Contents

Special issue on Novel Solutions for Next Generation Services

maja.matijasevic@fer.hr
imre@hit.bme.hu

As the convergence of "traditional" telecommunications and the Internet is gaining momentum, new network capabilities are being designed to support future services over Internet Protocol (IP) based network infrastructure. The papers selected for this special issue aim to address fundamental issues in creating an environment for better service support, including next generation broadband access networks, IP security and transition to IPv6, Session Initiation Protocol (SIP) signaling, VoIP performance, and new collaborative service provisioning models in mobile networks.

This special issue of the Infocommunications Journal comprises five original peer-reviewed research papers, based on those initially presented at the *10th International Conference on Telecommunications – ConTEL 2009,* held in Zagreb, Croatia, in June 2009. They were selected from a total of 67 papers accepted for publication at ConTEL 2009, and they were significantly extended before submission for potential journal publication. All submitted papers were reviewed by at least three independent reviewers for assessment. Authors had to satisfactorily respond to the reviewers' comments and to revise the papers accordingly before the final acceptance.

The first paper is "The Bridging Virtualization Approach to Next Generation Broadband Access Networks" by *J. Matias, E. Jacob, M. Aguado and J. Astorga* of the University of the Basque Country, Spain. The authors propose a new network convergence approach in next generation broadband access networks (NGBAN), called the bridging virtualization, which uses the concept of instances to deal with service requests. They introduce a secure instantiation mechanism for NGBAN and a profile-based configuration service based on XML profiles.

The second paper, entitled "VoIP Performance with IPsec in IPv4-IPv6 Transition Networks", is authored by *R. Yasinovskyy, A. L. Wijesinha and R. Karne* of Towson University, Maryland, USA. The paper presents and discusses the results of an extensive series of experiments in a LAN environment with the goal to determine the impact of IPsec on VoIP performance in IPv4 to IPv6 transition networks. The performance parameters of interest include packet inter-arrival time, jitter, packet loss, and throughput for voice packets. The overall voice quality was evaluated by using Mean Opinion Score (MOS), and the session performance metrics include IPsec key exchange time and call setup time when using SIP.

The third paper, entitled "Nature Inspired Self-healing Model for SIP-Based Services", by *Z. Rusinovic* of Ericsson Nikola Tesla Zagreb, Croatia and *N. Bogunovic* of the Faculty of Computing and Electrical Engineering, University of Zagreb, Croatia, deals with improving the performance of Session Initiation Protocol (SIP) servers. The authors present a self-healing SIP model capable of recognizing and restarting failed SIP services without losing active SIP dialogs. The proposed approach results in swift problem detection and faster recovery as demonstrated by performance tests.

In the fourth paper entitled "Space-efficient signaling scheme for IP prefix and realm information in Virtual Networks", *A. Mäkelä* from Aalto University, Espoo, Finland, and *J. Korhonen* of Nokia Siemens Networks, Helsinki, Finland, present a Mobile IP based approach to create virtual private IP networks. The proposed approach uses a scheme which compresses information on IPv4 network prefixes and realms, thus reducing the size of signalling messages and improving signaling efficiency. An extension of the proposed approach onto IPv6 network prefixes is considered.

Finally, the fifth paper entitled "Achieving Collaborative Service Provisioning for Mobile Network Users: The CollDown Example", by *V. Podobnik, I. Bojic, L. Vrdoljak and M. Kusek* of the University of Zagreb, Croatia, Faculty of Electrical Engineering and Computing, presents a concept and a model of collaborative service provisioning in mobile networks. A prototype proof-of-concept service, called Collaborative Downloading or CollDown for short, demonstrates the proposed approach.

We would like to thank the authors who submitted articles for this issue, and the reviewers for providing a constructive and timely feedback. Our thanks also go to *Csaba Szabó,* the Editor-in-Chief of Infocommunications for offering this special issue, and to the editorial staff for their help.

Maja Matijasevic
University of Zagreb,
Faculty of Electrical Engineering and Computing,
Croatia

Sándor Imre
Budapest University of Technology and Economics,
Department of Telecommunications,
Hungary

Guest Editors

MAJA MATIJASEVIC is a Professor in the Department of Telecommunications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. Her main research interests include quality of service for advanced multimedia services and networked virtual environments and the network functionality for supporting them in converged IP-based networks, with particular focus on session negotiation, adaptation, and mobility. She is a principal researcher in a national research project and a national research program, and she has been a principal investigator in several research projects in collaboration with industry. She has over 60 journal and conference publications and several book chapters. She has been involved in over 30 international conferences and workshops in the role of a Program Chair, Publicity Chair, Publication Chair, TPC member, and Organizing Committee member. She received her Ph.D. degree in Electrical Engineering from the University of Zagreb and the M.Sc. in Computer Engineering (1997) from the University of Louisiana at Lafayette, LA, USA. She is a member of IEEE and ACM.

SÁNDOR IMRE was born in Budapest in 1969. He received the M.Sc. degree in Electrical Engineering from the Budapest University of Technology (BUTE) in 1993. Next he started his Ph.D. studies at BUTE and obtained Dr.Univ. degree in 1996, Ph.D. degree in 1999 and DSc degree in 2007. Currently he is carrying his activities as a Professor and a Head of Department of Telecommunications at BUTE. He is a member of Telecommunication Systems Committee of the Hungarian Academy of Sciences. He participates in the Editorial Board of two journals: Infocommunications Journal and Hungarian Telecommunications. He was invited to join the Mobile Innovation Centre as R&D director in 2005. His research interests includes mobile and wireless systems. His main research interests and contributions are in the areas of various wireless access technologies, mobility protocols and reconfigurable systems.

List of reviewers

László Bokor, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Ádám Máté Földes, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Peter Fülöp, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Győző Gódor, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Ivan Gojmerac, FTW – Telecommunications Research Center Vienna, Austria
Gábor Gulyás, Budapest University of Technology and Economics, Hungary
Vlasta Hudek, Croatian Telecom Inc., Croatia
Gábor Jeney, Budapest University of Technology and Economics, Hungary
Zoltán Kanizsai, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Mario Kusek, University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
Ignac Lovrek, University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
Maja Matijasevic, University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
Miljenko Mikuc, University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
Szabolcs Novaczki, Budapest University of Technology and Economics, Dept. of Telecommunications, Hungary
Robert Schulcz, Budapest University of Technology and Economics, Hungary
Sándor Szabó, Budapest University of Technology and Economics, Hungary

Infocommunications Journal

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest, Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu, Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications
Tel.: +36 1 463 3261, E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: Regiszter Kft.

The bridging virtualization approach to Next Generation Broadband Access Networks

JON MATIAS, EDUARDO JACOB, MARINA AGUADO, JASONE ASTORGA

University of the Basque Country, Department of Electronics and Telecommunications, Spain
{jon.matias, eduardo.jacob, marina.aguado, jasone.astorga}@ehu.es

Keywords: broadband access networks, NGN, network virtualization, Ethernet services, Carrier-Grade Ethernet, security

Next Generation Broadband Access Networks (NGBAN), the next evolutionary step of current broadband access, have experienced a great evolution in the last few years. The NGBAN architecture is based on the reference model introduced by the Broadband Forum in TR-144, while the Next Generation Services (NGS) are based on the Ethernet Services defined by the Metro Ethernet Forum in its 6.1 technical specification. In this context, network convergence means that the same network must be capable of transporting all the existent telecommunication services (voice, video and data). This paper introduces Carrier-Grade Ethernet as transport technology to achieve convergence in provider networks. A new approach for network convergence is also presented, the bridging virtualization, which uses the concept of instances to deal with service requests. Then, a secure instantiation mechanism for NGBAN is explained, which is based on the EAPoL protocol (IEEE 802.1X). Finally, a profile-based configuration service is introduced, which defines the services through XML profiles inspired on the MEF specifications (UNI, EVC per UNI and EVC).

1. Introduction

Provider networks have experienced significant improvements due to the developments carried out around the concept of *Next Generation Networks* (NGN). Architectural evolutions in access and core networks are the main results of all this effort. There have been different approaches to NGN, but convergence, security, ubiquity and mobility are constants in all of them.

The concept of network convergence involves drastic changes of the traditional way of thinking. This concept implies that the same network must be capable of transporting all the provided services, instead of using different network architectures and technologies for each service or type of services. Even though circuit-switched services are not taken into account, the challenge is not trivial, since each service has its own particularities.

Currently, IP/MPLS and *Carrier Ethernet* [1,2] are the main alternatives for achieving network convergence. This paper has been focused on Ethernet proposals, specifically on *Carrier-Grade Ethernet* (CGE) [3,4] developments in which Ethernet is used as transport technology (opposite to Ethernet as service or interface). In fact, multiple technological components [5] are involved, such as Provider Bridges (PB), Provider Backbone Bridges (PBB), Provider Backbone Bridges – Traffic Engineering (PBB-TE) or Shortest Path Bridging (SPB).

All of them are IEEE standards (the last two in draft status), where PB and PBB address scalability and management issues, PBB-TE adds traffic engineering and SPB contributes with a link-state protocol [6] approach. However, there are also a few other standards covering different aspects that Ethernet technology (which emerged from LAN environments) must fulfill if it is to succeed as a transport technology. This is the case of Operations and Maintenance (OAM) capabilities, which are introduced by Connectivity Fault Management standard [7] (IEEE 802.1ag) and ITU-T Y.1731 [8] recommendation. Connectivity verification, rapid recovery and performance measurement are some of their improvements, essential procedures as carrier class technology.

In order to improve the coexistence of all these proposals, we have analyzed virtualization techniques to achieve a complete platform, in which real benefits can be obtained from this cooperation between different technologies, and we have developed a testbed based on Click tool [9] with promising results. The prototype consists of several layer 2 nodes (bridges), each with multiple simultaneous instances of different CGE technologies running on the same machine (or split up in several machines). This is what we call the *bridging virtualization* approach.

The bridging virtualization approach was introduced at ConTEL 2009 [10]. Based on this initial work, further improvements are presented at this paper. Our most relevant contributions are: the concept of Next Generation Services based on the Ethernet Services defined by the Metro Ethernet Forum [11] and the complete description of the *secure instantiation process*. In the later contribution we also introduce the *service port* concept, as well as the single-step and two-step AAA instantiation, and the profile-based configuration process.

The scope of this paper covers both access and aggregation networks, where Broadband Forum is the main meeting point for vendors and service providers. The aim of this organization is to assure development and deployment of broadband networks. There are se-

veral technical recommendations [12] from this forum with special relevance that must be taken into account: Multi-Service Architecture & Framework Requirements (TR-058), Migration to Ethernet-Based DSL Aggregation (TR-101) and Using GPON Access in the context of TR-101 (TR-156). The last two establish the introduction of Ethernet into provider networks, as well as practical aspects for QoS, Multicast, OAM and security.

Future Internet (FI) related initiatives are another important source of contributions to this new perspective of what broadband networks should be. In some cases, the proposals deal with Post-IP scenarios, where new network architectures are designed. All these approaches arise from the necessity of new paradigms for current Internet in order to overcome its limitations. The recently created Future Internet Assembly [13] (FIA), promoted by the European Commission, collects all those restrictions as conclusions in the Bled Manifesto [14].

Regarding the worldwide activities, different research programs about Future Internet have been promoted by both the National Science Foundation (NSF, in USA) and the European Commission. There are also several initiatives in Asia, led by Japan and Korea, which revolve around Future Internet. GENI (Global Environment for Network Innovations) and FIND (Future Internet Design) are the main programs funded by the NSF, whereas the AKARI Project is the most important one in Japan; meanwhile 4WARD, DICONET, FEDERICA and EIFFEL are some of the projects funded by the 7th European Commission Framework Program (FP7), which are focused on similar issues.

The structure of the paper is as follows. Firstly, Section 2 introduces the architecture of the system; a solid alternative to achieve network convergence by using Carrier Grade Ethernet; and the concept of Next Generation Services. Afterwards, Section 3 presents our proposal, the bridging virtualization approach, which is based on virtualization techniques and implemented with Click tool. Finally, Section 4 specifies how we have defined the secure instantiation process for service authentication and authorization, and Section 5 summarizes the paper with some conclusions.

2. Next Generation Broadband Access Networks

2.1 System architecture

Even today, it is quite common to find solutions where clients access each service (or each group of the same type of services) through different networks. This is the case of telephone service (voice), Internet access (data) or TV broadcast systems (video). Since the beginning, the three of them have been operated as diverse business models, where they must also deal with connectivity issues.

Nowadays, both telcos and cable companies offer bundled telecommunication services, which include voice, video and high speed data. Introduced as triple-play, the idea behind this concept is the provisioning of several broadband services over a single broadband connection.

This type of paradigm has three different actors, with three specific functions to cover. On the one hand, the client (C) is the final user that wants to access a service. On the other hand, the service provider (SP) is the entity that offers a telecommunication service to its customers. Between them, the connectivity provider (CP) gives an added value to services, such as security, ubiquity, mobility, multicast or QoS. The real challenge for CPs is to provide all the services through the same network efficiently and in a cost effective way.

The work done by the Broadband Forum must be considered as reference point, since it has defined a complete architecture which covers both retail and wholesale scenarios with a complete study of current and future possible alternatives. This forum has gener-

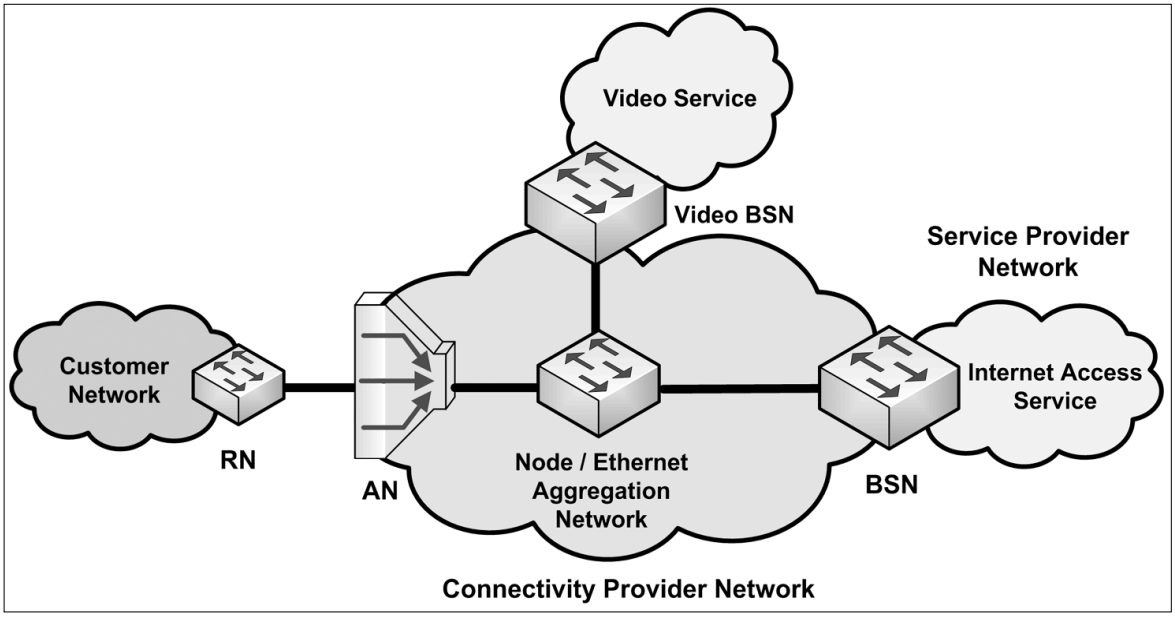


Figure 1. NGBAN architecture

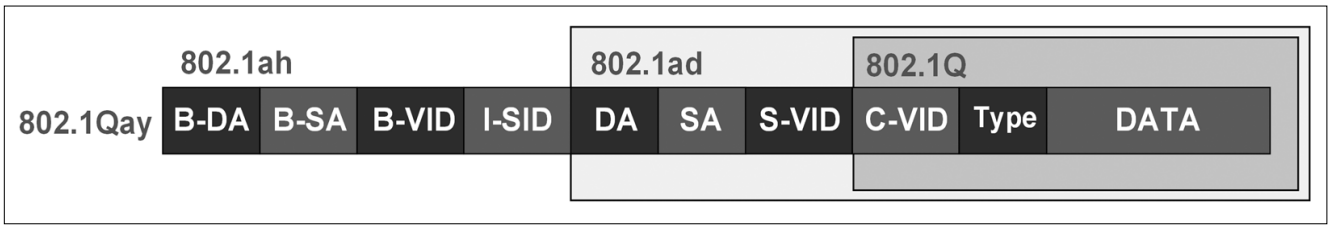


Figure 2. Carrier-Grade Ethernet frame format

ated several technical recommendations which have been turned into de facto standards by the telecom industry.

Regarding the architecture, there are two basic technical reports: TR-058 (Multi-Service Architecture and Framework Requirements) and TR-144 (Broadband Multi-Service Architecture and Framework Requirements). The TR-058 addressed the evolution from previous deployed DSL architectures to actual Multi-Service DSL architectures, but on top of an existing legacy ATM access network. Afterwards, the TR-101 (Migration to Ethernet-Based DSL Aggregation) was introduced as the next evolutionary step in the process of upgrading the access network to support Ethernet transport and switching capabilities. Therefore, the TR-144 extends the scope of TR-058 from a DSL centric architecture to a generic converged Broadband Multi-Service network architecture, which is exactly the aim of NGBAN.

As previously mentioned, there are three main entities in the reference architecture (*Fig. 1)*: the customer, the network provider and the service provider. The system allows the customer to access a set of service providers through the infrastructure and resources offered by multiple network providers.

The customer network has two types of elements: the Customer Premises Equipment (CPE) and the Residential Node (RN). The RN is a layer 2 device which carries out some essential adaptation functions between both entities. A layer 2 RN is one of the alternatives proposed by the TR-144.

The first element of the network provider is the Access Node (AN). The connection between the RN and the AN is known as the First Mile, local loop or access network. There are multiple technologies (i.e. xDSL, cable or FTTx) available in current deployments. Beyond the AN, the aggregation network is the part of the network provider which aggregates the traffic coming from multiple Access Nodes, and it is crucial for network convergence. It is very challenging for the transport technology to deal with the diverse nature of diverse service traffic. The Broadband Service Node (BSN) is the last element of network provider. It is also a layer 2 device and covers some essential adaptation functions.

The Service Provider is the entity which manages its customers (identity and credentials) and provides the services. A service is an agreement between the customer and the service provider. This means that the network provider must not impose any restriction over the service definition. The concept of Next Generation Services (NGS) is introduced as a new way of understanding what a telecommunication service could be. A NGS is a Layer 2 connection between a customer and a service provider (end-to-service), totally independent of Layer 3. This means that the connectivity provider relies only on Layer 2 (Ethernet) to join customers and services. Following this approach, Internet access (or IP connectivity) is just another service, whose multiple instantiations do not cause any collision in connectivity provider networks since IP Layer is transparent for them. Therefore, IPv4 and IPv6 (or multiple IP address schemes) coexistence becomes a reality, just as new paradigms where IP is not present (such as High Definition video directly over Ethernet).

2.2 Network convergence

Network convergence is one of the main challenges that Next Generation Networks must face. There are several initiatives and proposals dealing with this issue and all of them share some points: convergence, security, ubiquity, mobility and quality of service.

There are some different definitions for convergence; in this case, the concept is applied to the network scope: the same network to transport all the services (NGS) between the final client and the service provider. With the system architecture in mind, the convergence has a direct effect on the connectivity provider. This means that the architecture and technology selected by this provider must assure the multiplicity of services over the same physical network, what results in a multi-service and multi-provider solution.

For this purpose, IP/MPLS and Carrier Ethernet are the most realistic alternatives. The first one has been discarded because it prevents Next Generation Services (Layer 3 dependence). Moreover, IPv4 and IPv6 (or multiple IP address schemes) coexistence is not trivial. However, Carrier Ethernet fits right in with the requirements imposed by NGS (Ethernet services) and relies on a Layer 2 alternative endorsed by Metro Ethernet Forum [11] (MEF), later covered.

There are three main options for Carrier Ethernet as transport technology: Ethernet over SONET/SDH, Ethernet over MPLS, and Carrier-Grade Ethernet (CGE). The CGE technology is the best one to fulfill all the requirements imposed by Next Generation Services, and has overcome several significant challenges that traditional Ethernet (as LAN technology) evidences as carrier technology. Actually, CGE involves multiple technologies to accomplish all these challenges, such as: IEEE 802.1ad (PB), IEEE 802.1ah (PBB), IEEE 802.1Qay (PBB-TE), IEEE 802.1aq (SPB), or IEEE 802.1ag (OAM).

The necessity for network differentiation has firstly emerged in LANs environments, where companies want to isolate the traffic of each department. Virtual LANs standard [15] (IEEE 802.1Q) defines a new frame format *(Fig. 2)* that allows to differentiate Ethernet frames through Q-tag (12 bits) in order to split up the network in a logical way.

However, the same necessity appears at connectivity providers, which motivates the development of Provider Bridges standard [16] (IEEE 802.1ad), also known as Q-in-Q. This solution provides a new level of hierarchy, where customers' and providers' identification tags coexist in the same frame (Fig. 2) by encapsulating client tags (C-VID) in service tags (S-VID). Apart from this new technique, Spanning Tree Protocol (STP) and Independent VLAN Learning (IVL) are still used, limiting the scalability of developments based on PB. This restriction is motivated by a shared and flat MAC addressing scheme and the restriction of a maximum of 4096 service instances due to only 12 bits capacity in VID tags.

Because of these scalability restrictions, a new standard has been developed: Provider Backbone Bridges [17] (IEEE 802.1ah) or MAC-in-MAC. PBB overcomes PB's restrictions by encapsulating 802.1ad frames (Fig. 2) into a new provider's MAC header. In this case, instead of using a 12 bit Q-tag, a new field of 24 bits called I-SID (I-tag) is used to differentiate the services; achieving wide deployment scalability. The forwarding is based on the new header's fields (B-DA, B-SA and B-VID), totally isolated from customer's addressing scheme. So, PBB improves PB through scalability and isolation, but it maintains flooding and STP mechanisms.

Both 802.1ad and 802.1ah rely on the Spanning Tree Protocol [18] (IEEE 802.1D) to avoid loops. However, STP is not a suitable protocol for provider environments, because its goal is to get a loop-free topology by disabling those links that are not part of the tree. The generated final situation is very inefficient because it causes congestion on certain links, while others are not used at all. As an alternative, Multiple Spanning Tree Protocol (MSTP) could be introduced to get a better load balancing, but the limitation still remains.

Provider Backbone Bridges – Traffic Engineering [19] (IEEE 802.1Qay) improves CGE through traffic engineering capabilities.It is based on the MAC-in-MAC encapsulation (Fig. 2) defined in PBB, but operationally differs from it. PBB-TE disables some well known mechanisms of Ethernet like flooding, broadcasting or MAC learning, and also ignores STP associated states. On doing this, another mechanism is needed to fill the forwarding tables and assure a loop-free topology. The answer is a management system. PBB-TE achieves a connection-oriented behavior from a packet switched network by exploiting bridging forwarding mechanisms.

The forwarding decision is made according to the destination MAC address and VLAN ID (60 bits), providing great capacity to traffic engineering. The local scope of VLAN ID (B-VID) is the main difference from traditional VLAN schemes where this ID is global. This fact allows the reutilization of identifiers, which can obtain a global meaning by adding the destination address.

Shortest Path Bridging [20] (IEEE 802.1aq) is another recent development that proposes an alternative to STP dependence. SPB is a draft standard that uses PBB data plane combined with the well-known link state protocol IS-IS [6]. This enhancement adds carrier-grade any-to-any infrastructure capabilities by using the shortest path from any source to any destination.

Regarding quality of service, it is supported over 802.1p [18] efforts (included in IEEE 802.1D), and is a DiffServ based approach to provide QoS. There are eight different priorization schemes, which are included in a specific field of the VLAN tag (3 bits).

There are several developments regarding management capabilities, namely IEEE 802.1ag [7], which provides a mechanism for service fail proactive signaling; IEEE 802.3ah, which defines OAM capabilities for the first mille; IEEE 802.1AB, which allows topology discovery; ITU-T G.8031, which adds Ethernet protection mechanisms; and ITU-T Y.1713 [8], which gives additional management capacities to 802.1ag.

Definitely, Carrier-Grade Ethernet is supported by all these improvements to become the transport technology for connectivity providers. Some of the characteristics that CGE has acquired are future proof capacity for multimedia, quality of service support, scalability and hierarchical solutions, OAM capacities, and cost-effectiveness. Therefore, access and aggregation networks can be faced by native Ethernet solutions.

2.3 Next Generation Services

As previously introduced, a Next Generation Service (NGS) has been defined as an Ethernet Service. In this case, the work done by the MEF must be taken into account. This forum develops technical specifications and implementation agreements which could be considered as reference model for NGS.

The previously introduced architecture defines an end-to-end Ethernet scenario in which multiple broadband technologies could be used for service delivery. There are three different visions of Ethernet: an interface between two nodes, a service or a transport technology. Ethernet as a service means that all the Ethernet frames that enter a network provider must be delivered unmodified when leaving the provider, whereas Ethernet as a transport means that Ethernet technology is used to deliver the packets across the network provider. The former is the basis of NGS, while the latter is addressed by Carrier-Grade Ethernet (CGE).

NGS is focused on the Ethernet Services definition done by the MEF. Therefore, some terminology of MEF must be introduced. A User to Network Interface (UNI) is a physical interface or demarcation between the network provider and the customer or subscriber (located between RN and AN). An Ethernet Virtual Connection (EVC) is a logical representation of an Ethernet service as defined by the associate between two or more UNIs. The most common way of implementing an EVC is

through an S-VLAN ID of IEEE 802.1ad. A fundamental characteristic of this definition is that multiple EVCs can be multiplexed on the same UNI, which is essential to enable broadband service multiplicity over the first mile.

Three types of EVCs have been defined. The E-Line Service Type is a point-to-point EVC connection between two UNIs. In this regard, site-to-site layer 2 VPNs or Ethernet Internet access are some examples. The E-LAN Service Type is a multipoint-to-multipoint EVC connection among multiple UNIs (two or more). Multi-site layer 2 VPNs or Transparent LAN Service are examples. Finally, the E-Tree Service Type is a rooted multipoint connection (or point-to-multipoint) among multiple UNIs (two or more). E-Tree defines two different roles for a UNI: root or leaf. Each leaf is able to connect with all the root UNIs, whereas the connectivity between leaves is not allowed.

The MEF Ethernet service definition framework specifies the Ethernet Service attributes and parameters which define the UNI and EVC requirement for each Ethernet Service Type.

MEF Services are classified into two categories: port-based and VLAN-based. Port-based category implies a single service instance per UNI, which means that the network resource is dedicated to the same EVC. Consequently, this type of services can be identified on a per port basis. On the other hand, VLAN-based category implies multiple service instances per UNI, which means that the network resource is shared among multiple EVCs. Therefore, a new mechanism is needed to differentiate the services. MEF proposes the use of VLAN tags at data layer, which enables service differentiation on a per C-VLAN ID (Customer VLAN identifier) basis.

VLAN-based services make use of multiplexing attribute previously introduced, which allows multiple EVCs on the same UNI. On the other hand, port-based services make use of a special case of bundling attribute, the all-to-one bundling. The bundling service attribute enables two or more C-VLAN IDs to be mapped to a single EVC at a UNI. Moreover, both types of services are able to use two additional attributes: the C-VLAN ID preservation and the C-VLAN CoS preservation. Both preservation attributes define whether the C-VLAN ID or C-VLAN CoS is preserved unmodified across the EVC. This four attributes give great flexibility to the final system.

Another significant contribution of MEF is the complete definition and classification of Ethernet Services based on three set of attributes. The UNI attributes specify the physical interface capabilities, the service multiplexing capability or the C-VLAN bundling capability. The EVC per UNI attributes specify the C-VLAN mapping to EVC or the ingress and egress quality of services parameters (CIR, CBS, EIR and EBS). Finally, the EVC attributes specify the EVC type, the list of UNIs, the VLAN or CoS preservation or the service frame delivery behavior.

The NGS definition profile is composed of UNI, EVC per UNI and EVC attribute definitions, plus a new attribute which defines the Service. This last element has not been defined by the MEF and is an agreement between the service provider and the customer.

3. Bridging virtualization

3.1 Network virtualization

This section introduces a new approach to network convergence, which is based on virtualization techniques. Current developments on systems virtualization have allowed a new approach: the achievement of network convergence through network virtualization *(Fig. 3)*.

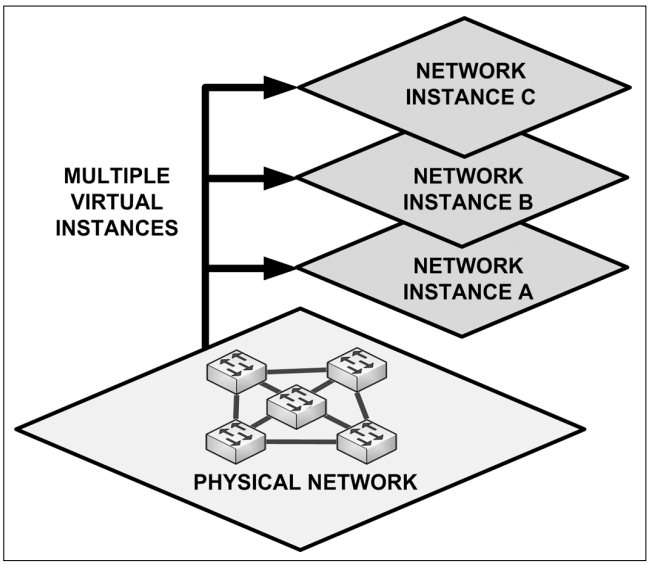


Figure 3. Network virtualization

The idea behind any kind of virtualization is the sharing of the same resources (hardware and/or software) by several instances. In the case of network, this means that the physical links (bandwidth), network interfaces and network devices (bridges in this case) are shared by different virtual network instances which use the same network infrastructure at the same time.

Sometimes it is difficult to get the same network fulfilling all the different requirements imposed by each service. Carrier-Grade Ethernet, for instance, faces this challenge by developing several different solutions like PB, PBB, PBB-TE, or SPB. In this case, all of them are proposed by the same organism that has taken care of making them compatible with each other. However, this is not always possible.

Next Generation Services will demand different behaviors from the connectivity network, and also complete isolation between services. This type of service does not depend on Layer 3, instead end-to-end (customer-to-service) connectivity relies exclusively on Layer 2 (CGE). This way, any Layer 3 protocol could be used transparently for connectivity provider network. Therefore, the final architecture would consist entirely of bridges (as has been described in Section 2).

The improvement of network efficiency, the reduction of capital and operational costs (CAPEX and OPEX), or the enhancement of provider agility are some of the ideas behind network virtualization. Resource sharing could be applied to get a network that has different behaviors depending on the instance in which the service resides. In this case, the network resources (like links, bandwidth, or equipment) are shared between all the Next Generation Services, where each instance is isolated, secured and managed by a different virtualization process.

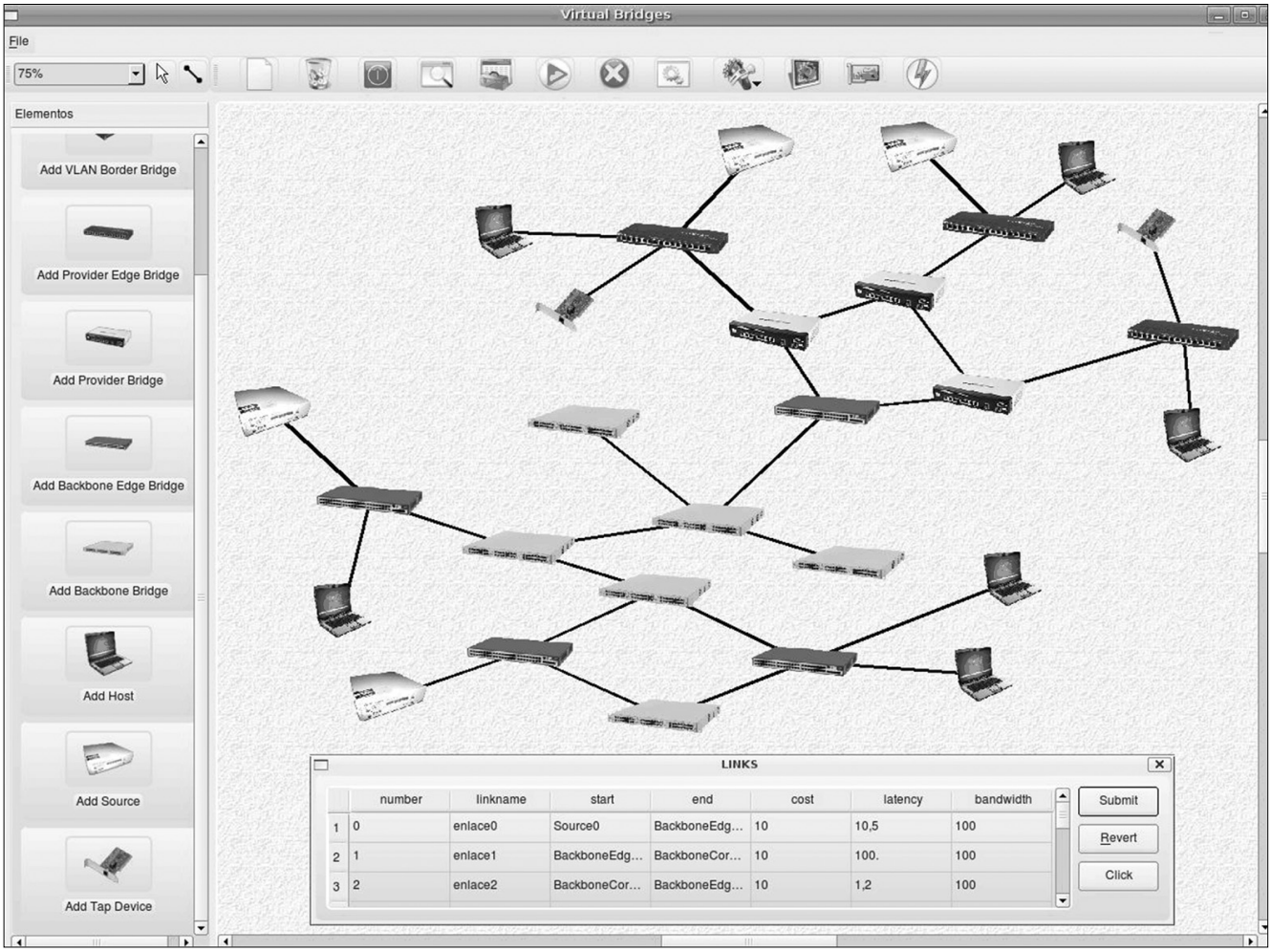
This paper proposes a solution where connectivity provider network achieves a convergence model by bridging virtualization. The data plane of each instance would be differentiated from others by using the VLAN identifier present in all Carrier-Grade Ethernet packets, which means that each VID can be associated with a different instance of the network. The control plane of each instance would manage the behavior of the bridges depending on the associated virtualization process that rules the forwarding engine. One instance could be controlled by Provider Backbone Bridges, other by PBB-Traffic Engineering, other by Shortest Path Bridging, other by a new proprietary development, and so on.

The improvement of this model is that any other solution, apart from IEEE proposals, could be addressed; the only restriction lies on using the same data plane present in CGE. This means that any new scheme that enhances the performance of a service could be easily incorporated just by adding the new virtual instance. More than in deployment scenarios, this feature is lightly valuable for testing new developments; but deployment networks could also be benefitted by the system's modularity.

All these concepts have been proved in a testbed. For the implementation, the Click tool [9] (originally developed at MIT) has been used. This tool allows network emulation with real interaction between Click and network nodes. The main benefit of using Click is that lots of developments can be reused (or improved) to get a really complex device just by adding modular components, where some might need to be implemented from scratch. This is the case of Provider Bridges (802.1ad), Provider Backbone Bridges (802.1ah) or PBB-Traffic Engineering (802.1Qay), which have been developed as new components. All of them have been implemented starting from the Virtual LAN (802.1q) component.

