ICT-601102 STP TUCAN3G

Wireless technologies for isolated rural communities in developing countries based on cellular 3G femtocell deployments

M43

Multiple flows control management

Contractual Date of Delivery to the CEC:	**31th December 2013**
Actual Date of Delivery to the CEC:	**31st January 2013**
Author(s):	**O. Muñoz, A. Pascual, J. F. Rubio (UPC), A. G. Marqués, C. Figuera (URJC), K. Kilgour, A. Ratanov (IPA), O. Tupayachi, A. Ueunten (TdP)**
Participant(s):	**UPC, URJC, IPA, TdP**
Workpackage:	**WP4**
Est. person months:	**5,4**
Security:	**Int**
Dissemination Level:	**CO**
Version:	**f**
Total number of pages:	**39**

Abstract:

This is an internal intermediate document towards D43 defining the requirements and procedures for the control of flows, including the sensing of backhaul relevant quality parameters (needs to be interfaced with 5A2), offloading architectures and key performance indicators. These activities will feed WP6. Basically two topics will be addressed: traffic offloading and user scheduling. The investigation of traffic offloading techniques will allow reducing the congestion of the backhaul, a crucial aspect in TUCAN3G. On the other hand, with user scheduling and admission control we tackle the optimization of the access network in the short-term. Nevertheless, this latter aspect should also consider the required backhaul quality and the general design should take into account the proper interface between the access network and the transport network should be considered.

Keyword list: traffic offloading, VSAT cache, scheduling, admission control, backhaul monitoring, energy consumption

Document Revision History

DATE	ISSUE	AUTHOR	SUMMARY OF MAIN CHANGES
24th Oct 2013	a	UPC	Proposal of a preliminary ToC
16th Dec 2013	b	UPC, URJC	Integration of preliminary contributions
28th Dec 2013	c	UPC, IPA	Update of the contributions
15th Jan 2014	d	UPC, URJC, IPA	Revision of contributions
31st Jan	e	UPC, TdP	Incorporation of contributions from TdP, and definition of a workplan for D43
17th Mar 2014	f	URJC, UPC	Review and final version

Deliverable D43 scheduled work plan

DATE	DESCRIPTION
3rd- 7th March 2014	Phone Call for discussing the current state of the work for D43 and refine the ToC using M43
18th March 2014	1st call for contributions to D43
25th March 2014	Partners in charge incorporate new contributions and send the updated document to UPC
31st March 2014	UPC generates the first version of the D43
14rd- 19th April 2014	Phone Call for discussing the current state of the work for D43
28th April 2014	2nd call for contributions to D43
5th May 2014	Partners in charge incorporate new contributions and send the updated document to UPC
12th May 2014	UPC generates the second version of the D43
1st June 2014	2rd call for contributions to D43
13th June 2014	Partners in charge incorporate new contributions and send the updated document to UPC
20th June 2014	UPC generates a final version of the document and it is distributed to the technical coordinator
30st June 2014	Document is delivered to the EC

Executive Summary

This is an intermediate document towards the elaboration of deliverable D43 in which the main objective is to describe the techniques defined and analysed in the activity 4A3, tackling optimization of the access network in the *short-term* and the reduction of the required backhaul traffic. In this regard, the following topics are going to be investigated:

- Traffic offloading
- Scheduling and admission control
- Interactions between the access network and the backhaul

In each case the scenarios, the techniques to develop and the working plan are sketched. The final document (D43) will describe how 4A2 and 4A3 interface and how the conclusions will influence the activities in WP6.

DISCLAIMER

Table of Contents

References

[Ameigeiras03] P. Ameigeiras, *Packet scheduling and Quality of Service in HSDPA*, PhD Thesis, Dept. of Communication Technology, Aalborg University, Denmark, Oct. 2003

[Elliot02] R. C. Elliot, and W.A. Krzymien, “Scheduling Algorithms for the cdma2000 packet data evolution’, in *Proc.of the IEEE Vehicular Technology Conference* (VTC), Vancouver, Canada, vol. 1, pp. 304–310, September 2002

[Fernandez13] J. Fernandez-Bes, A. G. Marques, and J. Cid-Sueiro, “Battery-aware selective communications in energy-harvesting sensor networks: Optimal solution and stochastic dual approximation,” in Proc. of 10th Intl. Symposium on Wireless Commun. Systems, Aug. 2013

[Gatsis10] N. Gatsis, A. G. Marques, and G. B. Giannakis, “Power Control for Cooperative Dynamic Spectrum Access Networks with Diverse QoS Constraints”, IEEE Trans. Commun., vol. , pp., 2010

[Goldsmith05] A. Goldsmith, *Wireless Communications*, Cambridge Univ. Press, 2005.

[Holma06] H. Holma and A. Toskala, *WCDMA for UMTS – HSPA EVOLUTION* AND LTE, 4th Edition, John Wiley & Sons, Ltd.

[Huang06] J. Huang, R. A. Berry, and M. L. Honig, “Distributed interference compensation for wireless networks,” IEEE J. Select. Areas Commun., vol. 24, no. 5, pp. 1074–1084, May 2006

[IETF-RFC2205] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, ReSerVation Protocol (RSVP) Version 1 Functional Specification, RFC 2205, September 1997.

[IETF-RFC2474] K. Nichols, S. Blake, F. Baker and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers”, RFC 2474, December 1998.

[IETF-RFC3031] E. Rosen, A. Viswanathan, R. Callon, “Multiprotocol Label Switching Architecture”, RFC 3031, January 2001.

[IETF-RFC2638] K. Nichols, V. Jacobson and L. Zhang, “A Two-bit Differentiated Services Architecture for the Internet”, RFC 2638, July 1999.

[Khan11] Khan, M. F., Khan, M. I., & Raahemifar, K.; "Local IP Access (LIPA) enabled 3G and 4G femtocell architectures". In Proceedings of the 24th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1049-1053, May 2011

[Lee10] Lee, K., Lee, J., Yi, Y., Rhee, I., & Chong, S.; "Mobile data offloading: how much can WiFi deliver?". In Proceedings of the 6th International COnference (p. 26). ACM, November 2010

[Lin10] Lin, Y. B., Gan, C. H., & Liang, C. F.; "Reducing call routing cost for femtocells". IEEE Transactions on Wireless Communications, pp. 2302-2309, vol. 9, no. 7, July 2010

[Palomar06] D. Palomar and M. Chiang, “A tutorial on decomposition methods for network utility maximization,” IEEE J. Select. Areas Comm., vol. 24, no. 8, pp. 1439–1451, Aug. 2006

[Stanczak06] S. Stanczak, M. Wiczanowski, and H. Boche, Resource Allocation in Wireless Networks: Theory and Algorithms. Berlin, Germany: Springer, 2006

[SCF13] Small Cell Forum Release Two documents, available online at http://www.scf.io/en/index.php?utm_campaign=Release%2520Two

[TUCAN3G-D21] ICT-601102 STP TUCAN3G, "Socio-economic scenarios, technical specifications and architecture for the proof of concept", deliverable D21, May 2013, Available online: www.ict-tucan3g.eu.

[TUCAN3G-D22] ICT-601102 STP TUCAN3G, "Operational Technical Handbook", deliverable D22, May 2013, Available online: www.ict-tucan3g.eu.

[TUCAN3G-D41] ICT-601102 STP TUCAN3G, "UMTS/HSPA network dimensioning", deliverable D41, November 2013, Available online: www.ict-tucan3g.eu.

[TUCAN3G-M52] ICT-601102 STP TUCAN3G, "Transport network architecture and interface to the access network", milestone M52, January 2013, to be published on-line.

[TS 23.107] 3GPP Technical Specification TS 23.107 Quality of Service (QoS) concept and architecture

[TS 23.207] 3GPP Technical Specification TS 23.207 End-to-end Quality of Service (QoS) concept and architecture

[TS 23.829] 3GPP Technical Specification TS 23.829 Local IP Access and Selected IP Traffic Offload (LIPA-SIPTO) - Release 10

[TS 25.444] 3GPP Technical Specification TS 25.444 Iuh Data Transport

[TS 25.467] 3GPP Technical Specification TS 25.467 UTRAN Architecture for 3G Home Node B (HNB)

[Wang07] X. Wang, G.B. Giannakis, and A.G. Marques, "A Unified Approach to QoS-Guaranteed scheduling for channel-adaptive wireless networks", in *Proceedings of the IEEE*, vol. 95, no. 12, Dec. 2007

[Zdarsky11] Zdarsky, F. A., Maeder, A., Al-Sabea, S., & Schmid, S.; "Localization of data and control plane traffic in enterprise femtocell networks". In Proceedings of the 73rd IEEE Conference on Vehicular Technology (VTC Spring), pp. 1-5, May 2011

List of abbreviations & symbols

ABI	AN-BH interface
AMC	Adaptive Modulation and Coding
AN	Access Network
AP	Access Point
AWGN	Additive White Gaussian Noise
BH	Backhaul
BLER	Block Error Rate
CQI	Channel Quality Indicator
CS	Circuit Switched
CDMA	Code Division Multiple Access
CPICH	Common Pilot Channel
DiffServ	Differentiated Services
DSCP	Differentiated Services Code Point
GGSN	Gateway GPRS Service Node
HMS	HNB Management System
HNB	Home NodeB
HNB-GW	Home NodeB Gateway
HSDPA	High-Speed Downlink Packet Access
IntServ	Integrated Services
LIPA	Local IP Access
MPLS	Multiprotocol Label Switching
PDP	Packet Data Protocol
QoS	Quality of Service
RAB	Radio Access Bearer
RSVP	Resource Reservation Protocol
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTT	Round Trip Time
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SIPTO	Selected IP Traffic Offload
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
UE	User Equipment (the mobile terminal and SIM Card)
SF	Spreading Factor
SNIR	Signal-to-Noise plus-Interference Ratio
TDMA	Time Division Multiple Access

1 INTRODUCTION

The present document describes the work to be carried out in activity 4A3. The main objective is the optimization of the access network (AN) in the *short-term* (scheduling period), in contrast with the work developed in activity 4A1 of WP4 [TUCAN3G-D41], where the network planning is exposed, and also in contract to the work carried out in activity 4A2 of WP4 [TUCAN3G M42], where the access network is optimized and monitored in the *long-term* (multiple scheduling periods).

The *short-term* access network optimization deals with topics like channel state-aware packet scheduling, congestion and admission control. Nevertheless, all these topics also depend on the quality of the backhaul, demanding certain kind of coordination between transport network and the access network.

Designing the backhaul is an aspect of paramount importance in the project. The envisioned network architecture depicted in Figure 1 has been introduced in [TUCAN3G-D21]. It should be noted that linear-like architecture of the transport network has higher probability to produce congestion in the links close to the backhaul GW especially when the traffic demand increases. In this respect, there are several investigations on different techniques aiming at the reduction of the backhaul load by means of traffic offloading and local cache for satellite links.

