congestion are in the focus of MEVICO WP4, but although LEDBAT is announced to be active, the working group did not meet since mid 2010. The activity until 2010 produced an informational RFC 6297 on a survey of lower than best effort protocols and a remaining working group draft on Low Extra Delay Transport, which is currently expired. So it seems questionable whether the working group will proceed.

## B.8 Datagram Congestion Control Protocol (DCCP)

The IETF DPPC working group started in 2002.

### B.8.1 Results and current status

The Datagram Congestion Control Protocol (DCCP) [RFC4340-RFC4342] is a transport protocol that provides bidirectional unicast connections of congestion-controlled unreliable datagrams. DCCP is suitable for applications that transfer fairly large amounts of data and that can benefit from control over the tradeoff between timeliness and reliability.

The DCCP protocol has been established in a series of half a dozen RFCs in the Proposed Standards Track since 2006. The DCCP working group is still active but only on extensions and did not participate in IETF meetings in 2011 and 2012.

## B.9 Network Virtualization Overlay (NVO3)

The IETF PCN working group started in May 2012.

### B.9.1 Description of IETF Working Group NVO3 <datatracker.ietf.org/wg/nvo3/charter/>

Data centers supporting virtualized hosts and machines from different parties by a virtualized environment for each of the parties are the main subject of the working group. Therefore the three key requirements are:

- traffic isolation, so that one party's traffic is not visible to any other,
- address independence, so that one party's addressing scheme does not collide with another within the data center and
- support for the placement and migration of virtual machines anywhere within the data center, without being limited by data center network constraints such as the IP subnet boundaries.

The basic view of the framework draft <tools.ietf.org/pdf/draft-lasserre-nvo3-framework-03.pdf> on the data center architecture is shown in Figure 8.

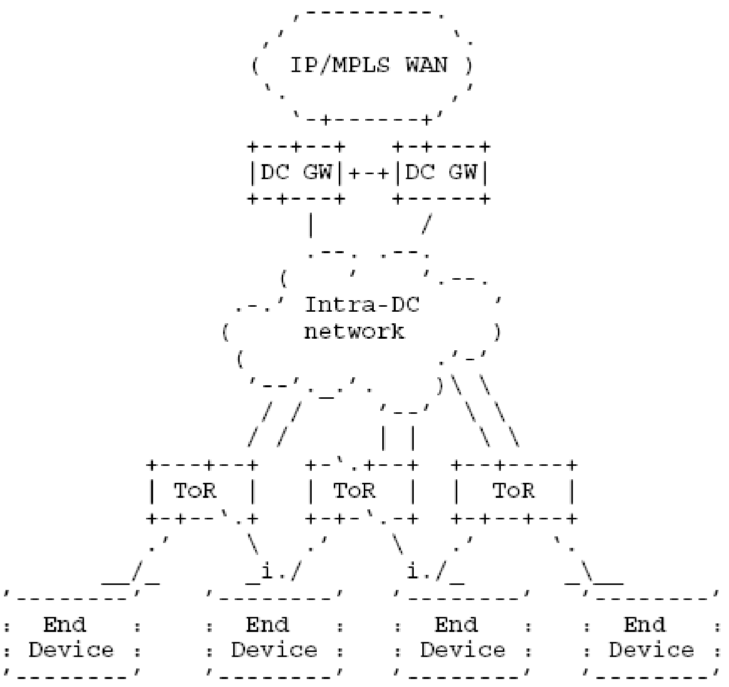


Figure 8: Generic architecture for data centers (DC) (ToR: Ethernet aggregation switch)

### B.9.2 Current status and relevance of the IETF NVO3 work in MEVICO WP4

The working group is in its starting phase without having produced working group of RFC documents. About a dozen individual drafts are already proposed. The role of data centers in the mobile core is of importance for MEVICO WP4 and virtualization approaches are especially relevant in MEVICO+.

## B.10 Distributed Mobility Management (DMM)

The IETF DMM working group started in March 2012.

### B.10.1 Main goals of IETF DMM <datatracker.ietf.org/wg/dmm/charter/>

A charter is not jet officially published. The architecture draft proposes a distributed architecture of mobility management in terms of abstracted logical functions <draft-chan-dmm-architecture-00>. Mobility management functions are abstracted into different logical functions:

- allocation of home network prefixes or home addresses to mobile nodes,
- location management which includes managing the IP addresses and locations of the mobile nodes,
- and mobility routing which includes intercepting and forwarding packets.

A distributed architecture can be constructed by providing the mobility routing functions in multiple networks in the data plane and a distributed database is used to host the location management function. This generalized architecture enables different distributed mobility designs using primarily the existing mobility protocols (MIP and PMIP) and their extensions. Several existing distributed mobility management proposals are briefly reviewed using this framework. It is expected that the different proposals, when expressed in terms of the generalized framework of logical functions, can interwork with each other as well as with the existing hierarchical deployments.

### B.10.2 Relevance of the IETF DMM work in MEVICO WP4 results and current status

The DMM working group is still in the starting phase. There are no working group drafts and RFCs adopted by the working group, but already around two dozen individual drafts submitted. Prospective working group drafts are prepared for the framework and gap analysis as well as for requirements.

- <draft-chan-dmm-framework-gap-analysis-02>
- <draft-ietf-dmm-requirements-01>

In addition, a draft <draft-demaria-dmm-dimensioning-considerations-00> on dimensioning and cost aspects of distributed versus centralized mobility management is proposed by Telecom Italia. This draft describes a general method to calculate the costs of the centralized and distributed scenarios. A simple model is introduced in order to calculate the costs of mobility management options based on data available to an operator. Such work is related to approaches in MEVICO WP4 and WP6, although the draft seems preliminary in the current version without much detail.