Click tool is able to emulate multiple devices over the same hardware or distribute them over several machines. In order to validate the functionality of the proposal, it is required to test different architectures in access and aggregation networks. Therefore, a new graphical tool *(Fig. 4)* has been developed to make this process

Figure 4. Graphical interface for Click



more efficient. This new implementation is focused on network interconnection, while each node is based on previously described components. All the network devices of this graphical tool could be sniffed (being one of its strengths), assuring that everything is working correctly.

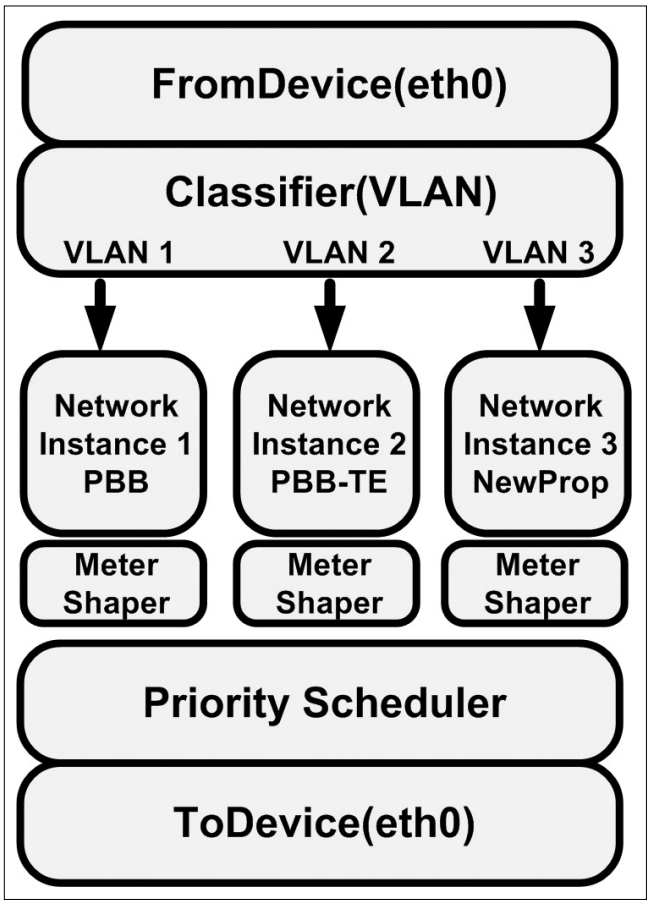


Figure 5. Bridging virtualization Click design

Fig. 5 introduces the click design that supports bridging virtualization, and therefore the network virtualization model for next generation network's convergence. Each packet goes through a classification process while entering the device through a "physical" port (it could be a real or virtualized port).

This classification determines the instance to which the packet belongs. As previously mentioned, this is done by associating VLAN identifiers and bridging instances, where each instance implements one possible CGE solution. On the other hand, the outcome of each packet goes through a prioritization process according to 802.1p developments. Previous to enforce this process, the definition of a QoS policy is needed to determine the interrelation between all the instantiated processes.

The flexibility of the model allows even the coexistence of a bridge based instantiation and a router based instantiation over the same node. This is a consequence of its modular design, where the forwarding decision of each packet is made after the classification process, and it could be taken at any level.

3.2 Click prototype

The prototype has three different parts: the click bridges, the GUI tool and the bridging virtualization technique. Each of them is analyzed in this section with more detail.

The basic element that emulates the behavior of PB, PBB and PBB-TE has been defined by creating new modules for Click. The design phase has been crucial to obtain the best reusability of the developed subsystems.

It is important to notice that in all the previously mentioned technologies, there is a remarkable difference between edge and core nodes. This results in each technology having two differentiated behaviors. The edge nodes have to encapsulate and forward packets, whereas core nodes only have to forward packets depending on the specific technology implemented. The encapsulation process is not as easy as it could seem, because a previous management process must configure the values of the new fields depending on the data that arrives in the packets. On the other hand, the forwarding process affects both the edge and the core nodes, which means that this module could be reused.

The development of the basic elements has been done in C++, a restriction imposed by the Click tool. Previous work and development of the VLAN modules have been taken as the starting point.

Apart from these new basic elements, tap devices (Layer 2 virtual network devices) have been used to achieve network connection virtualization between elements. A new tap interface is created when a new virtual link is defined between two nodes. Because of this, traffic can be sniffed from every port of every node. Having interactivity with real nodes is as easy as changing a tap device by a real interface. The real interfaces can be useful to test the platform with real nodes or to split up the platform in multiple machines.

Once those basic click elements are developed and tap devices are created, the bridges are composed by the creation of new .click files in which the interaction between both is defined. This .click files use specific nomenclature and are launched by the Click tool (that must be previously installed on the target machine) as independent processes. Afterwards, each process can be accessed by telnet to control and manage its behavior.

The problem is that the definition of those files could become very complex depending on the specific network architecture. Moreover, if this architecture changes, it could affect a huge amount of files which must be correctly configured in order to not have unstable schemes.

Because of this all, a new graphical tool has been developed. This new GUI is not only able to make network scheme composition more efficient, but also control the consistence of the final .click files (one per node) and the creation of all the required taps. The GUI is also very useful to define multiple different network schemes, without the complexity of doing all this by hand.

Apart form the bridge nodes, the GUI is also able to introduce source nodes (which generate traffic with specific characteristics and frame format) and capture devices.

Therefore, the GUI has several important tasks. First of all, it must create the tap devices needed to achieve network virtualization. Then, it must compose the .click files (one per node/element) depending on the graphically defined connections (or links) among them. Afterwards, it must launch the click processes, each of them associated with a virtual node defined by the .click files. Finally, the GUI tool has the ability to interact with the virtual nodes once they are running. This capability relies on specific handlers used by click, which have been defined in the previous mentioned design process of each element.

The last part is the one that introduces the bridging virtualization techniques. As it has been presented, it is a new radical approach for network convergence that can be useful for both research and deployment scenarios.

These techniques have a direct impact on .click files composition rather than basic elements development. This means that all the previous work defining CGE technologies can be reused only by extending the way in which the packets are handled when they come into or go outside the click processes.

It has been defined a data plane based on Ethernet, which uses 802.1ad or 802.1ah frame format. This means that all the packets carry a VLAN identifier which can be used to differentiate the specific instance of each of them. So, it is a restriction that any new definition of network behavior must support this frame format.

Focusing on the prototype, the VLAN identifier must be classified when a packet enters the bridge node (click process) in order to determine to which network instance it belongs. Once the packet is processed, the forwarding process of the specific CGE technology determines the outside interface. Since all the interfaces are shared by all the instances, a priority schedule is needed to manage the order in which the packets are sent through each interface. It is important to remember that all those interfaces can be real network devices or virtual tap devices.

The GUI tool must be also adapted to support the definition of this new type of nodes through the modification of previously used .click template files.

4. Secure instantiation process

This section is focused on the functional aspects of the service instantiation process to securely control the service delivery. There are two methods defined: the two-step AAA and the single step AAA. The latter one is common, and will be described in detail. The main difference between them is that the single step AAA only authenticates and authorizes the services, whereas the two-step AAA authenticates and authorizes both the network and the services. Both support multiple simultaneous processes of authentication and authorization of services.

A new extension of the Port-Based Network Access Control standard [21] (IEEE 802.1X) is also introduced, which defines the EAPoM (EAP over MAN or EAPoL-in-EAPoL) protocol. This extension adapts the 802.1X standard to a new scenario in which multiple services must be controlled, and is closely related to the *service port* concept.

4.1 Service port

The main restriction of IEEE 802.1X is that the standard can only control the access to the network, instead of being able to control the access of each customer to each service. It is an all or nothing access control to the network, while more granular control per service is needed.

Originally, the IEEE 802.1X standard defined the physical port of a device as the resource to be controlled; each physical port of a bridge must be authenticated by the supplicant (which is the client in an EAPoL scenario) that wants to access the network. Afterwards, with the development of IEEE 802.11i (security for WiFi networks), the physical port control was turned into a logical port control definition. With the introduction of the logical port, the IEEE 802.1X standard still remains the same, since each authentication process is univocally identified by the customer's MAC, and each customer traffic can be easily distinguished by this MAC even all the traffic goes through the same physical port.

In this new proposal, this logical port concept must be extended to address the new requirements. Following this tendency, the service port definition *(Fig.6)* is introduced as the basic element that enables the operation of the EAPoM protocol. The service port splits up the logical port into additional new ports, each of which has its own associated authentication process that rules the access to each service. This multiplicity of authentication processes is supported by EAPoM, which is able to differentiate multiple EAP processes from the same supplicant (or subscriber).

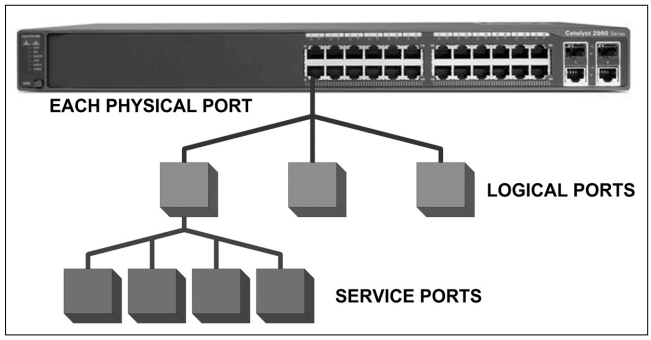


Figure 6. Service port

Another important definition associated with service ports, is the differentiation and identification of each service flow. Two are the main alternatives for this task: source-destination MAC address or VLAN identifier. The service flow identification based on source-destination MAC addresses could be used on point-to-point services or when the service is based on multicast traffic (and a multicast MAC address is defined). In both cases the access control definition is well restricted (source and destination); while in a multipoint scenario a cer-

tain variety of Access Control Lists (ACL) should be used (a list of multiple sources and destinations).

The second choice relies on IEEE 802.1Q standard [15] to identify each service flow through the use of different VLAN identifiers per service. In this case, the subscriber equipment must tag the traffic. Optionally, the Residential Node could make this tagging when customer's devices do not support it.

4.2 Two-step/single-step AAA instantiation

The two-step AAA instantiation process has two phases: the AAA process with the network provider and the multiple AAA processes with the service providers.

In the first AAA process, the subscriber is authenticated and authorized by the network provider. This procedure is similar to the standard IEEE 802.1X process. The only difference is that the EAP exchange between the subscriber and the Access Node is transported by the new EAPoM protocol, which uses a different Ethertype and MAC group address. This first network instantiation is demanded by the customer previous to any service instantiation procedure. The subscriber's identifier and associated credential are supplied by the network provider, and the concrete EAP method depends on its security policy. If this network's AAA process finalizes successfully, the subscriber is correctly configured to access the network; then, a new logical port associated with the subscriber is ready. This logical port is further divided into service ports by the subsequent service AAA processes.

As previously introduced, this first network AAA process could be omitted depending on the network's or the wholesale system's security policy. Each subsequent service AAA processes are equal, and the same as a single step AAA instantiation.

In the single step AAA Instantiation process *(Fig. 7)*, each service is consciously requested by subscribers each time they want to access a specific service. The subscriber's AAA client sends a type 5 EAPoM packet, which means that an EAPOL-in-EAPOL frame is encapsulated in the packet. The inner part has a unique service identifier associated to this AAA process with packet type 1, EAPOL-Start. From this moment, the same outer part and the same service identifier in the inner part are used for every exchange associated to this process.

This EAPOL-Start encapsulated inside the EAPoM packet goes through the Residential Node and the access network, and reaches the Access Node. The Access Node, as the first device at the network provider's premises, captures the EAPoM packet based on Ethertype or MAC group address filter rules. Then a new service port is instantiated and associated to the subscriber's MAC and the service identifier that goes inside the EAPoM packet. After that, there is a common IEEE 802.1X exchange between the subscriber and the Access Node with the only particularity that all the messages are encapsulated inside EAPoM packets with the associated service identifier. This involves a variable group of EAP request-response exchanges, starting by the identity request-response.

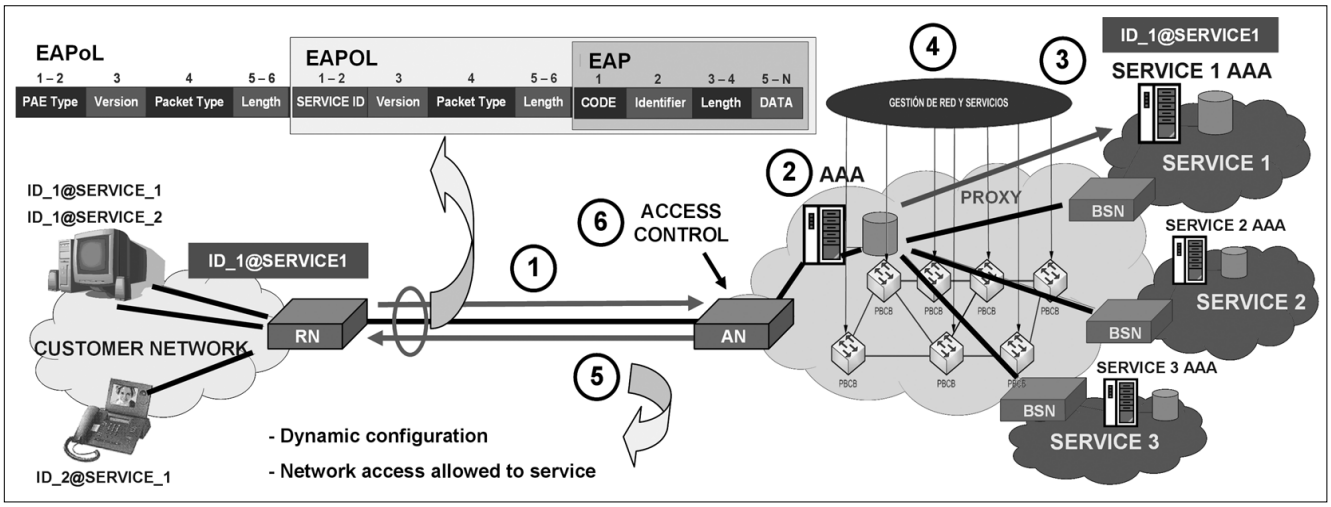
The subscriber identifier has two parts, the service provider identifier and the unique name that the customer has for this provider; this looks like ID@SERVICE (and is similar to other name schemes like e-mail or SIP ones).

Once the Access Node receives the identity response, the EAP packet is sent to the network authentication server using another protocol, in this case RADIUS. The network authentication server analyzes the subscriber identifier and depending on the service provider identifier part, it proxies the RADIUS packets to their respective service provider's authentication server. This means that the EAP exchange is made between the subscriber and the service provider; and therefore, the service provider is responsible for the authentication and authorization of the subscriber.

4.3 Profile-based configuration

Apart from AAA functions, this process is adapted to enable a configuration process. In this context, there is a secure relationship between the subscriber, the network provider and the service provider. This asso-

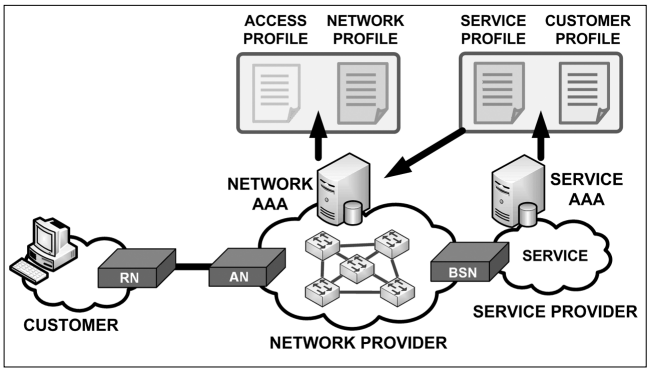
Figure 7. Secure instantiation of NGS



ciation can be used to define a set of configuration options that depends on the subscriber's identity.

This configuration process is based on the definition of profiles *(Fig. 8)*, which are XML files. This assures interoperability and a platform independent solution for configuration. Each provider has to create its own profiles associated with the service. Afterwards, the profiles are distributed together with the EAP success packet, once the AAA process ends successfully. The distribution of these profiles is done through new extended attributes of the RADIUS protocol.

Figure 8. Profile based configuration



The service provider defines two profiles: the service profile and the client profile. The service profile has all the parameters associated with the service and the subscriber's identity that must be interpreted by the network provider. The client profile has the parameters that must be configured by the customer.

The network provider also defines two profiles: the network profile and the access control profile. Both are dynamically synthesized from the service profile. The network profile defines the specific configuration that must be applied to the network in order to be able to assure the service requirements. The access control profile defines de access control policy that must be applied to the corresponding service port instantiated by the subscriber in the Access Node.

In case of using the IEEE 802.1ad standard (Provider Bridges), which has been recently adopted by the Broadband Forum, all the parameters of C-VLAN and S-VLAN associated to the service will be dynamically created and distributed to the Access Node through the configuration profile's mechanism.

A working prototype with all this concepts has been implemented in a Linux based environment. A standard development of IEEE 802.1X has been modified to support the new EAPoM protocol. The changes have been done quite easily because of its similarities with the standard, and both a supplicant and an authenticator with EAPoM support have been released. The authenticator has also been modified to support the new service port instances and a control access scheme associated with this concept. Finally, new RADIUS attributes have been defined in order to transport the previously introduced configuration profiles after a successful AAA exchange.

5. Summary

This paper introduces a new approach based on Carrier-Grade Ethernet to provide network convergence in NGBAN. The architecture recommended by Broadband Forum for Multi-Service is the reference point for the system architecture, where all the nodes are defined as bridges. Several Carrier Ethernet alternatives have been presented, among which Carrier-Grade Ethernet has been selected as transport technology for connectivity provider networks.

After all this introduction of current technology, a new approach for NGBAN has been presented. It is called *bridging virtualization,* and two new developments around it have been shown: the implementation of a prototype that validates bridging virtualization, which is made by using Click; and a graphical tool that is able to compose all the nodes (real or virtual) to get a virtualized network. The functional validation of the approach has been carried out by the implementation of these tools.

This paper introduces a new Multi-Provider and Multi-Service framework, where subscriber's access to services is granted depending on the result of an AAA process. Therefore, security improvements are evident for both network and service providers, since only previously authenticated and authorized traffic gains access to the network. This control is made per service and based on *service ports* instances, which is an evolution of standardized IEEE 802.1X logical ports. The AAA exchanges between the subscriber and each service provider are carried by the new EAPoM protocol. Apart from security aspects of the AAA exchange, a profile based configuration procedure has been associated with it. This means that configuration process depends on subscriber ID and is done in a secure context.

Another important achievement is the nomadic access to services, which means that subscribers can access their services with any location or network provider restriction. The nomadism is supported by AAA proxy mechanisms and the dissociation between identities (customer and service) and network parameters.

Finally, some remarks about future work are presented. Currently, we are working on new models for network virtualization at data-plane, which are based on the MAC addressing scheme instead of the VLAN identifier. A detailed definition of the AAA policy and obligations (for policy enforcement) are also needed to complete the security proposal. Apart from this, the Openflow technology has been considered to implement the bridging virtualization approach instead of (or in addition to) Click tool.

Authors

JON MATIAS received his B.Sc. and M.Sc. Degrees in Telecommunication Engineering from the University of the Basque Country (UPV/EHU, Spain) in 2003. He currently works as a part-time lecturer and full-time researcher in the Department of Electronics and Telecommunications at the Faculty of Engineering of Bilbao (UPV/EHU). He is also pursuing the Ph.D. degree in Telecommunication Engineering at the same University focused on access networks and security. His research interests include Computer Networks, Broadband Access Networks, Wireless Networks, Services Provisioning and Security.

EDUARDO JACOB (IEEE Member): After obtaining his BSc and MSc Degrees in Electric Engineering in 1987, he spent a few years as network manager in an architecture firm. Later, he worked as R&D project leader in Teletek (now Robotiker Tecnalia) in the field of Telecomm Engineering. He came back to the Faculty of Engineering of the University of the Basque Country in Bilbao where he received his Ph.D. degree in 2001. He is actually a professor at the same University and teaches degree courses on Mobile Networks & Services and Ph.D. courses on Cryptography in Communications and Security in Wireless Systems. He also is the head of the I2T Research Lab at his university where has directed several public and private R&D projects and acted as reviewer for projects and articles. His research interests are security in distributed systems and experimental platforms for network research. He has been appointed as ICT expert in the Advisory Council of the Basque Data Protection Agency and currently holds its Presidency.

MARINA AGUADO received her B.Sc. Degree as Telecommunication Engineer from University of the Basque Country (UPV/EHU) in 1992 and her M.Sc. in Management of Manufacturing Systems from Cranfield University, England, a year later. She concluded her European Ph.D. degree in Telecommunications Engineering in 2009. She worked as a trainee in Ford Motor Company, UK in 1993. From 1994 to 2003, she worked at the traffic control center in several railway companies in Brazil (CVRD, MRS Logistics & BrasilFerrovias), first as network support analyst, and finally, as R&D manager responsible for IT projects on railway operation. At present she works as Assistant Lecturer at ETSI (UPV/EHU) and researcher in the I2T (Engineering and Research on Telematics). Her current research lines are Broadband Wireless access technologies, Ethernet and Mobile WiMAX networks and handover related issues in the transportation scenario.

JASONE ASTORGA received her B.Sc. and M.Sc. Degrees in Telecommunication Engineering in 2004 from the University of the Basque Country (UPV/EHU). She received another M.Sc. in Information and Communication Systems in Wireless Networks in 2008, also from the UPV/EHU. Currently she is a Ph.D. student in the Department of Electronics and Telecommunications in the Faculty of Engineering of Bilbao in UPV/EHU, where she also works as a full-time researcher and part-time lecturer. Her research interests include networking, security in distributed environments and mobility management.

References

[1] Minolti, D., Johnson, P., and Minolti, E., Ethernet-Based Metro Area Networks. Planning and Designing the Provider Network, McGraw-Hill, 2002.

[2] Kasim, A., Delivering Carrier Ethernet: Extending Ethernet Beyond the LAN, McGraw-Hill, 2008.

[3] Meddeb, A., "Why Ethernet WAN Transport?," IEEE Com. Magazine, November 2005, pp.136–141.

[4] Allan, D., Bragg, N., McGuire, A., and Reid, A., "Ethernet as Carrier Transport Infrastructure," IEEE Com. Magazine, February 2006, pp.134–140.

[5] Fedyk, D., and Allan, D., "Ethernet Data Plane Evolution for Provider Networks," IEEE Com. Magazine, March 2008, pp.84–89.

[6] Allan, D., Ashwood-Smith, P., Bragg, N., and Fedyk, D., "Provider Link State Bridging," IEEE Com. Magazine, September 2008, pp.110–117.

[7] IEEE Std. 802.1ag, Virtual Bridged Local Area Networks – Connectivity Fault Management, 2007.

[8] ITU-T Y.1731, OAM functions and mechanisms for Ethernet based networks, 2006.

[9] Click Project, *http://read.cs.ucla.edu/click/*

[10] Matias, J., Jacob, E., Aguado, M., and Astorga, J., "Enhancing NGN's Versatility for Multi-Service Support: the Bridging Virtualization Approach," The 10th International Conf. on Telecommunications, ConTEL 2009.

[11] MEF TS 6.1, "Ethernet Services Definitions – Phase 2," Technical Specification, April 2008.

[12] Broadband Forum Technical Reports, *http://www.broadband-forum.org/technical/download*

[13] Future Internet Assembly, *http://www.future-internet.eu/*

[14] Bled Declaration, *http://www.future-internet.eu/fileadmin/documents/bled_documents/Bled_declaration.pdf*

[15] IEEE Std. 802.1Q, Virtual Bridged Local Area Networks, 2005.

[16] IEEE Std. 802.1ad, Virtual Bridged Local Area Networks: Provider Bridges, 2006.

[17] IEEE Std. 802.1ah, Virtual Bridged Local Area Networks: Provider Backbone Bridges, 2008.

[18] IEEE Std. 802.1D-2004, MAC Bridges, 2004.

[19] IEEE 802.1Qay, Virtual Bridged Local Area Networks: Amendment: Provider Backbone Bridge Traffic Engineering, 2009.

[20] IEEE 802.1aq/D2.5, Draft Standard, Shortest Path Bridging, 2010.

[21] IEEE Std. 802.1X-2004, Port-Based Network Access Control, 2004.

Call for Papers

Prospective authors are invited to submit original research papers for publication in the upcoming issues of our Infocommunications Journal.

Topics of interests include the following areas:

Data and network security
Digital broadcasting
Infocommunication services
Internet technologies and applications
Media informatics
Multimedia systems
Optical communications
Society-related issues
Space communications
Telecommunication software
Telecommunications economy and regulation
Testbeds and research infrastructures
Wireless and mobile communications

Theoretical and experimentation research results achieved within the framework of European ICT projects are particularly welcome.

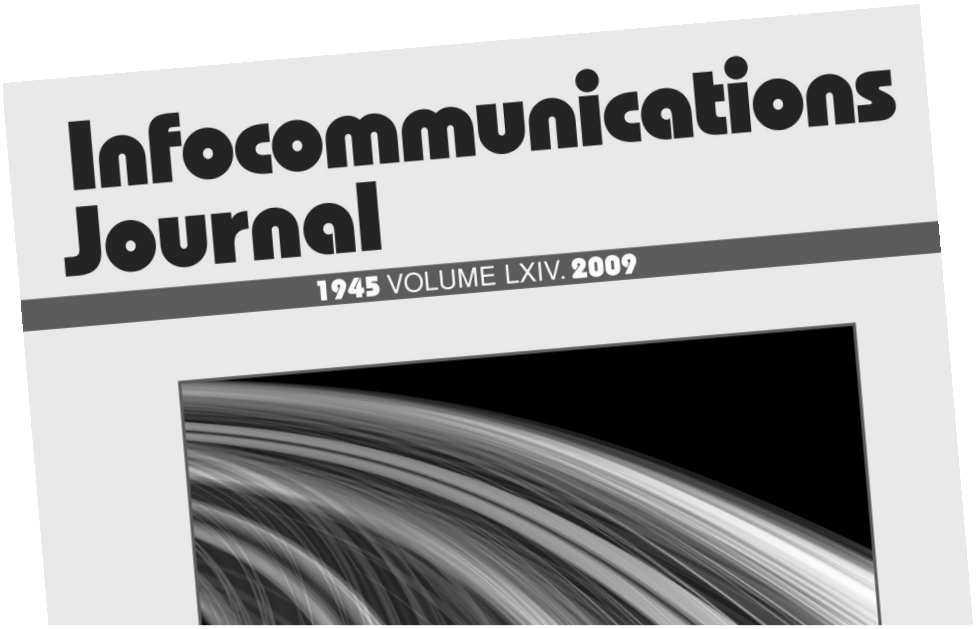
From time to time we publish special issues and feature topics so please follow the announcements. Proposals for new special issues and feature topics are welcome.

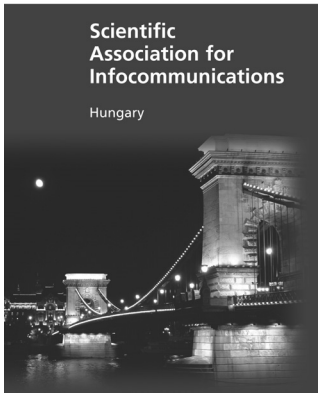
Our journal is currently published quarterly and the editors try to keep the review and decision process as short as possible to ensure a timely publication of the paper, if accepted.

As for manuscript preparation and submission, please follow the guidelines published on our website:
http://www.hiradastechnika.hu/for_our_authors

Authors are requested to send their manuscripts via electronic mail (preferably) or on a CD by regular mail to the Editor-in-Chief:

Csaba A. Szabó
Dept. of Telecommunications, Budapest University of Technology and Economics
2 Magyar Tudósok krt., Budapest 1117, Hungary
E-mail: szabo@hit.bme.hu





VoIP performance with IPsec in IPv4-IPv6 transition networks

ROMAN YASINOVSKYY, ALEXANDER L. WIJESINHA, RAMESH KARNE

Towson University, Maryland, USA
{ryasinovskyy, awijesinha, rkarne}@towson.edu

Keywords: VoIP, IPsec, IPv6, Teredo, 6to4

We conduct experiments in a LAN environment to determine the impact of IPsec on VoIP performance during the IPv4 to IPv6 transition period. VoIP performance with IPsec is measured in the presence of varying background traffic with IPv4, IPv6 and 6to4. To study the effect of NAT traversal, we also use Teredo. The IPsec scenarios that are evaluated include no-security (in which traffic bypasses IPsec), network-to-network (in which traffic is tunneled between IPsec gateways), and client-to-network (in which traffic is tunneled between the client and IPsec gateway). We use the popular Openswan implementation of IPsec and focus on ESP with the authentication option. The measures used for evaluating VoIP performance are delta (packet inter-arrival time), jitter, packet loss, throughput, and Mean Opinion Score (MOS). We also determine the time for the IPsec key exchange and call set up using SIP. Our results indicate that VoIP performance with IPsec in IPv4-IPv6 transition networks is not significantly different from that in today's IPv4 networks.
In particular, we find that 1) performance with IPv4, IPv6, or 6to4 is similar; 2) the overhead due to NAT traversal with Teredo is comparable to that when using NAT with 6to4 on the edge device; 3) performance degrades significantly when the amount of background traffic exceeds network capacity regardless of whether IPv4, IPv6, 6to4, or Teredo is used.

1. Introduction

VoIP continues to grow in popularity due to its low cost and convenience. While VoIP is primarily used today over IPv4 networks, VoIP calls in the next-generation Internet are likely to be placed between devices on IPv6 networks. However, the IPv4 to IPv6 transition is expected to last for several years due to the vast base of installed IPv4 networks. During this period, communication between many IPv6 networks will only be possible using existing IPv4 connectivity. The recommended interim measure to address this issue is 6to4 encapsulation [1]. 6to4 enables IPv6 traffic to be carried over IPv4 transit networks with minimal changes to the current infrastructure. Unfortunately, most IPv4 networks employ NAT, and 6to4 does not work with NAT unless the NAT box also serves as a 6to4 router. The situation is more complex if multiple NATs are traversed.

While Teredo [2,3] offers a solution to the NAT traversal problem, it requires additional infrastructure including servers and relays, as well as Teredo-aware clients. Teredo is proposed and supported by Microsoft, with Vista and Windows7 having Teredo enabled by default; there is also a Linux implementation [4]. Thus, a study of VoIP performance in IPv4-IPv6 transition networks needs to consider 6to4 as well as Teredo.

In this paper, we conduct experiments in a test LAN to evaluate the impact of IPsec with 6to4 and Teredo on VoIP performance. Although other approaches to VoIP security exist, IPsec VPNs were used in our study since they are frequently used to protect all IP traffic in IPv4 networks. The basic question we address in this research is the following. To what extent does the overhead added by an IPsec VPN impact VoIP performance during the IPv4 to IPv6 transition period? To assess this impact, we evaluate VoIP performance in an IPsec VPN over IPv4, IPv6, and 6to4, and when Teredo is used for NAT traversal. We also measure the time to complete the IPsec key exchange and for call set up using SIP.

The experiments with IPv4, IPv6 and 6to4 use a test LAN with 6to4, and IPsec gateways that allow VoIP quality to be evaluated using three IPsec scenarios: no-security in which traffic bypasses IPsec; network-to-network in which traffic is tunneled between IPsec gateways; and client-to-network in which traffic is tunneled between the client and IPsec gateway. All IPsec experiments used ESP tunnel mode with the authentication option [5]. To evaluate the effect of NAT on VoIP performance with IPsec, we configured a 6to4 gateway to serve as a NAT box, and also set up a server enabling Teredo to be used for NAT traversal by the clients.

In our experiments, VoIP traffic is transmitted through Linux routers on the test LAN together with data traffic at various rates. Congestion is introduced by using a 100 Mbps transit network to carry traffic from a gigabit Ethernet. VoIP performance is then studied by measuring values of delta (packet inter-arrival time), jitter, packet loss, and throughput using Wireshark. Finally, voice quality is estimated by computing the MOS (Mean Opinion Score).

The values of delta (packet inter-arrival time) reflect delay in the network, but do not estimate the actual end-to-end delay. Auxiliary measurements we conducted to estimate the end-to-end delay in our network showed

that it is well within the commonly accepted 150 ms limit except when the network is unstable at very high loads. Also, since Wireshark was run on a separate machine and not on the softphones, the measured values do not consider jitter buffer effects, and decoding and decryption delays at the receiver prior to playback. Thus, it is possible that the values of the reported measures might not represent the actual voice quality experienced by the receiver. To address this issue, we computed an average MOS based on MOS values assigned by human listeners. Since this average MOS correlated well with the MOS by using values measured by Wireshark, we believe that the measured values accurately reflect actual call quality.

The main contributions of this paper are results demonstrating that 1) the popular IPsec-based VPN technology used in IPv4 networks today can continue to be used during the IPv4 to IPv6 transition period with no significant impact on VoIP quality; 2) the additional overhead due to 6to4 or Teredo processing has a negligible effect on VoIP quality with IPsec if the network capacity is not exceeded. This paper is an extended version of [6]. The differences are the inclusion of VoIP performance measurements with IPsec when Teredo is used for NAT traversal; and the determination of delays for VPN establishment, and for user registration and call set up via SIP.

The rest of this paper is as follows: In Section 2, we briefly discuss related work. In Section 3, we describe the test network, and in Section 4, we present the results. In Section 5, we present the conclusion.

2. Related work

In a previous study on IPsec with IPv6 using real traffic [7], hosts with an Intel Pentium II 450 MHz processor and 128 MB memory running Free BSD 2.2.8, and routers with an Intel Pentium III 500 MHz processor were used. Their study compared the end-to-end throughput for IPv4 and IPv6 without IPsec, with only AH, with only ESP, and with both AH and ESP. The application used for the study was digital video. The experiments showed that for large amounts of data, the use of authentication and encryption reduces the throughput by 1/9. In this case, the throughput was about 10 Mbps for UDP and 6 Mbps for TCP. Their study demonstrated the feasibility of securely transmitting video using IPsec over IPv6 with ordinary hardware. However, their study does not apply to VoIP and it did not specifically consider IPsec scenarios that are common to today's VPNs using modern implementations on Linux systems that are popular today.

The overheads of an IPSec VPN server with IPv4, and performance improvements are studied in [8,9]. The studies use Openswan, were mainly concerned with the overhead due to the IKE/ISAKMP key exchange, and show that it is much larger than the ESP overhead. In [10], performance of voice and video in an IPsec VPN for videoconferencing is analyzed and it is concluded that the VPN cannot meet QoS requirements under heavy loads. Studies have also examined the IPsec overhead with IPv4 for email and Web applications [11], and Web servers with IPv4 and IPv6 [12]. The performance of 6to4 without IPsec for TCP traffic is evaluated in [13] and it is found that the additional overhead due to tunneling is minimal.

An evaluation of IPsec with 6to4 is done in [14], but the study does not address VoIP performance. In [15], the authors describe the implementation of an IPsec VPN using IPv6, discuss the tradeoffs, and perform testing. VoIP performance over IPv6 and IPv4 without IPsec or 6to4 is compared in [16], and it is shown that the difference in VoIP call quality due to the different IP versions is negligible. In [17], the impact of IPv6 on SIP is studied considering both 6to4 and Teredo. Our study is similar, but also takes varying levels of background traffic into account. The focus in [18] is on comparing 3G UMTS network performance over IPv6 with IPv4 and tunneled IPv6 for multimedia systems; it does not deal with VoIP performance over IPv6 with Teredo or 6to4.

The main difference between this study and the previous studies is that we focus on VoIP performance with IPsec in IPv4-IPv6 transition networks. To this end, we study VoIP performance in IPsec VPNs over IPv4, IPv6 and 6to4, and by using Teredo for NAT traversal. VoIP performance is measured by making calls using softphones and sending the VoIP traffic and variable amounts of other UDP data traffic through a LAN with several routers.