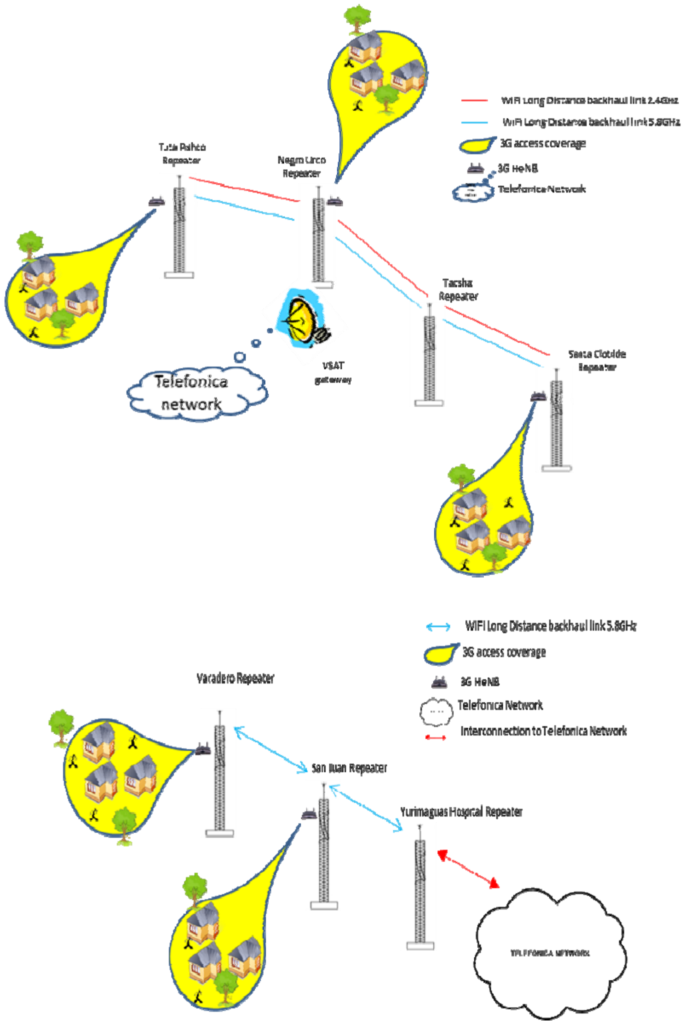

Figure 1. Network architectures considered in: left) Napo river and right) Balsapuerto

1.1 Objectives

- Evaluate different possibilities for transport network offloading.

- Propose procedures allowing effective control of multiple flows when the access network is connected to a variable quality transport network.
- Definition of the interface between the access network (AN) and the backhaul (coordination with work-package WP5)

1.2 Organization

This document is organised as follows. Section 2 investigates different traffic offloading strategies with the objective of reducing the use of the backhaul. The investigated techniques will consider 3GPP and non-3GPP based solutions and will take into account the legal requirements of each solution. Section 3 addresses the scheduling and admission control and it is divided into two main sub-sections. In sub-section 3.1 the channel state-aware packet scheduling is considering when there is just one serving Home Node B (HNB) or when there are multiple HNBs. Energy consumption will be an aspect to be considered for the decisions. Admission control procedures are also investigated in this section. On the other hand, if this interface is investigated in sub-section 3.2 it appears to be part of the scheduling and admission control, rather than offloading. Finally, section 4 provides a summary of the techniques investigated and could be considered for implementation.

2 TRAFFIC OFFLOADING

Femtocells or Home Node B were originally conceived to improve 3G/4G indoor coverage for home subscribers. In that scenario, HNBs relay the traffic towards the core network using a private (typically xDSL) backhaul network.

In TUCAN3G, femtocells are used to provide outdoor coverage. On the other hand, while classical macrocells or NBs (Node B) are connected to the core network using an optical fiber, HNBs in TUCAN3G are connected to the core using a WiFi multihop network. In this configuration, the traffic load is expected to be higher than in the indoor configuration due to higher number of users, while the capacity is severely constrained than the original scenario due to use of limited WiFi. This problem is even more acute in deployments where the connection with the core network also includes a satellite segment. A solution to reduce the load of the backhaul network and avoid outages associated to overloads is to implement offloading techniques allowing smart routing of local traffic [Lee10].

Section 2.1 will consider different alternatives to implement offloading mechanisms and evaluate the benefits for the network performance. Section 2.2 will analyze how the offloading mechanisms can be implemented in actual 3GPP networks as well as potential recommendations of the standards.

2.1 Network architectures and benefits of traffic offloading

In this activity several alternatives to implement offloading techniques for the considered scenarios will be analyzed, which will take into account the current standardizations efforts (LIPA/SIPTO solutions, enterprise architectures…) and will consider both voice and data offloading. The activity will conclude with some recommendations regarding the configuration of offloading schemes.

The alternatives analyzed in this document are (see Figure 2):

a) installing a gateway/proxy in the backhaul,
b) local switching of users served by the same HNB, and
c) shortest path.

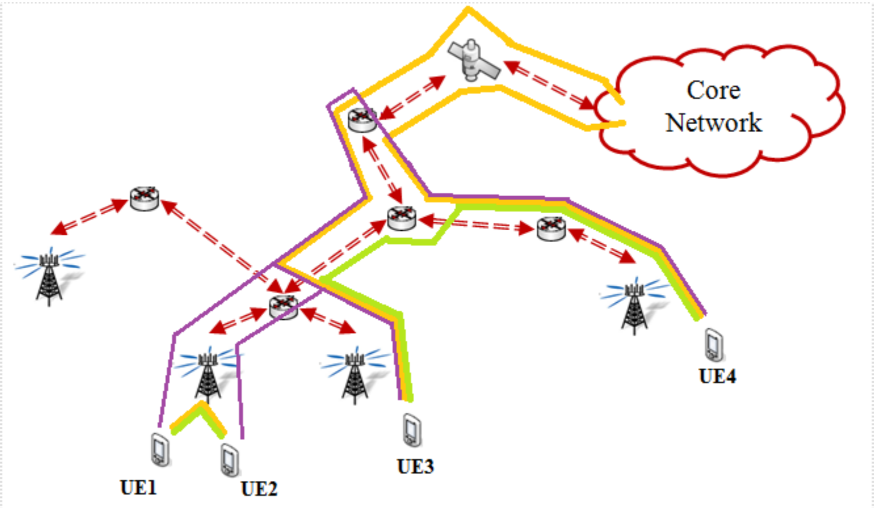

Figure 2. Different alternatives for local offloading. UE 1 is connected to UE 2 and UE 3 is connected to UE 4. The purple route corresponds to alternative a (installing a gateway/proxy in the backhaul); the yellow route corresponds to alternative b (local switching of users served by the same HNB); and the green route corresponds to alternative c (shortest path). Note that the yellow line connecting UE3 and UE4 though the core network is the conventional route we target to shorten for a better usage of the backhaul.

Alternative a) consists in installing a gateway in the node that connects the backhaul to the core network [Lin10][Zdarsky11]. This avoids sending local traffic to the core network, which is a good solution for deployments where the connection between the backhaul and the core is limited or expensive (for example, when it is implemented using a satellite link).

Alternative b) consists in connecting the users served directly by the same HNB [Khan11]. This option is appropriate to access to local data and to voice calls between close-by users.

Alternative c) consists in routing all local traffic through the backhaul network using the route with the shortest number of hops. Although it is difficult to implement in practice, this solution will serve as a benchmark to assess the benefits of implementing local switching.

The performance will be evaluated using simulations. The focus will be placed first on evaluating the gains in terms of rate increase and congestion reduction. If the results are satisfactory, then the impact on energy consumption will be evaluated too. Local offloading on both voice and data traffic will be considered (first each of them separately and then jointly).

The key aspects of the simulations will be: 1) network topology; 2) link configuration (terminals and effective throughput); and 3) traffic models. If the simulations suggest that traffic offloading mechanisms should be implemented in practice, then more features such as channel variability (due to fading, weather conditions, …) and solar refill patterns will be taken into consideration too. Each of the key aspects is described in more detail next.

1) The topology of the backhaul network will have a critical impact on the performance gains associated with local offloading. Clearly, long networks with many hops will benefit more from smart routing techniques. The two scenarios considered in TUCAN3G (Balsapuerto – San Juan & San Gabriel- and Napo –Santa Clotilde, Negro Urco & Tutapisco-) will be analyzed. If convenient, topologies not explicitly considered in TUCAN3G but that may be relevant for rural deployments will be analyzed too.

2) Some requirements of the different technologies in the access and backhaul network will be taken into account in the simulations to measure the key performance indicators. Among those are: a) the actual technology implemented (WiLD, WiMAX satellite link,..); b) HNB can serve a relative small number of users; or c) IP encapsulation of voice traffic in the uplink is different from that in the downlink; just to name a few.

3) One of the most important aspects to evaluate the benefits is to estimate the amount of local traffic. Although some models exist for voice traffic, there is little information about data traffic. To mitigate that problem, simulations will consider a broad range of scenarios ranging from low to very high local traffic. Regarding the total traffic consumed by the users, the simulations will consider the patterns of traffic along the day provided in [TUCAN3G-D41] (population, penetration, itinerancy, traffic per user).

2.2 Implementations in 3GPP networks

The aim of this subtask is to analyse how the schemes studied in 2.1 can be implemented in actual 3GPP networks and, in particular, in the network deployed in TUCAN3G. Two types of solutions will be analysed: solutions that comply with the 3GPP standard (releases 10 and later) and solutions that do not (pre-standard or non-standard). Corresponding recommendations, based on the findings in this subtask, will be provided to WP6.

2.2.1 Offloading implementations complying with the 3GPP standard

3GPP has standardised Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO) mechanisms since 3GPP Release 10, gradually adding more functionality. Recently, ip.access has carried out work within the Small Cell Forum to look at enterprise data and offload architectures. Some of the results are available in the references [SCF13]. These include the introduction of intermediate gateways for data and voice offloading, and mobility with reduced signalling towards the core network.

LIPA was designed as a service to allow a radio bearer access through a specific local IP address that is not publicly accessible – e.g. it is not on the public internet. The original use cases include access to enterprise intranets, home media storage, as well as printers [TS 23.829]. SIPTO is a similar offloading optimisation allowing a radio bearer towards a public address to be offloaded without traversing the operator core network, unlike LIPA.

In the case of TUCAN3G deployment scenarios, the backhaul transportion to the Core Network is expected to have a high latency in addition to existing limited bandwidth problem which probably causes multiple congestions due to the deployment conditions. Consequently, it is highly beneficial to use a technology to reduce the load (either signalling or user data) on the backhaul. In the future, with an expected growth of the population of the communities will entail a significant increase of the local traffic, including voice (VoIP). The reference architectures will take advantage of this. Such techniques may also improve the user experience.

The deliverable D43 will describe the examined architectures and, for each one, will look at the potential benefits of the remote deployment scenarios for both public and local data and will seek to make use of analysis carried out in section 2.1. In addition, the qualitative gains to the user experience will also be described.

2.2.2 Non-standard offloading implementations: Data traffic caching over satellite

In addition to the ones in 3GPP standards, there are pre-standard or non-standard implementations, including some on the border of content delivery, that provide alternative offload mechanisms. When analysed in detail, it can be seen that some of the mechanisms are simpler to implement and adapt to the specific scenarios considered in TUCAN3G. Although such techniques may provide considerable benefits to the user experience, some of the proposed implementations will also have disadvantages in terms of reliability, flexibility, modifications to existing protocols, security, and exposure of the internal structure of the core network (e.g. IP addresses usage).

Deliverable D43 will further analyse a number of these non-standardised implementations identified in the literature and describe their benefits and disadvantages. Special emphasis will be placed on local cache techniques for satellite links as they are one of the important and problematic part of TUCAN3G due to their high cost and at the same time, the only solution to backhaul in certain isolated areas.

Internet content keeps expanding rapidly with richer contents and web sites which include more and more visual materials – static and motion based (e.g. Adobe Flash technology, java scripts). Nowadays, typical web pages are larger than ever (e.g. www.youtube.com). They can easily host more than 250 and be accessible through over 240 HTTP requests, while their contents are also very dynamic requiring preloading. In addition, Web sites become very dynamic with content preloading every time the user moves their cursor over an object or link. On the other hand, home and business networks have far more networking devices than ever, with laptops, tablets, iPhones and many other devices connecting via 3G networks or WiFi. In Peru, considering rural information from Fitel´s – TdP BAS (“Wide Band for Isolated Localities”) Project, the most visited webpages are Google, Youtube, Windows Update and Facebook.

This explosion Internet of contents makes the access through VSAT really challenging. Throughout the past decade, VSAT equipment manufacturers were able to improve web performance using various web acceleration technologies:

- **Content cache:** Onboard cache memory for recently or frequently accessed documents, objects or scripts and prefetch documents that are likely to be accessed, DNS caching.

- **HTTP Compression:** Use of encoding methods.
- **Code Optimization:** Optimizing java or http code to send the less information possible, this feature includes “White Space Removal” and “compressing images”

Among the three, only the first requires additional network elements and is relevant to the proposed TUCAN3G architecture.

Although traditional hierarchical web cache control technologies provide tremendous improvement in web user experience, satellite-cache distribution technologies could extremely improve the results. This progress can be obtained by utilizing a large on-board multi Gigabyte memory in the central cache. This way, a large amount of web content could be stored and by taking advantage of the inherent one-to-many satellite network topology, to distribute and populate content on the cache memory of all VSATs in a particular network.

Each object in a web page has a date of "expiration" which is used by cache devices to understand if the object can be saved and how long. The use of cache devices reduces the number of requests by providing memory local cache, previously captured data and by improving the response times (in average) and the use of bandwidth

2.2.2.1 Content caching

Caching is not new to Internet technology. Caching technology like other acceleration methods is designed to reduce bandwidth and enhance the user experience. The main idea of content caching is to store popular content which is frequently accessed by the users. Whenever a user desires to access the required content, the content is fetched immediately from the local cache storage instead of fetching it from the remote web server.

Today cache servers are deployed throughout the Internet at either the core (carriers), the edge Internet Service Providers (ISPs) or the end user locations. Caching is also available on web browsers. In low roundtrip delay networks, the combined use of caching by the ISP and caching at the browser will provide enough performance improvement to end users. However, when ISPs and end users are separated by ~600msec of round trip delay and an expensive transmission path (space segment), the performance improvement is diminished.

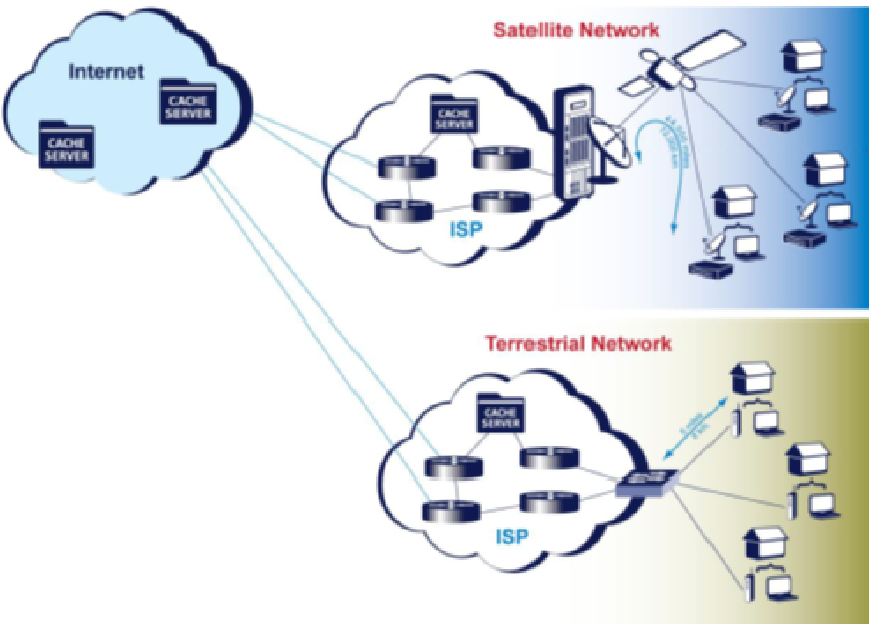

Figure 3. Web Cache-Architecture

As shown in Figure 3 cache servers at the hub location are typically installed by satellite ISPs in order to reduce the terrestrial backhaul usage and to bring content closer to the user. Nevertheless, the traffic still needs to travel more than 72,000 km between the hub and the end user. The objective of satellite cache technology is to bring the content within a shorter distance (hence with lower delays) to the end users based on Distributed Cache architecture located at every site in the network. In this architecture each cache stores the content that has been accessed by local users and then common content is also shared with all other cache elements in the network. As a result each cache device contains all of the content that has been accessed by any user in the network. This architecture is mostly suitable for VSAT networks because of the inherent one-to-many broadcast characteristics of satellite. Content needs to be transmitted only once over the satellite to be received and stored by each of the satellite network nodes.

2.2.2.2 Content caching tests in Telefónica del Perú

In February 2013, Telefonica Peru and GILAT satellite networks carried out a 5-days tests of internet traffic download. Optimization was done using web caching features on VSAT modems. During the tests, Web Enhance VSAT (manufactured by GILAT) is used for distributed caching, and Skymon software is used to measure the traffic offloading optimization. The main focus of the analysis was the satellite bandwidth optimization, as well as the improvement of the user experience with faster downloads. The tests were performed in two different scenarios:

- Controlled environment with one VSAT
- Multiple VSATs

One VSAT working in a controlled environment

The purpose of this experiment is to test VSAT cache capability directly and under controlled conditions. The simplified architecture of the test environment is given in Figure 4, where there is one VSAT placed in the lab and two computers connected to the VSAT.

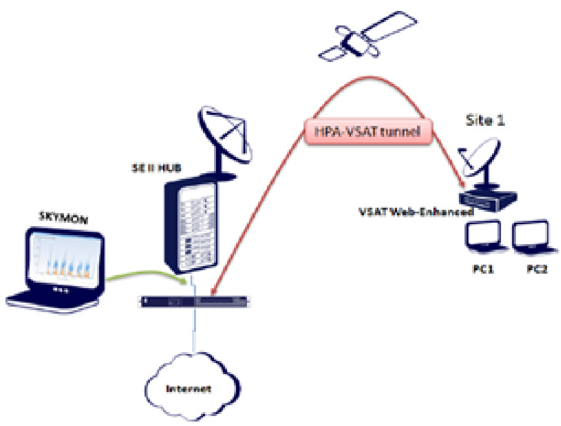

Figure 4. Simplified architecture for the experiments with one VSAT

For the test, a brief list of webpages is selected for the tests, then they opened in PC1. Since the VSAT cache was empty at the beginning, all the page details had to be downloaded entirely through the outbound. These pages were opened from the PC1. Since the VSAT cache was empty, all the pages were entirely transmitted through the Outbound. The next step consisted to access the same list of pages, but this time from the PC2. In this case, many of the objects of those pages were downloaded directly from the VSAT cache.

Figure 5 depicts the Outbound kpbs graph associated to the VSAT that was monitored from the Skymon. This figure includes both downloads, from the PC1 (all pages and items are served entirely in the Outbound), and from the PC2. In the left side of Figure 5 we observe the outbound traffic during the time all pages are downloaded from the PC1, while the right side of Figure 5 shows the outbound traffic during the download from the PC2 (using the VSAT cache). As expected, the outbound traffic for the PC2 is relatively lower and VSAT caching succeeded to reduce in more than 20% the outbound usage.

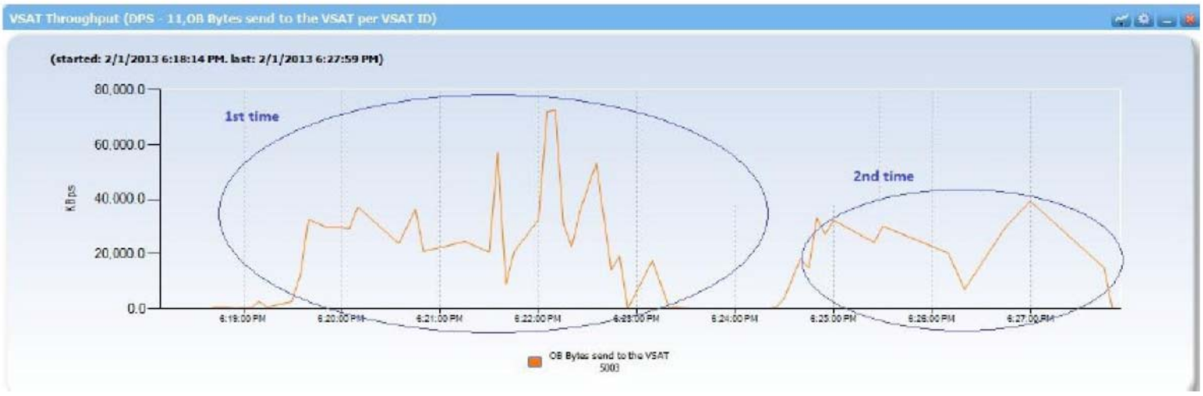

Figure 5. Outbound graph associated to the VSAT

On the other hand, we evaluated the download time for both situations as given in Table 1. The first column shows the selection of the web pages used during the tests, while the second and the third column show download durations for the attempts from PC1 and PC2 respectively. Except for a couple of particular cases, the second attempt has a lower download time.

URL List	Intento 1	Intento 2
	05:30 p.m.	06:30 p.m.
http://www.hotmail.com	18.281	11.843
http://www.hi5.com	17.398	15.125
http://www.rpp.com.pe	31.109	26.226
http://www.elcomercioperu.com.pe	25.046	13.781
http://www.youtube.com	6.257	6.765
http://www.wikipedia.org	5.179	2.015
http://www.facebook.com	17.953	9.304
http://www.sunat.gob.pe	15.609	25.757
http://www.peru21.pe	37.968	18.039
http://www.claro.com.pe	40	40
http://www.viabcp.com	0.023	0.007

Table 1. Website download time comparison (in seconds)

Multiple VSATs

In the configuration shown in Figure 6, 9 operational VSAT are installed in Internet cafes. The objective of this second phase consists of determining the outbound savings for a long term traffic by adding the cache capability of representative sites while they work all day long. This will give a good indication of quantitative values associated to total bandwidth saving on the outbound if it is assumed to have similar traffic pattern and similar number of computers in most of the VSAT locations.

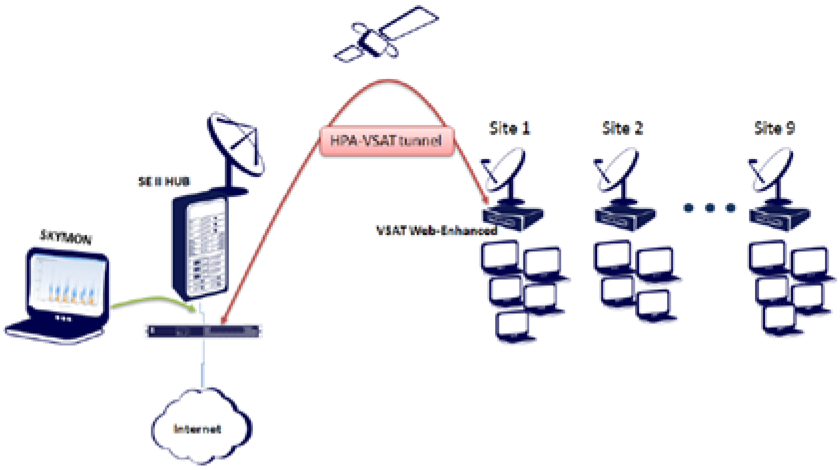

Figure 6. Architecture for the experiments with multiple VSAT

For the representation of the results, unlike the phase 1 tests, outbound Kbps graphs are not going to be used to be able to determine the caching functionality in a direct and precise way. Instead, specific graphs are going to be used to represent the Kbps if no caching was used and the number of bps transmitted over the satellite when finally caching capability is applied. Outbound Kbps graphs are not going to be used for this phase in the VSAT, as it was used in phase.

Not only HTTP traffic will be used. With the kind of representation shown below, it is not possible to determine the HTTP traffic percentage that was saved because of the caching functionality in a direct and precise way. In contrast, specific graphics directly associated to the Web Enhance capabilities are used. These graphs represent the number of Kbps that would have been transmitted by the outbound (if there is no cache in the VSAT) versus the Kbps that are finally transmitted over satellite (which correspond to all the objects that have not been provided since the VSAT to client computers).