When VoIP traffic passes through a VPN tunnel, all IP payloads carrying the voice traffic (including UDP and RTP headers), the ESP trailer, and the message authentication code are encrypted. These fields (excluding the authentication code field) plus the ESP header can also be authenticated. Furthermore, the inner IP header is also encrypted and could optionally be authenticated. However, since there is no IPsec protection within the network sites, a protocol such as SRTP [19] would be needed for end-to-end VoIP security.

3. Network and experimental setup

Fig. 1 shows the test LAN for the network-to-network IPsec scenario. In this scenario, router #1 and router #4 act as both 6to4 and IPsec gateways. Router #4 can also serve as a NAT box. The test LAN for the client-to-network scenario is the same except that IPsec is enabled at client #1 instead of router #1. To test Teredo, NAT is enabled on router #1 and router #4, and a Teredo server is deployed on the network between router #2 and router #3.

Calls using Linphones [20] (softphones) are made between the two clients. We use Linphones for consistency and convenience as they exhibited stable behavior and were easy to configure with either IP version. MGEN [21] is used to generate UDP background traffic. One cli-

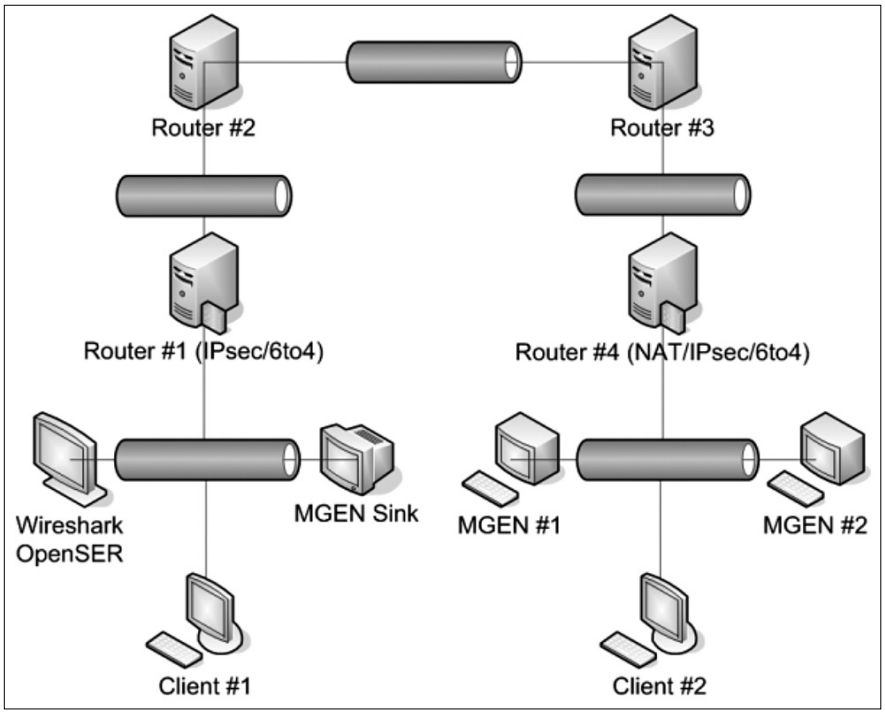
ent (client #2) and the traffic generators (MGEN#1 and MGEN #2) are located on a gigabit Ethernet. During a call, the VoIP data consisting of 20 ms voice packets generated by a Linphone is collected for 2 minutes (i.e., 2-minute conversations) and the results for the second minute only are used (to eliminate any startup effects).

We use the Openswan implementation of IPsec with IKEv1 [22] since it is presently used in most VPNs. The newer Strongswan implementation with IKEv2 [23] addresses several security issues with IKEv1 but has yet to be widely deployed.

In the network-to-network scenario for example, the background traffic (MGEN UDP data at rates of 50, 100, 150 and 200 Mbps) first passes through the 6to4 router #4. It then becomes IPsec ESP traffic and passes through three 100 Mbps networks connected by 2 routers (router #3 and router #2 respectively) before entering the destination network via router #1 that acts as an IPsec/6to4 endpoint gateway as shown in the figure. Although packet loss is possible in this network, packets cannot arrive out of order. When IPsec processing is necessary, it is always done first. For IPv6 packets, this is followed by 6to4 processing.

ESP encryption and authentication are then applied by router #4 to IPv6 packets carrying the VoIP data with the addition of an ESP header and an outer IPsec header (tunnel mode). The resulting packets are then prefixed with an IPv4 header (6to4 encapsulation) and forwarded to the destination through the intermediate IPv4 networks and routers. At the destination network, router #1 decapsulates the received 6to4 packet, and does IPsec authentication and decryption before forwarding the IPv6 voice and data traffic to their respective destinations.

Figure 1.
Test LAN with IPSec/6to4 the case of a site-to-site VPN



Wireshark [24] running on the destination network captures the voice traffic delivered to the client by port mirroring at the switch and reports values of delta, jitter, packet loss, and throughput that are used to measure VoIP performance. The machine running Wireshark doubles as an OpenSER SIP server [25] for setting up the calls.

The specifics of hardware, software and MGEN traffic used for the experiments are as follows:

Hardware: Router/Server/MGEN: Dell Optiplex GX260 (Pentium 4, 2.4 GHz, 512 Mb RAM, Intel PRO/1000, 3Com 10/100); Client: Dell Optiplex GX270 (Pentium 4, 2.4 GHz, 2048 Mb RAM, 3Com 10/100); Switches: Cisco Catalyst 2950, Netgear GS108 (1000), Netgear FS308 (10/100), Trendnet TE100-S55E (10/100).

Software: CentOS 5 (2.6.18-92.1.22) (Routers, SIP Server, NTP Server, Wireshark), Windows XP (SP3) (Generators + Sink), Fedora 10 (2.6.27.9-159) (Clients), Linphone 2.1.1-1 (ITU-G.711 codec), Wireshark 1.0.3, MGEN 4.2b4, OpenSER 1.3.4-1, Openswan 2.6.14-1.

MGEN Background Traffic: n streams are used to generate 5n Mbps of background traffic, where n=10, 20, 30, 40.

4. Results

Each experiment is run several times and the results shown are averages over three runs.

A) Delta (Packet Inter-arrival Time)

We consider the maximum (max) and mean values of delta (shown in *Fig. 2*) and its relative frequency distribution.

In general, it is not possible to directly relate the value of delta to the actual delay.

Max Delta: When there is no background traffic, there is no packet loss and max delta is about 40 ms for all four IPsec scenarios with either IP version or 6to4 encapsulation. When there is 50 Mbps of background traffic, there is still no packet loss, and max delta increases slightly, but again the differences due to IPsec scenario, IP version and 6to4 encapsulation are insignificant.

When background traffic is at 100 or 150 Mbps, larger increases in max delta are seen but the values are not significantly different for the no-security, and client-to-network scenarios with either IP version or 6to4 encapsulation. When background traffic is increased to 200 Mbps, delta for some packets exceeds 100 ms for no-security with 6to4 and client-to-network with either IP version.

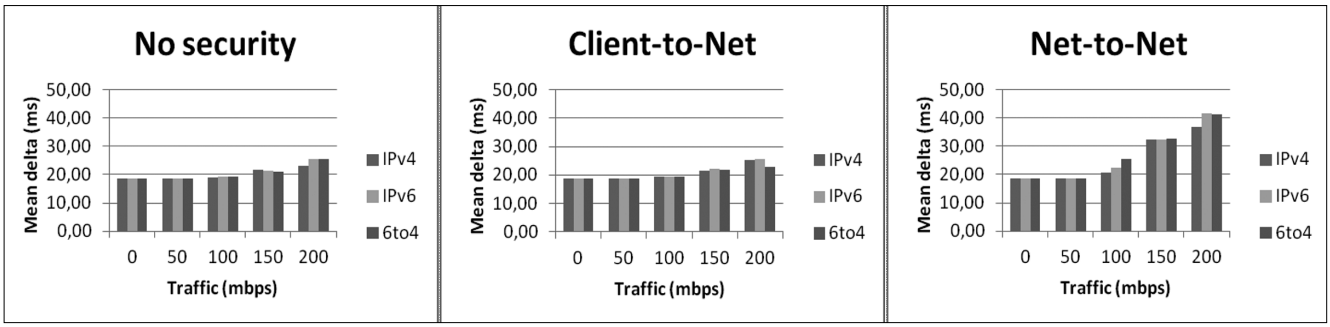


Figure 2. Mean delta

For the network-to-network scenario, it is possible for some packets to have extremely large delta values at 100 Mbps with 6to4, at 150 and 200 Mbps with IPv4, and at 200 Mbps with IPv6 (max delta exceeds 800 ms in these cases). These large delta values for the network-to-network scenario at high background traffic rates are due to packet loss and increased delays.

Mean Delta: In contrast to max delta, the values of mean delta are stable. For the no-security and client-to-network scenarios, there is very little difference in mean delta values with either IP version or 6to4 encapsulation, and it varies from 19-26 ms (the mean is 26 ms at 200 Mbps for no-security with IPv6 and 6to4, and for client-to-net with IPv4 and IPv6). For the network-to-network scenario at background traffic rates of 100 Mbps or less, mean delta values are similar to those for the other scenarios and there is at most a small difference in values with either IP version or 6to4 encapsulation.

At a background traffic rate of 150 Mbps, mean delta values are 33 ms, and at 200 Mbps it is 37 ms with IPv4 and 42 ms with IPv6 and 6to4. For the no-security and client-to-network cases, the standard deviation of delta varies from 8-16 ms when the background traffic is increased from 0-200 Mbps, and there is little difference in the standard deviation with either IP version or 6to4 encapsulation. For the network-to-network scenario, there is more variability in the delta values: at 0 and 50 Mbps, the standard deviation is 8 ms with either IP version or 6to4 encapsulation; at 100 Mbps it is 11 ms with IPv4 and IPv6, and 24 ms with 6to4; at 150 Mbps it is 21 ms with IPv6 and 6to 4 and 29 ms with IPv4; and at 200 Mbps it is 30 ms for 6to4 and approximately 40 ms for IPv4 and IPv6.

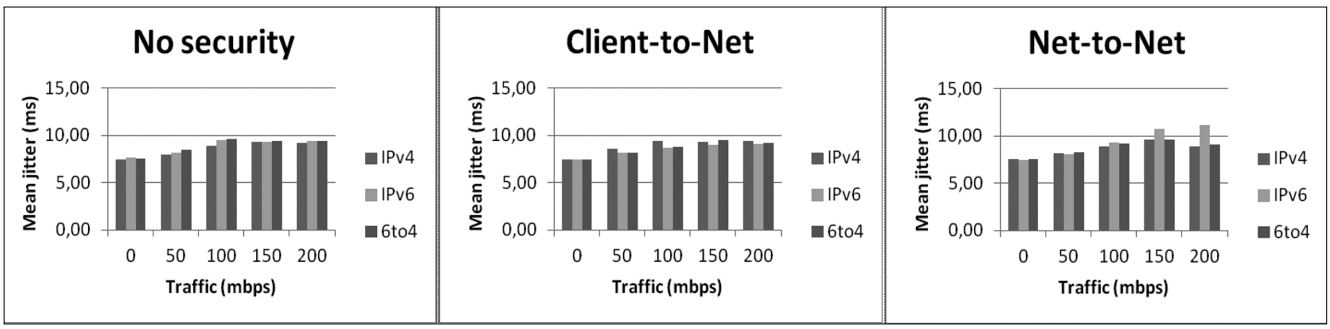
Relative Frequency Distribution: The relative frequency distribution of delta provides more details concerning the actual values of delta that are obtained. At background traffic rates of 0 and 50 Mbps, for all IPsec scenarios with either IP version or 6to4 encapsulation, approximately 70-80% of packets have delta values between 0-24 ms and the rest have delta values between 25-49 ms. The same is true at 100 Mbps, for the no-security and client-to-network scenarios, with either IP version or 6to4 encapsulation, except that a very small number (less than 1%) of packets have delta values between 50-74 ms.

For all three IPsec scenarios with either IP version or 6to4 encapsulation at 150 and 200 Mbps, at most 6% of packets have delta values between 50-74% and a very small number (at most 1.5%) have delta values of 75 ms or more. For the network-to-network scenario, the delta distribution has more variability: with either IP version or 6to4 encapsulation at 0 and 50 Mbps, the delta distribution is similar to the other scenarios; at the higher rates of background traffic, the distribution is also similar to the other scenarios except that the percentages of packets having delta values respectively between 50-74 ms and 75 ms or more increases (for instance, the percentage of packets having delta values between 50-74 ms varies from about 5-30% and the percentage of packet having delta values of 75 ms or more varies from about 5-15%.

B) Jitter

Max Jitter: Max jitter ranges from 13 ms for IPv4 with no security to 24 ms at 150 Mbps for the network-to-network scenario with IPv4. In the case of IPv6, max jitter ranges from 13 ms for the client-to-network or network-to-network scenarios with no background traffic to 21 ms for the network-to-network scenario with background traffic at 200 Mbps. With 6to4, max jitter varies from 13 ms

Figure 3. Mean jitter



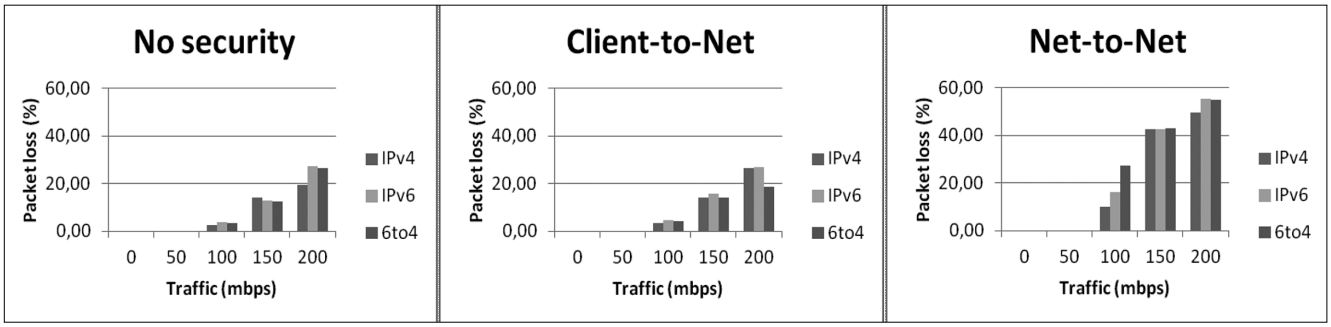


Figure 4. Packet loss

with no security and no background traffic to 26 ms for the network-to-network scenario with background traffic at 100 Mbps. However, max jitter sometimes reached 27 ms even with no security and no background traffic. Thus, it is important to examine both maximum and mean jitter.

Mean Jitter: Mean jitter values are shown in *Fig. 3.* The values range from 7-10 ms for IPv4, from 7-11 ms for IPv6, and from 7-10 ms with 6to4. The results show that mean jitter is not affected significantly by IPsec and 6to4 processing.

C) Packet loss

Packet loss percentages are shown in *Fig. 4.* Wireshark calculates these percentages by using RTP sequence numbers to determine missing packets. We note that there is no packet loss when background traffic is at 0 or 50 Mbps for all IPsec scenarios with either IP version or 6to4 encapsulation. At 100 Mbps of background traffic, packet loss with IPv4 varies from 2% for no-security to 10% for the network-to-network scenario. For IPv6, the range is from 4-16%, and for 6to4, it is from 4-27%. The highest packet loss percentage is 55% for the network-to-network scenario at 200 Mbps with either IPv6 or 6to4.

D) MOS

The maximum MOS of 4.41 is obtained with 0 or 50 Mbps of background traffic regardless of the IPsec scenario and regardless of whether IPv4, IPv6 or 6to4 encapsulation is used. At 100 Mbps of background traffic, MOS values are good with a slight drop for the client-to-network and network-to-network IPsec scenarios with either IP version or 6to4 encapsulation. As expected, at 150 or 200 Mbps of background traffic, the MOS drops to unacceptable levels.

E) Throughput

The throughput is the number of bits transferred per second considering only the voice packets. Throughput is shown in *Fig. 5.* If there is no packet loss, it is easily verified that the expected throughput with 20 ms voice packets is 0.4*s, where s is the total size in bytes of the voice packets including all headers and data. The packet size with IPv4 is 218 bytes, and with IPv6 and 6to4 is 238 bytes (since there are 160 bytes of voice data, 12 bytes of RTP header, 8 bytes of UDP header, either 40 bytes of IPv6 header or 20 bytes of IPv4 header, and 18 bytes of Ethernet header plus trailer). This gives expected throughput rates of 87.2 kbps with IPv4, and 95.2 kbps with IPv6 or 6to4.

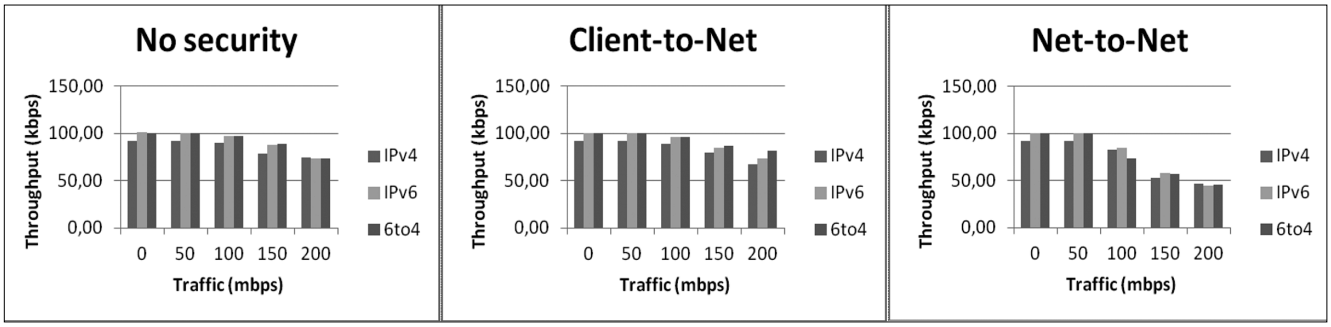
As expected, the measured and expected throughput is similar with background traffic rates up to 50 Mbps but differ when the rates are 100 Mbps or higher. This is because there is packet loss and increased delays at higher background traffic rates, which lowers the number of packets received per second. In general, the throughput for all IPsec scenarios is about the same for a given rate of background traffic with either IP version or 6to4. This implies that the additional overhead due to the extra headers and processing with IPsec and IPv6 or 6to4 does not significantly affect the voice throughput.

Note that size of a packet captured by Wireshark does include the 4-byte CRC at the end of an Ethernet packet. It computes the throughput by multiplying the number of packets received during the measurement interval by the observed packet size.

F) NAT with 6to4 or Teredo

To determine the additional overhead on the IPsec/6to4 router when it is using NAT to handle traffic from IPv4

Figure 5. Throughput



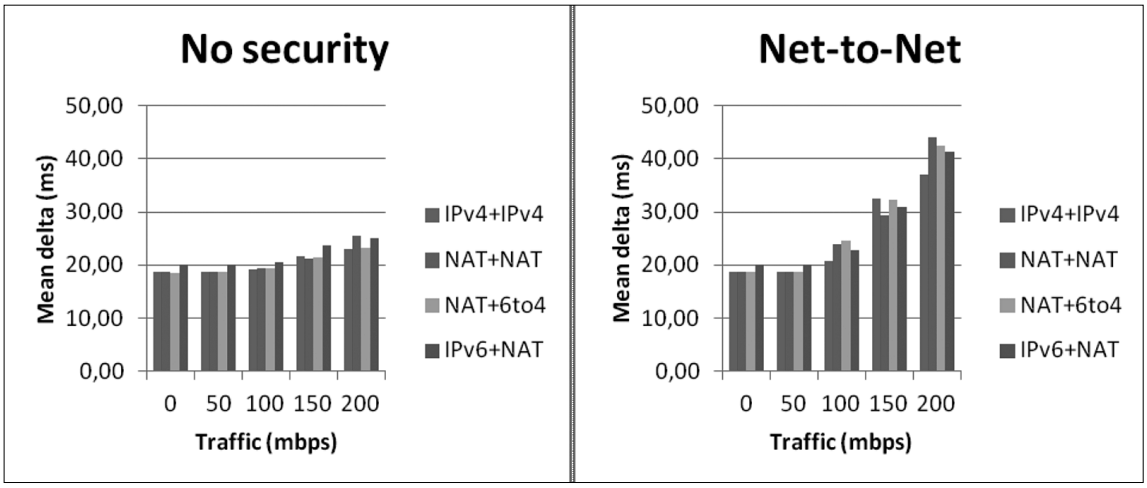


Figure 6. Mean delta (NAT with 6to4)

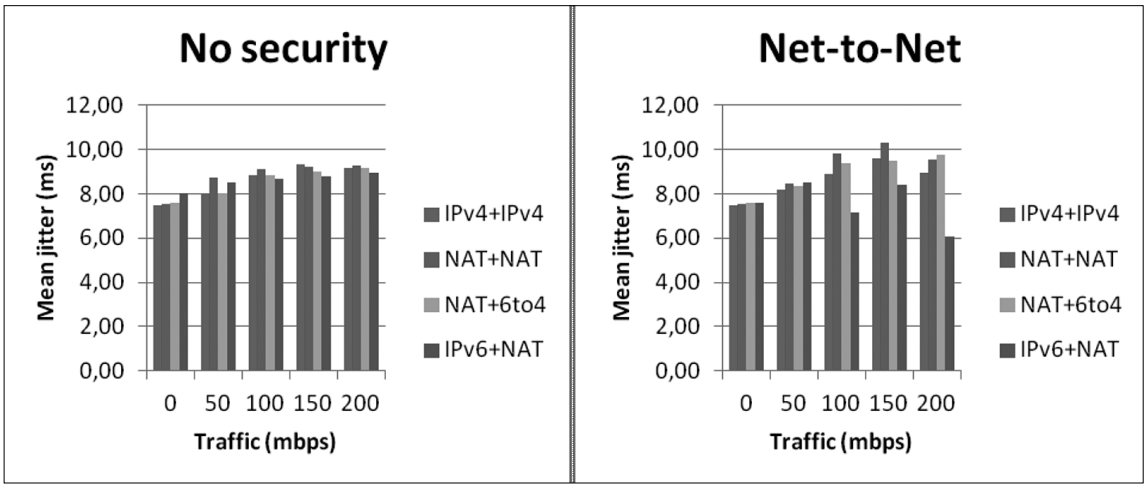


Figure 7. Mean jitter (NAT with 6to4)

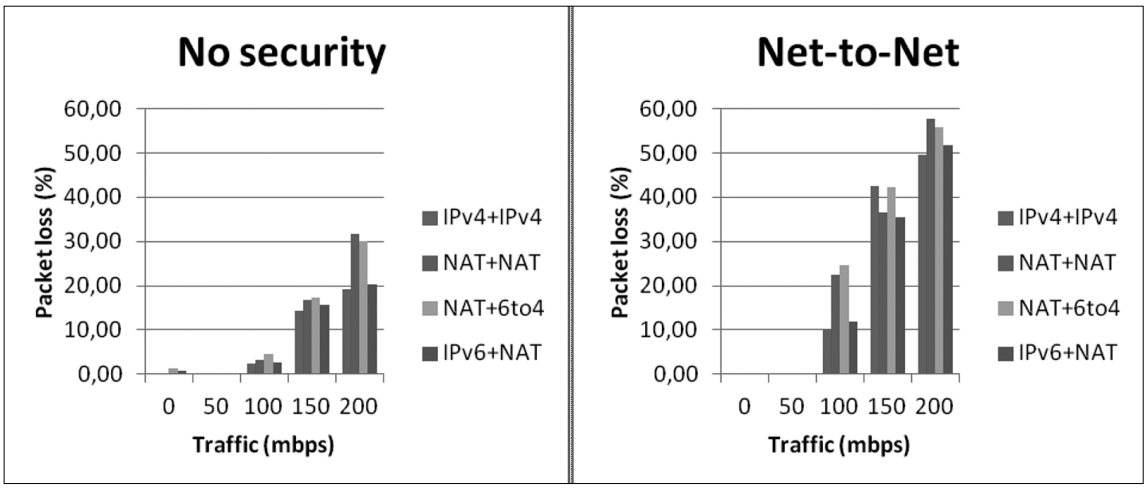


Figure 8. Packet loss (NAT with 6to4)

subnets, we repeated the previous experiments after enabling NAT. The difference in the values of all measures of interest due to NAT processing was found to be negligible regardless of the IPsec scenario. For reasons of space, we only show mean delta, mean jitter and packet loss for the no security and network-to-network IPsec scenarios with NAT *(Fig. 6-8)*.

In the figures, IPv6+NAT means the clients are IPv6 and the background traffic is IPv4 with NAT, NAT+6to4 means clients are IPv4 with NAT and background traffic is 6to4, and so on. However, as noted previously, NAT traversal over 6to4 is not possible unless the NAT box and 6to4 edge router are co-located and the NAT box is capable of performing the necessary 6to4 functions.

Teredo required the exchange of 6 pairs of router solicitation and router advertisement pairs between the Teredo client and the Teredo server over IPv4 before voice packets were exchanged over IPv6. Each solicitation message is an 89-byte IPv4 datagram that carries a UDP-encapsulated IPv6 message prefixed with a 13-byte Teredo authentication header. The IPv6 message contains an 8-byte ICMPv6 message. The advertisement message is a 125-byte IPv4 datagram that carries a UDP-encapsulated IPV6 message prefixed with a 13-

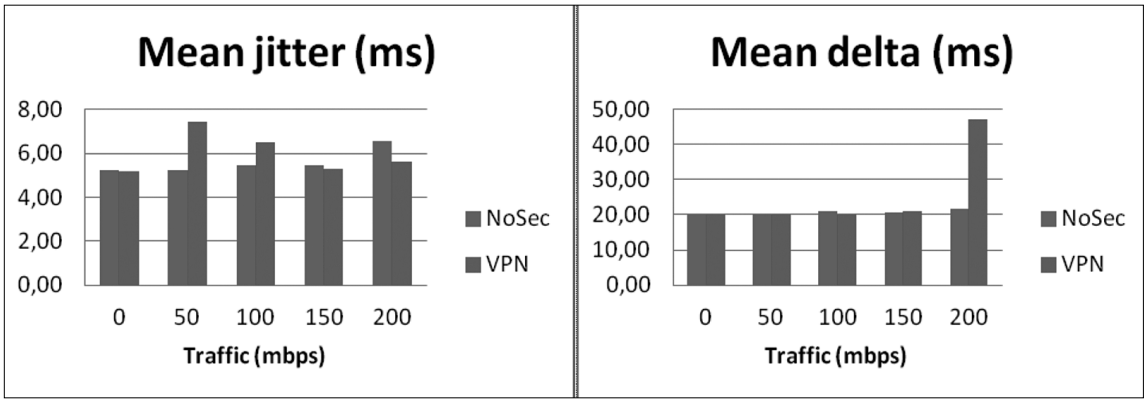


Figure 9. Mean jitter and mean delta (Teredo)

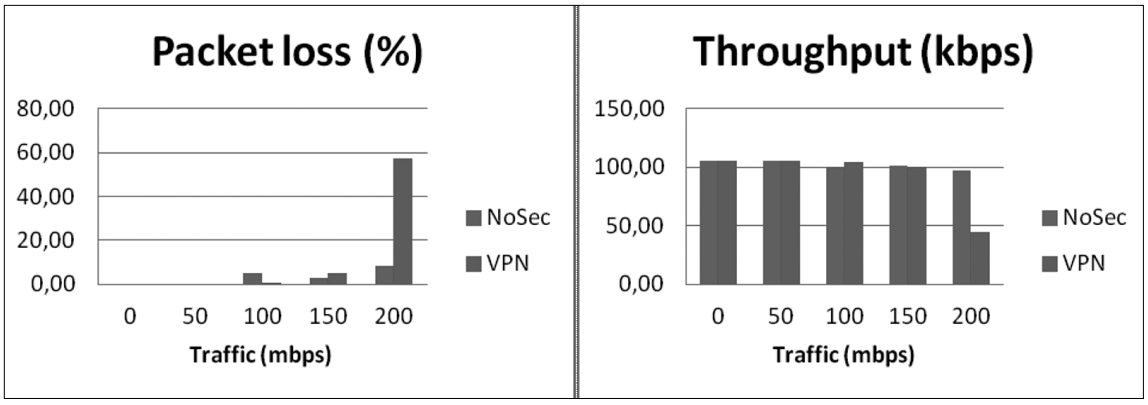


Figure 10. Packet loss and throughput (Teredo)

byte Teredo authentication header and an 8-byte Teredo indicator. The IPv6 message contains a 56-byte ICMPv6 message. Each voice packet over Teredo is a 248-byte IPv4 datagram that consists of a UDP-encapsulated IPv6 message containing 160 bytes of voice data, a 12-byte RTP header, and an 8-byte UDP header. In contrast, a 6to4 voice packet has 8 bytes less due to not requiring the extra UDP header. The only IPsec scenarios possible with Teredo are no security and network-to-network. Comparing *Fig. 2, 3 and 9* we see that that the mean jitter and mean delta values for Teredo and 6to4 are similar.

Fig. 10 contains packet loss and throughput for Teredo. The packet loss values for Teredo are slightly better than those for 6to4 in Fig. 4. The expected throughput value for Teredo computed using the formula given earlier is 0.4*266=106.4 kbps, and this value is close to that achieved for background traffic at 150 Mbps or less.

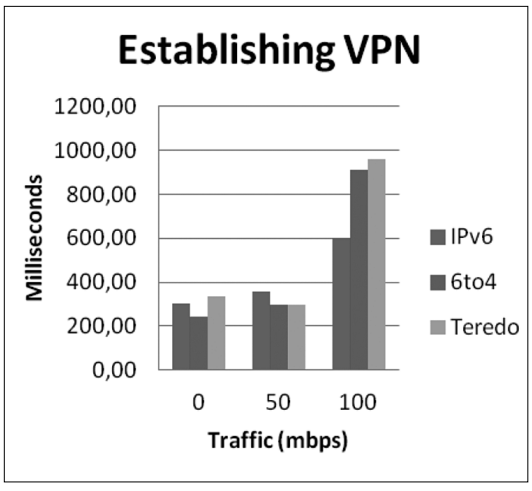


Figure 11. IPsec key exchange delay

When background traffic is at 200 Mbps, throughput in the VPN scenario drops drastically due to the high percentage of lost packets.

G) IPsec key exchange and SIP call set up times

We also determined the delays for the key exchange to initially set up the VPN using the IKEv1/ISAKMP handshake.

Fig. 11 shows that the delays during the initial key exchange are between 200-400 ms with background traffic levels of 0-50 Mbps regardless of whether IPv6, 6to4 or Terdeo is used. When the traffic level exceeds 100 Mbps, delays are unpredictable due to a backoff algorithm that doubles the delays between retries. This delay would be a factor in evaluating overall VoIP performance if the handshake is repeated several times during a call to provide additional protection against key compromise or staleness.

Fig. 12 reports the delay between sending a SIP registration request (required of all clients prior to call set up) and receiving the 200 OK message. It is seen that delays for the network-to-network scenario with no security are comparable for IPv6, 6to4, and Teredo (around 50 ms) when background traffic levels are 0-50 Mbps. With a VPN tunnel, the delays appear to be unpredictable even with background traffic at 50 Mbps.

Finally, *Fig. 13* measures the SIP call set up delay, which is the time between sending the invite message and receiving the 180 ringing message. With 0-50 Mbps of background traffic, the call set up delays are less than 20 ms for IPv6, 6to4 and Teredo with or without a VPN tunnel. When the background traffic level rises to 100 Mbps, delays are on the order of 120 ms.

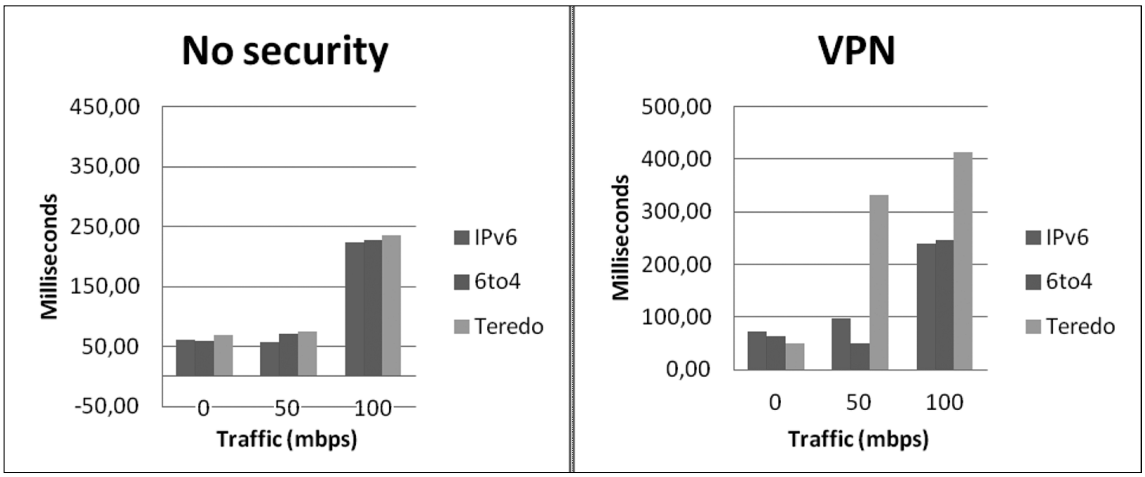


Figure 12. SIP registration delay

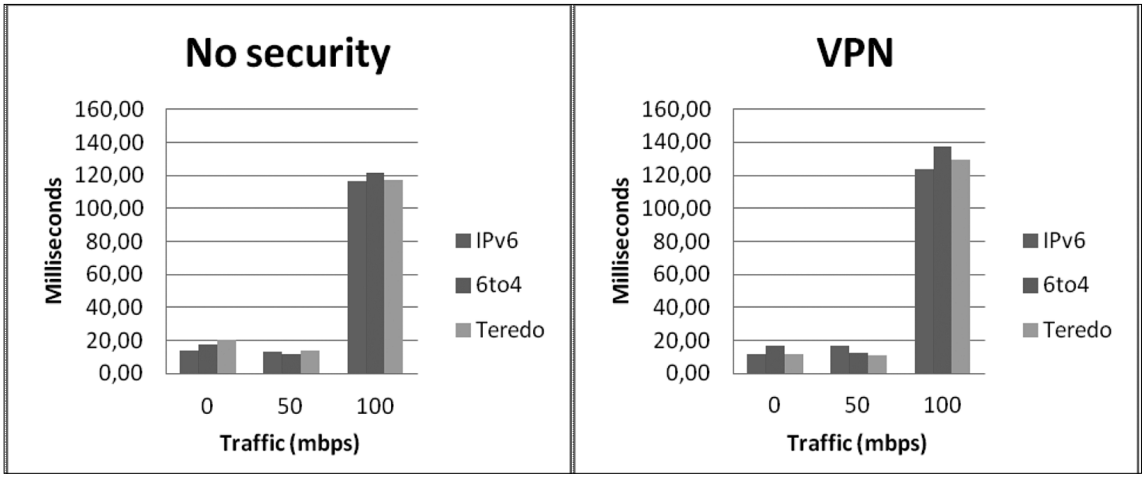


Figure 13. SIP call set up time

5. Summary

We conducted a study to evaluate VoIP performance with IPsec in IPv4, IPv6 and 6to4 networks, and also when using Teredo for NAT traversal in a test LAN. The experiments used softphones to make calls and generated background traffic to create congestion on the links and routers. The results demonstrated the feasibility of using a single Linux box to handle IPsec, 6to4 and NAT processing. It was found that voice quality is acceptable as long as the traffic does not exceed network capacity.

The study showed that VoIP performance with IPsec is not adversely affected by the overhead due to 6to4 or Teredo. Future studies should evaluate the impact of multiple calls and the use of IPsec with both Teredo and 6to4 in a more complex test network.

Authors

ROMAN YASINOVSKYY received his B.Sc. degree in computer science and the M.Sc. degree in computer control systems and technologies from the National University of Kyiv-Mohyla Academy, Ukraine in 2002 and 2004 respectively. He received his D.Sc. degree from Towson University in 2009. His dissertation dealt with VoIP performance over IPv6. He is currently a lecturer in the Department of Computer and Information Sciences at Towson University. His research interests include operating systems security, networking, and parallel/distributed computing.

ALEXANDER L. WIJESINHA is an associate professor in the Department of Computer and Information Sciences at Towson University. He holds a Ph.D. in computer science from the University of Maryland Baltimore County, and both a M.S. in computer science and a Ph.D. in mathematics from the University of Florida. He received a B.S. in mathematics from the University of Colombo, Sri Lanka. His research interests are in computer networks including wireless networks, VoIP, network protocol adaptation for bare machines, network performance, and network security.

RAMESH K. KARNE is a professor in the Department of Computer and Information Sciences at Towson University. He obtained his Ph.D. in Computer Science from the George Mason University. Prior to that, he worked with IBM at many locations in hardware, software, and architecture development for mainframes. He also worked at the Institute of Systems Research at University of Maryland, College Park as a research scientist. His research interests are in bare machine/dispersed operating system computing.

References

[1] Carpenter, B. and Moore, K.,
"Connection of IPv6 Domains via IPv4 Clouds",
RFC 3056, February 2001.

[2] Huitema, C.,
"Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)",
RFC 4380, December 2006.

[3] Huang, S-M., Wu, Q., and Lin, Y-B.,
"Tunneling IPv6 through NAT with Teredo Mechanism,"
In Proc. 19th International Conference on Advanced Information Networking and Applications (AINA'05), Vol. 2, pp.813–818, 2005.

[4] Miredo – Teredo for Linux and BSD,
http://www.remlab.net/miredo/

[5] Kent, S.,
"IP Encapsulating Security Payload (ESP)",
RFC 4303, December 2005.

[6] Yasinovskyy, R., Wijesinha, A., and Karne, R.,
"Impact of IPsec and 6to4 on VoIP Quality over IPv6,"
In Proc. 10th International Conference on Telecommunications (ConTEL'09), pp.235–242, 2009.

[7] Ariga, S., Nagahashi, K., Minami, M., Esaki, H., and Murai, J.,
"Performance Evaluation of Data Transmission using IPSec over IPv6 Networks,"
In Proc. INET'2000, July 2000.
http://www.isoc.org/inet2000/cdproceedings/1i/1i_1.htm

[8] Shue, C., Shin, Y., Gupta, M., and Choi, J.Y.,
"Analysis of IPSec overheads for VPN servers,"
In Proc. 1st IEEE ICNP Workshop on Secure Network Protocols, pp.25–30, 2005.

[9] Shue, C., Gupta, M., and Myers, S.A.,
"IPSec: Performance Analysis and Enhancements,"
In Proc. IEEE International Conference on Communications (ICC), Glasgow, Scotland, June 2007.

[10] Perez, J.A., Zarate, V., Montes, A., and Garcia C.,
"Quality of Service Analysis of IPSec VPNs for Voice and Video Traffic,"
In Proc. International Conference on Internet and Web Applications and Services/Advanced International Conference on Telecommunications (AICT-ICIW), pp.43–43, 2006.

[11] Hadjichristophi, G.C., Davis IV, N.J., and Midkiff, S.F.,
"IPSec overhead in wireline and wireless networks for web and email applications,"
In Proc. 22nd IEEE IPCCC, April 2003.

[12] Meenakshi, S.P. and Raghavan, S.V.,
"Impact of IPSec Overhead on Web Application Servers,"
In Proc. International Conference on Advanced Computing and Communications (AdCom), pp.652–657, 2006.

[13] Liu, L. and Gao, W.,
"Building IPsec VPN in IPv6 Based on Openswan,"
In Proc. IFIP International Conference on Network and Parallel Computing Workshops (NPC), pp.784–787, 2007.

[14] Zeadally, S. and Raicu, I.,
"Evaluating IPv4 to IPv6 transition mechanisms,"
In Proc. 10th International Conference on Telecommunications (ICT), pp.1091–1098, 2003.

[15] Mujinga M., Muyingi, H., and Rao, G.S.V.R.K.,
"IPSec overhead analysis in dual stack IPv4/IPv6 transition mechanisms,"
In Proc. ICACT, 2006.

[16] Yasinovskyy, R., Wijesinha, A., Karne, R., and Khaksari G.,
"A Comparison of VoIP Performance on IPv6 and IPv4 Networks,"
In Proc. 7th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA'09), pp.603–609, 2009.

[17] Hoeher, T., Petraschek, M., Tomic, S., Hirschbichler, M.,
"Evaluating Performance Characteristics of SIP over IPv6,"
Journal of Networks, Vol. 2, No. 4, pp.40–50, 2007.

[18] Bokor, L., Kanizsai, Z., and Jeney, G.,
"Performance Evaluation of key IMS operations over IPv6-capable 3G UMTS networks,"
In Proc. 9th International Conf. on Networks, 2010.

[19] Baugher, M., McGrew, D., Naslund M., Carrara E., and Norrman K.,
"The Secure Real-time Transport Protocol (SRTP),"
RFC 3711, March 2004.

[20] Linphone,
http://www.Linphone.org/

[21] The Multi-Generator (MGEN) Version 4.2,
http://pf.itd.nrl.navy.mil/mgen/mgen.html

[22] Openswan,
http://www.Openswan.org/

[23] strongSwan,
http://www.Strongswan.org/

[24] Wireshark,
http://www.Wireshark.org/

[25] OpenSER SIP server (now OpenSIPS),
http://www.OpenSER.org/

Nature inspired self-healing model for SIP-based services

ZORAN RUSINOVIC
Ericsson Nikola Tesla, Zagreb, Croatia
zoran.rusinovic@ericsson.com

NIKOLA BOGUNOVIC
Faculty of Computing and Electrical Engineering, University of Zagreb, Croatia
nikola.bogunovic@fer.hr

Keywords: autonomic computing, self-healing, SIP, nature inspired computing

The performance characteristics of Session Initiation Protocol (SIP) servers determine user-perceived quality of the services supported by SIP networks. SIP servers therefore must be able to provide service with appropriate reliability.
We present the self-healing SIP model capable of recognizing and restarting failed SIP services without losing active SIP dialogs. Novel approach to an evaluation of the SIP server healthiness has been presented that enables rapid problem detection and consequently quick recovery. Tests show that a proposed model exhibits very promising results with respect to number of successful SIP requests during SIP server operation.

1. Introduction

Session Initiation Protocol (SIP) is a controlling protocol for initiating, managing and terminating IP-based multimedia services across packet networks. In addition to being the controlling protocol between different nodes in IP Multimedia Subsystem (IMS) network, SIP provides building blocks for new media-blending applications and is used for enterprise multimedia applications, multimedia sessions, instant messaging and gaming.

SIP is standardized by the Internet Engineering Task Force (IETF), and has been adopted by 3rd Generation Partnership Project (3GPP) and 3rd Generation Partnership Project 2 (3GPP2) for IMS. An example of the SIP infrastructure is shown in the *Fig. 1.*

We assume the reader to be familiar with SIP, and present here some SIP characteristics only to pinpoint important elements during SIP messages processing.

1.1 SIP characteristics

SIP is an application-layer signaling protocol that can initiate, modify, and terminate interactive multimedia sessions over IP between intelligent terminals. It shares lots of the features that made the HyperText Transfer protocol (HTTP) a success: it is a clear text client/server protocol using Uniform Resource Locators (URL) for addressing. SIP goes beyond the scope of Voice over IP (VoIP) to provide building blocks for new enterprise communication applications:

1. Powerful addressing schemes (URLs) for user-centric services.

2. Features and media negotiation for improved plug-and-play, easily upgraded media-blending applications and terminals.

3. Seamless integration with existing enterprise IP networks and applications: integration with network Domain Name Servers (DNS) and with the corporate directory using the Lightweight Directory Access Protocol (LDAP).

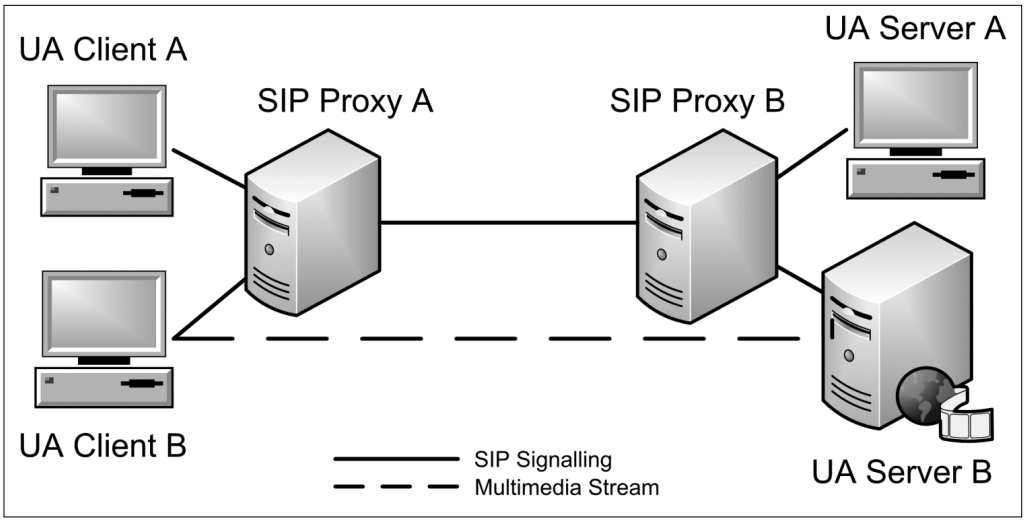


Figure 1. Example of SIP signaling through proxy. User Agent (UA) Server is a SIP service server

4. Built-in extensibility to other information technologies used in enterprises: e-mail, documents transported as Multipurpose Internet Mail Extension (MIME) attachments, etc.

5. Subscription/notification mechanism suitable for transporting user presence and terminal information.

In *Fig. 1* User Agent Client A (UAC A) tries to set up a session with User Agent Server B (UAS B) using SIP. It can be seen that SIP signaling and the multimedia data are completely separated, as SIP protocol is used only to set up or tear down multimedia session. As a result of the separation of SIP signaling traffic and the associated multimedia data stream it might happen that multimedia stream isn't available simply because of the problem with SIP stack and consequently with SIP session establishment, leading to data being unavailable. This will happen even in cases when the data for multimedia services itself is available at the (possibly different) service-providing servers, but because of the SIP signaling problems no connection can be established.

In Fig. 1 this would correspond to the SIP signaling path (marked with full line) between SIP Proxy A and SIP Proxy B being broken and multimedia stream path (marked with broken line) between UA Client A and UA Server B being fine. Although UA Server B itself works fine and despite the fact that no problems exist for providing multimedia stream, as a result of the SIP signaling problems no connection can be established and no service can be provided. This leads to a decreased user-perceived quality or even failure of possibly critical services.

2. Self-healing model for a single network element

When problems occur, traditional approaches for troubleshooting are based on the knowledge and experience of system administrators to discover problems and find ways to correct them. Unfortunately that approach, in addition to being laborious, is time-consuming and can lead to the SIP service being unavailable for a long period of time.

Self-healing is the property of any device or system to recognize that it is not operating correctly and to make any necessary adjustment needed to restore itself to normal operation. In a previous paper [2] the authors discussed the SIP service self-recovery model based on the monitoring of SIP messages exchange between SIP agents that was focused on the ability to efficiently utilize self-healing environment for SIP-based services within a single SIP network element. This article extends that work in such a way that it addresses cross-server healing between multiple network elements in the SIP-based networks. Together with the self-protecting SIP stack capability described in [4] we believe this to be a step towards an autonomic environment for SIP-based services.

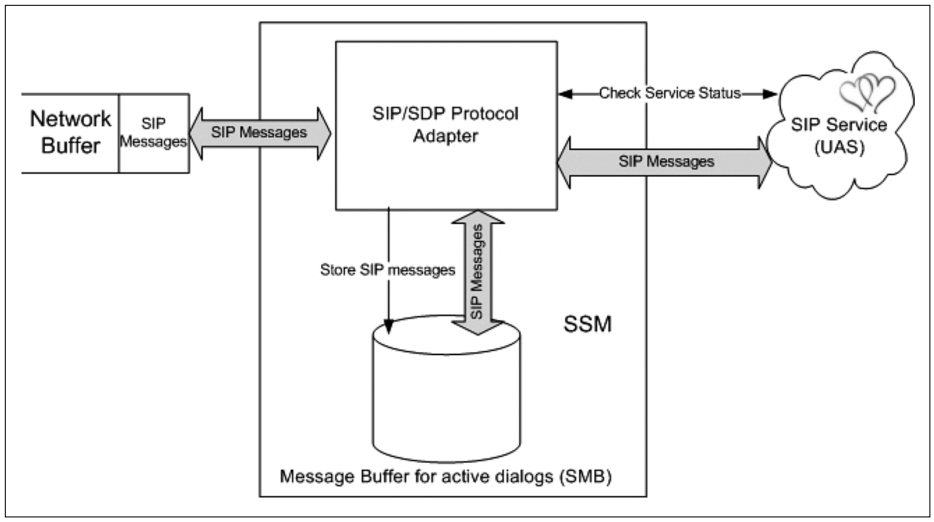
Our approach is based on the heartbeat monitoring with a purpose to detect whether monitored services are working or not. Heartbeat monitoring is an approach that can be seen as type of environment awareness since it provides awareness about the health status of system parts [5]. When the SIP Service fails, its peer SIP User Agent (i.e. peer SIP node) will detect this by expiration of timers defined as the part of SIP protocol, however for the failed service to continue its operation local monitor is needed that can detect that the SIP service has failed and that will try to restart it. The following sections will describe self-healing model based on the Windows Server 2008 operating system, but most conclusions therein can be applied to other operating systems as well.

2.1 SIP Service Monitor (SSM)

Fig. 2 depicts high-level relationships among the framework's main components. SIP services Monitor (SSM) is the component that implements monitoring and recovering for SIP services. If the service is recognized as failed it can be terminated and restarted. In the Windows environment the common approach to detect if the application is blocked is to send WM_NULL message to the suspected application by using the API call SendMessageTimeout().

The WM_NULL message performs no operations, and the recipient will ignore this message, however if the target application is non responsive the API call TerminateProcess() can be used to kill the hung instance of the application. This method is used, for example by the Windows Task Manager to recognize not responding applications and is also used as a base for some auto-

Figure 2.
Relationships among the main components of the SIP self-healing model



nomic self-healing tools frameworks [6]. There are, however, few problems with this approach that makes it inapplicable for SIP self-healing framework (as well as for many other cases).

In Windows environment there are two kinds of threads:

1. User interface threads that create its window (which can be invisible) and have GetMessage() loop which is used to respond to user actions.
2. Worker threads which do not create its own window and are used to do a compute intensive job in the background.

Windows applications usually have a single thread used for all user interface components which creates one or more worker threads. User interface thread typically runs with higher priority than worker threads, so that user interface is responsive to the user while the working thread is doing a background job. Typical example for this are SIP applications which usually create a worker thread for parsing SIP messages and running SIP protocol state machines in the background during the times when there is no user input. Being a preemptive OS, upon reception of the WM_NULL message Windows will suspend the lower-priority worker thread (if still running) and assign the CPU to the higher-priority user interface thread that responds to SendMessage Timeout(). For those reasons if it happens that SIP message parsing or state machine handling fails, using the above described approach, SIP application will appear healthy despite the fact that there is a problem with SIP protocol handling thread, and that the only healthy part of an application is an user interface.

Another problem lies in the fact that this method is not applicable to applications running as Windows services which by default don't have message loops and run in different desktop. This is because services run in their own security context and not the context of the user that is logged in. Therefore services also run in their own virtual desktop which means that they cannot display any blocking user interface nor can they interact with the user (since normally hooks are only global within a single desktop).

In this work we propose another approach based on monitoring applications ability to process SIP protocol messages. In addition to avoiding all the problems related to the first method, this approach can monitor changes in processing times of messages and reflect more accurately applications behavior and its health.

2.2 SIP flow monitoring in SSM

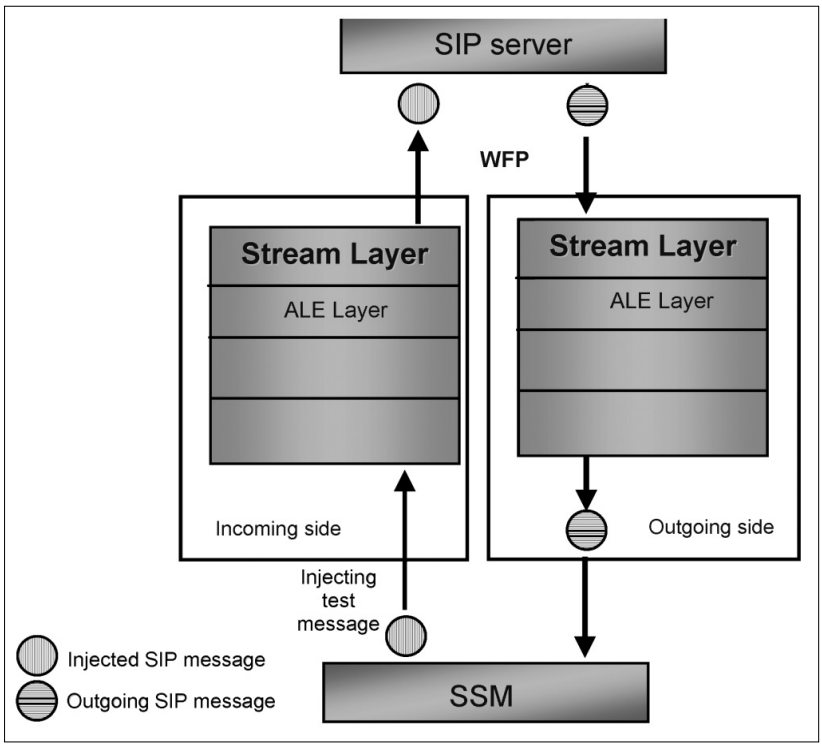
With the introduction of the new networking stack in Windows Vista and Windows Server 2008 with a single transport and framing layer, a new API known as a Windows filtering Platform (WFP) has been introduced. WFP allows, among others, to filter and modify packets, monitor and authorize connections at different layers in TCP/IP processing path.

By providing a simpler development platform, WFP is designed to replace previous packet filtering technologies such as Transport Driver Interface (TDI) filters, Network Driver Interface Specification (NDIS) filters, and Winsock Layered Service Providers (LSP).

As seen in *Fig. 3* on the incoming side all packets arriving at SIP port are monitored on the Stream Layer in the packet processing path using management API functions (FWPM) for the managing of the filter engine (FWPM_LAYER_STREAM_V4 and FWPM_LAYER_INBOUND_TRANSPORT_V4 for TCP and UDP respectively). Upon the reception of the SIP message a WFP callout function is called which determines if the received message is a request or a response (callout functions provide functionalities that extend the capabilities of the WFP, and can be registered at any layer). If a message does not begin with "SIP/" prefix, then the message is assumed to be a request which (even if incorrect one) requires a final response. SSM can operate in two modes, each providing different levels of recovery. In the basic mode WFP is used to monitor outgoing SIP traffic. When there is no outgoing traffic for a predefined minimum time period, SSM will use WFP to inject a testing SIP INVITE message on receive path, consisting only of mandatory headers excluding Call-Id header. If SIP UA server is responsive it should reply with "400 Bad Request" response, which will be caught by SSM at the SIP UA server's outgoing queue.

To monitor health of the SIP UA server more thoroughly, processing times of the injected messages can

Figure 3.
SSM operates on the WFP Stream layer



be stored and compared (since processing time should be in a microsecond range this can be done using timestamp counter in a CPU). If despite this testing SIP INVITE message being injected no outgoing SIP message is generated from the SIP UA server, SSM will conclude that server has failed and will try to restart it. Otherwise SIP UA server is assumed to work correctly, and the generated response is simply dropped in SSM.

In the advance mode in addition to the basic operation, each incoming SIP messages are cloned into the SSM Message Buffer (SMB) together with the hash value of the corresponding dialog as seen in *Fig. 4.* During the SIP UA server restart SSM is used to listen on the server's socket and to store all incoming requests into the SMB. Dialog hash value (to be used for comparing dialogs by numbers) is calculated using the 64-bit FNV1 algorithm using Call-Id value, remote tag and local tag of the SIP request as an input. Since initial INVITE messages do not have remote tag (which is provided by the SIP UA server itself in the response) the hash value of such request is calculated after the response has been generated. SIP requests are held in the SMB for the whole duration of the dialog, and are deleted from there only after the final response on BYE request has been sent by the SIP UA server. If sometime during the operation SSM detects problems with SIP UA server, in addition to restarting the server, SIP messages from SMB are used to re-initialize the dialog and transaction states of the server into states server had before restart. During this re-initialization phase initial INVITE requests from SMB are injected into the incoming buffer. Callout function is used to detect the new local tag value from the first non-100 provisional response generated by the SIP UA server. If it differs from the old one (from the pre-restart phase), all other SIP requests of the same dialog are injected into the incoming buffer with this new local tag. In addition to this, all responses to injected requests are dropped by SSM, to prevent SIP UAC receiving responses to old requests.

Reason behind this is that remote SIP UAC will have dialog state machines driven by old tag values and the local SIP UA server has them driven by new local tag values. Therefore it is necessary to make an appropriate mapping for all new SIP requests.

During the advance mode of operation prior to buffered INVITE requests being injected into the incoming buffer, each SIP request from SMB is internally checked in the SSM. This is done in the SSM's SIP message recognizer, whose purpose is to prevent sending incorrect SIP message to SIP server. Recognizer is a piece of code that scans and parses the input and recognizes whether it is in the language of the grammar, but does not produce an abstract syntax tree or any other form of output that represents the contents of the input. Because of this intertwining of lexical analysis and parsing of SIP messages, instead of having separate scanner, integrated recognizer has been used. This approach has a number of advantages including discarding of the lexical disambiguation by means of the context in which a lexical token occurs. Consequently the possibly complex interface between scanner and parser is removed, and both lexical and syntax checking are integrated into a single analysis phase. This approach is sometimes called scannerless parsing; however this term is somewhat misleading since all of the characters are anyhow scanned from the input buffer. Implementation difference is that instead of having one scanning function that accumulates characters into tokens, there are multiple functions that can read characters and then try to match it against some grammar construct.

If the SIP message is recognized as a correct SIP message it is then injected into the incoming buffer, otherwise it is dropped in the SSM itself and the "400 Bad Request" response is sent to the peer UA directly from the SSM. This way we prevent SIP server from going into cyclic restarts if the particular SIP request is causing server restart.

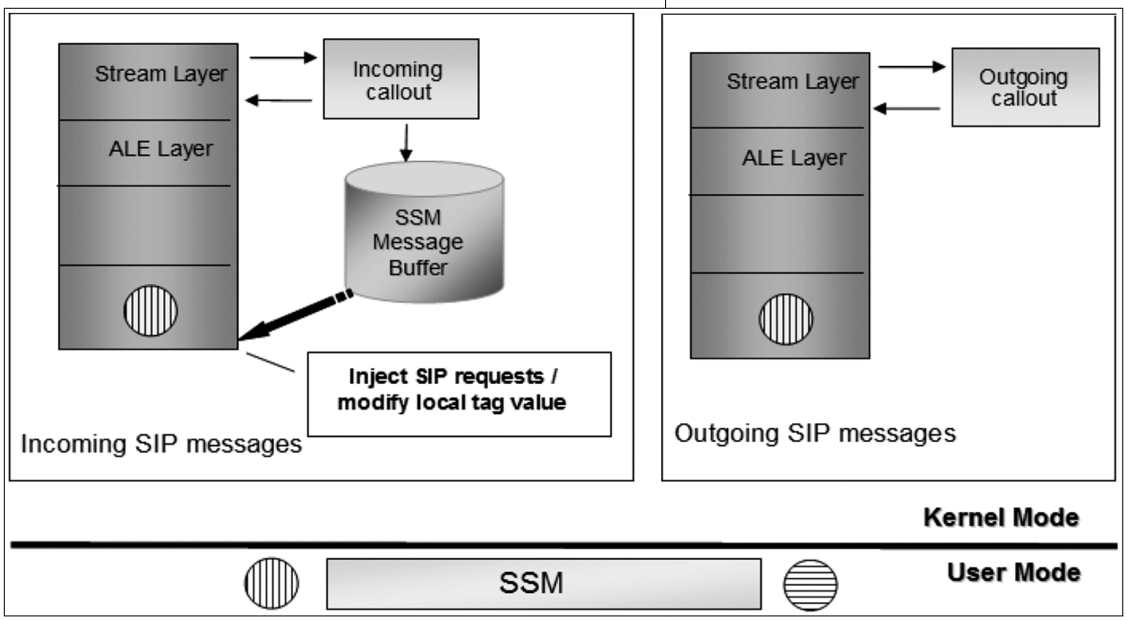


Figure 4. SSM operating in the advanced mode. Each incoming SIP request is cloned into the SSM Message Buffer.

The cost of this virtually fault free operation is increased processing time at both SIP message receiving and sending sides. Calculating dialog hash value requires not only message cloning but also finding required SIP message elements during message receiving and sending. This is done using WFP callout functions in Windows kernel mode.

3. Self-healing in the SIP-based networks

In the previous section we have described a model of self-healing within a single SIP-based network element. In this section we will extend that model to the SIP-based network architecture.

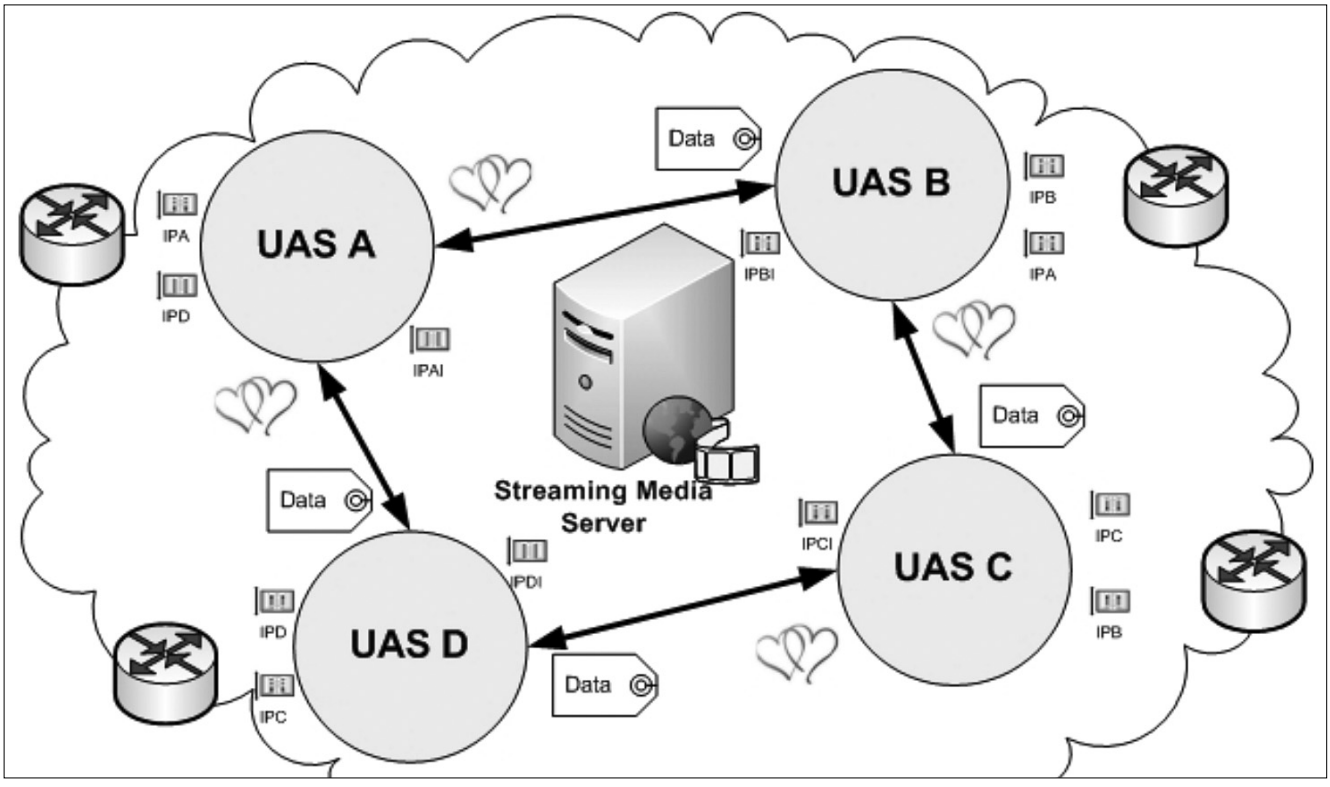
To solve the problem of service availability the current practice (as implemented by some providers) is to dedicate another SIP node as a backup of a primary SIP node in such a way that either each primary SIP node has its backup node (2*N model) or in a way that a set of SIP nodes is dedicated as backup set to relay the primary ones out of service. The problem with both of these approaches is that in addition to requiring large number of nodes to being deployed, the service becomes unavailable during the transition phase needed for the backup server to become operational. Moreover, neither of these approaches is capable of preserving SIP sessions that were in a set-up phase at the moment when a failure has happened [1].

We have based our approach loosely on the unilateral mode of anycast-based model for service continuity in IMS networks [1]. In this mode public interface of the UAS consists of two IP addresses, a primary and a secondary IP address. This secondary address is the primary address of another UAS. Problem with that model is that in case of node failure it does not preserve SIP sessions in progress. Consequently the service will be disrupted for all sessions in progress when a given UAS encounters failure. With our approach UAC perceives undisrupted level of service despite the UAS failure.

To ensure this, similar principle as described in the single network element scenario has been used. Unlike with single network element where the heartbeat monitor is implemented in the element itself, in this configuration the heartbeat monitor is implemented in the partner UAS, that is called guardian UAS (conversely the UAS being monitored is called guarded UAS). The configuration of such self-healing network is shown in the *Fig. 5.*

The network model in Fig. 5 is configured as an NK network with K=1, which means that every UAS sends heartbeats to only one partner UAS, namely the guardian one (in the Fig. 5. UAS B monitors heartbeats of UAS A, UAS C of UAS B, and so on). In addition to two public addresses each UAS has a dedicated inner interface (IPxI) that is used to convey parsed content of SIP messages, in the internal format, to its guardian node. Each UAS announces its primary and secondary address to

Figure 5. Configuration of the self-healing network model corresponds to a NK network (K=1). Heartbeat monitor is implemented in the guardian UAS. Each UAS has its own guardian UAS (in this case UAS B is a guardian UAS of UAS A, UAS C is a guardian UAS of UAS B and so on). Each UAS announces its address to router over OSFP.



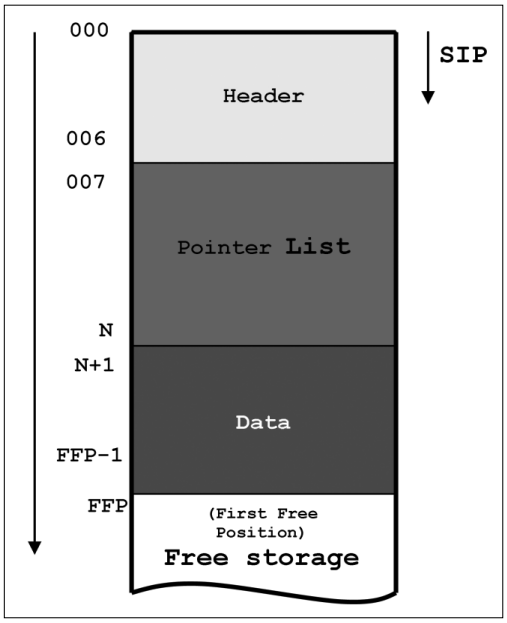


Figure 6.
Internal format used to convey parsed SIP messages between the guardian and guarded UAS. It consists of the Header part that conveys dialog and transaction hashes and Pointer lists that points to headers which are stored in "Data" part.

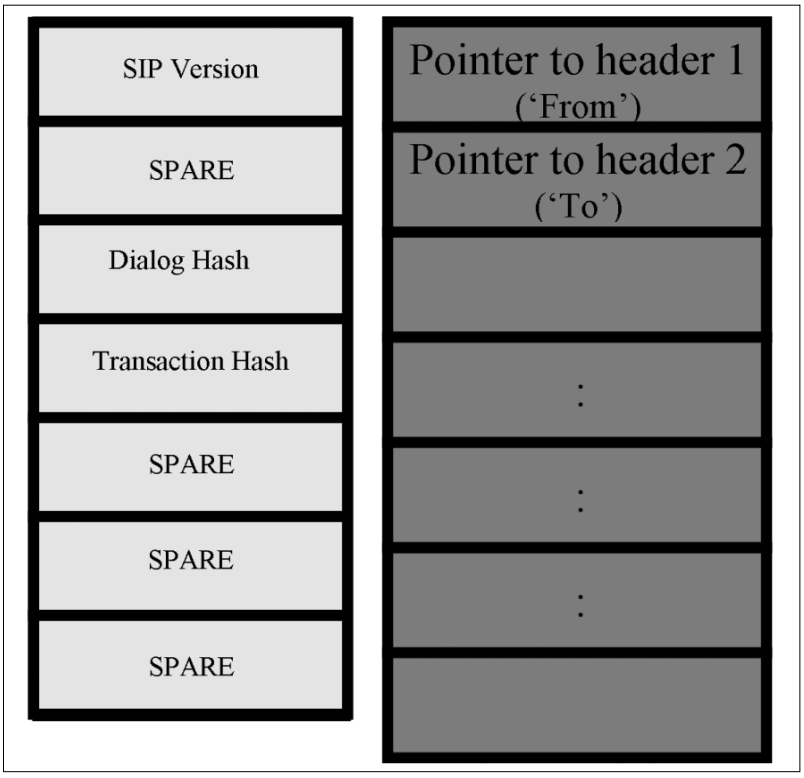


Figure 7. Header part of internal format (left) and Pointer list (right)

the router using Open Shortest Path First (OSPF) protocol (or another IGP protocol), so if the server dies, the router will remove it from an announcement.

3.1 Heartbeat monitoring in network

Heartbeat monitoring works as follows: when the UAS receives message from the UAC the message is parsed and converted into the internal format shown in the *Fig. 6* and *Fig. 7.* Internal format consists of the fixed header which conveys hashes of the particular dialog and transaction and a fixed number of pointers that point to the particular SIP message headers which are stored in the data part of the internal format. As a consequence of using an internal format a guardian node does not need to reanalyze the whole message but can directly use parsed data from the internal format in case that it must take over the function of its guarded node (this mechanism is similar to the one in MPLS where edge devices use labels instead of IP addresses to further forward them). The message is then sent to its guardian node in its internal format using nodes inner interface.

If the SIP method received in message is BYE or CANCEL guardian node will use dialog hash value to find and delete this session from the list of active sessions, otherwise it will simply store the received message in the SMB. Those messages will then be used to re-initialize dialog at UAS (in case of the failure of the guarded node) in the same ways as it is described for the single network element.

Previously described mechanism for heartbeat pacing, when no data is received from the guarded UAS for a certain minimum time period, is applied here as well. Guardian UAS will use SSM to send a testing SIP INVITE message to the guarded UAS, consisting only of mandatory headers, excluding Call-Id. If alive guarded UAS will response with "400 Bad Request" that will be sent to back to the guardian UAS and ignored. When the guarded UAS (e.g. UAS A in Fig. 5.) gets out of service secondary addresses of the guardian UAS (e.g. UAS B) are announced in IGP (Interior Gateway Protocol) and, as already described for a single network elements, SIP messages from SMB (of the guardian UAS) are used to re-initialize the dialog and transaction states of the UAS for active dialogs.

In this configuration UAC that were attached to the failed UAS will be dispatched to the UAB in a transparent way. For the period during which the failed UAS is out of service a guardian UAS might temporarily encounter a burst of SIP messages. To compensate that we propose the usage of simple cellular automata (CA) based model.

3.2 CA model of networks dynamic

We assume the reader to be familiar with Cellular Automata, and present here only some basic elements.