Figure 7 and Figure 8 depicts the outbound HTTP traffic (in Kbps) related only to the VSATs cache capability. In addition, any other traffic from VSATs with no caching capability is discarded. The blue line corresponds to the results when no caching in the VSAT was considered, while the orange line denotes the results when caching is available.

From Figure 7 and Figure 8 we can remark the following observations:

- Cache-based solution saves 15-20% of required throughput
- In the busy hours, those savings are up to 40%, which corresponds to 1-2 Mbps for the 9 sites

2.2.2.3 Conclusion

Based on the results of these tests, TdP considers that the approach of data traffic optimization for Internet services over satellite provides significant gains, which are accumulative to those approaches based on centralized caching or based on data compression, in terms of backhaul capacity savings and faster downloads. TdP recommends the use of this approach for saving capacity in the satellite backhaul and improving the time of downloads in TUCAN3G.

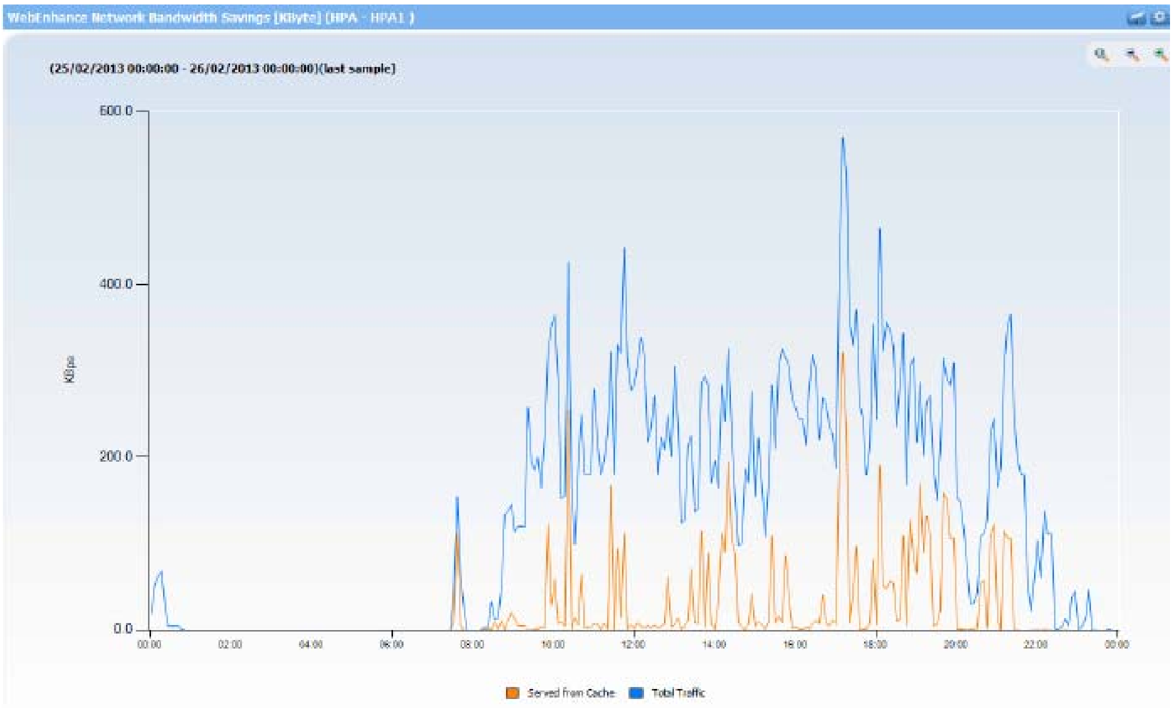

Figure 7. HTTP traffic in KBps during averaged over a 4-day test

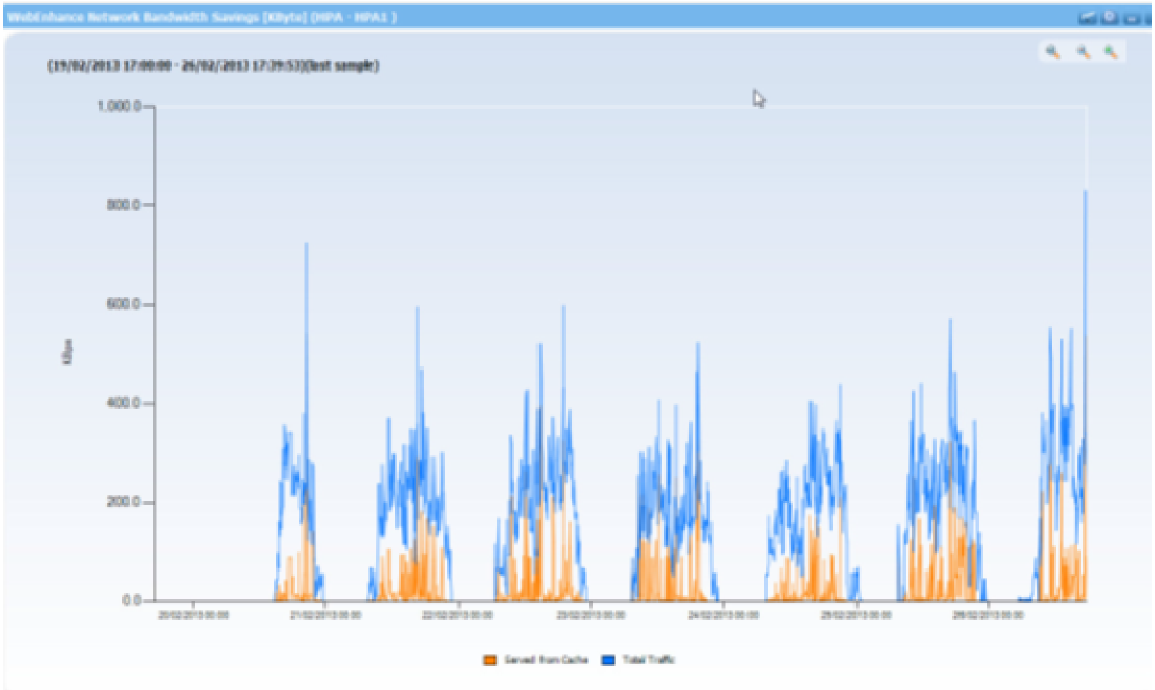

Figure 8. HTTP traffic in KBps during 7-day test

2.2.3 Regulatory issues

One important issue appears on local voice offloading; as standardized methods, neither LIPA nor SIPTO offers voice offloading through operator's core network by their definition. Therefore, in case an operator has a regulatory requirement which requires to intercept the user data (also known as “Lawful Intercept” [TS 23.829]), alternative mechanisms to gather this data are needed if the implementation is to be viable. Similar issues also occur in the related area of non-standard solutions (content caching and local switching in general). An understanding of these issues is important for any operator considering a remote deployment.

In D43, a description of types of Lawful Intercept requirement will be given and the range of mechanisms (or none) to facilitate these requirements for both standardised and non-standardised architectures in rural scenarios. Suggestions will be made on regulatory and operator approaches that may allow deployment of these architectures.

2.3 Expected outputs

The main outputs (documents) to be generated at the end of the activity will be the following:

- Benefits of traffic offloading. The different offloading solutions presented briefly in section 2.1 will be described in detail. The benefits in terms of power consumption and traffic congestion will be analysed via simulations. Different scenarios will be analysed (regarding network topology, local traffic patterns). The results will be analysed and recommendations for the other tasks/substasks will be provided.
- Architectures and solutions for local offloading complying with the 3GGP standard. Different architectures will be described, their strengths and weaknesses will be identified, and their convenience/compatibility to TUCAN3G deployment scenarios will be discussed.
- Architectures and solutions for local offloading suitable for TUCAN3G that do not comply with the 3GPP standard. Different architectures suitable for TUCAN3G will be described, and their strengths and weaknesses will be identified. A methodology for the selection of one or another will be described.
- The Lawful Intercept problem in voice and data traffic, so that offloading schemes can be actually deployed.
- Recommendations for WP6.

2.4 Activity work plan

According to [TUCAN3G-D22], all the activities to be developed in the framework of traffic offloading will be developed during the period from 1st June 2013 to 30th June 2014, with corresponding details provided in the following table.

Task	Description	Input	Output	To	PM	Partner
4A3.1	Benefits of traffic offloading	Local traffic demand models and long haul traffic demand models (4A1)	Evaluation of backhaul congestion and recommendations of per-hop backhaul capacity needs as a function of traffic evolution	WP4	1.5	URJC
4A3.2a	Proposal of traffic offloading techniques and architecture elements	Procedures and evaluation results of 4A3.1	Recommendations for implementation and standardization	WP6, WP7	0.5	IPA

Table 2. Activities related to the elaboration of D43

3 SCHEDULING AND ADMISSION CONTROL

This section investigates the optimization of the access network in the *short-term*. First, section 3.1 addresses the optimization of radio resources by means of user scheduling and admission control taking into account the current state of the channel, the interference level and the current state of the backhaul. In this regard, section 3.2 investigates mechanisms to exchange control and state information between the access network and the backhaul.

3.1 Channel state-aware packet scheduling

Dynamic schemes that adapt the available resources (codes, power levels, user-HNB scheduling) to the instantaneous (channel) state information are a key element to mitigate fading and guarantee QoS in wireless networks [Goldsmith05]. The design of such schemes has been a very active research topic. As a result, most contemporary standards include advanced mechanisms to implement fast channel-aware resource allocation. Originally, the schemes were typically designed in a heuristic manner, under reasonable assumptions. Nowadays, the schemes are typically designed as the (approximate) solution to judiciously formulated optimization problems [Stanczak06], [Palomar06]. The selection of the formulation is indeed critical. On the one hand, it must be sophisticated enough to incorporate the operating conditions of the network as well as the QoS requirements of the related applications. On the other hand, it must give rise to simple solutions that can be easily implemented in practical systems (e.g., by entailing low computational complexity and being amenable to distributed implementations).

In this section channel state-aware packet scheduling techniques are considered firstly for a single HNB and secondly for multiple HNB. While in the first case the interference is considered as fixed, in the second case the schemes to be designed will consider the interference generated by other HNBs. In any case, the traffic and the classes of services demanded by the active users demand will be taken into account.

3.1.1 Single HNB techniques

This section addresses channel state-aware packet scheduling techniques for single HNB (UL and DL), while considering the interference generated by other cells as fixed. Furthermore, the impact of the scheduling on the other cells (i.e. interference generated) is not considered.

3.1.1.1 DL scheduling

Firstly, we will consider downlink scheduling. Given a power budget for the HNB, the goal is to optimize the number of codes and the power allocated to each user to maximize a utility function that depends on the user data throughputs, while guaranteeing the voice service to each user.

To address the problem we will consider a set of voice users, K_v, and a set of data users, K_d, already admitted in the system (admission control will be considered in a different section). In case that a certain user has a data connection and a voice connection simultaneously, it will be treated as two independent users, one for each connection (even though the HNB limitations apply to individual users, even if they bear two types of service). If the HNB limits the total number of users in the system, then the number of elements in both K_v and K_d, i.e. $|K_v|$ and $|K_d|$, will be both less than this maximum number. In case that the HNB limits the number of data connections and voice connections independently, we will establish a different limit for the number of elements in K_v and K_d.

We will consider that each data user may receive several channels within a pool of N_d channels (in HSDPA the maximum number of HS-DPSCH channels, i.e., the physical channels used for data transmission, is less than or equal to 15). We will consider that each voice user is assigned one dedicated physical channel from a pool of N_v voice channels (each one considering, for instance, a spreading factor (SF) of 128), with $N_v \geq |K_v|$. If the data channels and the dedicated channels for voice use the same carrier, the power of the carrier and the code tree is shared between both types of channels. This imposes a common constraint over the total number of available data codes, n_D. As the number of available codes with greater spreading factors depends on how many codes with shorter spreading factors are used, the total number of available channels for data, n_D, will be a function of the number of voice channels in use, i.e., $n_D(|K_V|)$. If the power of the carrier is shared among data and voice channels, there is a common constraint in the HNB power as well.

Denoting p_i the power assigned to each of the codes given to the *i*-th user:

- The sum of the total power assigned to data channels, $\sum_{i\in\{K_D\}} n_i p_i$,
- The total power assigned to voice channels, $\sum_{i\in\{K_V\}} p_i$, must be less than or equal to the total power available at the HNB, P_T.

All of the above may summarize mathematically in the following constraints:

$$\sum_{i\in\{K_D\}} n_i \leq n_D(|K_V|)$$

$$\sum_{i\in\{K_D\}} n_i p_i + \sum_{i\in\{K_V\}} p_i \leq P_T \tag{1}$$

Let us define vector $\mathbf{r}_d$ as a vector containing the users data rates, r_i for $i \in \{K_D\}$. Within this section we will focus on maximizing a utility function that increases with the user data rates, $u(\mathbf{r}_d)$, while guaranteeing the quality of the voice calls. Such a problem may be formulated as follows:

$$\underset{\{p_i, n_i, p_{TD}\}}{\text{maximize}}\ u(\mathbf{r}_d)$$

$$\text{s.t.}\quad C1:\ r_i \leq n_i \log_2\left(1 + M_D \frac{\frac{p_i}{L_i}}{(\sum_{j\in\{K_V\}} p_j + p_{TD} - p_i)\frac{\beta_d}{L_i} + \sigma_i^2}\right), for\ i \in \{K_D\}$$

$$C2:\ M_V \frac{\frac{p_i}{L_i}}{(p_{TD} + \sum_{\substack{j\in\{K_V\}\\ j\neq i}} p_j)\frac{\beta_v}{L_i} + \sigma_i^2} \geq SNIR_0, for\ i \in \{K_V\}$$

$$C3:\ \sum_{i\in\{K_D\}} n_i p_i + \sum_{i\in\{K_V\}} p_i \leq P_T$$

$$C4:\ \sum_{i\in\{K_D\}} n_i \leq n_D(|K_V|)$$

$$C5:\ \sum_{i\in\{K_D\}} n_i p_i = p_{TD} \tag{2}$$

In problem (2), M_D and M_V denote the code gain for data and voice users respectively (it is different as the SF of voice and data channels may be different), β_d and β_v denote the orthogonality factor for data and voice users respectively (ideally this orthogonality factor is zero, in practice it will be a value less than 1) , L_i is the equivalent channel path-loss between HNB and the i-th user, and σ_i^2 represents the power of the noise-plus-intercell interference experienced by the *i*-th user. The rate constraint (C1 in (2)) is a limit for the rate of each data user (according to Shannon law), and the constraint C2 imposes a minimum received signal-to-noise plus-interference ratio (SNIR) for each voice user. Finally, note that we have introduced an additional variable, p_{TD}, that stands for the total power assigned to data users, i.e. constraint C5.

In problem (2), constraint C3 must be fulfilled with equality. Notice that if we find a solution for the previous problem, where equality C3 is achieved with strict inequality, we can always multiply all powers p_i by a common factor greater than 1 until equality C3 is fulfilled with equality. By doing this, the SNIR for both data users (see C1 in problem (2)) and voice users (see C2 in problem (2)) improves (the numerator increases more than the denominator, as the noise power is not multiplied by the common factor). Therefore, the SNIR constraint for voice users is still fulfilled and the rate for the data users may be increased which in turns will increase the cost function. On the other hand, the SNIR constraint for voice users, see C2 in problem (2), must be also achieved with equality, as any other solution that fulfills the SNIR constraint with strict inequality implies a power spending higher than necessary and, consequently, higher levels of interference to data users.

As the constraint C2 and C3 in problem (2) must be achieved with equality, the problem (2) can be actually separated in two parts:

$$P1:\ C2: M_V \frac{\frac{p_i}{L_i}}{(p_{TD} + \sum_{\substack{j\in\{K_V\}\\ j\neq i}} p_j)\frac{\beta_v}{L_i} + \sigma_i^2} = SNR_0$$
$$C3: p_{TD} + \sum_{i\in\{K_V\}} p_i = P_T \tag{3}$$

$$P2: \underset{\{p_i, n_i, p_{TD}\}}{\text{maximize}}\ u(\mathbf{r}_d)$$
$$\text{s.t. } C1:\ r_i \le n_i \log_2\left(1 + M_D \frac{\frac{p_i}{L_i}}{(P_T - p_i)\frac{\beta_d}{L_i} + \sigma_i^2}\right), for\ i \in \{K_D\}$$
$$C4: \sum_{i\in\{K_D\}} n_i \le n_D(|K_V|) \tag{4}$$
$$C5: \sum_{i\in\{K_D\}} n_i p_i = p_{TD}$$

The first part, P1 stated in (3), defines a set of linear equations system with $K_V + 1$ equations and $K_V + 1$ unknowns. By solving this set of equations, we may compute the power required for each voice user, p_i^* for $i \in \{K_V\}$, and the available power for data, p_{TD}^*, to guarantee the voice service.

The second part, P2 stated in (4), is an optimization problem that depends on the powers assigned to each code and number of codes assigned to each user.

To find a solution that solves both problems simultaneously, the strategy will be as follows: we compute first the power for each voice user to guarantee the voice service, $p_i^* \ for\ i \in \{K_D\}$ and the remaining power is then left for the data users, p_{TD}^*. Then, as a second step we distribute the number of codes and power available for data among the data users according to the chosen utility function, i.e., we will solve P2 for $p_{TD} = p_{TD}^*$ Note that we can use a greater time scale for the problem P1 than for the problem P2. This approach is actually close to a practical implementation as HSDPA uses the remaining power of the carrier, i.e., not used for dedicated channels. Furthermore, the resource allocation for HDSPA channels may be faster (2 ms) than for voice users.

To solve problem P2 we need to define a certain utility function. As a preliminary study we focus on maximizing a utility function that depends on the rate of data users, $\sum_{k \in \{K_D\}} \alpha_k r_k$. The weights α_k can be adjusted to minimize average delay, maximize fairness, etc. The problem can then be written as follows:

$$P3: \underset{\{p_i, n_i\}}{\text{maximize}} \sum_{i \in \{K_D\}} \alpha_i n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{(P_T - p_i)\frac{\beta_d}{L_i} + \sigma_i^2} \right) \quad (5)$$

$$\text{s.t.} \sum_{i \in \{K_D\}} n_i \le n_D \left(|K_V| \right).$$

with $|K_V|$ the total number of voice users.

Let us consider two simple scheduling strategies: TDMA and CDMA.

a) The TDMA approach consists in selecting in every scheduling period (for instance every 2 ms) one user and giving all the available codes and available power for data transmission.

It can be proved that this approach is the optimum solution of the following simplified problem, where the cost function is a lower bound of the previous one

$$P3a: \underset{\{p_i, n_i\}}{\text{maximize}} \sum_{i \in \{K_D\}} \alpha_i n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{P_T \frac{\beta_d}{L_i} + \sigma_i^2} \right) \quad (6)$$

$$\text{s.t.} \sum_{i \in \{K_D\}} n_i \le n_D \left(|K_V| \right).$$

If the user selected to get all the resources during the scheduling period fulfills the inequality $1 + \frac{p_{TD}^* k_1}{N} < \frac{\alpha_1 k_1}{\alpha_2 k_2}$, then the users are ordered (without loss of generality) as follows:

$$\frac{1}{\alpha_1 k_1} < \frac{1}{\alpha_2 k_2} < \ldots \frac{1}{\alpha_{K_D} k_{K_D}} \text{ with } k_i = \frac{M_D}{P_T \beta_d + \sigma_i^2 L_i}.$$

b) CDMA: The CDMA approach consists in dividing all the power among the available data channels, and deciding how many codes are to be assigned to every user.

$$P3b: \underset{\{n_i\}}{\text{maximize}} \sum_{i \in \{K_D\}} \alpha_i n_i \log_2 \left(1 + M_D \frac{\frac{p_{TD}^*}{N L_i}}{(P_T - \frac{p_{TD}^*}{N}) \frac{\beta_d}{L_i} + \sigma_i^2} \right) \tag{7}$$

$$\text{s.t. } \sum_{i \in \{K_D\}} n_i \le n_D \left(|K_V| \right).$$

The solution for problem (7) is actually to give all N codes to the user with the highest value for

$$\alpha_i \log_2 \left(1 + M_D \frac{\frac{p_{TD}^*}{N L_i}}{(P_T - \frac{p_{TD}^*}{N}) \frac{\beta_d}{L_i} + \sigma_i^2} \right).$$

In practice, the CDMA approach is closer to a practical implementation. With HSDPA, variable SF and fast power control are disabled and replaced by adaptive modulation and coding (AMC) and extensive multi-code operation. In HSDPA the CQI (channel quality indicator) defines the coding and modulation combination supported by the terminal for a BLER equal to or lower than 10%, according to the E_c/I_o measured in the CPICH channel. The idea in HSDPA is to enable a scheduling such that, if desired, most of the cell capacity may be allocated to one user for a very short time, when channel conditions are favorable.

The total number of channelisation codes with SF 16 is respectively 16 (under the same scrambling code). In the code domain perspective, the SF is fixed; it is always 16, and multi-code transmission as well as code multiplexing of different users can take place. The maximum number of codes that can be allocated is 15, but depending on the terminal capability, individual terminals may receive a maximum of 5, 10 or 15 codes**.** The Transmission Time Interval (TTI) or interleaving period has been defined to be 2 ms which is shorter compared with the 10, 20, 40 or 80 ms TTI sizes supported in Release 99 [Holma06].

The assignment of all the resources to a certain user must consider the channel conditions in order to achieve a good spectral efficiency. However, buffer status, fairness, etc. should play a significant role as well. According to [Holma06], the scheduler in the Node B evaluates for different users what are the channel conditions, how much data are pending in the buffer for each user, how much time has elapsed since a particular user was last served, for which users retransmissions are pending, and so forth. The exact criteria that have to be taken into account in the scheduler is a vendor-specific implementation issue. Several publications list different HSDPA packet scheduler options, including [Elliot02],[Ameigeiras03]. In the final version of the deliverable, different scheduling approaches will be tested and compared. More specifically, we plan to extend an algorithm derived at UPC to manage the queues with the purpose of minimizing the average delay experienced by each of the multiple flows managed by the HNB, and compare with other approaches that work with a linearization of the average delay ([Wang07]).

3.1.1.2 UL scheduling

Similarly to the work done for the DL, allocation of codes and power to each active user will be considered in this section. The main difference is that now there is a power budget constraint per user, not per HNB.

Following the notation used in the previous subsection for the DL case, let K_V and K_D denote the set of voice and data users, respectively. Also, p_i represents the power allocated to each of the codes used by user i. For the case of voice users, the SNR has to fulfill the following condition:

$$M_V \frac{\frac{p_i}{L_i}}{\sum_{j\in\{K_D\}} n_j \frac{p_j}{L_j} + \sum_{\substack{j\in\{K_V\} \\ j\neq i}} \frac{p_j}{L_j} + \sigma^2} = SNR_0 \qquad \forall i \in \{K_V\} \tag{8}$$

where we emphasize that the target SNR has to be achieved with equality (any other solution that fulfills the SNR constraint with strict inequality implies a power spending higher than necessary and, consequently, higher levels of interference). The power of the noise plus inter-cell interference represented by σ^2 is equal for all the users since now the HNB plays the role of the receiver and is the same for all the users and connections. Note also that the orthogonality factor does not appear in the previous expression since in UL we assume that users are not synchronized in time. As in the DL case, the number of data codes has to fulfill the following condition (following the same notation):

$$n_i \leq n_D \tag{9}$$

Note that, differently from the DL case, in UL each user has an independent constraint related to the maximum number of data codes to be used as each user is allocated a different scrambling code.