Cellular automata, firstly introduced by Ulam and Von Neumann [8], are a special class of finite automata that can be described by the 3-tuple of *Eq. (1).* They contain large numbers of simple identical components with only local interconnections.

$$A = (S, N, \delta) \tag{1}$$

In the above equation S is a nonempty set, called the state set, $N \subseteq \mathbf{Z}^2$ is the neighborhood, and $\delta : S^N \rightarrow S$ is the local transition rule.

A lattice of N identical finite-state machines (i.e. cells), each with an identical pattern of local connections to other cells for input and output, is called a cellular space. Each cell is denoted by an index i and its state at time t is denoted s_i^t (where $s_i^t \in S$). Cell i together with the cells to which cell i is connected is called the neighborhood η_i^t of the cell i. Local transition rule $\delta: S^N \rightarrow S$ gives the update state s_i^{t+1} for each cell i as a function of η_i^t. Typically CA works in a discrete manner. That is to say time goes step by step and a global clock provides an update signal for all cells.

The proposed models consists of a one-dimensional automata with three cells per each UAS and is similar to the Hodgepodge Machine of Gerhardt and Schuster which was used to simulate oscillating chemical reactions [9].

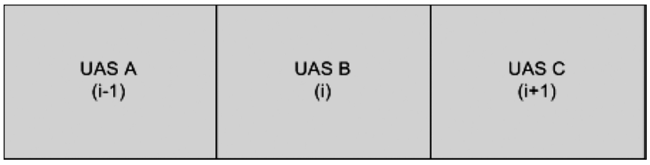


Figure 8.
One-dimensional neighborhood of the CA consisting of the UAS itself (UAS B in the Fig. 5) and its guardian server (UAS C in the Fig. 5) and guarded server (UAS A in Fig. 5)

Neighborhood of the CA consist of the UAS itself, and its guardian and guarded UAS, as displayed in the *Fig. 8*. The updating of cell sites is done asynchronously during the heartbeat monitoring at each heartbeat. Each cell can be in one of the five states:

- 0 = healthy
- 1 = infected
- 2 = ill with high load
- 3 = ill with overload
- 4 = death

Cell (UAS) is healthy if it receives traffic only on its primary IP address. If the secondary IP address is active as well we consider such UAS infected since it has to deal with excessive traffic of its guarded UAS. However in this state UAS handles this additional traffic in such a way that a regular traffic has not escalated into the high-load or overload traffic. In states 2 and 3 UAS has to handle excessive traffic but in such a way that this additional traffic is causing a high-load or overload, respectively (difference being that in high-load situation UAS receives more messages that it can process within a given period of time but it retains control of how to handle them, while in the overload situation messages are lost without control). Cell is dead if the corresponding UAS does not return heartbeats.

Formally we define this CA as follows:

$$\begin{gathered} A = (S, N, \delta) \text{ with } S = \{0,1,2,3,4\} \text{ and} \\ N = \{c_{i-1}, c_i, c_{i+1}\} \\ c_{i-1} = \text{guardian UAS} \\ c_i = \text{observed UAS} \\ c_{i+1} = \text{guarded UAS of the } c_i \end{gathered} \tag{2}$$

Transition rule is defined as follows:

$$\delta(c_i^t) = \begin{cases} 0 & if\ c_{i-1}^t = 0\ and\ c_{i+1}^t = 0 \\ 1 & if\ c_{i-1}^t \geq 1\ and\ c_{i+1}^t = 0 \\ 1 & if\ c_{i-1}^t = 4\ and\ c_{i+1}^t \neq 4 \\ c_i^t & otherwise \end{cases} \tag{3}$$

If all the UAS from the CA neighborhood at time t are healthy (meaning that they are in state 0) then at time t+1 they will remain in state 0 and no action is taken. However if at time t a guarded UAS is not healthy and provided that a guardian UAS itself is healthy then at time t+1 a guardian UAS will change its state into the state 1 and will take over part of the traffic of its guarded UAS (this transition takes place regardless of the state of a guardian UAS when the guarded UAS is dead and a guardian UAS is not). In all other cases the state of the guardian UAS remains the same.

This transition rule ensures that for situations where failing of the guarded UAS (e.g. UAS A in Fig. 5) causes excessive traffic at the guardian UAS (e.g. UAS B in Fig. 5) its guardian UAS (UAS C in Fig. 5) jumps in by announcing in IGP its secondary IP address and provisioning that some UAC traffic is dispatched to it. Naturally, if the excessive traffic persists too long because of (multiple) nodes failure, the risk of overload remains.

4. Testing and results

To test the efficiency of the described self-healing solution as well as the cost of SIP message preprocessing we have tested solution in our local lab environment. Test was run on an isolated Ethernet network using dual-core AMD Opteron processor running at 2.4 GHz with 8 GB of RAM for running SIP Service and SSM. From the SIP traffic generator SIP requests are sent according to a standard proxy 200 scenario. For testing self-healing within the network configuration as depicted in Fig. 5 has been set up.

Request intensity was 10, 50, 70 and 100 requests per second respectively, with each of these four different SIP loads running for half an hour. SIP UA server was modified to fail (enter an infinite loop) every 10 minutes. *Fig. 9* shows number of successfully handled requests without using SSM, using SSM in basic mode, and using SSM in advance mode. For the network self-healing testing a random UAS would fail every 10 minutes.

As seen in Fig. 9 using SSM even in the basic mode increases percentage of successful SIP requests to ~99.998 percent. The difference to 100% is lost on requests being processed or just sent from SIP UAC at the time of SIP UA server restart. In the advance mode, all the SIP requests being processed or sent from the SIP UAC were available locally in the SMB and were used to reinitialize all UA server state machines to the pre-restart baseline bringing successfulness to 100 percent.

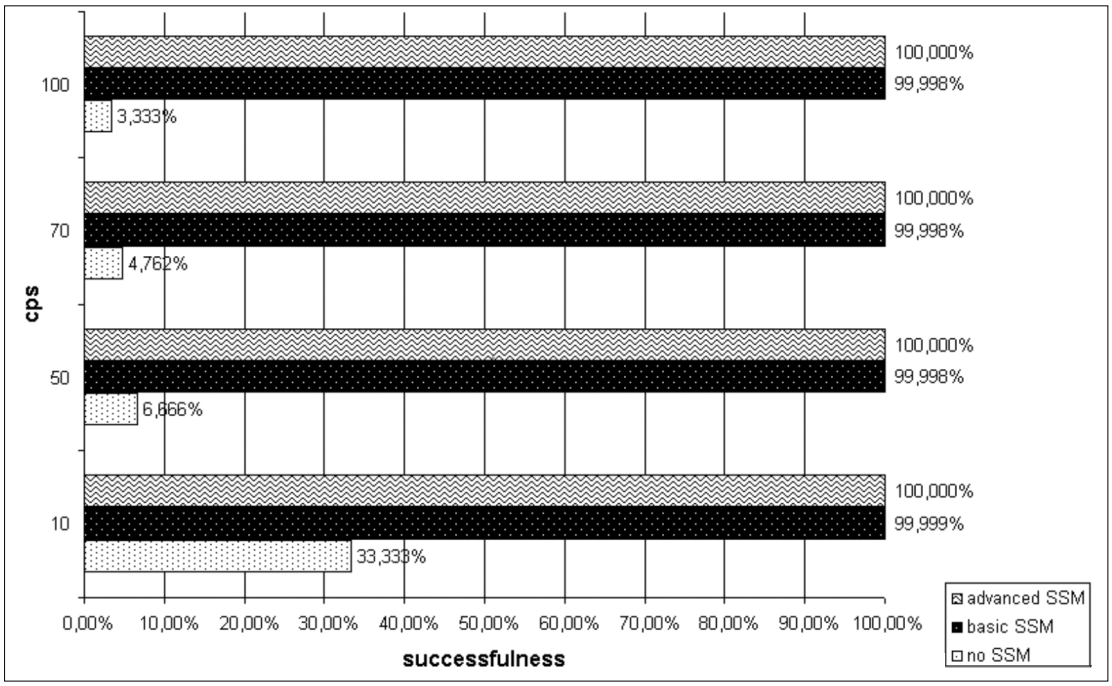


Figure 9. Results showing number of successfully processed SIP request without SSM, with SSM working in basic mode and with SSM working in advanced mode

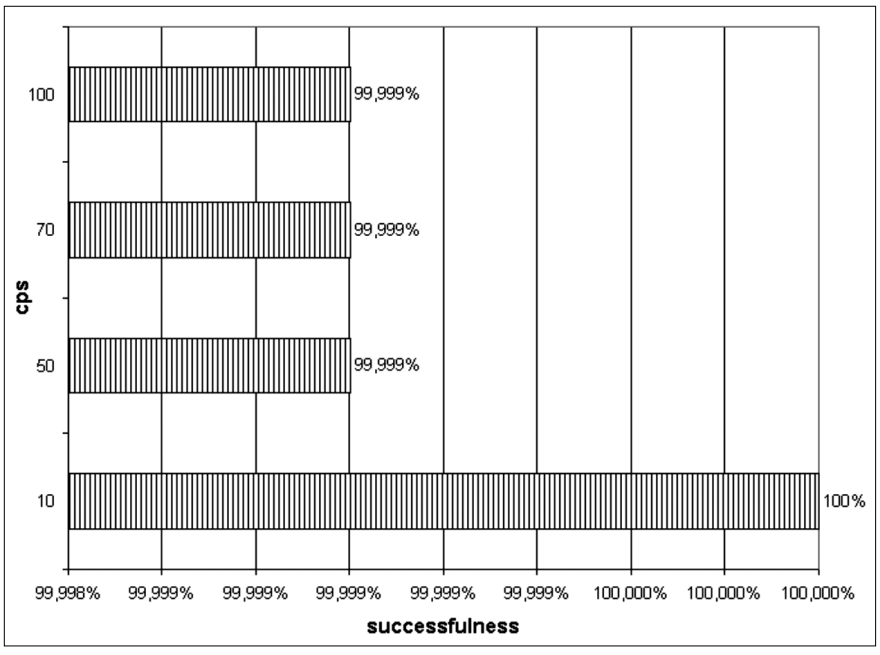
Results for the self-healing recovery in the network are shown in *Fig. 10.* In this case during high-load traffic this approach reaches five nines of successfulness. We attribute this to small portion of SIP messages being inevitably lost during route reconfiguration after UAS has failed.

The cost of the SSM operation is shown in the *Table 1.* Comparing the time needed to process a SIP message without SSM active and with SSM active it can be seen that in the basic mode which provides health-monitoring and service recovery (but no recovery of active dialogs and requests that were in progress during restarts) SSM adds ~10% of overhead to the processing times.

Message	Bytes	SDP	No SSM	Basic SSM	Advance SSM
INVITE	931	yes	8.4 μs	9,1 μs	11,8 μs
100 Trying	320	no	2.9 μs	3,1 μs	4,4 μs
183 Progress	854	yes	7.8 μs	8,4 μs	15,1 μs
PRACK	386	no	3.5 μs	3,9 μs	5,4 μs
200 OK	330	no	3.0 μs	3,4 μs	4,3 μs
UPDATE	731	yes	6.6 μs	6.9 μs	10.1 μs
200 OK	721	yes	6.6 μs	6.8 μs	10,2 μs
183 Progress	622	no	5.4 μs	6,1 μs	8,2 μs

Table 1. Messages decoding times

Figure 10. Results showing number of successfully processed SIP request in the network using the SSM



However in the advance mode which provides recovery of all dialogs the overhead is ~50% because of the additional time needed to copy each request into the SMB and to calculate corresponding hash value. Despite the somewhat increased processing times, numbers from Table 1 demonstrate the capability of our technique to handle SIP processing requirements of non-trivial size for real SIP-based systems. It is worth to mention that the demonstrated efficiency could further be improved by optimization that should be applied to memory handling routines.

Similar results are obtained for network recovery, shown in *Table 2,* which demonstrates that the app-

Message	Bytes	SDP	Processing time w/o heartbeat monitoring	Processing time with heartbeat monitoring
INVITE	931	yes	8.2 μs	12,1 μs
100 Trying	320	no	2.9 μs	4,6 μs
183 Progress	854	yes	7.8 μs	15,8 μs
PRACK	386	no	3.7 μs	15,7 μs
200 OK	330	no	3.2 μs	4,9 μs
UPDATE	731	yes	6.8 μs	10.9 μs
200 OK	721	yes	6.5 μs	10,9 μs
183 Progress	622	no	5.5 μs	9,1 μs

Table 2. Messages decoding times for network

roach with internal format distribution between two UAS is very efficient and introduces a very slight overhead.

Finally, the effectiveness of the CA model dynamics to compensate the traffic outbursts is shown in *Fig. 11.* UAS were dimensioned to handle 120 calls per second (CPS) in high-load traffic without messages being dropped. As seen in the picture after the guarded UAS fails without CA assistance almost 30% of messages were lost during 70 CPS traffic and almost 90% during 100 CPS traffic. However with the CA assistance the number of dropped messages decreases to ~3% and ~9% for 70 and 100 CPS respectively. Such results are understandable since with CA assistance excessive message traffic will be dispatched between several UAS in a transparent way.

5. Summary and conclusion

In this work we have presented an approach to self-healing SIP networks. New measure for evaluating SIP nodes health, based on a SIP requests processing capabilities, has been proposed. By experimental measurements it is shown that the proposed solution is very efficient in self-healing for both single network element, as well as in the SIP-based network and the obtained results are very promising. With the proposed approach service providers can ensure that outputs are not fuzzy and is always within service level agreements Further self-healing capabilities could evolve under this model including media server recovery to provide self-healing not only for signaling traffic, but for media traffic as well.

Authors

ZORAN RUSINOVIC received M.S. in Computer Science from the Faculty of Computing and Electrical Engineering of Zagreb University in 2004 and is currently a Ph.D. candidate at the Computer Science department. In 2004 he joined Ericsson Nikola Tesla company in Zagreb, Croatia, where he is currently a senior consultant for enterprise IT solutions. He has been engaged in many R&D projects in Croatia, Sweden and Germany, working in the area of network system engineering, network evolution, QoS/ TE managements and service architecture. For the past few years his focus has been IMS architecture and end-to-end QoS for multimedia delivery in SIP-based networks. His research interests include next-generation networking, biologically inspired computing, self-star and autonomic systems.

NIKOLA BOGUNOVIC graduated in 1967 from the Faculty of Electrical Engineering, University of Zagreb where he was awarded M.Sc. and Ph.D. degrees in 1971 and 1984, respectively. From 1968 until 1971 he was a research assistant at Institute Rudjer Boskovic, Zagreb. In late 1971, he was assigned a position of visiting research associate to UKAEA, Cullham Lab., England. Upon his return to Rudjer Boskovic in 1972 he was a principal investigator in various projects. At Vanderbilt University, USA. he was engaged as full time visiting associate professor in 1985. In 1990 he was assigned a position of co-associate professor to Faculty of Electrical Engineering and Computing, University of Zagreb. In 1996 he was appointed head of a Division of Electronics at Rudjer Boskovic, and occupied a position of scientific advisor in 1998. In 1999 he was assigned a position of full professor at Faculty of Electrical Engineering and Computing where he headed the Department of Intelligent Systems from 2005 to 2008. Professor Bogunovic is a full member of the Croatian Academy of Technical Sciences and has published over 100 scientific and professional papers. His scientific interests include computer based instrumentation systems, intelligent systems and methodologies for complex computer system design.

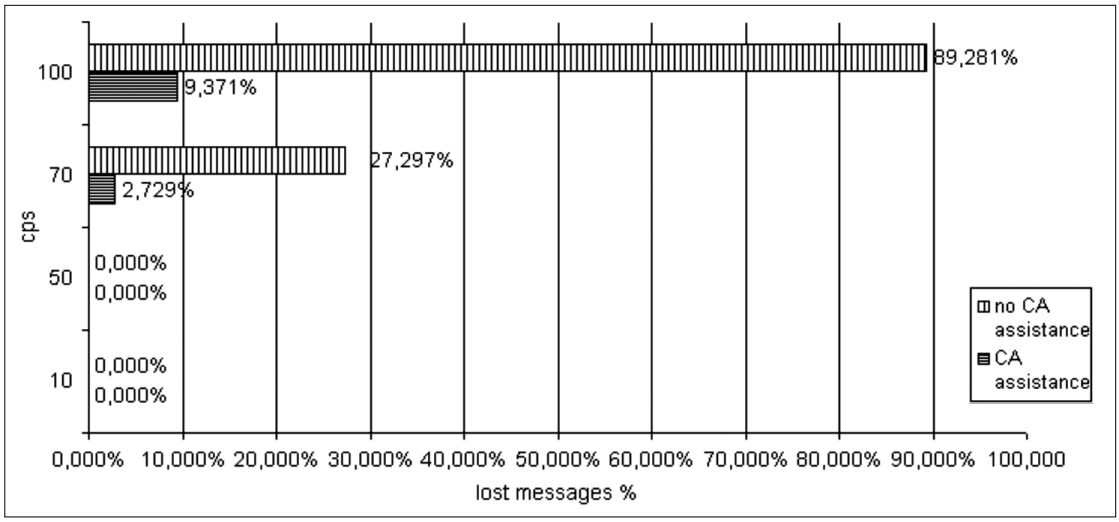


Figure 11. Percentage of lost messages with and without CA assistance during guarded node failure. UAS were dimensioned to be able handle 120 CPS in high-load traffic.

References

[1] Boucadair M.,
"Introducing Autonomous Behaviors into IMS-Based Architectures",
In: Autonomic Computing and Networking, M.K. Denko, L. Tianruo Yang, Yan Zhang (Eds.), Springer-Verlag, Berlin Heidelberg, pp.155–178, 2009.

[2] Rusinovic, Z., Bogunovic N.,
"Self-healing Model for SIP-Based Services",
In Proc. of the 10th Int. Conf. on Telecommunications (ConTEL 2009), pp.375–379, 2009.

[3] Rosenberg, J., Schulzrine, H., Camarillo, G.,
"SIP: Session Initiation Protocol", RFC 3261.

[4] Rusinovic, Z., Bogunovic N,
"Self-Protecting Session Initiation Protocol",
In: Lecture Notes in Artificial Intelligence, Vol. 5177, Proc. of the 12th International Conf. Knowledge-Based and Intelligent Information and Engineering Systems, I. Lovrek, R.J. Howlett, L.C. Jain, (Eds.), Springer-Verlag, Berlin Heidelberg, pp.717–724, 2008.

[5] Sterritt R., Bustard D.W.,
"Towards an Autonomic Computing Environment",
In Proc. of the 14th Int. Conf. on Database and Expert Systems Applications (DEXA), pp.699–703, 2003.

[6] Sterritt R., Chung S.,
"Personal Autonomic Computing Self-Healing Tool",
In Proc. of the 11th IEEE International Conference on the Engineering of Computer-Based Systems ECBS, pp.519–527, 2005.

[7] Kuthan, J.,
"Accelerating SIP,"
SIP 2002, Paris, France, 2002.
http://www.iptel.org/

[8] J.V. Neumann,
"The Theory of Self-Reproducing Automata",
A.W. Burks (Ed.), Univ. of Illinois Press, Urbana and London, 1966.

[9] Gerhardt, M., and Schuster, H.,
"A cellular automaton describing the formation of spatially ordered structures in chemical systems",
In: Physica D, Vol. 36, pp.209–221, 1989.

Space-efficient signaling scheme for IP prefix and realm information in Virtual Networks

ANTTI MÄKELÄ
Aalto University, Department of Communications and Networking, Espoo, Finland, antti.makela@tkk.fi

JOUNI KORHONEN
Nokia Siemens Networks, Helsinki, Finland jouni.korhonen@nsn.com

Keywords: compression, Mobile IP, virtual operator, prefix, realm

As residential Internet access has become increasingly commoditized, the incentives to lower costs in enterprise and similar networks with service level agreements have grown as well. Simply switching to a VPN provided over Internet brings cost savings but at the same time loses any guarantees of service level. We have devised a Mobile IP-based approach to virtualize networks without neglecting Quality of Service. As part of this specific approach, the signaling traffic levels go up as complexity of the network grows. To mitigate this issue, we have created a simple yet effective scheme to compress information on IPv4 network prefixes and realms. In this paper, we present analysis of our scheme's effectiveness and feasibility using both generated and real-world test material. We also consider the extensibility to IPv6 network prefixes.

1. Introduction

Various broadband Internet access technologies have become commoditized in recent years. However, while the commoditized access for public Internet works more than adequately for residential and individual purposes, corporations and other organizations require more customized service. Instead of basic Internet access, organizations often have needs for such services as private site-to-site connectivity and authentication of users – all provided with quality assurances. These, in turn, are not commodities, but may be prohibitively expensive especially for smaller entities.

A virtual service operator model attempts to provide benefits of both commoditized access and customized services. The model is based on the concept of leveraging the commoditized access and implementing an additional, virtual layer providing the required services without costly infrastructure investments. A very basic example is providing site-to-site connectivity with a Virtual Private Network [1] instead of more traditional dedicated line, Frame Relay, ATM or MPLS connections. The traditional and VPN-based site-to-site connectivity are illustrated in *Fig. 1.*

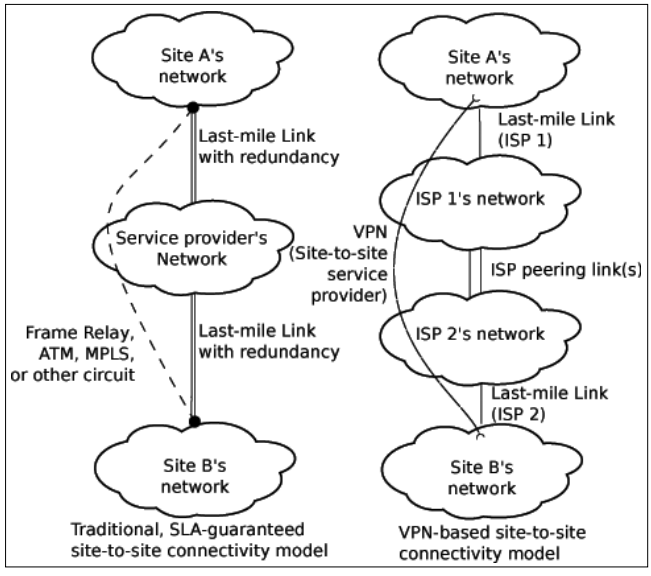
A problematic issue with VPNs as a technology has been lack of flexibility – capabilities to do dynamic reconfiguration of the virtual network when underlying topology changes: e.g., links go up and down, IP allocation changes, and similar issues. Furthermore, redundancy in case of outages is often limited.

The other alternative to VPNs – a traditional multi-homing setup – is also not possible with commodity connections, since in practice you cannot set up routing protocols through such connection. Even if you could, the convergence time of multi-homing can be in the order of minutes.

Allowing sites to have multiple inexpensive, dynamically reconfiguring, redundant connections, possibility to have the sites mobile, and providing all this with at least limited Quality of Service requires new approaches.

As one possible approach to address the issues mentioned earlier, we have proposed a solution based on Mobile IPv4 Network Mobility protocol [2]. The Mobile IPv4 Network Mobility takes care of near-instant switchovers in case of outages. Mobile IP can address the dynamic reconfiguration issue as well; Mobile IP is, by de-

Figure 1. Traditional operator and VPN models



sign, always actively striving for a functional point of attachment. The primary advantage of Mobile IP as a technology compared to e.g. routing protocols is extremely lightweight yet comprehensive signaling, allowing for the fast reaction times to changing conditions. All signaling transactions consist of single request and response messages, yet the message format facilitates complex information structures due to extensibility. Thus, reactions to topology changes can be as fast as single round-trip-time (RTT). Furthermore, Network Mobility allows routing information to be disseminated from a single, centralized point and does not require any sort of routing protocols to be set up for internal use.

Mobile IP Network Mobility has several immediately obvious challenges, especially when the number of sites and site-to-site connections increase. With the basic IPv4 Network Mobility, the Home Agent acts as a topological anchor for all data plane traffic to and from sites, creating a bottleneck at the center of the virtual topology. Therefore, our proposed extensions [3] to the basic Mobile IPv4 Network Mobility protocol add Route Optimization functionality. The intention is to offload most of traffic away from the Home Agent and allow direct and optimal paths between sites, thus conserving Home Agent's bandwidth and avoid the Home Agent becoming the bottleneck.

Route optimization in case of Mobile IPv4 introduces another operational challenge. Assuming that the number of sites and their connectivity is not fixed and varies over time, then a mechanism should exist for each site to learn about each other without extensive preconfiguration. To address this issue, our proposed extensions allow a Home Agent to distribute information about existing Mobile Routers to their peer routers. With this Home Agent assisted Route Optimization ("HAaRO") approach, Mobile Routers, either when registering to the Home Agent or when updating their mobility bindings, will learn of each other, thus being able to route traffic directly between them instead of via the central Home Agent.

Depending on the number of sites and the networks located behind the Mobile Routers the overhead of Mobile IPv4 signaling grows significantly. The signaling overhead may not be a show-stopper in general for the HAaRO approach but needs to be addressed as networks grow.

In order to reduce the signaling overhead when our proposed approach is used, the extensions include simple encoding algorithms for the route optimization information in the Mobile IPv4 signaling messages. There are two algorithms: one for compressing the subnet routing-related IP network prefix information, and one for compressing the administrative scoping-related realm information. In this paper we evaluate the usefulness of the compression algorithms and study their performance in various scenarios. The algorithms have been designed having low computation and memory footprint requirements in mind. Additionally, our algorithms do not expect past history information of previous messages and each message is self-contained compression-wise. This design sacrifices efficiency over simplicity to some degree. The compression of route optimization information is an optimization for the HAaRO approach, not the core functionality of the whole solution. Furthermore, the algorithms may have applications in other areas where IP topology information needs to be transmitted between nodes that have both limited processing power and bandwidth, allowing for a cost-effective means for communication.

This paper is based on earlier work [4] published in ConTEL 2009 conference. New additions are focusing on our efforts to extend the functionality to IPv6 network addresses. Thus we are also studying the algorithms' effectiveness on IPv6 network prefixes, although the original design was based on IPv4 network addressing. Besides that, we have made some clarifications to the text in general and our earlier results.

The rest of the paper is structured as follows. Section 2 has a more in-depth details on the Mobile IP-based Virtual Operator model. Section 3 provides an introduction to the two compression algorithms. Section 4 details our experimentation set-up and testing procedure, with corresponding results of our study. Finally, Section 5 concludes this paper with Section 6 containing some possibilities for future work.

2. Virtual operator model utilizing network mobility and route optimization

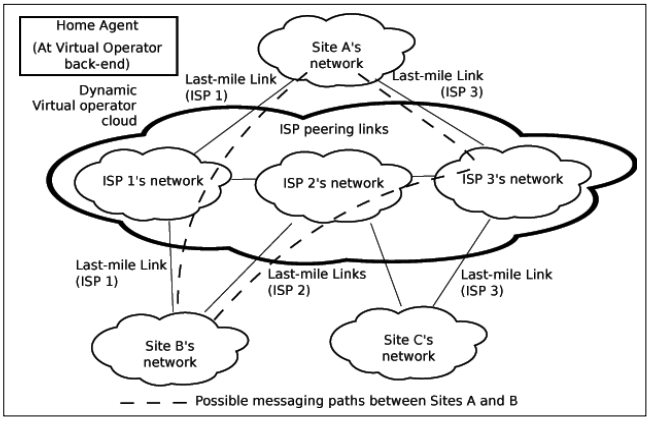
This paper frequently uses terms such as an *IPv4 prefix* [6], *IPv6 prefix* [7], and a *realm* [8].

The realm within this paper defines an administrative domain. Realms are named in similar fashion to Internet domain names, such as *"foo.example.com"*.

The IPv4 prefix corresponds to an IPv4 network address, in form of *a.b.c.d/yy,* where *yy* is the prefix length, in bits. In the same vein, the IPv6 prefix is an IPv6 network address, in form *aaaa:bbbb:cccc:dddd::/yy*. For more details on these concepts and how they relate to our algorithms see Section 3.

Fig. 2 shows an example virtual operator deployment, where the customer entity, such as an enterprise, has

Figure 2. Mobile IP Nemo/Route optimization model



three regional sites (sites A, B and C). Each site is connected to two regular ISPs with regular, commodity Internet connections. For example, Site A-to-B connection has two possible paths – one directly through ISP 1 and other utilizing ISP's 2 and 3. The gateway router at each site connecting to the Internet also acts as a Mobile Router. The regional site networks, in effect, become Mobile Networks. Unlike in the usual case where the mobility processing (handovers) occur when the network physically moves, in this case the mobility is triggered by changing conditions of the ISP's networks. As an example, assume that Site A's primary link is provided by ISP 1. When Site A's link to ISP 1 goes down, the inter-site connectivity remains up – the Mobile IP process simply changes the network's point of attachment to ISP 2 and connectivity is preserved. The scenario is more fully outlined in [5].

As noted earlier, utilizing Mobile IPv4 Network Mobility bestows heavy throughput requirements on the Home Agent and the virtual operator's own back-end network. The HAaRO approach [3] proposes an extensive use of route optimization, where Mobile Routers at customer sites would exchange traffic directly as much as possible without routing traffic via the Home Agent. The HAaRO approach also attempts to solve route optimization related management issues, such as discovering peer Mobile Routers, by having the Home Agent distribute prefix and realm information to Mobile Routers piggybacked in the Mobile IPv4 registration signaling. In this way the virtual operator can provide almost self-organizing composition of dynamically changing mobile networks without extensive configuration management. One area worth mentioning is that full convergence is not required; Triggering of Route Optimization can start at any Mobile Router. Non-optimized traffic is simply forwarded via the Home Agent as a fallback measure. However, assuming that all Mobile Routers have uniform Mobile IPv4 registration refresh time $refresh_{MIP4}$, it is possible to totally distribute the changes in the virtual network approximately within $refresh_{MIP4}$ time.

In IP communication protocols, a very common model is to separate control plane functionality from data plane. Furthermore, the signaling messages that are transmitted on the control plane typically have several constraints. Typical constraint is preference for low bandwidth. Another common constraint is a signaling format requiring all information to be transmitted within single Packet Data Unit (PDU). This is especially the case in request-response type protocols where each request warrants only a single response.

A large enterprise might have tens or hundreds of regional sites with a varying number of IP subnets due to incremental nature of deploying networks. A Home Agent distributing large amount of customer site prefix and realm information easily increases the PDU size beyond what is considered reasonable. While the transport protocol, UDP, allows PDU sizes up to 64 kilobytes, the underlying IP layer is responsible for fragmentation which is typically not desired or at least should be kept to a minimum.

It becomes evident that the growth of the signaling message size must be addressed. This paper evaluates the usefulness of the compression algorithms we have developed for reducing the signaling message size and studies their performance in various scenarios.

3. Space efficient encoding of prefix and realm information

A) Introduction

As stated in Section 2, IP network prefixes correspond to IP network address, in form of *a.b.c.d/yy* (IPv4) or *aaaa:bbbb:cccc:dddd::/yy* (IPv6).

An IP address is divided into *network* portion and *host* portion. In case of IPv4, the highest *yy* bits are considered the network portion identifying a subnet, while lower bits are for distinguishing individual hosts within the subnet. Networks are generally [9] allocated according to the number of expected hosts in the subnet: e.g. a network with prefix length of *yy*=28 consists of 16 host addresses. The first address of the network is the network itself, and last is reserved for broadcast, thus allowing for 14 true hosts. The network may be further subnetted, e.g. to two *yy*=29 networks, each accommodating up to six hosts. The network prefix length can be 0-30, network size /31 having no room for true hosts and /32 being a single IP address.

In case of IPv6, the lowest 64 bits of 128-bit IPv6 address are considered to be the host portion and highest 64 bits are the network portion – these allocations do not change. Thus, for purposes of network prefixes, only the first 64 bits have significance. The network prefix length can thus be 0-64, and primary consideration on prefix length is further subnetting. The prefix length has no effect on number of hosts that can be accommodated in the network.

The compression algorithms for both prefix and realm compression had the following as guiding design goals:

- The algorithms are for generic optimization purposes only, not core functionalities of the HAaRO approach.
- The algorithms should have low computational requirements.
- The algorithms should have a minimal memory footprint as e.g., a Home Agent may be serving thousands of sites, thus memory consumption is an issue.
- The algorithm should not require maintaining state between messages i.e., each compressed message should be self-contained.

Given the above design goals we can already see that the compression efficiency is not going to be among best available algorithms.

There are existing approaches to address IPv4 network prefix compression, for example in the context of MANETs [10]. However, the design parameters of our proposals are different from MANET.

IPv6 has significantly more development. Recently 6LOWPAN [11] has standardized a frame format for the

transmission of IPv6 packets over low-power personal area IEEE 802.15.4 networks [13]. Since the IEEE 802.15.4 maximum physical layer packet size is only 127 octets with 25 octets frame overhead that leaves only 102 octets for the media access control layer. Therefore 6LOWPAN has gone to extreme on defining a space efficient encoding of IPv6 and transport layer headers [12]. From the 6LOWPAN we could actually take advantage of its novel approaches for compressing IPv6 addresses. First, add a specific handling for well-known IPv6 prefixes effectively replacing the static known part up to the first 64 bits of the IPv6 prefix with one bit or maximum two-three more bits. Second, use of *contexts*. In our concept, a *context* is actually close to *Master Prefix,* which is further discussed in Section 3.B. The difference is that in our algorithm the context size has been one, whereas in 6LOWPAN there can be up to 16 contexts.

Besides 6LOWPAN, the IPv6 specification [7] includes zero compression for text formatted IPv6 addresses. However, the zero-compression serves only notational purposes of simplifying human-readable addresses and has no direct significance for our purposes.

B) Prefix compression Algorithm

1. IPv4 prefix compression

The prefix compression algorithm is fully presented in [3], Section 4-1. The basis for the algorithm's design is the assumption that transmitted prefixes will be relatively close to one another – in best case, sequential, (e.g. 1.1.1.0/24, 1.1.2.0/24,...). The assumption stems from common design where an organization is assigned a single IP block which is then subnetted and distributed amongst organization's network sites and the use case with Mobile Routers.

Other design parameters were simplicity in the common case of a single prefix per Mobile Router and possibility to extend the functionality later. One such extension is the IPv6 prefix compression illustrated in next section.

The presented prefix compression algorithm does not define compressor/decompressor implementation; only the data format. There may be more efficient compression implementations (See Section 6) compared to our approaches. The implementation-specific differences stem from the order the prefixes are processed by the algorithm.

The compression algorithm is based on concept of *Master* and *Delta* prefixes. At least one Master Prefix is sent before any Delta Prefixes. After the initial Master Prefix is sent, the following prefixes can be either Master or Delta.

Master prefixes are encoded asis, except tailing octets comprising wholly of zeroes are dropped. Thus e.g. 1.2.3.0/24 is encoded with 3 octets and 1.2.0.0/16 with 2 octets; However, 1.2.0.0/20 is also encoded with 3 octets since prefix length of /20 extends to the third octet, even though the contents of the third octet is zero.

Delta prefixes are always encoded as a single octet; the 8 least-significant bits of the prefix are included. Delta prefixes can be used if the difference between the delta prefix and the master prefix fits to the 8 encoded octets. If the Master Prefix is longer than the current prefix, the tailing bits are not compared due to them implicitly being zeroes. Thus even significantly different prefix may still be compressed if the length is shorter, e.g. if 1.2.3.4/28 is the master prefix, 1.6.0.0/16 can be encoded with a single byte.

Content	O	D	M	Plen	Data
Bits	1	1	1	5	0,8,16,24 or 32

Figure 3. Protocol structure concerning prefixes

When decoding a Delta-encoded prefix, the prefix is formulated as follows:

– Fill the right-hand side of the prefix with zeroes until Prefix Length (Plen) is reached.
– Fill the 8 bits, starting from Plen, from the received Delta prefix.
– Fill the remaining bits from Master Prefix.

The repeatable protocol structure defined in the [3] related to prefixes is presented in *Fig. 3.* Of the three flags, the 'O' concerns network topology and is of no significance to the compression. Flags 'D' and 'M' determine the contents of the Data field.