Let us define the following variable that takes into account the noise plus inter-cell interference in addition to the received power corresponding to the data connections:

$$\sigma_{nd}^2 = \sum_{j\in\{K_D\}} n_j \frac{p_j}{L_j} + \sigma^2 \tag{10}$$

According to this, the set of equations presented in (8) can be written in matrix form as follows (each row corresponds to each of the voice users that are assumed to be numbered with the following order: $i = 1, 2, ..., |K_V|$, being $|K_V|$ the total number of active voice users):

$$\begin{bmatrix} M_V & -SNR_0 & -SNR_0 & \cdots & -SNR_0 \\ -SNR_0 & M_V & -SNR_0 & \cdots & -SNR_0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -SNR_0 & -SNR_0 & -SNR_0 & \cdots & M_V \end{bmatrix} \begin{bmatrix} \frac{p_1}{L_1} \\ \frac{p_2}{L_2} \\ \vdots \\ \frac{p_{|K_V|}}{L_{|K_V|}} \end{bmatrix} = \sigma_{nd}^2 SNR_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{11}$$

Note that all the previous equations are completely symmetric with respect to users. That means that all the users will fulfill that the power allocated to the voice connection is inversely proportional to the path loss (or, in other words, the powers received at the HNB from all the voice users will be equal):

$$\begin{aligned} p_i &= \alpha \cdot L_i \qquad \forall i \in K_V \\ \alpha &= \frac{\sigma_{nd}^2 SNR_0}{M_V - SNR_0\left(\left|K_V\right|-1\right)} \end{aligned} \tag{12}$$

As commented at the beginning of this section, in the UL scenario, the transmit power constraints are individual, i.e., on a per-user basis. That means that if $p_i = \alpha \cdot L_i > P_T^{(i)}$ for some voice user (where $P_T^{(i)}$ represents the maximum transmission power for the *i*-th user), then the SNR constraints cannot be fulfilled and some users should be dropped off from the system. Note that this implies that

$$\begin{aligned} \alpha &= \frac{\sigma_{nd}^2 SNR_0}{M_V - SNR_0\left(\left|K_V\right|-1\right)} = \left(\sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} + \sigma^2 \right) \frac{SNR_0}{M_V - SNR_0\left(\left|K_V\right|-1\right)} \le \min_{i \in \{K_V\}} \frac{P_T^{(i)}}{L_i} \\ \sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} &\le \left(\min_{i \in \{K_V\}} \frac{P_T^{(i)}}{L_i} \right) \frac{M_V - SNR_0\left(\left|K_V\right|-1\right)}{SNR_0} - \sigma^2 \end{aligned} \tag{13}$$

if we want that all the voice users are served properly. It is interesting to emphasize that the previous condition is just a constraint over the maximum total received power at the HNB corresponding to the data connections.

Using the previous partial result concerning the voice users, we can now formulate the resource allocation problem (in terms of number of codes and powers) for the data users. As in the DL case, the rate for the *i*-th data user can be upper-bounded as (the following expression has been obtained by combining properly (10) and (12)):

$$\begin{aligned} r_i &\le n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{\sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} + \sum_{j \in \{K_V\}} \frac{p_j}{L_j} - \frac{p_i}{L_i} + \sigma^2} \right), \qquad \forall i \in \{K_D\} \\ &= n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{\sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} + \frac{\sigma_{nd}^2 SNR_0 \left|K_V\right|}{M_V - SNR_0\left(\left|K_V\right|-1\right)} - \frac{p_i}{L_i} + \sigma^2} \right) \\ &= n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{c_V \sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} - \frac{p_i}{L_i} + c_V \sigma^2} \right) \\ &\text{where } c_V \text{ is defined as } c_V = 1 + \frac{SNR_0 \left|K_V\right|}{M_V - SNR_0\left(\left|K_V\right|-1\right)} \end{aligned} \tag{14}$$

Now, the final resource allocation problem for the data users can be written as follows (again, we use the same formulation as in the DL case unless stated otherwise):

$$\begin{aligned}
&\underset{\{p_i, n_i\}, i \in \{K_D\}}{\text{maximize}} \quad u(\mathbf{r}_d) \\
&\text{s.t. } r_i \le n_i \log_2 \left(1 + M_D \frac{\frac{p_i}{L_i}}{c_V \sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} - \frac{p_i}{L_i} + c_V \sigma^2} \right), \ \forall i \in \{K_D\} \\
&C1: \sum_{j \in \{K_D\}} n_j \frac{p_j}{L_j} \le \left(\min_{i \in \{K_V\}} \frac{P_T^{(i)}}{L_i} \right) \frac{|K_V|}{c_V - 1} - \sigma^2 \\
&C2: n_i \le n_D \qquad \forall i \in \{K_D\} \\
&C3: n_i p_i \le P_T^{(i)} \qquad \forall i \in \{K_D\}
\end{aligned} \tag{15}$$

The previous problem has to be solved resorting to numerical suboptimum solutions that involved a linearization of the rate constraint for data users as this constraint is not convex in general. Particular utilities functions (including the weighted sum-rate) and strategies to solve the problem for these specific utility functions will be included in the final version of the deliverable.

3.1.1.3 Admission control

In previous sections we have considered a set of K_V and K_D users already admitted in the system, with the only limitation imposed by the maximum number of users (total or voice and data users). However, there are other limitations such as the power budget or QoS requirements. If the power budget is not enough to guarantee the voice service, it means that the batteries are very close to run down. On the other hand, if the QoS for data traffic is below an acceptable quality, some of the users should not be admitted in the system. Techniques for user selection are to be preempted will be explained in the final version of the deliverable. These techniques will consider the presence of a bandwidth limited backhaul with average transmission rate constraints.

3.1.2 Multiple HNB techniques

In this subtask, the focus is on designing dynamic resource allocation schemes (both uplink and downlink) that account for the topology of the access network. This implies that users of different cells will be jointly optimized (in contrast with the previous subtask, where the optimization was carried out only for users within the same cell). The main modifications in the problem formulation are:

a) The objective now needs to consider all users and HNBs jointly. This can be achieved by using a multiobjective optimization approach or just by redefining the objective as weighted sum.

b) Users must be allowed to be served by different HNBs. This way one of the variables to design (optimize) are the scheduling variables that indicate which HNB is serving a specific user. (This problem is somehow related to the classical problem of assigning a channel/frequency to a specific user in orthogonal access networks [Goldsmith05].) It will be shown that the decision of assigning a specific HNB to serve a specific user does not depend only on the channel conditions of the user, but also on the interference generated by the other users and the overall load of the HNB.

c) Another important issue is whether the solution is going to be found in a centralized or in a distributed manner. The approach in this subtask will be first to solve the problem assuming that all information is available and then to develop (semi-)distributed solutions [Stanczak06], [Huang06], [Gatsis10]. Key for this will be the implementation of dual approach to solve the optimization problem and the definition of auxiliary variables that will facilitate the separability of the optimization problem in the dual domain (dual decomposition [Palomar06]).

To incorporate into the optimization some of the particularities of the networks in TUCAN3G, the problem will account explicitly for the battery levels and the power consumption model. The idea is that if a HNB is short of battery, users should be served by other close-by HNBs. To do so, we will follow the approach in [Fernandez13], which deals with energy-harvesting wireless sensor networks, and adapt it to the operating conditions of our networks. Initially, only the power consumed by the HNBs serving the user will be considered, but later on, the power consumption of the entire route from the femto to the core network will be considered too (this clearly corresponds to design resource allocation schemes for the access network that account for the state of the backhaul network, and will be subject of research in the corresponding subtasks).

The fact of femtos being able to serve a relatively small number of users simultaneously (see [TUCAN3G-D41]) will also be accounted for in the formulation. This implies that mechanisms for admission control have to be considered too. As a first step, users should be rejected only if the maximum number of simultaneous active connections has been reached. However, one can also envision scenarios where HNBs that are severely limited (e.g. when the energy stored in the battery is very low) reject connections that have a low priority (or that require a very high power to be served) so that future high priority requests can be served.

Simplified example:

To illustrate better the approach to be followed in the task, we present a simplified formulation that would give rise to a preliminary design.

Let p_i and γ_i denote the transmit-power and SNIR of user *i*, $u_i(\gamma_i)$ a SINR utility function for user *i* (the higher the SNIR, the more satisfied the user), $j_i(p_i)$ a power cost function for user *I* (the higher the power, the higher the cost/unsatisfaction of the user). The power consumed by the HNB *l* is P_l, the associated cost is $J_l(P_l)$. Moreover, let $h_{i,l}$ denote the power channel gain between user *i* and HNB l and $w_{i,l}$ denote a Boolean variable which is one if user *i* is assigned to HNB *l* and zero otherwise.

With *t* denoting the time instant, the optimal power and scheduling can then be found as the solution to the following problem

$$\begin{aligned}
\text{maximize} \quad & \sum_{t=1}^{T}\sum_{i,l} w_{i,l}(t)\Big[u_i\big(\gamma_{i,l}(t) - j_i\big(p_i(t)\big) - J_l\big(P_l(t)\big)\big)\Big] \\
s.t. \quad & C1: \sum_{l} w_{i,l}(t) \le 1 \\
& C2: \sum_{l} w_{i,l}(t)\gamma_{i,l}(t) \ge \gamma_i^{\min} \\
& C3: \gamma_{i,l}(t) \le \frac{p_i(t)h_{i,l}(t)}{\sigma_l + \sum_{m\neq i} p_m(t)h_{i,l}(t)} \\
& C4: \frac{1}{T}\sum_{t=1}^{T} P_l(t) \le \frac{1}{T}\sum_{t=1}^{T} Ein_l(t) \\
& C5: P_l(t) = f\left(\sum_{i} w_{i,l}(t)p_i(t)\right)
\end{aligned} \qquad (16)$$

where $Ein_l(t)$ is the energy stored in the battery at HNB *l* and $f(\cdot)$ is a function that relates how much power consumes the HNB to receive and forward the information sent by the users.

The formulated problem is not convex and its solution is non-trivial (see, e.g., for a related problem in a different context [Gatsis10]). However, it can be shown that the optimum scheduling follows a short-term greedy strategy, a strategy that assigns the user to the HNB that maximizes the short-term utility of the system. Such an utility not only accounts for the instantaneous terms in the objective, but also for the interference generated to other users, the power consumed by the HNB and the corresponding battery level.

3.1.3 Expected outputs

The main outputs to be generated at the end of the activity will be the following:

- Development of channel state-aware packet scheduling techniques for a single HNB, management of multiple flows and admission control, given the power budget and the state of the battery of the HNB.
- Development of channel state-aware packet scheduling techniques for multiple HNBs, including optimum association of the users to one HNB. The interference caused to other HNBs, the power consumption and the battery level of the HNBs will be considered for designing the suitable procedures.
- Methods for QoS-preserving scheduling and admission control with backhaul-aware constraints will be proposed for the 3G access network as a function of the quality and saturation level of both the access network and the backhaul.