The data field can contain either a Mobile Router address (if M=1; in this case value of 'D' will be ignored) or a *Master* or *Delta* prefix. The algorithm always maintains the current master prefix; If a new Master prefix is received, the new one will replace the existing one.

Thus, the first time the structure is received, it always has M=1 and gives the Mobile Router the following prefixes are bound to; The next one always has D=0 and M=0, and provides the first Master Prefix. After that all combinations are permissible. Note that Delta prefixes will *always* apply to the current Master Prefix, even if the Mobile Router has changed. Thus, the value of M bit also has no significance to the compression.

Prefix Length is included in the separate 5-bit Plen field, allowing for values of 0-31.

2. Refinements concerning IPv6

The original header octet only allows for prefix length sizes of 0-31, which is too small for IPv6 prefixes. Thus the header octet has been redesigned and is slated to be included in the next version of [3]. The revised version, shown in *Fig. 4,* allows for more signals for Mobile Router purposes and also allows transmitting for IPv6 prefix lengths.

Content	O	D	Plen or indication	Data
Bits	1	1	6	0,8,16,24 or 32

Figure 4. Protocol structure concerning prefixes

We take advantage of the fact that besides IPv4 prefixes of length 0-31, we can handle the cases without prefix information, or with special prefix lengths, by overloading the Plen field. Whether the Plen field corresponds to prefix length or one of the special cases is now identified by a new I-bit (Indication), and old 'M'

and 'O' bits have been removed. The special cases, where I is set to 1, affect interpretation of Plen field as follows:

- When the address is a prefix of length /64, the Plen is set to zero. This effectively extends Plen field to 7 bits to allow for the /64.
- When the prefix length is IPv6-compatible representation of IPv4 prefixes [7], in effect a prefixes of size /96-/127, the Plen is set to a value from 32 to 63, again simply extending the Plen field to 7 bits. When the prefix is a Master Prefix, only 0-4 octets are encoded since the highest 96 bits are well-known.
- When the address is a Mobile Router, that is, corresponding to the old header's case M=1, the Plen is set to 1. As an extended Plen field, this would correspond to a prefix size /65, which is an illegal prefix length for both IPv4 and IPv6.
- When the address is a prefix of length /128, corresponding to a single host, the Plen field is set to 2. As with the case of Mobile Router, this would correspond to a prefix length /66, which is also considered illegal.
- Since the old 'O' field is related to Mobile Router instead of individual prefixes, it can also be indicated with a specific bit within Plen field if I=1. For our efficiency study purposes, we did not assign a specific value.
- Further signaling can be added as necessary.

Since the prefix size variation can be larger with IPv6, it might be beneficial to use more than one single delta byte. However, currently we can only indicate the presence of delta with a single bit, not delta sizes. Additional bits to indicate presence of multiple delta octets would also require an extra octet, thus the effectiveness is questionable.

C) Realm compression Algorithm

In the context of Mobile IP, the term *Realm* is used to signify an administrative domain. "Realm name" can be considered a Domain Name System's (DNS) domain name; Similar structure is evident. The compression algorithm for realm names is inspired by the original domain name compression presented in RFC 1035 [14]. However, our algorithm has been enhanced further to a full-fledged dictionary-based system with a simple, computationally lightweight encoder and decoder. The algorithm is designed to perform well in the common cases of relatively few (maximum 128) separate realm labels. A maximum length of a realm or a domain name is less than 256 octets. Assuming a naive implementation of the dictionary that makes separate copies of the stored strings and some indexing overhead, the maximum dictionary memory usage is around 40 kilobytes (i.e., 128* (255+*overhead*) bytes).

The algorithm works on a label level. The dictionary is updated dynamically with one or more labels from the input "realm name" to the algorithm. Every time a single label or a suffix of a "realm name" is not found in the dictionary, one or more labels are inserted to the dictionary and encoded into the output octet stream using the encoding shown in *Fig. 5.* Of every input "realm name" the suffixes that were not found in the dictionary are also inserted into the dictionary.

7	6	5	4	3	2	1	0	1-127 Octets
0	**Label Length**							**Label**

Figure 5. Bit and octet encoding of a new label

When a label or a suffix of a "realm name" is found in the dictionary, the index to the dictionary is encoded to the output octet stream as shown in *Fig. 6.* The dictionary holds maximum 128 strings. When the 129*th* string would be inserted into the dictionary, the dictionary gets reset and the new string will become the 1*st* string in the dictionary.

Figure 6. Bit and octet encoding of a dictionary index

7	6	5	4	3	2	1	0
1	**Index**						

Consider an example of the input realm *"foo.bar.example.com"*. The compression algorithm searches the dictionary with the following strings:

- *"foo.bar.example.com",*
- *"foo.bar.example",*
- *"foo.bar", and*
- *"foo"*.

The longest found string is encoded. If no matching string was found, the label *"foo"* is encoded and inserted into the dictionary. The process is repeated until the whole input realm has been encoded (i.e. *"bar.example.com"* would be the input realm for the second iteration).

The algorithm also keeps track of the longest suffix per input realm that only consist of encoded labels. Every time an index to the dictionary can be encoded the tracked longest suffix gets reset. Once the input realm has been processed, the longest tracked suffix is inserted into the dictionary.

For example, if *"bar.example.com"* is the longest tracked suffix of *"foo.bar.example.com"*, then the algorithm will insert *"bar.example.com"* and *"example.com"* into the dictionary. The realm compression algorithm implements a greedy approach of encoding matches. The search algorithm and the dictionary update function does not try to optimize updates and immediately encodes the first found match. Lazy evaluation algorithms are known to gain better compression than greedy ones [20].

The octet value 0x00 is used to terminate the encoded stream of input realm. The encoding of both label or index cannot output the value 0x00 and the Mobile IPv4 extensions defined for HAaRO make use of it in header encoding.

The realm compression algorithm has also other interesting properties. The compressor/decompressor state is actually the dictionary. This allows easy and efficient handling of the packet data. There is no need to keep past or future input data available for the algorithm, only the very short piece of data that needs to be com-

pressed or decompressed at time. Furthermore, the entity doing the compression may at any time switch off the compression algorithm without the decompressor noticing it or causing any encoding size penalty compared to completely uncompressed data. This is due the design of the whole encoding of the HAaRO messages. The uncompressed form of the realm information is as compact as the original realm octet string.

1. Realm-compression Algorithm as a basis for dictionary-based IPv6 prefix compression

Due to the limitation of single-byte-deltas, we wanted to study the feasibility of dictionary-based approach for compressing IPv6 prefixes. We adjusted our realm-compression algorithm to make it more suitable for prefix compression while still maintaining same low memory footprint requirements. The changes are based on the idea that single octet of an IPv6 address can now be considered a "label".

The encoding is changed so that in Fig. 5, the "label length" now corresponds to "number of labels following". Since label length is always one octet, there is no need to identify each previously unknown label separately. Furthermore, single octets are not stored in the dictionary as was the case with realms, rather, all sub-prefixes (in contrast to suffixes in realms) are stored instead. Also, instead of using a terminating octet `0x00`, the total length of prefix is encoded as-is before each prefix.

As an example, consider two IPv6 prefixes, 2001:0000:00a9::/55 and 2001:0000:00c0::/42. When encoding the first prefix, the dictionary is empty.

The encoding for the first prefix (besides prefix length) is the following: `0x07 0x20 0x01 0x00 0x00 0x00 0xA9 0x00`. This consists of number of labels following (7 octets, since all labels are unknown), and the prefix asis. After the prefix has been encoded, the following sequences are inserted into dictionary:

2001:0000:00a9:00
2001:0000:00a9
2001:0000:00
2001:0000
2001:00
2001

The shortest sub-prefix, 20, is not inserted, as there is no gain in encoding an octet with another octet.

The encoding of the second prefix is now `0x82 0x01 0xc0`, where `0x82` is reference to the third dictionary entry of "2001:0000:00". `0x01` corresponds to single octet that follows, which is encoded asis (`0xc0`). After encoding, the dictionary gets one additional entry: 2001:0000:00c0, as the previous sub-prefixes are already in dictionary and single octets are not stored.

4. Testing setup and results

A) Prefix compression

Two implementations of the prefix compression algorithm, one for IPv4 and one for IPv6, were written in C, utilizing standard Socket API. These consisted of requester and responder components, with the requester acting in the role of a Mobile Router and responder acting in a role of Home Agent. The responder initially reads in and processes a list of network prefixes associated with Mobile Router addresses, and then starts waiting for a request. The requester contacts the responder with a request message, and the responder returns a list of prefixes, compressed in accordance to the scheme. This followed standard Mobile IP processing, although true Mobile IP headers were not used.

The implementations offers the possibility to send data either uncompressed or compressed. The uncompressed option was used to obtain the baseline comparison for each test case, where neither delta-compression or removing tailing zeroes are used. The compressed option causes the responder to proceed as follows:

1. Read in all the prefixes and Mobile Router addresses associated with each prefix.
2. Group all prefixes managed by single Mobile Router together.
3. For each Mobile Router, sort all prefixes in order; The IP addresses are simply considered unsigned 32-bit integers, or in case of IPv6, unsigned 64-bit integers. Prefix lengths are not taken into account at this point.
4. Merge the prefixes from Mobile Routers back together in the input order.
5. Process prefixes in accordance to the compression algorithm:
   For each prefix, check whether the prefix can be encoded as a delta of previous prefix;
   If yes, encode as delta prefix (1 octet),
   if no, encode as new master prefix (0-8 octets, depending on the prefix length and IP protocol version). If Mobile Router changes, encode the new Mobile Router's address.
6. Save the result into a buffer, waiting to be sent to the requester.

In encoding, we used the formats presented in Section 3.B, the original format for IPv4 prefixes and revised version for IPv6 prefixes.

To come up with realistic test cases, the following considerations were taken into account:

– Intended usage and design:
  Prefixes are likely to be numerically close to each other, including the completely sequential case.
– The algorithm is designed to provide efficient compression while prefix length variations are small. When prefix lengths have greater variation, the delta may no longer fit into a single octet and a new master prefix has to be set, consuming space. This is especially a concern with IPv6.
– Number of prefixes may vary depending on organizations size.
– Number of Mobile Routers may vary depending on organizations size.

With this in mind, a prefix generator was implemented that can generate list of IPv4 network prefixes in *a.b.c.d/yy* format or IPv6 network prefixes in *aaaa:bbbb:*

cccc:dddd::/yy format, and associate each prefix with a Mobile Router address. This prefix list could then be fed to the responder component of the implementation, allowing for a quick testing of various cases. Generator can provide selectable number of prefixes of various lengths around specified base prefix length (*/yy*) with selectable degree of randomness induced both to prefix length and prefix sequentialness. The random number generator used was standard ISO C rand() function which provides 32-bit random integers with uniform distribution. For IPv6 prefixes, two 32-bit integers were concatenated to form a single 64-bit prefix. The random number generator was initialized using seed derived from system time before each new batch of prefixes.

In the cases where randomness plays a part – e.g. when testing the performance where the compressed prefixes are not completely sequential or prefix lengths vary – each test case was conducted 10 times and results averaged. Minimum and maximum of each case was also recorded to check for outliers.

Besides these generated test cases, a more realistic source for real-world network prefixes was used for IPv4 – the global Internet BGP routing table. In case of IPv4, picking an A-class network of appropriate size to correspond for each generated test set allowed a comparison of the generator to real-world subnetting schemes.

The test matrix chosen includes 10, 500 and 5000 prefixes, in sequential, near-sequential or totally random order, shared equally between either 1 or 10 Mobile Routers. Furthermore, each case had the prefix length either as static (/24), or varying to a smaller (<8 bits) or greater (<16 bits) degree. The "totally random" case attempted to utilize entire (IPv4 or IPv6) address space, to create appearance of totally unrelated network prefixes. "Near-sequential" simply means occasionally skipping the next network in sequence, and is our closest expectation of real-world use-case.

In addition, three test cases for IPv4 were taken from global BGP routing Table [15]. The routing table was based on data on 28th of August, 2008. The chosen /8 networks and their sizes are listed in *Table 1*.

It should be noted that when conducting IPv6 tests, lessons learned from IPv4 testing could already be taken into account, and some IPv6-specific tests concerning the increased address space were added. These differences are further detailed in Section 4.G.

B) IPv4 network prefix compression results

Before measuring compression efficiency, a baseline had to be established. The baseline for IPv4 prefix compression – simply not compressing data at all – can be seen in *Table 2*.

The non-compressed size for each case shows that as number of prefixes grow, the number of Mobile Routers has less and less proportional effect to the overall size – each new Mobile Router adds a static five octets (1 header octet, 4 octets for address) to the data. Thus, for further observations, the case with 10 MRs is not significantly different from the case with a single MR as an additional MRs simply increases data size by constant 5 octets each.

The effect of sequentialness to the compression is shown in *Table 3*, based on the case where prefix lengths do not vary. The compression factors in the table are based on the average compression in each of the randomized cases. As can be seen, the performance is best when prefixes are sequential and worst when random. The real-world networking data, shown in the BGP column, appears to reflect the sequential/near-sequential cases more than the completely random case, which is encouraging.

As mentioned, all test cases with random elements were ran 10 times. Under no circumstances did the individual test runs significantly differ from the average.

Table 1. Real-world BGP routing table entries

Network	Number of Prefixes	Purpose
56.0.0.0/8	10	Comparison for 10 prefixes
129.0.0.0/8	497	Comparison for 500 prefixes
72.0.0.0/8	4534	Comparison for 5000 prefixes
64.0.0.0/8	5907	Comparison for 5000 prefixes

Table 2. IPv4 Network Prefix compression test cases, uncompressed data sizes

Prefixes/MRs	Uncompressed size
10/1	55
10/10	100
500/1	2505
500/10	2550
5000/1	25005
5000/10	25050

The results in *Table 3* are based on the case where the prefixes are all of same length; In this case, /24. The effect of varying the prefix lengths are shown in *Table 4*. The values in percentages show compressed data length compared to uncompressed data length with same number of prefixes. The effect of adding varying prefix lengths to a small degree (cases with /8, prefix lengths from /20 to /27) causes what was expected compared to static

Table 3. Effect of IPv4 prefix sequence to compression. Percentages in parenthesis are the remaining data compared to uncompressed data.

Count	Sequential	Near-seq	Random	BGP
10	27 (49%)	27 (49%)	45 (88%)	(49%)
500	1009 (40%)	1017 (41%)	1973 (79%)	(45%)
5000	10045 (40%)	10124 (40%)	19697 (79%)	(41/42%)

Table 4. Effect of varying prefix lengths on compression. Case indicates Number of Prefixes / Maximum variation.

Case	Uncompr	Sequential	Near-seq	Random
10/8	55	27 (49%)	29 (53%)	48 (87%)
500/8	2505	1027 (41%)	1225 (49%)	1973 (79%)
5000/8	25005	10241 (41%)	12261 (49%)	18934 (76%)
10/16	55	27 (50%)	32 (59%)	44 (80%)
500/16	2505	1030 (41%)	1379 (55%)	1793 (72%)
5000/16	25005	10237 (41%)	13643 (55%)	16226 (65%)

case – the compression factor decreases. Remaining data length is around 50% of the original compared to 40% in the non-varying case.

Increasing the variability to a larger degree (cases /16, prefix lengths from /16 to /31) causes more interesting results. In the sequential case, results are identical with the smaller variation, which was expected. With near-sequential cases, the resulting data size is slightly larger but not significantly, which is expected. However, in the totally random case, the compressed data size is actually smaller than with smaller variance.

The reason for the better compression ratio in totally random case, especially with the full 5000 prefix set, might stem from the limits of IPv4 address space. As our definition of "random" is filling the entire 32-bit address space, and the data set includes short prefixes, the individual prefixes in the data set may not be so unrelated after all – a very short prefix may include significant portion of the entire address space.

However, since the effect also appears with cases of 500 and 10 prefixes, the more likely reason is the algorithm's ability to drop tailing zeroes and make the delta to higher-end bits of the master prefix. When prefix length variation is small, almost all prefixes become master prefixes. When prefix length variations are higher, the tailing zeroes effect kicks in; e.g. a /28 prefix followed by /16 one. The latter prefix quite likely has identical or near-identical most significant bits, and the tailing zeroes can be left out; Thus it becomes a delta-prefix. e.g. when Master Prefix being 1.12.23.16/28 the prefix 1.13.0.0/16 can be expressed as a delta, while 1.13.34.32/28 would be encoded as a new Master Prefix.

C) Realm compression

Our C++ implementation of the realm compressor/decompressor was tested against the test material containing 500 *Fully Qualified Domain Names* (FQDN) generated from the zone *"example.com"* (total size of 11188 octets). In reality a MIP4 message would contain in maximum few dozens of realms. The material was generated using a list of English dictionary words of 3 letters and higher, a subzone label (*aaaa...xxxx*) repeating 25 times and suffix *"example.com"*.

Finding a comparable real-world case to determine the accuracy of our generated examples is hard in this case, as the usage pattern is different from regular DNS. However, regular DNS deployment can still give some indication on the algorithms effectiveness on domain/realm names. Thus, we retrieved the all FQDNs from zone *"cs.tut.fi"*, consisting of 1285 entries and size of 26403 octets. A real HAaRO deployment is not intended to be used with so many different entries.

The ordering of realms in the input data was expected to have an impact on the compression efficiency. If realms that share common suffixes are close to each other, then the dictionary gets updated less frequently. Low number of updates to the dictionary implies fewer dictionary resets, which in turn was expected to have positive effect on compression. Each dictionary reset causes the compressor to lose its current context. It will take a while for the compressor to re-populate the dictionary and provide good compression again.

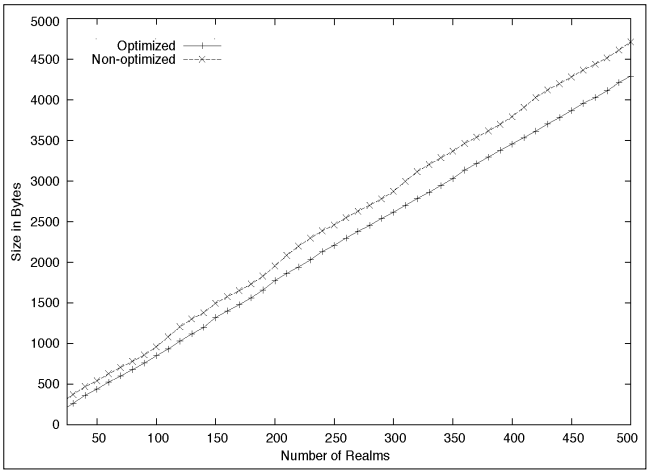
D) Realm compression results

As mentioned, our realm compression test material is a generated list of FQDNs (i.e., realms in this paper) from *"example.com"* domain and containing up to 500 realms consisting total of 11188 characters. In optimal order, with least number of dictionary resets, the algorithm compresses it to 4294 octets, compressing the data to 38% of the original size. The optimal order in this case being one where all realms within the same subzone (e.g. *"aaaa.example.com"*) are all presented in sequence.

With a harder, non-optimal order, the compressed data contains 4716 octets, leaving 42% of original length. As the dictionary has 128 entries, even non-optimal ordering does not cause significant efficiency losses until after first reset. The "non-optimal" order is one where the adjacent subzones are always different (e.g. *"bbbb.example. com"*, *"cccc.example.com"*, ...) until after 25 cycles.

The "bumps" caused by dictionary resets can be seen in *Fig. 7*.

Figure 7. Realm compression results



Approximately after every 100*th* input of non-optimally ordered realms, the dictionary gets reset. For optimally ordered realms the resets are less frequent and less visible due the dictionary learning faster the current context. In overall we are satisfied of the result given the simplicity of our algorithm.

Comparing the results from generated data and our chosen real-world example, the zone "cs.tut.fi" compressed from 26403 octets to 13783, leaving 52% of original data. Considering that the example is from a standard DNS deployment where host names vary significantly, slightly worse compression factor is to be expected.

E) Combined IPv4 prefix and realm compression

Most interesting results are expected when both prefix and realm compression are used together, the reason being that both prefix and realm input data compresses differently depending on the order the individual prefixes and realms are stored in the input data.

In the HAaRO specification, there is a dependency between prefixes and realms. Each prefix may be associated with zero or one realms. A realm cannot be encoded on its own without a prefix. However, prefixes can be encoded without realms.

The optimal ordering of prefixes may not favor realms, and vice versa. Therefore, in the combined case we examine what is the impact of favoring either prefixes or realms optimal ordering during the compression. We also present results of unrealistic case where both prefixes and realms are ordered optimally for a comparison purposes (this is not supported by the current signaling message format although network design may take this into account).

F) Combined IPv4 prefix and realm compression results

Fig. 8 shows the results of combined prefix and realm compression. We can make two obvious conclusions. First, the compression has significant effect on the message size and given the simplicity of our algorithms, their use should be encouraged. Second, the ordering of the HAaRO information has an impact on the compression efficiency. We can see that ordering by a prefix will give the best results due the nature of our algorithms – the prefix compression algorithm is a simple delta coding, which effectively only has a history of the one immediately preceding prefix. Therefore it is more sensitive to sudden changes in the input data than the realm compression algorithm which has a history of up to 128 preceding labels and realms. Thus, it's better to optimize the ordering by the algorithm that is most sensitive to the input data and it that way gain the best overall result.

It would be possible to design the Mobile IPv4 message extensions in such way that optimal ordering of both prefixes and realms were possible. However, this would imply increase in extension header overhead, which would effectively void the slightly better gains in compression efficiency. Furthermore, the message extension handling would complicate as two distinct sets of data should be kept in memory until the decompression has completed in order to allow merging of prefix and realm information.

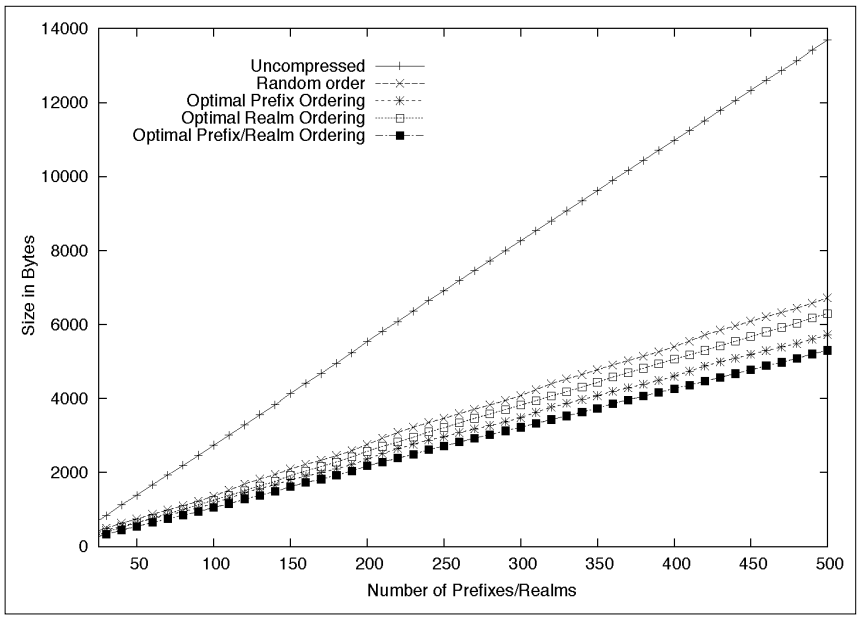
G) IPv6 prefix compression considerations

As stated in Section 3.B-2, IPv6 network prefixes have certain differences to IPv4 network prefixes, which affects test case planning. Furthermore, certain lessons from IPv4 tests could be taken into account.

Since one of IPv6's goals is simplifying routing tables and preventing fragmentation of IPv6 address space, the BGP table for IPv6 is not such a good source for real-world network prefix allocation, as routing strategies prefer to have only Provider Aggregated (PA) addresses to appear in global routing tables [16]. In IPv4, Provider Independent(PI) addresses assigned to organizations are widely used and thus appear in global BGP tables, but they are not considered a scalable solution for IPv6 multihoming [17]. The BGP table for IPv6 thus shows only aggregated address spaces allocated for service providers – the subnetting of the address space is not visible.

IPv6 deployment is still in early phases, thus no "best practices" based on operational experience have emerged. There are a number of IPv6 related recommendations and guidelines for various cellular networks produced for example by IETF. These recommendations mainly concern what is the link model of certain cellular technology, how it shows to the IPv6 stack, what is the recommended addressing for the link model and

Figure 8.
Results of combined prefix and realm compression



so on [18,19]. One particularly interesting point regarding addressing is that in certain cases, IETF has recommended the use of a separate /64 network prefix for each user (or rather mobile host), even if the network contains only a single host. This addressing model has actually been adopted by all major cellular macro networks such as WiMAX and 3GPP.

Also, in more traditional environments, the allocated prefix sizes for a site can be anything from /32 to /64, although the most common allocations appear to be /48, as recommended by the specifications. However, due to the static-sized "host portion", for optimal allocation the prefix size per a site should not be concerned with number of hosts, but number of individual sub-networks. Thus a site with three networks should be allocated a /62. However, an ISP may choose a more relaxed allocation posture and e.g. allocate a /56 or even /48 even though only fraction of the possible subnets will be used.

For IPv6 prefixes, we replicated same test cases as we did with IPv4: 10, 500 and 5000 prefixes, with varying prefix lengths around /48 (comparable to IPv4's /24). Additional cases stem from the mentioned IPv6's address allocation characteristics: Due to the abundance of address space, the prefix lengths might not be fine-tuned to match the needs and even allocated in a lazy fashion. Thus, we added the case where the prefix lengths vary even more, up to 32 bits (causing prefix lengths from /32 to /64 to be generated).

As stated in Section 3.C-1, we also checked the possibility to use dictionary-based algorithm, based on our realm compression work, to compress prefixes. This was also tested with each case.

H) IPv6 network prefix compression results

As stated earlier, the IPv6 work was only commenced as IPv4 work had been completed. Therefore, some specific test cases could be omitted from the start, namely omitting any considerations for the number of Mobile Routers – as with IPv4 prefixes, these would only increase the payload by constant amount each. As with IPv4, a baseline with no compression needed to be established. Affecting considerations for valid baseline is the "zero compression" for addresses mentioned in IPv6 specifications [7].

Although compressing the largest block of zeroes in an address is beneficial, the approach is purely for human-readable text notation. The network equipment still treats IPv6 addresses as 128-bit numbers, even if they primarily consist of zeroes. From signaling perspective, however, we are already utilizing zero compression – tail-dropping zeroes that exceed prefix length.

While a full IPv6 address is 128 bits (16 octets), the latter 8 octets are never considered to be part of a prefix. Thus, instead of considering "uncompressed" data to consist of full 128-bit IPv6 addresses, we are in this case using a value of 9 octets per prefix, counting 1 octet for encoding prefix length and 8 octets for the prefix (highest 64 bits of an IPv6 address). The uncompressed sizes for various data set sizes are shown in *Table 5*.

Prefixes	Uncompressed size
10	90
500	4500
5000	45000

Table 5. IPv6 network prefix compression test cases, uncompressed data sizes

The effect of sequentialness to the compression is shown in *Table 6,* in the case where prefix lengths do not vary. The test case that the table illustrates is functionally identical to the IPv4 Table 3. However, it should be noted that as the base prefix length of /48 was chosen, it is clear that in random case the compression is achieved solely because the tail-dropping of zeroes, in contrast to the IPv4 case where the scarcity of address space allowed for at least some compression. However, percentage-wise, the sequential or near-sequential cases are still impressive. As with IPv4, individual test cases did not significantly differ from the average even in extreme cases.

When varying prefix lengths, things get more interesting. In all cases, the results seem to get better with 500 prefixes than with 10 prefixes, implying that the algorithms effectiveness increases with larger data sets as effect of encoded Master Prefix decreases. In the cases where there sequence has gaps, the efficiency worsens the greater the variance in prefix sizes is, which is

Table 6.
Effect of IPv6 prefix sequence to compression. Percentages in parenthesis are the remaining data compared to uncompressed data.

Count	Sequential	Near-seq	Random
10	25 (28%)	25 (28%)	70 (78%)
500	1010 (22%)	1030 (23%)	3500 (78%)
5000	10100 (22%)	10290 (23%)	35000 (78%)

Table 7. Effect of varying prefix lengths on compression

Count	Uncompr	Sequential	Near-seq	Random
10/8	90	25 (27%)	25 (28%)	74 (82%)
500/8	4500	1046 (23%)	1461 (32%)	3682 (82%)
5000/8	45000	10459 (23%)	14655 (33%)	36859 (82%)
10/16	90	26 (28%)	39 (43%)	74 (82%)
500/16	4500	1051 (23%)	1736 (39%)	3689 (82%)
5000/16	45000	10465 (23%)	17576 (39%)	36889 (82%)
10/32	90	37 (41%)	48 (53%)	82 (91%)
500/32	4500	1064 (24%)	2109 (47%)	3487 (77%)
5000/32	45000	10545 (23%)	21078 (46%)	33474 (74%)

somewhat expected. As before, in the totally random cases the efficiency depends solely on the tail-dropping of zeroes. In random cases, space consumed is greater than with identical prefix lengths case with smaller variances, but with the greatest variance tail-dropping nearly half of the prefix (for /32's) kicks in.

1.Comparison to dictionary based algorithm

As stated in Section 3.C-1, we also tested the feasibility of dictionary-based compression on batches of IPv6 prefixes. Overall, the results are no better than with the delta-compression algorithm, and in most cases worse. For comparison purposes, cases with 5000 prefixes, where any local anomalies have leveled off, are presented in *Table 8*. The uncompressed data in this case is always 45000 bytes.

Case	Delta-compr	Dictionary-compr
Single length, sequence	10100 (22%)	20023 (50%)
Single length, near seq	10296 (23%)	20062 (51%)
Variable length/8, seq	10459 (23%)	20964 (50%)
Variable length/16, partial seq	17576 (39%)	22543 (78%)
Variable length/32, partial seq	21078 (47%)	24393 (54%)

Table 8.
Comparison of algorithms in cases with 5000 prefixes

As can be seen, even though the delta-compression only has a single delta-byte, the dictionary-based approach has significantly poorer performance, even in the most optimal case of sequential networks. The differences are probably stemming from the fact that in dictionary compression there is no concept of "Master Prefix" as such – there are number of "Master Prefixes" in the dictionary, but no implicit one – so master has to always be explicitly encoded, even if it is simply a single label.

5. Conclusions

Although the presented Prefix and Realm compression algorithms are not designed to provide the most efficient compression factor possible, considering their simplicity and the specific application they perform remarkably well. Our generated test-material also appears to be accurate representation of real-world networking in Prefix Compression case, at least for IPv4.

Prefix and Realm Compression algorithms presented allow for a simple and efficient way to transmit prefix and realm information over standard Mobile IPv4 signaling messages. The effectiveness depends on the order the items are fed to the algorithm; We conclude that optimal prefix ordering should be prioritized over realm ordering. The use cases also support this approach, as number of realms is usually lower than the number of prefixes. On a typical case, we were able to compress the data to approximately 40% of original size.

When extending the prefix compression algorithm to IPv6, it scales surprisingly well provided that the network address allocation strategy conserves addresses. Within a single organization this can well be the case. If address space is used in a more erratic fashion, the efficiency suffers.

While these algorithms provide a quick and simple approach to reduce signaling traffic, a different scheme might be more desirable for larger networks. However, this approach can be implemented with relatively low effort and appears to scale up to several hundred prefixes and realms before even requiring the problematic fragmentation of signaling PDUs.

6. Future work

The compression algorithms presented in this paper are only very basic, simple methods for compressing the realm and prefix information according to the specification. Optimizing the prefix- and realm compression order may provide higher gains. Areas to improve is the significance of Mobile Router compression order; Also, the Mobile Router address itself might be eligible for compression. In case of IPv6, it may also be possible to add variable number of delta octets for added benefit, if prefix length can somehow be more efficiently encoded. Using relative prefix lengths might be a possibility, since it's quite likely that prefix length variation will be limited.

For realm compression algorithm, a more intelligent label and suffix replacement algorithm instead of a complete dictionary reset would most probably better the realm compression significantly.

Acknowledgment

Antti Mäkelä acknowledges TEKES GIGA Program WISEciti project and FI ICT SHOK programs under which this work was conducted.

Authors

ANTTI MÄKELÄ is a post-graduate student and researcher at Aalto University, Department of Communications and Networking. His work history also includes 9 years at largest telecommunications operator in Scandinavia, Teliasonera corporation, where he participated in several research projects covering IPv6 deployment and emerging security challenges before moving to academia. His current areas of interest include virtual operators as business model, IP mobility and routing, and future Internet architectures in general. Mäkelä received his MSc (2002) in Computer Science from Tampere University of Technology, Finland, and is now working on his Ph.D dissertation. He is an active participant in IETF standardization organization. As an additional recognition, he has obtained a Cisco Certified Internetwork Expert (CCIE) certification and is recognized as an inventor in a number of patents in his areas of interest.

JOUNI KORHONEN is a Senior Specialist, Internet standards, at Nokia Siemens Networks. His current areas of interest and responsibilities include Internet technologies at large, wireless cellular network architectures, and the development of a future Internet combined with wireless broadband. He currently participates and leads several projects related to a future (wireless) Internet and also IPv6 deployment challenges. Korhonen received both his Ph.D (2008) and M.Sc. (1998) in Computer Science from the University of Helsinki, Finland. He is an active participant in IETF and 3GPP standardization organizations. He currently serves as the co-chair of the Diameter Maintenance and Extensions (DIME) working group in IETF. He has authored a number of publications, IETF RFCs and 3GPP standards. He also holds a number of patents on his areas of interest.

References

[1] Z. Zhang, X. Chu, B. Li, Y-Q. Zhang,
An overview of Virtual Private Network (VPN):
IP VPN and Optical VPN,
Photonic Network Communications, Vol. 7, Nr. 3,
May 2004.

[2] K. Leung, G. Dommety, V. Narayanan, A. Petrescu,
Network Mobility (NEMO) Extensions for Mobile IPv4,
IETF RFC 5177, April 2008.

[3] A. Mäkelä, J. Korhonen,
Home agent assisted route optimization for
mobile IPv4,
IETF Draft draft-ietf-mip4-nemo-haaro-00.txt,
(Work in progress), 2010.

[4] A. Mäkelä, J. Korhonen,
Space-efficient signaling scheme for
Home Agent Assisted Route Optimization for use
in Virtual Networks,
In Proc. of the 10th International Conference on
Telecommunications (ConTEL 2009), June 2009.

[5] A. Mäkelä,
Concept for providing guaranteed service level over
an array of unguaranteed commodity connections,
In Proc. of the 25th Symposium on Applied Computing
(ACM SAC 2010), March 2010.

[6] Y. Rekhter, T. Li,
An Architecture for IP Address Allocation with CIDR,
IETF RFC 1518, 1993.

[7] R. Hinden, S. Deering,
Internet Protocol Version 6 Addressing Architecture,
IETF RFC 4291, February 2006.

[8] B. Aboba, et al,
Review of Roaming Implementations,
IETF RFC 2194, September 1997.

[9] R. Bush, et al,
IPv4 Address Allocation and Assignment Policies
for the RIPE NCC Service Region,
RIPE Document ID 492, February 2010.

[10] T. Clausen, C. Dearlove, J. Dean, C. Adjih,
Generalized MANET Packet/Message Format,
IETF RFC 5144, February 2009.

[11] G. Montenegro, N. Kushalnagar, J. Hui, D. Culler,
Transmission of IPv6 Packets over
IEEE 802.15.4 Networks,
IETF RFC 4944, September 2007.

[12] J. Hui, P. Thubert,
Compression Format for IPv6 Datagrams
in 6LoWPAN Networks,
IETF Draft draft-ietf-6lowpan-hc-07.txt,
(Work in progress), 2010.

[13] IEEE Computer Society,
IEEE Std. 802.15.4-2003, IEEE, October 2003.

[14] P. Mockapetris,
Domain names – Implementation and Specification,
IETF RFC 1035, 1987.

[15] BGP Update Reports,
IPv4 address pools, Advertised.,
http://bgp.potaroo.net/ipv4-stats/prefixes_adv_pool.txt
(Fetched on August 28th, 2008)

[16] G. Van de Velde, C. Popoviciu, T. Chown,
O. Bonness, C. Hahn,
IPv6 Unicast Address Assignment Considerations,
IETF RFC 5375, December 2008.

[17] G. Huston,
Architectural Approaches to Multi-homing for IPv6,
IETF RFC 4177, September 2005.

[18] M. Wasserman,
Recommendations for IPv6 in 3rd Generation
Partnership Project (3GPP) Standards,
IETF RFC 3314, September 2002.

[19] B. Patil, F. Xia, B. Sarikaya, J.H. Choi, S. Madanapalli,
Transmission of IPv6 via the IPv6 Convergence
Sublayer over IEEE 802.16 Networks,
IETF RFC 5121, February 2008.

[20] R.N. Horspool,
The effect of non-greedy parsing in Ziv-Lempel
compression methods,
IEEE Data Compression Conf., pp.302–311, 1995.

Achieving collaborative service provisioning for mobile network users: the CollDown example

VEDRAN PODOBNIK, IVA BOJIC, LUKA VRDOLJAK, MARIO KUSEK

University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{vedran.podobnik, iva.bojic, luka.vrdoljak, mario.kusek}@fer.hr

Keywords: mobile user, collaborative service, energy-efficiency, semantic matchmaking, user profile, implicit social network

This paper proposes a distributed model of service provisioning process for mobile network users who share a common interest in a service content offered in the network. The model is based on an idea that such users could individually acquire disjunctive parts of service content from a remote server (via a wide area mobile network), and then subsequently exchange them among themselves in an ad hoc local network in a peer-to-peer fashion, thus allowing each user to reassemble the entire content for her/his own use. The benefits of this approach include lowering the load on the "expensive" link to the telco's network and saving the energy for wide-area communication on users' mobile devices. The key questions to solve include recognizing the "common interest" of users and achieving their collaboration as described above. As an example, we show a service in which mobile users form an implicit social network and collaborate while downloading a selected multimedia content from the server under the control of the telco. A proof-of-concept implementation, named Collaborative Downloading (CollDown for short), is evaluated by using real mobile phones in a real network.

1. Introduction

Simultaneous development of mobile devices and telco infrastructure resulted in the network ability to provide more complex mobile services [1]. In this paper we study a particular type of mobile services, namely, *collaborative services* for *mobile users*. A mobile user is defined as a person who uses a mobile device (e.g., a mobile phone) which has the ability to attach both to a public land mobile network and an ad hoc or other local wireless network (e.g., Bluetooth, IEEE 802.11) operating in the unlicensed spectrum. A collaborative service [2,3] is defined as a service in which a set of users, sharing a common goal (e.g., downloading the same video clip from a service server), work together as a group in order to achieve that goal. The main idea behind the collaborative service is that the network operator can build an implicit social network of mobile users based on information provided by the users at the time of subscription. The social network is built with respect to user interests, their mobile devices' characteristics and the context in which they find themselves while requesting a service. The implicit social network of mobile users is built autonomously, without the interference of users themselves and in order to provide useful information for telcos [4].

This paper presents an extended version of our previous work [5], in which our main contribution was an implementation of an agent-based [6] middleware for group-oriented mobile service provisioning, while this paper focuses on an evaluation of the collaborative service called *Collaborative Downloading* (CollDown for short). The evaluation, focusing on overall download time and energy consumption, is done by using a real world network environment and Sony Ericsson mobile phones. The rest of this paper is organized as follows. Section 2 describes the idea of a three-layered architecture for enabling collaborative services. Section 3 elaborates upon a proof-of-concept collaborative service called *CollDown*. Section 4 concludes the paper and proposes some directions for future work.

2. Architecture for enabling collaborative services

The proposed model for enabling the collaborative service provisioning has a three-layered architecture, based on three views on the mobile users [7]. These layers are the *physical layer,* the *ontology layer* and the *social layer,* as shown in *Fig. 1*.

The **Physical layer** observes mobile users as persons physically situated in a mobile network environment and using mobile devices. Mobile users are connected to a telco's network via physical links and communicating with the telco's *base stations* via *wireless links*. Mobile users' devices are also equipped with Bluetooth and/or Wi-Fi technology which enable ad hoc connections among mobile devices. The base stations are interconnected using *wired links* in the telco's core network.

The **Ontology layer** observes mobile users through their semantic profiles, which all refer to an ontology representing domain knowledge. We use the World Wide Web Consortium (W3C) Composite Capabilities/Preferences Profile (CC/PP), an RDF-based specification which

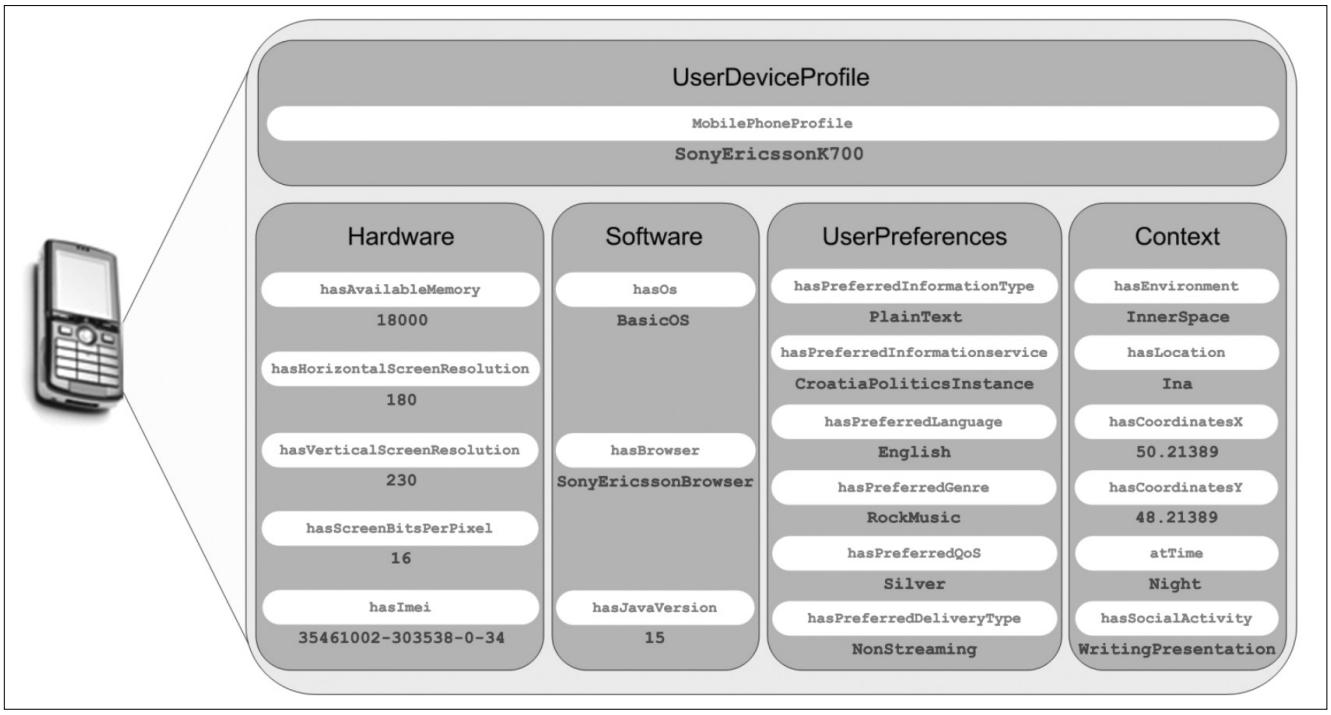
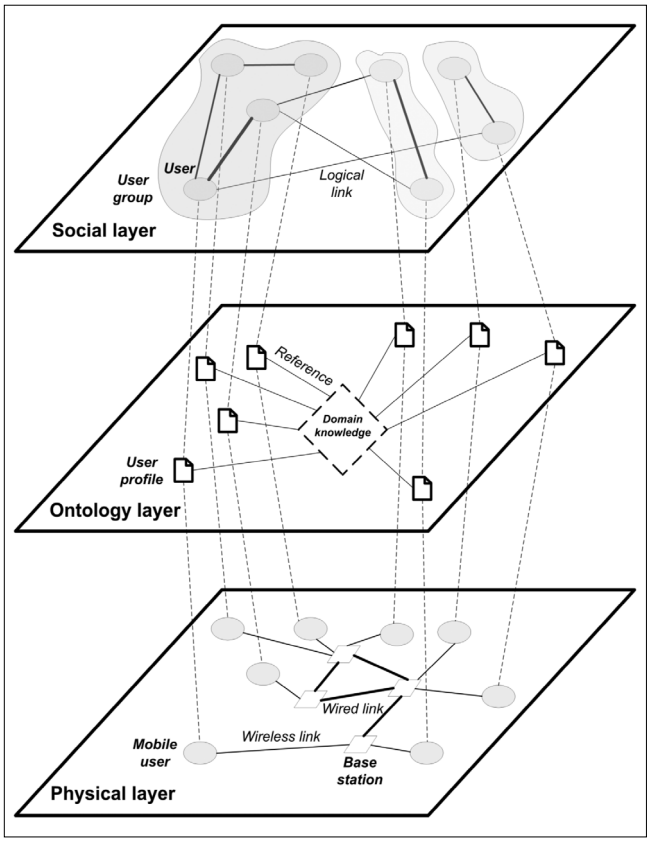


Figure 2. User profile

describes device capabilities and user preferences, and apply it for adapting the content to the user's mobile device. Such an approach maps each user with a mobile device (from the physical layer) to a user profile (within the ontology layer).

Figure 1.
Three-layered architecture for enabling collaborative service provisioning



The user profile (see *Fig. 2*) consists of five parts:
- *UserDeviceProfile,* describing the type of user mobile device (e.g., a mobile phone);
- *Hardware,* describing the configuration of user mobile device (e.g., screen vertical resolution of 230 pixels);
- *Software,* describing the software installed on user mobile device (e.g., an operating system);
- *UserPreferences,* describing the preferences and the interests of a mobile user (e.g., English as the preferred language);
- *Context,* describing the mobile user's context (e.g., the "time of day" is "night").

The **Social layer** observes mobile users through an implicit social network. Mobile users' profiles are initially filtered according to a certain criterion (e.g., location). Profiles of filtered users are then compared by semantic matchmaking between all pairs of those mobile users. Specifically, the context part of user profiles enables telcos to calculate the physical distance between users. Both the semantic similarity and the physical distance can be useful in order to determine the potential target set of users for a specific telecom service. Thus, we can create a social network and identify its subnets with respect to the combination of the user physical location and the semantic similarity among the user profiles.

Semantic matchmaking is the process of comparing two objects represented through semantic profiles, resulting with a number within a certain interval (i.e., a number between *0* and *1*, where the larger number reflects a higher level of similarity between profiles). A detailed description of semantic matchmaking of user profiles (structured according to Fig. 2), as well as building an implicit social network of mobile users, is given

in [4], while the procedure for grouping users who form a social network into clusters is described in [8].

One of the practical issues that need to be addressed when considering the proposed model implementation is how the privacy of mobile users' profiles is handled. For the purposes of this work, we consider the role of a telco as being both a network provider and a service provider. As such, it provides a content of interest (e.g., video clips) to mobile users who subscribe to a related service. We are currently considering only the collaboration among subscribers having the same operator, but the model could be extended to have more than one telco. From the technical point of view, introducing more than one telco would be fairly straightforward, while from business point of view it would require collaboration not just among users, but among their respective telcos as well.

We also considered an approach with a third party service provider, with which a telco would have a service agreement in place, but decided to exclude it for the time being due to unsolved privacy issues (sharing user profiles with third party entities as something that users have not agreed on). In the described approach, telcos do not share user profiles with third parties, but directly provide services to their own subscribers (based on the information that the users have willingly provided upon subscription).

3. Proof-of-concept service: CollDown

A collaborative proof-of-concept service we developed is called *Collaborative Downloading* (CollDown for short). CollDown aims to make the mobile service provisioning process more efficient for both mobile users and telcos. It lowers battery consumption for mobile users' devices which is of great importance because limited energy supply is one of the main obstacles for using advanced mobile services. It also has several benefits for telcos: their mobile network resources are less loaded, and the overall CPU load and memory consumption on their service (content) servers are lower, meaning that the servers consume less energy and are thus more environmentally friendly.

The proposed approach moves a step further from the "traditional" content download service model we have today. Namely, when a mobile user wants to download a certain multimedia content to her/his mobile device, s/he is not aware of other mobile users interested in the same content. Consequently, each user communicates only with the service (content) server and downloads the complete desired content via a mobile network (e.g., General Packet Radio Service; GPRS, or Universal Mobile Telecommunications System; UMTS). We refer to this standard approach to mobile service provisioning as *individual approach*.

The basic idea of CollDown is that mobile users, who are physically located at a relatively small distance from each other (e.g., 10 meters – the Bluetooth range), and who are interested in the same content, collaborate and download the desired content together. Two typical examples of such situations are rock concerts and sport events, where users may want to view close-up views of performers or in-game action replays during the event. To allow collaborative download, the media content on the server is divided into a number of disjunctive parts and each part can be downloaded independently from the server via a mobile network and afterwards exchanged with other mobile users via an ad hoc network (e.g., Bluetooth or Wi-Fi). Exchange of content parts is performed as an "auction" where mobile users compete on the "market" to determine which parts of a requested content to download directly from the service server and which to exchange with other users.

Further details about the procedure how mobile users from a targeted group accept or decline to participate in a particular CollDown scenario are given in [5], while the auction, which manages the exchange of content parts, is presented in [10]. In this article, we refer to such a new mobile service provisioning approach as a *collaborative approach*.

Mobile users who collaborate can form either *flexible groups* (in which mobile users can leave a collaborating group before the service provisioning is completed) or *fixed groups* (in which mobile users cannot leave a collaborating group before the service provisioning is completed). In the present application, we assume fixed groups, in which users may move or change their individual location only as long as they remain "within range" for *ad hoc* communication. This assumption may be considered realistic for the case where the content download takes less time than the usual time period for one (or more) users to leave the group. However, in future work we may consider different flexibility models which would allow users to leave their groups before the service provisioning is completed.

3.1 The system for CollDown service provisioning process

The system model for CollDown service is illustrated in *Fig. 3.* Let $\mathbf{I}=\{i_1, ..., i_N\}$ denote the set of N mobile users who are subscribers of a certain service offered by a telco and let $\mathbf{J}=\{j_1, ..., j_M\}$ denote a subset of mobile users who form a mobile ad hoc network. Set **J** is always a rather small subset of set **I** ($\mathbf{J}\subset\mathbf{I}$), while $M=|\mathbf{J}|$. The calculation of subset **J** from set **I** is based on a three-layered architecture for enabling collaborative service provisioning (presented in previous section), namely on grouping of mobile users who form a social network into clusters. The content on the service server is divided into smaller disjunctive parts. Let $\mathbf{K}=\{k_1, ..., k_p\}$ denote the set of content parts on the server, which the users may download individually and reassemble by sharing them through an ad hoc network.

In the system shown in Fig. 3, there are two modes in which mobile devices may communicate:

a) the wide-area mobile network, shown as GPRS; and

b) ad hoc communication, shown as Bluetooth.

Networks	Send (*μJoule/bit*)	Receive (*μJoule/bit*)
Bluetooth	0.064	0.064
GPRS	15.647	1.422
IEEE 802.11a	0.028	0.022

Table 1.
Energy consumption in GPRS, Bluetooth and Wi-Fi networks [11]

Mobile phones communicate over GPRS with the service server by using Session Initiation Protocol (SIP) as the control protocol and Hypertext Transfer Protocol (HTTP) for the content download. The direct communication among mobile phones takes place in a mobile ad hoc network and uses Bluetooth-based communication.

The energy consumption while sending and receiving data within GPRS, Bluetooth and Wi-Fi networks is compared in *Table 1* [11].

It may be noticed that Bluetooth-based communication consumes much less energy than GPRS-based communication. For the purpose of comparison, we also cite the energy consumption for a wireless LAN (IEEE 802.11a), which is even lower than Bluetooth.

3.2 The CollDown performance measurements

The CollDown service was evaluated in a real network environment, by using a prototype service server developed in Java, and SonyEricsson W910i mobile phones as mobile users' devices. (The CollDown server was accessible through a real GPRS network of a mobile operator.) SonyEricsson W910i phones support Java platform 8 (JP8), as well as different Application Programming Interfaces (APIs) required for CollDown service execution on mobile devices: JSR180 (SIP API), JSR82 (Bluetooth API) and JSR256 (Mobile Sensor API). The CollDown service is developed for mobile devices supporting Connected Limited Device Configuration (CLDC) 1.0 and Mobile Information Device Profile (MIDP) 2.0.

We have compared our *collaborative approach* based on CollDown with the *individual approach* for a video clip downloading in mobile networks. We measured the time and the energy consumption in both cases. *Fig. 4* presents the comparison of time required for a system composed of one, two, three and four mobile users/phones to download the whole video clip by using the individual approach and by using the collaborative approach. *Fig. 5* shows the respective energy consumption.

As shown in Fig. 4, the low data rate on Bluetooth links presents a serious limitation in this particular implementation of the collaborative approach. The time duration for the four-mobile-users scenario in the individual approach is 74.31 s, while in collaborative approach it takes 187.25 s. For practical purposes, a faster and more efficient ad hoc or (if available) a local area network (Wi-Fi) should be used.

Results in Fig. 5 show energy consumption in the system. In our system the overall energy consumption is calculated as a size of chunk multiplied by mobile devices' characteristic energy consumption per bit shown in Table 1. Chunk is the information unit which includes only data packets for individual approach and both data packets and gossip messages for collaborative approach. Gossip messages are used to propagate content availability information between users in an ad hoc network. While the gossip messages clearly introduce a certain overhead, the presented collaborative approach

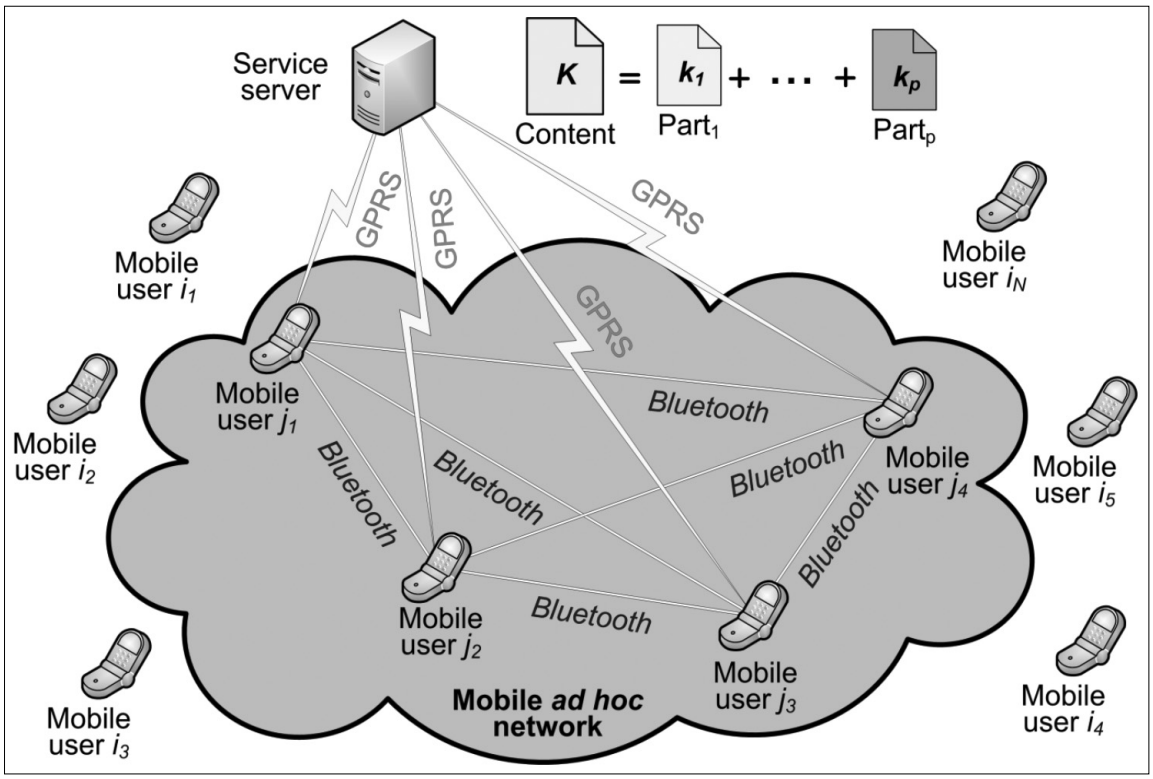


Figure 3.
System model for CollDown

consumes only 0.61 J, and the same four-mobile-users scenario for individual approach consumes 1.91 J of mobile users' phone battery energy.

3.3 The CollDown performance evaluation

Using collaborative download has many benefits, but also some drawbacks. The main advantage for mobile users is saving the mobile device battery energy, since the GPRS-based communication consumes more energy than the Bluetooth-based communication (see Table 1). Consequently, if some parts of content are exchanged with other mobile users via Bluetooth instead of being downloaded using GPRS, a significant amount of energy may be saved.

Additionally, redirecting some of the network traffic from GPRS to Bluetooth can free up some of the scarce access network resources and allow telcos to make more profit by offering these resources to other users and/or using them for other services. Collaborative downloading also reduces the total load on telcos' service servers since (ideally) each content part would only be downloaded once and service servers would have to process fewer users' queries. The servers' CPU and memory consumption would be lower, meaning that they would consume less energy and be more environmentally friendly.

Last but not least (when considering replacing Bluetooth with Wi-Fi where available), Wi-Fi technologies

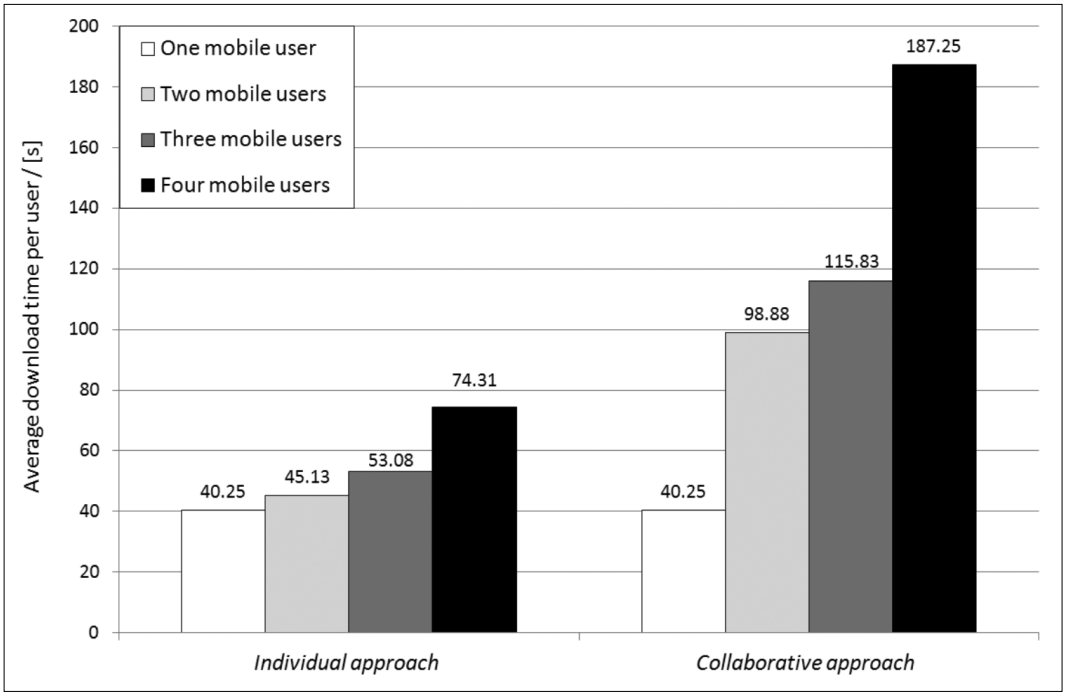


Figure 4. Comparison of time for the individual and the collaborative approach

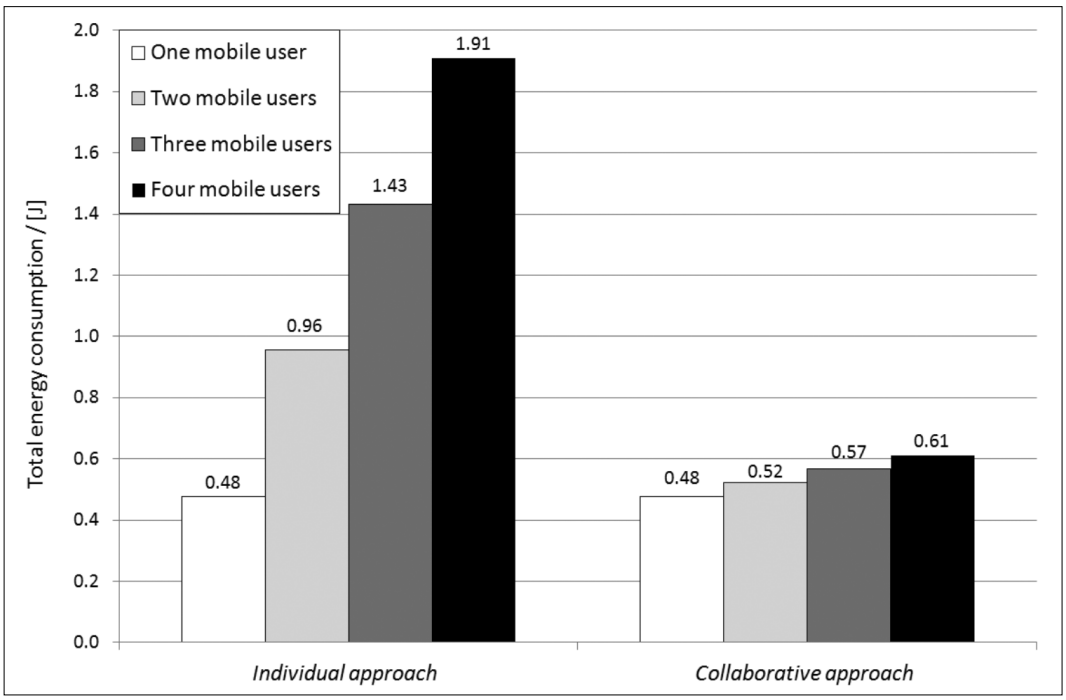


Figure 5. Comparison of energy consumption for the individual and the collaborative approach

are considered a non-risk for human health. The official stance of the Health Protection Agency is that *"there is no consistent evidence to date that Wi-Fi and WLANs adversely affect the health of the general population"* [12].

The main disadvantage of our approach is that the service response time for collaborative approach can be longer than for individual approach. Currently, Bluetooth-based communication consumes more time than GPRS-based communication. Although according to the specifications the declared speed for Bluetooth is higher, the test results reported in [13], as well as those we obtained, show data rates of only 10-11 KB/s. In future, Bluetooth v3.0 should be able to deliver up to 3 MB/s of data throughput [14].

Additionally, Bluetooth communication could be substituted with Wi-Fi communication, in which case the transmission speed would be higher, and energy consumption still lower than in GPRS network (see Table 1). Beside the transmission speed, the processes of semantic matchmaking and group formation, required for enabling user collaboration, introduce an additional delay in the service provisioning process.

4. Conclusion and future work

Mobile users are becoming more and more involved in the process of telecom service provisioning – they are evolving from individual users towards collaborative users. The collaborative users want to gain added value from the mobile service provisioning by collaborating during the provisioning process with other mobile users who share a common goal.

In this paper we described and evaluated *Collaborative Downloading* (CollDown), an energy-efficient Bluetooth-based collaborative service. The simulation results show a significant decrease of energy consumption of mobile users for the *collaborative approach* based on CollDown compared to the *individual approach* based on individual downloading via a mobile network. This is of prime importance since limited battery energy is one of the greatest challenges faced by both mobile users and telcos.

For future work we plan to implement CollDown service using Wi-Fi technology with the aim of improving our service's time efficiency. Moreover, we intend to extend our model with group flexibility to see how our approach can be used if a user leaves her/his group during the collaboration process.

Finally, we will use the three-layered architecture for enabling collaborative service provisioning to model other innovative examples of collaborative services, such as collaborative mobile learning and collaborative recommenders. The former service enables electronic learning experience for mobile users based on peer collaboration, while the latter provides recommendations from users with similar interests and preferences.

Acknowledgments

This work was carried out within research project 036-0362027-1639 "Content Delivery and Mobility of Users and Services in New Generation Networks", supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

Authors

VEDRAN PODOBNIK is a Research and Teaching Assistant at the Department of Telecommunications at the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia. He received his M.Eng. (2006) and Ph.D. (2010, Computer Science) degrees from the University of Zagreb. His research interests include multi-agent systems, social networks, electronic markets, context-aware services, business process automation and the semantic web. He has 29 publications in journals and conferences and 4 book chapters. He received the annual national award SCIENCE in the year 2007, and several student awards for academic excellence. He is a member of IEEE and KES Int.

IVA BOJIC is a Research Associate at the Telecommunication Department of the Faculty of Electrical Engineering and Computing (FER), University of Zagreb, Croatia. She received her Master in Computer Science degree from FER. Her diploma thesis was entitled "Application of Self-Organized Software Agents in Telecommunication Service Provisioning". As a student she received several awards for academic excellence, the most prestigious being the "Stanko Turk" Annual Award for the best diploma thesis in the field of the Computer Sciences. She also received the "Best Student Research Paper Award" at the 4th International KES Symposium on Agents and Multi-agent Systems – Technologies and Applications and "Exceptional outstanding paper" award at the 32nd International Convention on Information and Communication Technology, Electronics and Microelectronics, 2009. She is currently working towards her Ph.D. degree in the area of semantic web and user profiling. She is a member of IEEE and KES International.

LUKA VRDOLJAK received his Master in Electrical Engineering degree in 2009 from the University of Zagreb, Faculty Electrical Engineering and Computing. While he was a student, he received several awards in student competitions as well as for academic excellence, the most prominent being the Annual Rector's Award of the University of Zagreb in 2009. His diploma thesis was "Agent System based on Semantic Reasoning for Creating Social Networks of Telecommunication Service Users". As of late 2009 he works at the Erste&Steiermärkische Bank, Zagreb, Croatia as an IT Application Programmer, and he is also a part-time Ph.D. student at the Univ. of Zagreb.

MARIO KUSEK is an Assistant Professor at the University of Zagreb, Faculty of Electrical Engineering and Computing, in Croatia. He holds an M.S. (2001) and Ph.D. (2005) in Electrical Engineering, major in telecommunications and informatics, from the same university. His main research interests include distributed systems, software agents in next generation networks, and converged services on mobile terminals. He has 46 scientific journal and conference publications. He has participated in scientific projects funded by the Croatian Ministry of Science, Education, and Sports and industrial projects funded by Ericsson Nikola Tesla, Zagreb. His contribution in these projects were related to design and implementation of converged services in next-generation networks using various service platforms (Parlay/OSA, Parlay X, Jain SLEE and IMS) and terminals (Java Micro, iPhone and Android). He was involved in several international conferences and workshops as a TPC member and Publication Chair, and he is a management committee member in COST Action 0801. He is a member of IEEE and KES International. He is currently the Chair of the IEEE Communications Society Croatia Chapter.

References

[1] Belavic, R., Basuga, M., Podobnik, V., Petric, A., I. Lovrek, "Agent-based social networking for mobile user", International Journal of Intelligent Information and Database Systems (in press), 2010.

[2] Nandan, A., Das, S., Pau G., Gerla M., "Co-Operative Downloading in Vehicular Ad Hoc Wireless Networks", In: Proc. of the 2nd Annual Conference on Wireless On-demand Network Systems and Services (WONS'05), St. Moritz, Switzerland, pp.32–41, 2005.

[3] Zhang, J., Zhang, Q., Jia, W., "A Novel MAC Protocol for Cooperative Downloading in Vehicular Networks", In: Proc. of the IEEE Global Telecommunications Conference (GLOBECOM'07), Washington, USA, pp.4974–4978, 2007.

[4] Podobnik V., Lovrek I., "Telco Agent: Enabler of Paradigm Shift towards Customer-Managed Relationship", Lecture Notes in Computer Science, Vol. 6276, pp.251–260, 2010.

[5] Vrdoljak, L., Bojic, I., Podobnik, V., Kusek, M., "The AMiGO-Mob: Agent-based Middleware for Group-oriented Mobile Service Provisioning", In: Proc. of the 10th International Conference on Telecommunications (ConTEL'09), Zagreb, Croatia, pp.97–104, 2009.

[6] Jennings, N., Sycara, K., Wooldridge, M., "A Roadmap of Agent Research and Development", J. of Autonomous Agents and Multi-Agent Systems, Vol. 1, No.1, pp.7–36, 1998.

[7] Jung, J. J., Euzenat, J., "Towards Semantic Social Networks", Lecture Notes in Computer Science, Vol. 4519, pp.267–280, 2007.

[8] Podobnik, V., Galetic, V., Trzec, K., Jezic, G., "Group-Oriented Service Provisioning in Next Generation Network", In: Innovations in Multi-Agent Systems and Applications, D. Srinivasan and L.C. Jain (eds.), Springer-Verlag, Berlin Heidelberg, pp.277–298, 2010.

[9] Pentikousis, K., "In Serach of Energy-Efficient Mobile Networking", IEEE Communications, Vol. 48, No.1, pp.95–103, 2010.

[10] Bojic, I., Podobnik, V., and Kusek, M., "Agent-enabled Collaborative Downloading: Towards Energy-efficient Provisioning of Group-oriented Services", Lecture Notes in Computer Science, Vol. 6071, pp.62–71, 2010.

[11] O'Hara, K.J., Nathuji, R., Raj, H., Schwan, K., Balch, T., "AutoPower: Toward Energy-aware Software Systems for Distributed Mobile Robots", In: Proc. of the 2006 IEEE International Conference on Robotics and Automation (ICRA'06), pp.2757–2762, 2006.

[12] WiFi Summary: *http://www.hpa.org.uk/radiation/understand/radiation_topics/emf/wifi.htmJ2* (last visited on 10th July, 2010)

[13] ME Bluetooth Programming: *http://www.nowires.org/Presentations-PDF/AndreKpresentasjon.pdf* (last visited on 10th July, 2010)

[14] Bluetooth v 3.0: *http://www.bluetooth.com/English/Technology/Pages/Full_Speed_Ahead.aspx* (last visited on 10th July, 2010)

Performance improvement of MC-CDMA microstatistic multi-user detection

in nonlinear fading channels using spreading code selection

JURAJ GAZDA, PETER DROTÁR, DUSAN KOCUR, PAVOL GALAJDA

Dept. of Electronics and Multimedia Communications, Technical University of Kosice, Slovak Republic
{juraj.gazda, peto.drotar, dusan.kocur, pavol.galajda}@tuke.sk

Keywords: HPA, MC-CDMA, multiuser detection, microstatistic filtering, nonlinear detection

In this paper, we present a novel method for joint suppression of multiple-access interference (MAI) and nonlinear distortion introduced by high power amplifiers (HPA) in multicarrier code division multiple access (MC-CDMA) transmission system performing over frequency selective channel. The proposed method combines a proper selection of the spreading code at the transmitter side and application of the advanced microstatistic multi-user detector (MSF-MUD) at the receiver side. This scheme exploits relatively low peak-to-average power ratio (PAPR) and orthogonality of Golay codes in conjunction with MSF-MUD ability of effective suppression of MAI and nonlinear distortion. In order to have a good reference for the performance results of the proposed scheme, conventional minimum mean square error multi-user detector (MMSE-MUD) is introduced as well. As it will be shown by means of computer simulations, MSF-MUD with Golay codes employed for spreading can clearly outperform the other tested spreading codes and MMSE-MUD. The performance improvement is more remarkable, if Saleh model of HPA is employed.

1. Introduction

Future wireless communication systems must be able to accommodate a large number of users and simultaneously to provide the high data rates at the required quality of service. MC-CDMA is taking the advantage of two advanced technological concepts of wireless communications such as orthogonal frequency division multiplex (OFDM) and the code division multiple access (CDMA), what results especially in high spectral efficiency, the multiple access capability, robustness in the case of frequency selective channels, simple one-tap equalization, narrow-band interference rejection and high flexibility of the MC-CDMA. The outlined potential properties of the MC-CDMA represent the fundamental reasons, why MC-CDMA has been receiving a great attention over the last decade (e.g. [1,2]) and has been considering to be a promising candidate for the future advanced wireless communication systems.