3.1.4 Effort distribution

According to [TUCAN3G-D22], all the activities to be developed in the framework of channel state-aware scheduling will be developed during the period from 1st June 2013 to 30th June 2014 with the following effort

Task	Description	Input	Output	To	PM	Partner
4A3.3	Single-HNB channel state-aware packet scheduling	Local traffic demand, state of the network (short term)	Procedures to be fed to task 4A3.4 and possibly to WP6	WP7	1	UPC
4A3.4a	Multiple-HNB channel state-aware packet scheduling (emphasis on scheduling and admission solutions)	Local traffic demand, state of the network (short term)	Procedures to be fed to WP6, with the comparison of those currently used in commercial HNB	WP7	1	URJC
4A3.4b	Multiple-HNB channel state-aware packet scheduling (emphasis on on/off and coverage provision)	Local traffic demand, state of the network (short term)	Procedures to be fed to WP6, with the comparison of those currently used in commercial HNB	WP7	2	UPC

3.2 Interactions between the access network and the backhaul

The mechanisms described in the previous sections require the backhaul (BH) and the access network (AN) to exchange control and state information. The Iuh standard was intended to cope with a reliable BH with relatively limited (by macro-cellular terms) capability, such as a DSL line, where only some jitter may be present. However, 3GPP protocols assume that the operator dimensions sufficient backhaul capacity to cope with the anticipated traffic load and QoS needed, whilst recognizing that certain unspecified congestion control mechanisms may be used in the case of overload. Hence, no procedure for interaction between both networks has been standardized by the 3GPP, and the HNB is designed to be backhaul-technology agnostic. Therefore, in the TUCAN3G scenario an interface between AN and BH shall be designed and possibly implemented.

In the case of a remote deployment such as TUCAN3G, potential congestion is a serious problem for the overall user experience and affects service dimensioning, and so acquires increased importance at the design level. Some form of traffic shaping – i.e. prioritizing different classes of user traffic, signaling traffic and management traffic in order to keep the system operational as well as managing the user experience – is likely to be needed.

Regarding the work distribution in the TUCAN3G, WP4 is devoted to the AN, while WP5 focuses on the BH. Therefore, the AN-BH interface (ABI) needs a coordination of both workpackages (section 5 of [TUCAN3G-M52]). Specifically, this section is organised as follows. First, section 3.2.1 provides a brief summary of the contents of section 5 in [TUCAN3G-M52]. Section 3.2.2 reviews the 3GPP standard for dealing with non-3GPP transport networks inside a 3GPP network. Furthermore, in section 3.2.3 we discuss the QoS requirements of the different traffic classes supported by the AN and the corresponding information that must be passed to the BH. The requirements for scheduling and admission control procedures are reviewed in Section 3.2.4. Section 3.2.5 analyses different methods for obtaining the BH information; and finally section 3.2.6 describes how the BH state can be incorporated in the optimization algorithms described in section 3.1.

3.2.1 Overview

The ABI shall enable interactions in both BH to AN and AN to BH directions. Not only the AN requires the state information from the other side, but also the BH needs to know not only the QoS requirements for each connection established by the AN. Moreover, if a joint optimization approach is also considered, control mechanisms (messages from the scheduler to the BH) should also be provided by the interface. At least three types of interactions are needed in the ABI, which are represented in Figure 9. Solid thick arrows represent data or signalling exchange, while slashed thin arrows represent logical interactions. These interactions are described next.

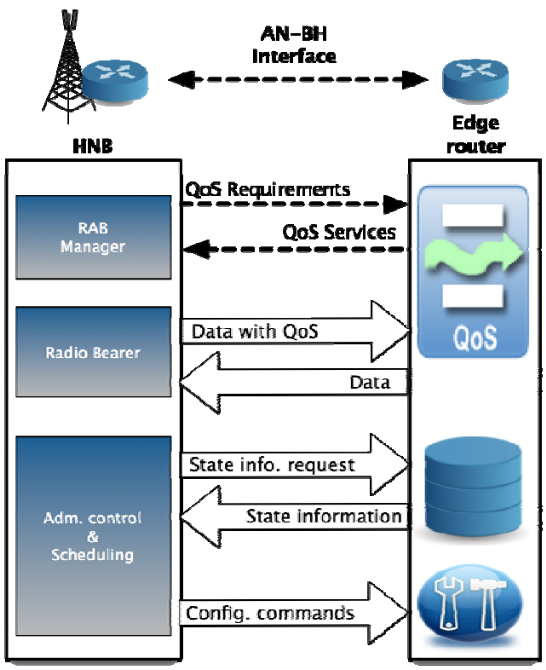

Figure 9. AN to BH interface (ABI), with main interactions.

3.2.1.1 QoS preservation in the BH

BH should guarantee QoS levels for the data exchanged between the HNB and the HNB-GW and congestion should be avoided. According to [TUCAN3G-D41], at least three QoS types must be supported by the BH, corresponding to three traffic classes: control traffic, voice traffic and data traffic. Each one requires different QoS levels in terms of rate, delay and jitter. To guarantee the appropriate QoS levels, three alternatives will be assessed in the BH:

1) Integrated services (IntServ) based on Resource Reservation Protocol (RSVP), [IETF-RFC2205],
2) Differentiated services (DiffServ) based on DSCP (Differentiated Services Code Point) marking [IETF-RFC2474],
3) Multiprotocol Label Switching (MPLS) [IETF-RFC3031]. This is further explained in [TUCAN3G-M52].

Convenience of the first and third options shall be investigated in WP5. DSCP is supported by all the BH technologies considered in this project, and then DSCP represents a suitable default option.

Regarding the congestion control, there exists two main procedures: TCP rate control and queuing. Both methods are designed to limit the throughput of TCP-based data streams (UDP data streams do not have embedded end-to-end flow control). On the one hand, a TCP sender will attempt to adapt its packet rate to the end-to-end time delay and packet loss condition. For this purpose, it will analyse TCP acknowledgement and sliding window advertisement packets from the TCP receiver in order to estimate round trip time (RTT), packet loss and the amount of data the TCP receiver can accept. The TCP rate control method manipulates these metrics by delaying acknowledgement packets and/or modifying the window size in its TCP packet headers.

On the other hand, the queuing methods delay and discard packets rather than modify TCP packets. The use of queuing is preferable over TCP rate control for the following reasons:

1) It can be applied to non-TCP traffic;
2) It is much more responsive as it does not rely on TCP flow control reaction time which is RTT dependent;
3) It is easier to implement and much less costly to execute (decreasing sliding window increases amount of packets to process);
4) It is more accurate, as the AP can discard just the right number of bytes to throttle the traffic rate to a configured threshold.

It is important that any shaping takes into account details such as:

- traffic class (management, CS voice, PS),
- number of AP users in CELL_DCH and RAB quality assigned to them, and,
- total available backhaul bandwidth.

3.2.1.2 BH state information for AN algorithms design

AN algorithms for packet scheduling and admission control shall take into account the BH state. The type of information needed is analysed later on in the deliverable. Moreover, different procedures can be proposed for collecting BH state information. Finally, typical HNB has limited processing capabilities. Hence, we will investigate the possibility of providing an external agent that implements some of the steps of the algorithms.

3.2.1.3 AN and BH joint optimization

The AN and BH joint optimization would require a method for transferring additional information (such as configuration commands) from the AN to the BH. Alternatives will be briefly discussed in [TUCAN3G-M52]. In order to implement an interface that is able to perform these interactions, a formal interface based on service primitives will be described in the deliverable. This formal interface will include commands, requests and responses, and detail exchanged information.

3.2.2 3GPP background

3GPP user plane backhaul is sent over UDP, which does not require acknowledgement by definition. As this is unacknowledged, it assumes reliable transport and relies on higher layers for traffic control. TCP/IP application traffic is greedy and fills up a channel, but at least measures itself and is self-limiting.

The QoS architecture for 3GPP mobile networks is described in [TS 23.107]. Furthermore, in [TS 23.207] the architecture is extended to provide end-to-end QoS. For this purpose several possible mechanisms are defined. Specifically, when resources not owned or controlled by the UMTS network are involved in the UMTS data transport, it is necessary to interwork with the network that controls those resources. The standard defines several approaches for this interaction:

a. Signalling along the flow path (e.g. RSVP); packet marking or labelling along the flow path (e.g. DiffServ or MPLS). According to the standard, both the GGSN and the UE should support DiffServ edge functionalities, but RSVP support is optional.
b. Interactions between both network management elements; and
c. Service level agreements.

The Annex A of the technical specification [TS 23.207] describes different scenarios of 3GPP networks that include a non-3GPP backbone network. The different scenarios correspond to different UE and GGSN capabilities. In all the cases, the non-3GPP network is in the core network and the AN is supposed to be UMTS-controlled. Then, the QoS in the AN is simply managed by the PDP Context signalling. The solutions for these scenarios comprise the combination of DSCP marking and RSVP with application layer signalling (mainly SIP and SDP).

In TUCAN3G scenario the non-3GPP network is placed between the core network and the HNB, and hence these solutions are not directly applicable. Moreover, as it is discussed in [TUCAN3G-M52], the IntServ option with RSVP is not supported by most of the equipment typically used in wireless backhauls. Finally, additional functionalities would be needed to provide backhaul-aware admission control and scheduling, as well as joint optimization. Hence, a different non-standard solution must be designed and deployed within the TUCAN3G project.

Any solution must also take into account different limitations on the UL and DL implementation: whilst an AP may be able to limit the amount of traffic it inserts in to the uplink, a HNB-GW does not have the resource to keep track of the traffic inserted into each downlink towards an AP.

In addition, the nature of the security model used is important as this impacts whether and how an AP can receive or derive information about the state of the backhaul and different approaches may be needed. Two possible scenarios are described next.

Scenario 1: Traditional Secure Deployment

The default Iuh reference model is as shown in Figure 10.

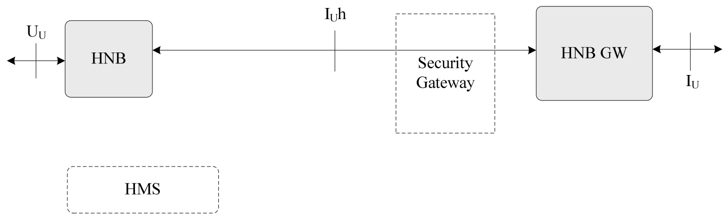

Figure 10. Default Iuh reference model (source [TS 25467] Release 9)

The default security model for standard deployments provides end-to-end security with the user plane and control plane protected by an IPSec tunnel, and the management traffic to the HMS also using the same or a different secure link. A HNB will not generally accept messages from a non-trusted source.

Consequently in this case the ability to monitor the state of the backhaul will be dependent on what the HNB can derive from the incoming traffic by itself, or indicators supplied by a trusted source – e.g. down the IPsec tunnel, or from an intermediate node via an additional secure link with a chain of trust.

Scenario 2: Non-secure deployment

In this scenario there is no secure link from the HNB to the HNB-GW deployed for user data and possible for control signaling as well. This has advantages like reducing the overall data rate demand on the backhaul, enabling easier communication of any external messages looking at the backhaul state and potentially allowing some routing optimization. However, by definition, it is not secure.

3.2.3 QoS requirements for AN traffic

The BH must preserve the QoS levels that are requested by the AN for each information flow. These QoS levels are characterized by quantitative and qualitative requirements in terms of QoS parameters (rate, delay, jitter, priority...). Some of these requirements (rate) have already been evaluated in [TUCAN3G-D41]. In D43, we will complete the profile of each QoS class from the requirements of the HNBs used in this project and the 3GPP standards. Specifically, the following parameters will be used as an input to WP5:

1) *Control traffic*: rate, maximum allowed delay (this will be a mandatory requirement for the BH), and maximum allowed jitter (if needed).
2) *Voice traffic*: rate (for one and more simultaneous voice connections, since the required bandwidth is not proportional to the number of calls), maximum allowed delay, and maximum allowed jitter (if needed). Moreover, circuit Switched (CS) resources for voice need to be protected because voice is more sensitive to packet loss, particularly bursty packet loss. In addition, and emergency call infrastructure has to be implemented.
3) *Data traffic:* rate (mean and maximum values), delay (mean and maximum values, if needed), and maximum allowed jitter (if needed).

Among the different alternatives for guaranteeing QoS in the backhaul, DiffServ is the most straightforward. In this case, the DSCP codes that are used in the HNB may or may not be standard. A precise list of their values for each QoS class will be provided to from the studies endeavoured in WP5.

3.2.4 Requirements for Access Network procedures

3.2.4.1 Information requirements

The algorithms for BH-aware admission control and packet scheduling require knowledge of the state of the BH. For this purpose, it is essential to define the (state) information required by the algorithms, and the way it should be provided.

First, admission control and scheduling procedures may need different types of information, like available rate, current delay and jitter, congestion level, status of the BH nodes' batteries, or even the network topology, which may occasionally change due to adverse atmospheric conditions, device failures or nodes that are switched off. Second, the state information may be needed with a different frequency and accuracy. Third, the information can be collected for individual links or for end-to-end path, and part of the information may be flow dependant or aggregated.

Therefore, a detailed characterization of the information needed by the AN procedures will be completed in this task. Table 3 and Table 4 show examples of the information that should be provided for the admission control and packet scheduling algorithms, respectively. Note that the content is just an example.

Parameter	Updating Frequency	Accuracy	Other considerations
Available rate	High	a) High b) Could be Yes/No	End-to-end Mandatory
Battery level	Low	Low (3 levels)	Per node Optional

Table 3. Information required for the admission control algorithm. The content of the table is just an example and must be completed in this task.

Parameter	Updating Frequency	Accuracy	Other considerations
Available rate	High	a) High	End-to-end Optional
Battery level	Low	Low (3 levels)	Per node Optional
Delay	High	a) High (milliseconds) b) Medium (queues length) c) Low (congestion yes/no)	End-to-end Mandatory

Table 4. Information required for the packet scheduling algorithm. The content of the table is just an example and must be completed in this task.

3.2.4.2 Requirements for AN procedures

Normal functioning of 3GPP signalling procedures imposes certain constraints to the ABI design. For example, if the admission control procedure needs to request instantaneous BH-state information from the backhaul, a very stringent upper bound on the response time is needed. Also, if a procedure is executed very frequently, the amount of information required from the BH could be too high, and this must be also taken into account in the interface design. Moreover, due to the HNB processing limitations, computational complexity of the proposed algorithms may have an impact on the ABI. Then, it is necessary to characterize the requirements for each procedure. Table 5 shows the requirements for both procedures. The content of the table is just an example and must be completed in this task.

Procedure	Maximum latency	Computational complexity	Periodicity
BH aware admission control algorithm	5 ms	Depends on the selected mechanism	Once for each connection: 5 times / minute in average
BH aware packet scheduling	1 ms	Medium to high. Depends on the algorithm	One for each frame of 10 ms

Table 5. Access network BH-aware procedures requirements. The content of the table is just an example and must be completed in this task.

3.2.5 Methods for obtaining BH state and comments on algorithms execution

The requirements described above affect the method used for collecting information from the BH. As an example, if the battery level of different HBNs must be known, some collaboration is needed from the BH, and an information exchanging mechanism must be implemented between them and the edge router. In addition, the method is also conditioned by the HNB limitations, which mainly are:

i) incapability to exchange non-standard information with a non-3GPP device, and

ii) very limited storing and processing capabilities.

In this section, we enumerate several alternatives to collect the information from the BH. Once the requirements of the AN algorithms are known, the best solution will be defined. The collection and aggregation of information in the BH and the way of moving it to the edge router are addressed in [TUCAN3G-M52].

Three possibilities are considered for obtaining the BH state information:

1) Non-collaborative methods. The AN collects the state information, or limits its usage, without the collaboration of the BH. For this purpose, two methods have been identified:

    a. *BH monitoring using RTCP* (Real Time Control Protocol) *headers*. In [TS-25.444] the transport mechanisms of the Iuh interface are described. If the interface is based on an IP network, RTP (Real Time Protocol) is used for the Circuit Switched (CS) domain. In this case, the use of RTCP to control the RTP connections is optional (but implemented by most vendors). RTCP headers contain time stamps that can be used to estimate end-to-end delay in the CS domain. Using this information, also the congestion level in the PS domain could be inferred.

    b. *Limiting aggregate traffic rate*. The HNB uses a centralized traffic shaping mechanism to bring packet loss and delay control ability from BH router to the HNB and take measures to protect delay sensitive traffic. The limits should be set on the total amount of transmitted or received traffic in order to match bandwidth available in the BH link. In addition, priority metrics should be assigned to incoming and outgoing traffic in the HNB according to its packet loss and delay requirements. The HNB should take into account the number of CELL_DCH users which can potentially be admitted and adjust the calculated bandwidth restrictions to ensure that voice packets are passed without interference. The number of CELL_DCH users should be adjusted in a way that it does not contradict available bandwidth configuration. The HNB reports the amount of packets carried through and the amount of packets discarded for every type of transferred and received traffic. Given the computational limitations in a HNB, using existing Linux traffic control mechanisms is advantageous. Precise mechanisms to be considered will include a Hierarchical Token Bucket mechanism, applied across queues of different traffic classes (indicated by DSCP values) and Ingress Policing. In this later case incoming traffic is selectively dropped according to a priority profile once the rate of traffic exceeds the configured limit. The Application Layer will then react to adapt to packet loss.

2) Collaborative methods. In this case, the BH must collect state information and pass it to the AN through the ABI. Three steps are needed: first, the information must be obtained in each node (and some information directly on the edge routers). Second, the information must be centralized and aggregated in an external agent, for example a bandwidth broker as defined in [IETF RFC2638]. These two steps are addressed in WP5. Third, the information must be passed to the AN. For this purpose several alternatives have been found:

    a. *Service primitives*: when an AN algorithm needs the BH state information, the HNB request it from the edge router which must have an updated version of the information or must request it to the bandwidth broker.

    b. *Always-updated*: if the latency constraints for admission control and packet scheduling algorithms cannot be fulfilled with the previous option, then the HNB must have an always-updated version of the BH state. This option can consume a lot of communication resources. For alleviating this problem, a differential approach can be used, which consists on updating the information only when the state of the network changes significantly.

3) Hybrid methods. With several methods based on signaling-along-the-path, certain amount of information could be collected. Although the collaboration of the BH is needed, no explicit interaction would be required in the design of the interface. A typical example of this is a HNB using RSVP to figure out if there exists enough bandwidth in the BH for a connection, and use this information in the admission control procedure.

The selection of the appropriate method will be based on the characterization of the AN algorithms requirements, as explained in section 3.2.4. Clearly, the selected approach will have a significant impact on the design of the AN algorithms, since they need to consider the accuracy and periodicity of the information to be collected. For example, if non-collaborative methods are used, the information would be frequently updated but will be very limited. However, with collaborative methods, if the updating rate is high, the collection of the state information would consume a lot of bandwidth in the BH.

Note that we have decided to start from the algorithms' requirements and then select the appropriate interface design. The opposite approach is also possible: start from the available options for the interface and then define the algorithms accordingly. Obviously, the first approach targets an optimal performance, while the second might be more realistic. If implementation constraints in the HNB do not allow using the first approach, the second will be used, and the algorithms would be adapted to the available information.

Finally, it is important to make some comments on the execution of AN algorithms. Due to storing and processing limitations of the HNB, an external processor to execute (partially or not) the admission control and packet scheduling algorithms may be required. This option would be appropriate if the latency constraints for these procedures are not too stringent, and the algorithms are too complex to be implemented in the HNB.

For the case of the admission control algorithm, a feasible solution is described next, as an example of the trade-offs described in this section. Let us suppose that the exchanging capabilities between the HNB and the edge router are limited. Also, assume that the information collection method is collaborative, and that the HNB is not able to perform a complex algorithm for admission control. Then, a feasible solution could be a procedure with the following steps:

1) Step 1. The HNB performs the admission control algorithm based on its own information (the easiest procedure would be to accept the connection if the number of transport connections is

less than the maximum allowed by the HNB). The output is a 0 (access denied) or a 1 (access accepted).

2) Step 2. The HNB sends a request to the primitive service for the edge router specifying the required bandwidth.
3) Step 3. The edge router gets the BH state information, runs a possibly complex admission control algorithm, and returns a 0 (access denied) or a 1 (access accepted).
4) Step 4. The edge router informs the HNB about its decision (0 or 1).
5) Step 5. The HNB multiplies both outputs and gets the final access response.

Note that the amount of information exchanged between both networks is small, and that all the complexity is assumed by the edge router. Note also that the algorithm is conditioned by the design of the interface.

3.2.6 Backhaul aware admission control and packet scheduling

The algorithms described in Section 3.1.2 will be extended to take into account the available BH-state information. The extension will depend on the available state parameters, the collection method, the updating frequency and the accuracy of the information. Regarding the available parameters, three scenarios are possible: only congestion or delay is known; congestion and available rate are known; additional parameters (jitter, battery level) are also collected. Regarding the updating frequency, long term (configuration of the BH), medium term (minutes or hours) and short term (seconds) will be considered. Finally, some parameters can be measured with high accuracy (an accurate numerical estimate of the parameter) or low accuracy (the parameters are quantized into two to a few levels). Several versions of the algorithms will be provided in order to consider different options.

3.2.7 Expected outputs

The approaches described above can be categorized along two axes: (i) the first by the security scenario, and (ii) by the approach adopted (learning from the traffic, configuration or with external inputs).

Specifically, the following outputs are expected:

1) Description of the architecture for QoS preservation in the BH.
2) Definition of the DSCP marks list.
3) Definition of the list of QoS requirements for each traffic class (for each DSCP mark).
4) Specification of the congestion control mechanism in the HNB or the edge router.
5) Definition of the information requirements for each AN algorithm.
6) Proposition of a method for monitoring the BH state in the AN. URJC will lead the area of work looking at external inputs, using its experience in WP5 and 5.2 above (collaborative methods), whilst IPA will lead the work looking at learning from the traffic alone (non-collaborative methods).
7) Specification of a formal interface, with service primitives and detailed exchanged information, which enables interaction between the AN and the BH.
8) Elaboration of a procedure for exchanging information between the AN and the BH.
9) Specification of a backhaul-aware admission control and packet scheduling algorithms.

3.2.8 Effort distribution

According to [TUCAN3G-D22], all the activities to be developed in the framework of interaction between the access network and the backhaul will be developed during the period from 1st June 2013 to 30th June 2014 with the following effort

Task	Description	Input	Output	To	PM	Partner
4A3.5a	Influence of transport network on the access network	Backhaul state acquisition (rate, delay)	Admission control decisions, others…	WP5 WP6 WP7	1.5	URJC
4A3.5b	Influence of transport network on the access network	Backhaul state acquisition (rate, delay)	Admission control decisions, others…	WP5 WP6 WP7	1.5	IPA

4 RECOMMENDATIONS TOWARDS OTHER WORK-PACKAGES

This section includes a tentative scheme of the impact of the different techniques within and outside the project. In D43, a list of detailed recommendations to WP5, WP6 and WP7 activities will be issued.

Technical contribution	Partner	Influenced WP	Enhancements to 3G standard	Applicability to IP.access equipment	Applicability to core network of Telefonica del Perú
Benefits of traffic offloading	URJC	WP6, WP7			Y
DL scheduling in a single HNB	UPC	WP6, WP7			
UL scheduling in a single HNB	UPC	WP6, WP7			
Admission control in a single HNB	UPC	WP6, WP7			
Dynamic Resource allocation and admission control in scenarios with multiples HNBs	UPC, URJC	WP6, WP7	Y	Y	Y
Interaction Between Access Network and Backhaul	URJC, IPA	WP5, WP6, WP7	Y	Y	Y
Backhaul monitoring	IPA	WP5, WP6	Y		Y
Contents caching	TdP				