One of the major requirements posed to the MC-CDMA is to reach the required date rate at the acceptable bit error rate (BER) and acceptable complexity for the defined number of the active users. It has been shown e.g. in [3-5], that in the case of MC-CDMA systems, the BER at the constant number of the active users is affected especially by nonlinear effects due to the HPA of the MC-CDMA transmitter, MAI resulting from cross-correlation properties of the spreading codes assigned to the particular users and by transmission channel complexity.

The analyses of the MC-CDMA signals have shown that due to their multi-carrier nature, the transmitted MC-CDMA signal is characterized by large envelope fluctuation [6]. This property of MC-CDMA signals forces the MC-CDMA transmitter HPA to operate with large input back-off (IBO) in order to keep the required BER and the out-of-band radiation below imposed limits. However, the large IBO will result in inefficient exploiting of HPA and consequently decreasing the coverage of the area of interest by acceptable MC-CDMA signals. As a consequence of this fact, it is crucial to minimize the impact of the nonlinear amplification on the transmission system performance at low IBO.

There are several metrics applied for signal envelope fluctuation quantifying [7]. Here, PAPR has been widely accepted for that purpose. The analyses of the PAPR of the transmitted MC-CDMA signals presented in [3,4] have shown that PAPR can be reduced by proper selection of the spreading codes. The alternative solutions of MC-CDMA performance improvement can be achieved by the application of additional methods of PAPR reduction (e.g. [8,9]) and by the compensation methods of the nonlinear distortion due to the nonlinear HPA. The nonlinear distortion compensation methods can be implemented at the transmitter or receiver side of MC-CDMA transmission system. Frequently used solutions at the transmitter side include pre-distortion [10], tone reservation [11], active constellation extension, selected mapping [9], different code allocation strategies [12], etc. The strategies applied at the receiver side usually combine iterative decoding [13] and nonlinear multi-user detection [14,15].

Because of the CDMA exploited in the MC-CDMA structure, the BER reached by the particular users is strongly dependent on MAI. The level of MAI is primarily determined by the spreading codes assigned to the particular users and the transmission channel properties [6]. As the spreading codes for MC-CDMA, Walsh codes, Gold and orthogonal Gold codes, polyphase Zadoff-Chu

codes as well as complementary Golay codes are mostly employed (e.g. [3-5]). In the case of wireless MC-CDMA systems, the transmitted signals are firstly nonlinearly distorted and subsequently they are affected by a fading channel. This effects can result in spreading code orthogonality loss and consequently, MAI increasing. Then, the level of the MAI can be reduced by the application of multi-user receivers [14,16].

As it follows from this short overview of the MC-CDMA performance degradation sources (PAPR, MAI, nonlinear distortion due to HPA), there are a number of approaches how to improve MC-CDMA performance (spreading code selection, PAPR reduction methods, multi-user receiver and nonlinear distortion compensation methods). Here, the application of the nonlinear multi-user receiver and simultaneously the properly selected set of spreading code application are considered as the very perspective solution.

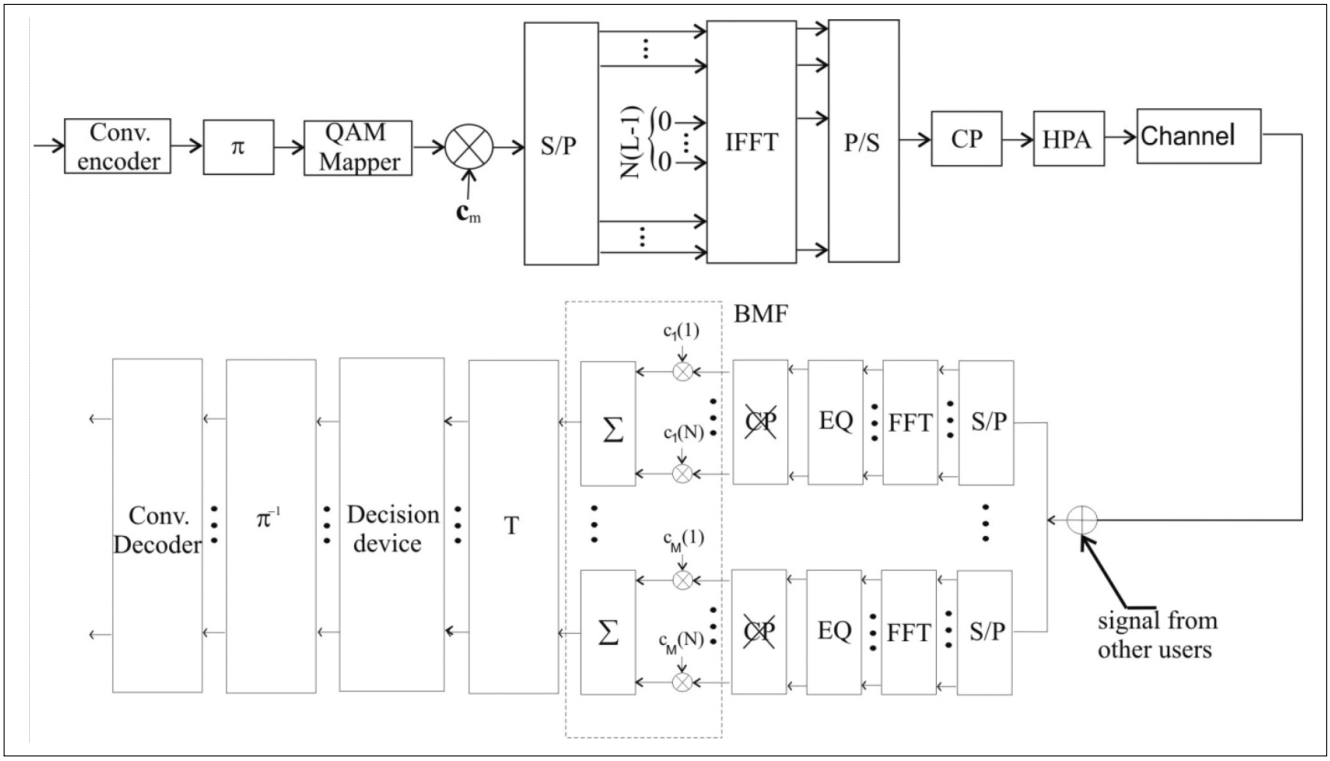
In this paper, we will deal with the performance analyses of MC-CDMA transmission system employing the nonlinear MSF-MUD and the different spreading codes. Originally, the MSF-MUD has been proposed in [14] as the multi-user receiver able to compensate the nonlinear distortion due to the HPA of transmitter. The performance properties of the MSF-MUD with regard to the different spreading codes for AWGN channel scenario have been discussed in [17]. This contribution is the extension of our previous study introduced in [17] to the analysis of MSF-MUD performance properties for the frequency-selective fading channel if Walsh codes, Gold and orthogonal Gold codes, polyphase Zadoff-Chu codes and complementary Golay codes are used as spreading sequences. As the nonlinear HPA model, Saleh and Rapp models have been taken into account. In order to illustrate the MSF-MUD performance, MMSE-MUD will be also applied as the MC-CDMA receivers. Because in [14,17], the design procedure of the MSF-MUD has been outlined only, the deeper description of the design procedure of the optimum MSF-MUD is included in this paper. The aim of our study is to find by means of computer simulations the pair of a receiver and a set of the spreading codes able to provide the best BER performance under the mentioned conditions. The simulation results show that the maximum improvement of the MC-CDMA performance compared to other system configurations (i.e. to all combination of MMSE-MUD with the different spreading codes) is achieved by MSF-MUD if the Golay codes are used as the spreading sequence. It will be also shown that this result is true especially for the scenario with strongly nonlinearly distorted MC-CDMA signals (i.e. for low values of IBO parameter).

The structure of our contribution is as follows. In the next section, the transmitter and channel model are briefly discussed. The MSF-MUD description and its design procedure are given in Section 3. The core of the paper is Section 4, where the computer simulations and obtained results are presented. Finally, some conclusions from the presented work are drawn in Section 5.

2. Transmitter and channel model

The block diagram of the simplified baseband model of MC-CDMA transmitter is depicted in *Fig. 1.* It can be seen from this figure that the information bits to be transmitted by a particular user are firstly fed to the block of non-

Figure 1. Block diagram of MC-CDMA communication system



recursive, non-systematic convolution encoder followed by the block of interleaving labeled by π in the presented scheme. Then, a baseband modulator transforms the encoded binary input to a multilevel sequence of complex numbers in M-QAM modulation formats.

The data obtained in such a way are spread by using a specific spreading sequence c_m, where c_m denotes the spreading sequence of the m-th user. As the spreading sequences, Walsh codes, Gold codes, orthogonal Gold codes, complementary Golay codes and polyphase Zadoff-Chu codes are considered [3]. The PAPR upper bound of the mentioned spreading codes is summarized in *Table 1* [3]. As it can be observed from this table, the PAPR bound of Golay codes and Zadoff-Chu codes is independent of the spreading code length L. If we assume that the subcarrier number N_c is a multiple of L, then PAPR of the Walsh codes is upper-bounded by $2N_c$ [3].

Spreading Codes	PAPR (upper bound)
Walsh codes	$\leq 2N_C$
Gold codes, orthogonal Gold codes	$\leq 2\left[t(m)-1-\frac{t(m)}{L}+\frac{2}{L}\right]$
Golay codes	≤ 4
Zadoff-Chu codes	2

Table 1.
PAPR bounds for MC-CDMA uplink signal

The spread symbols are modulated by the multi-carrier modulation implemented by the inverse fast Fourier transformation operation (IFFT). After parallel-to-serial (P/S) conversion, the cyclic prefix (CP) is inserted in order to mitigate the inter-symbol interference (ISI) caused by the frequency-selective fading channel. Finally, this resulting signal represents the HPA input signal.

Because of the high fluctuations of the envelope of the HPA input signal, the real HPA have to be modeled as nonlinear amplifiers. Following this idea and taking into account the recommendations concerning HPA modeling for OFDM communication systems presented in [18], Saleh and Rapp model of the HPA are assumed in our contribution [19,20]. It is well-known (e.g. [18]), that HPA for OFDM and MC-CDMA transmission systems can be modeled with advantage by the amplitude-to-amplitude *(AM/AM)* and the amplitude-to-phase *(AM/PM)* characteristics denoted as G_{u_x} and Φ_{u_x}, respectively.

Now, let us assume that the HPA input signal is given by

$$x(t) = u_x(t)e^{-j\phi(t)} \tag{1}$$

Then, the HPA output can be expressed by using and as

$$y(t) = G_{u_x}\left[u_x(t)\right]e^{-j\Phi_{u_x}\left[\phi(t)\right]} \tag{2}$$

In the case of Saleh model, AM/AM and AM/PM characteristics are given by [18,19]

$$G_{u_x} = \frac{\kappa_G u_x}{1+\chi_G u_x^2} \qquad \Phi_{u_x} = \frac{\kappa_\Phi u_x^2}{1+\chi_\Phi u_x^2} \tag{3}$$

In our considerations, the values of the Saleh model parameters of HPA have been set as follows: $\kappa_G = 2$, $\chi_G = \chi_\Phi = 1$ and $\kappa_\Phi = \pi/3$ what corresponds to the so-called Saleh model with simplified parameters [18]. On the other hand, Rapp model of HPA is described by the characteristics [20]

$$G_{u_x} = \frac{\kappa_G u_x}{\left[1+\left(u_x/O_{sat}\right)^{2s}\right]^{\frac{1}{2s}}} \qquad \Phi_{u_x} = 0 \tag{4}$$

Based on the recommendations of [18], the values of the Rapp model parameters such as $\kappa_G = 2$, $s = 3$ and $O_{sat} = 1$ are employed in our contribution.

The channel model considered in this paper is commonly uses the finite-length tapped delay line model of a frequency-selective multipath channel. In our simulations, we will consider the 6-tap multipath channel model where each tap is Rayleigh distributed. The cascade combination of the nonlinear HPA and the frequency-selective multipath channel can be considered to be the nonlinear fading channel.

3. Receiver structure

In this section, the receiver part of the MC-CDMA transmission system with multiuser detection is discussed. The block diagram of the baseband model of MC-CDMA receiver is given in Fig. 1. The receiver consists of the serial-to-parallel converter (S/P), blocks of the fast Fourier transformation (FFT), channel equalization (EQ) using zero forcing method, CP removal, bank of matched filters (BMF), block of linear or non-linear transformation (labeled as T), M-QAM demapper block and finally blocks of de-interleaving and convolution decoding. A more detailed description of the basic blocks of MC-CDMA receiver (FFT, EQ, CP removal and M-QAM demapper) can be found e.g. in [6].

The simplest receiver of MC-CDMA transmission system referred to as the single-user receiver can be obtained if the T-transformation block is represented by multiplication by a unit matrix. This kind of the receiver can provide acceptable BER if orthogonal spreading codes are used under the condition of linear HPA application and for AWGN channel scenario. If the transmission channel has to be modeled as a nonlinear fading channel, the single-user receiver cannot provide good performance. Because of the aforementioned properties of the transmission channel, the spreading code orthogonality is destroyed and consequently, MAI is increased. In order to avoid this performance degradation without requiring large back-offs in the transmitter amplifier, it becomes necessary to use multi-user detection techniques at the receiver side [6].

Conventional multi-user detectors such as MMSE-MUD e.g. are designed for linear environments and, as a result, might not exhibit enough performance improvement if the nonlinear models of HPA have to be taken into account.

It follows from the aforementioned facts that the MC-CDMA performance expressed by BER strongly depends on the spreading sequences, HPA model and the applied receiver. The spreading sequences selection depends on the PAPR of MC-CDMA signals as well as on the level of MAI due to cross-correlation properties of the particular spreading sequences. Because of the nonlinear distortion due to HPA and MAI inherency, it is expected that the application of nonlinear and multi-user receiver will overcome the linear receiver. Following this idea, a new nonlinear multi-user detector based on microstatistic filtering referred to as MSF-MUD has been introduced in [14]. The MSF-MUD uses piecewise linear filtering in conjunction with the threshold decomposition of the input signal, which introduces a nonlinear effect, to improve performance when a nonlinearity at the transmitter is present.

MMSE-MUD and MSF-MUD can be described by the same block scheme given in Fig.1. In the case of MMSE-MUD, the T-transformation block is represented by the multi-channel linear Wiener filter. The details concerning optimum MMSE-MUD design can be found e.g. in [6, 16]. On the other hand, in the case of MSF-MUD, the complex-valued multi-channel conventional microstatistic filter (C-M-CMF) is used as the T-transformation block. A simple outline of the optimum C-M-CMF design can be found e.g. in [14,17]. In the next section, we will present the design procedure of the optimum C-M-CMF for MSF-MUD in detail, originally presented in [21].

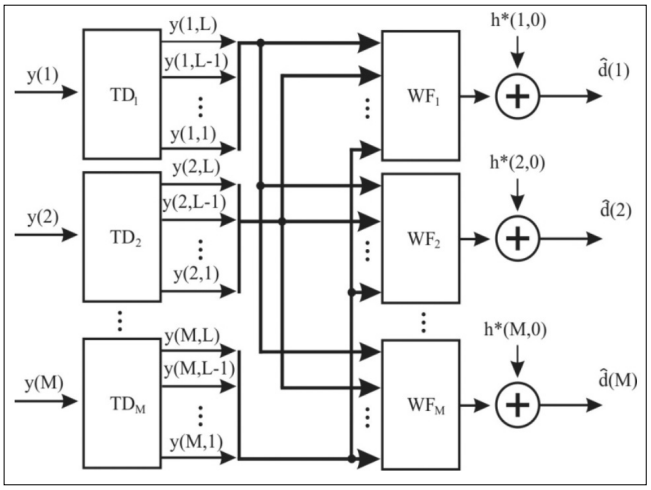


Figure 2.
Complex-valued multi-channel conventional microstatistic filter

3.1 Optimum C-M-CMF design procedure

A block scheme of the C-M-CMF is given in *Fig. 2.* Here, M, $y^{(i)}(n)$ and $\hat{d}^{(k)}(n)$ are the number of the input and output channels, the i-th input complex signal and the k-th output complex signal of the C-M-CMF, respectively. It can be seen from this figure that the C-M-CMF consists of M complex decomposers (TD) and the set of M complex multi-channel Wiener filter (C-M-WF).

Because the input signals to the TD are complex, a simple real-valued threshold decomposer (R-TD) applied in the M-CMF [22] cannot be used for their decomposition directly. In order to develop a suitable TD let us assume, that the complex signal $y^{(i)}(n)$ is expressed in the form

$$y^{(i)}(n) = Y^{(i)}(n)e^{j\phi(i,n)} \tag{5}$$

where

$$Y^{(i)}(n) = \left| y^{(i)}(n) \right| \tag{6}$$

and

$$\phi(i,n) = \arg\left[y^{(i)}(n) \right] \tag{7}$$

Generally, the performance of the i-th TD (TD_i) of C-M-CMF can be described as

$$D_C^{(i)}[y^{(i)}(n)] = \left[y^{(i,1)}(n) \ \ldots \ y^{(i,L)}(n) \right]^T \tag{8}$$

In this expression, $D_C^{(i)}[.]$ represents the complex threshold decomposition of the signal $y^{(i)}(n)$ due to TDi into a set of the L signals $y^{(i,j)}(n)$ and the superscript T signifies transposition.

The intention of the piecewise linear filter applications is e. g. to model a nonlinear system, where the nonlinearity is related especially to the magnitude of signals to be processed (e. g. AM/AM characteristic for the Saleh and Rapp models). On the other hand, the signal arguments $\phi(i,n)$ are usually uniformly distributed (e. g. in the case of the Rayleigh fading channel). Therefore, a decomposition of the argument of $y^{(i)}(n)$ does not result in any clear benefit. With regard to these considerations, the following decomposition of the complex signals can be used:

$$\begin{aligned} y^{(i,j)}(n) &= D_C^{(i)}[y^{(i)}(n)] = D_R^{(i)}[Y^{(i)}(n)]e^{j\phi(i,n)} = \\ &= \left[Y^{(i,1)}(n) \ \ldots \ Y^{(i,L)}(n) \right]^T e^{j\phi(i,n)} = \\ &= \left[y^{(i,1)}(n) \ \ldots \ y^{(i,L)}(n) \right]^T \end{aligned} \tag{9}$$

where

$$y^{(i,j)}(n) = Y^{(i,j)}(n)e^{j\phi(i,n)} \tag{10}$$

In (9), $D_R^{(i)}[.]$ represents the real-valued threshold decomposition of the positive real-valued signal $Y^{(i)}(n)$ due to R- TD_i into a set of the L signals given by

$$Y^{(i,j)}(n) = D_R^{(i,j)}[Y^{(i)}(n)] = \begin{cases} 0 & for \quad Y^{(i)}(n) < l_{j-1}^{(i)} \\ Y^{(i)}(n) - l_{j-1}^{(i)} & for \quad l_{j-1}^{(i)} < Y^{(i)}(n) \le l_j^{(i)} \\ l_j^{(i)} - l_{j-1}^{(i)} & for \quad l_j^{(i)} < Y^{(i)}(n) \end{cases} \tag{11}$$

for $1 \le j \le L$. The parameters $l_j^{(i)}$ constituting the vector $\mathbf{L}^{(i)} = \left[l_1^{(i)} \ l_2^{(i)} \ \ldots \ l_L^{(i)} \right]^T$ are the positive real-valued constants known as threshold levels of TD_i. The threshold levels are confined as $0 < l_1^{(i)} < \ldots < l_L^{(i)} = \infty$. The output signals of all TD are fed into the k-th M-WF (M-WF_k). Then, the k-th output of the C-M-CMF $\left(\hat{d}^{(k)}(n) \right)$ is given by the following expressions:

$$\hat{d}^{(k)}(n) = h_{(k,0)}^*(n) + \sum_{i=1}^{M} \hat{d}_k^{(i)}(n) \tag{12}$$

$$\hat{d}_k^{(i)}(n) = \sum_{j=1}^{L} d_k^{(i,j)}(n) \tag{13}$$

$$\hat{d}_k^{(i,j)}(n) = \mathbf{H}_k^{(i,j)H}\,\mathbf{Y}^{(i,j)}(n) \tag{14}$$

$$\mathbf{H}_k^{(i,j)} = \left[h_{(k,0)}^{(i,j)} \quad h_{(k,1)}^{(i,j)} \quad \dots \quad h_{(k,N)}^{(i,j)} \right]^T \tag{15}$$

$$\mathbf{Y}^{(i,j)}(n) = \left[y^{(i,j)}(n) \quad \dots \quad y^{(i,j)}(n-N) \right]^T \tag{16}$$

where the asterisk denotes complex conjugation, the superscript *H* denotes Hermitian transposition, the sequence $h_{(k,l)}^{(i,j)*}$ represents the part of impulse response of the M-WF$_k$ fed by signal $y^{(i,j)}(n)$. The constant term $h_{(k,0)}^{*}$ is applied in the C-M-CMF structure in order to obtain an unbiased C-M-CMF output. Now, let us define the block vector $\mathbf{H}_k^{(i)}$ containing the vectors $\mathbf{H}_k^{(i,j)}$ and the block vector $\mathbf{Y}^{(i)}(n)$ containing the vectors $\mathbf{Y}^{(i,j)}(n)$ as follows

$$\mathbf{H}_k^{(i)} = \left[\mathbf{H}_k^{(i,1)T} \quad \dots \quad \mathbf{H}_k^{(i,L)T} \right]^T \tag{17}$$

$$\mathbf{Y}^{(i)}(n) = \left[\mathbf{Y}^{(i,1)T}(n) \quad \dots \quad \mathbf{Y}^{(i,L)T}(n) \right]^T \tag{18}$$

Then, by using (17) and (18), the expression (13) can be obtained in this form

$$\hat{d}_k^{(i)}(n) = \mathbf{H}_k^{(i)H}\,\mathbf{Y}^{(i)}(n) \tag{19}$$

Finally, let us define the vector $\mathbf{H}_k$ and the vector $\mathbf{Y}(n)$ as follows

$$\mathbf{H}_k = \left[h_{(k,0)} \quad \mathbf{H}_k^{(1)T} \quad \dots \quad \mathbf{H}_k^{(M)T} \right]^T \tag{20}$$

$$\mathbf{Y}(n) = \left[1 \quad \mathbf{Y}^{(1)T}(n) \quad \dots \quad \mathbf{Y}^{(M)T}(n) \right]^T \tag{21}$$

Then, by using (20), (21) and (12), the *k*-th output of the M-CMF $\hat{d}_k(n)$ is given by

$$\hat{d}^{(k)}(n) = \mathbf{H}_k^H\,\mathbf{Y}(n) \tag{22}$$

In this expression, $\mathbf{H}_k^H$ represents the impulse response of M-WF$_k$. It can be seen from (22) that the C-M-CMF responses are given by a linear combination of vector elements obtained by the decomposition of the input signals of the filter. With regard to that fact, the C-M-CMF responses are still linear functions with respect to the C-M-CMF coefficients.

Let us assume that $y^{(i)}(n)$ (the input signals of C-M-CMF) and $d^{(k)}(n)$ (the desired signals) are stationary random processes. Because the C-M-CMF is a minimum mean square estimator, the set of parameters of the optimum time-invariant M-CMF given by

$$\mathbf{L} = \left[\mathbf{L}^{(1)T} \quad \dots \quad \mathbf{L}^{(M)T} \right]^T$$

and $\mathbf{H}_k$ is obtained as the solution that minimizes the cost functions

$$MSE(\mathbf{H}_k, \mathbf{L}) = E\left[e^{(k)}(n) e^{(k)*}(n) \right] = E\left[\left| e^{(k)}(n) \right|^2 \right] \tag{23}$$

where

$$e^{(k)}(n) = d^{(k)}(n) - \hat{d}^{(k)}(n) = d^{(k)}(n) - \mathbf{H}_k^H\,\mathbf{Y}(n) \tag{24}$$

In these expressions, *E*[.] denotes the expectation operator and (23) is the cost function referred to as the mean-square error of $d^{(k)}(n)$ estimation.

In order to minimize $MSE(\mathbf{H}_k, \mathbf{L})$, it is necessary to first estimate of the TD$_i$ parameters (**L** vector). Generally, for the L vector estimation, the scanning method, the genetic algorithm based method and the method of the cumulative distribution function can be applied [21]. The application of these method will lead onwards to an iterative procedure of the optimum C-M-CMF design [21]. On the other hand, it has been shown in [23] that if the C-M-CMF is used as the part of MSF-MUD applied for the M-QAM symbol detection, suboptimum parameters of C-M-CMF can be determined by a very efficient non-iterative method based on an analysis of the M-QAM symbol constellation diagram. It is well-known that the M-QAM symbols in the signal constellation are localized on a set of the circles with the center in the origin of the coordinate system. Then, the **L** vector can be estimated based on the estimation of the radiuses of the circles where the M-QAM symbols obtained at the output of BMF receiver are located. Because the mentioned radiuses can be estimated efficiently by using of the histogram of the modules of the complex M-QAM symbols detected by the BMF, the method under consideration is referred to as the histogram method. It has been shown in [23], that the histogram method can provide a robust estimation of the suboptimum parameters of the TDi providing very good performance of MSF-MUD.

If a suitable estimation of **L** is available, the optimum coefficients of the M-WFs can be computed as follows

$$\mathbf{H}_k^{opt} = \mathbf{R}^{-1}\mathbf{P}_k \tag{25}$$

where

$$\mathbf{R} = E\left[\mathbf{Y}(n)\mathbf{Y}^{H}(n) \right] \tag{26}$$

$$\mathbf{P}_k = E\left[d^{(k)}(n)\mathbf{Y}(n) \right] \tag{27}$$

In these expressions, **R** and $\mathbf{P}_k$ are the correlation matrix of vector sequence $\mathbf{Y}(n)$ and the cross-correlation vector of the desired signal $d^{(k)}(n)$ and vector sequence $\mathbf{Y}(n)$, respectively. **R** and $\mathbf{P}_k$ can be estimated by using a training sequence. Under the condition that $\mathbf{H}_k^{opt}$, **R** and $\mathbf{P}_k$ are computed according to (26) and (27), the mean-square error corresponding to optimum C-M-CMF is given by

$$\begin{aligned} &MSE(\mathbf{H}_k^{opt}, \mathbf{L}) = \\ &= \sigma_d^2 - \mathbf{P}^{H}\mathbf{H}_k^{opt} - \mathbf{H}_k^{optH}\mathbf{P} + \mathbf{H}_k^{optH}\mathbf{R}\mathbf{H}_k^{opt} = \\ &= \sigma_d^2 - \mathbf{P}^{H}\mathbf{H}_k^{opt} = \sigma_d^2 - \mathbf{P}^{H}\mathbf{R}^{-1}\mathbf{P}_k \end{aligned} \tag{28}$$

where

$$\sigma_d^2 = E\left[d^{(k)}(n) d^{(k)*}(n) \right] \tag{29}$$

By the evaluation of the vectors **L** and $\mathbf{H}_k^{opt}$, the design procedure of the optimum or suboptimum C-M-CMF and coincidentally MSF-MUD is got done.

4. Simulation results

In this section, we report on the simulation experiments that were carried out to study the effectiveness of the several variations of the described MC-CDMA transmission system performing through nonlinear fading channel. In our simulations, we consider the synchronous

uplink transmission (i.e. the transmission from the mobile terminal to the base station) for MC-CDMA systems employing N_c=128 subcarriers at the oversampling rate of 4, 16-QAM base-band modulation and 25% user load. For the sake of brevity, the perfect channel state information has been assumed at the receiver. For the channel equalization, zero forcing method has been applied. The convolution encoder with the coding rate R=1/3 followed by the interleaver with the block size of 32 has been used in order to improve the system performance.

The summary of the simulation parameters is listed in *Table 2*.

Parameter	Description
Modulation type	16-QAM
Cyclic prefix length	32 samples
Total number of sub-carriers	128
Oversampling rate	4
User load	25%
Interleaver block size	256
Channel model	6-tap multipath model, each tap Rayleigh distributed
Channel estimation	Perfect
Channel estimation	Perfect
Equalization	Zero forcing
Channel coding	Convolutional code with rate 1/3
Spreading codes	Walsh codes, Gold codes, orthogonal Gold codes with the period of L = 32 chips; Complementary Golay codes, polyphase Zadoff-Chu codes with the period of L = 31 chips
HPA	Saleh model, Rapp model
Receivers	MMSE-MUD, MSF-MUD
Simulation software	MATLAB

Table 2. Simulation parameters

In order to model the nonlinear HPA, Saleh and Rapp model of the HPA have been assumed. The particular parameters of the models have been listed in Section 2. The operating point of HPA has been given by the IBO parameter defined as $IBO_{dB}=10\log_{10}(P_{\max}/P_x)$, where $P_{\max}$ and P_x are the saturated power $P_{\max}$ and average input power P_x [14]. It is well known that the lower IBO is set, the higher nonlinear distortion due to HPA nonlinearity reveals. In our simulations, IBO has been set to 3 dB and 0 dB for Saleh model and 0 dB for Rapp model.

The channel model considered in our simulation has been represented by finite-length tapped delay line model of a frequency-selective multipath channel. Here, we used 6-tap multipath model where each tap is Rayleigh distributed. The intention of contribution is to study MSF-MUD performance properties if it is applied to a MC-CDMA system performing through frequency-selective fading channel and simultaneously, if Walsh codes, Gold and orthogonal Gold codes, polyphase Zadoff-Chu codes and complementary Golay codes are used as the spreading sequences.

In order to decrease the receiver complexity, very simple C-M-CMF with two-level TDi (L=2) has been exploited in the MSF-MUD structure. For the C-M-CMF design, the procedure described in the Section 3 has been applied. The matrix $\mathbf{R}$ and $\mathbf{P}_k$ have been estimated by using the training sequence consisting of 500 uniformly distributed of 16-QAM symbols. The training sequence has been transmitted before the information data transmission. For all TD of C-M-CMF, the same value of the threshold level $I_1^{(i)}$ has been applied. By using the histogram method mentioned in the Section 3, $I_1^{(i)}$ has been set to 0,6 for i=1,2,...,M. In order to illustrate the MSF-MUD performance, MMSE-MUD has been employed as the MC-CDMA receiver, as well. For the MMSE-MUD design, the procedure described in [6], exploiting the above mentioned training sequence has been applied.

The performance of the above described variations of the MC-CDMA systems (different HPA models, different receivers and different spreading codes) has been evaluated by *BER vs.* E_b/N_0 for the different values of IBO setting. The obtained results are given in *Figs. 3-5*.

Let us assume firstly that HPA is modeled according to Saleh model. Here, *BER vs.* E_b/N_0 for $IBO=0dB$ is given in *Fig. 3.* It can be seen from this figure that the minimum *BER* equal to 10^{-2} is reached at $E_b/N_0=40dB$. It indicates very clearly, that both receivers and all codes provide insufficient results for the analyzed scenario with strongly nonlinearly distorted MC-CDMA signals ($IBO=0dB$). In spite of that fact, it can be observed that the application of MSF-MUD in conjunction with Golay codes provides the evident improvement of the MC-CDMA performance in comparison with the application of the other codes in this scenario. On the other hand, it can be seen from the Fig. 3, that data receiving fails completely if MMSE-MUD is employed.

The simulation results for the Saleh model of HPA and for $IBO=3dB$ are presented in *Fig. 4*. It can be seen from this figure that the worst performance of MC-CDMA is provided if Walsh codes are employed, independently on the applied receiver. Here, this effect can be explained by the high PAPR of MC-CDMA signal (Table 1) in combining with the severe nonlinear distortion resulting in the loss of Walsh code orthogonality. On the other hand, the best performance is provided by joint application of MSF-MUD and Golay codes. This performance follows from relatively low PAPR of MC-CDMA signal (Table 1), when Golay codes are used and simultaneously, by MSF-MUD ability to compensate nonlinear distortion due to HPA. Fig. 4 indicates also that MMSE-MUD can provide the best results if the Golay codes are exploited. However, the results provided by the MMSE-MUD and the Golay codes are still worse than that provided by MFS-MUD what is due to by the nonlinear structure of the MSF-MUD. The joint application of the other codes (Gold codes, orthogonal codes and Zadoff-Chu codes) and both receivers provides the greater *BER* than that of Golay codes application but still better than that of the Walsh codes exploiting. The presented results confirm clearly that MSF-MUD overcomes still MMSE-MUD.

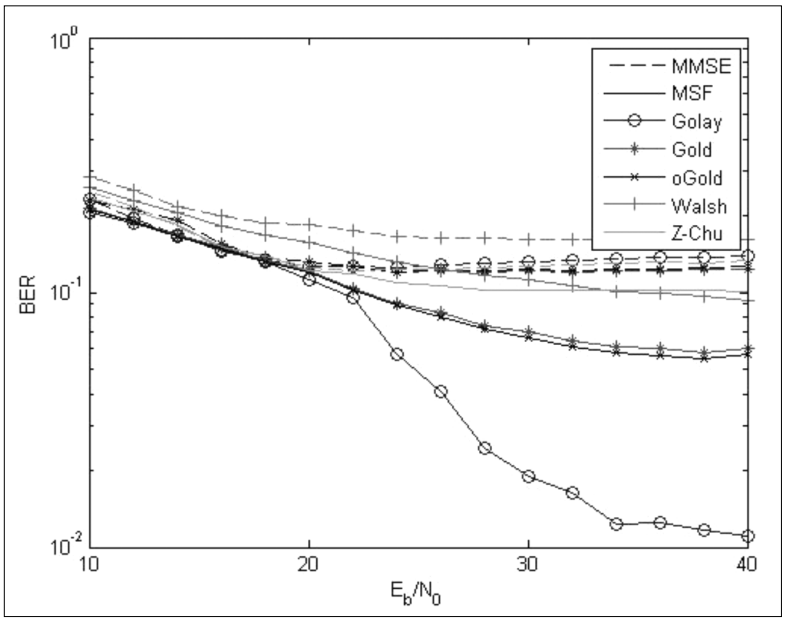


Figure 3.
Performance of the system for different spreading sequences, Saleh model of the HPA and IBO=0dB

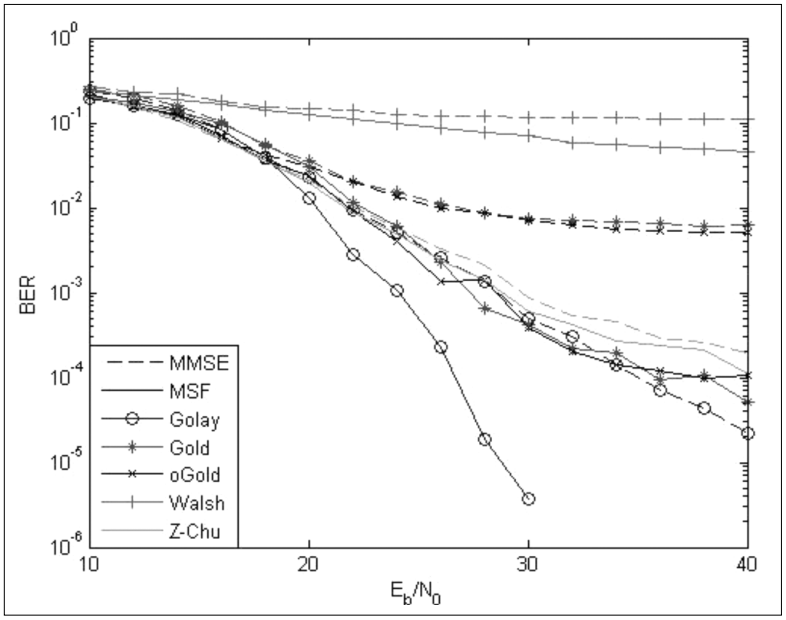


Figure 4.
Performance of the system for different spreading sequences, Saleh model of the HPA and IBO=3dB

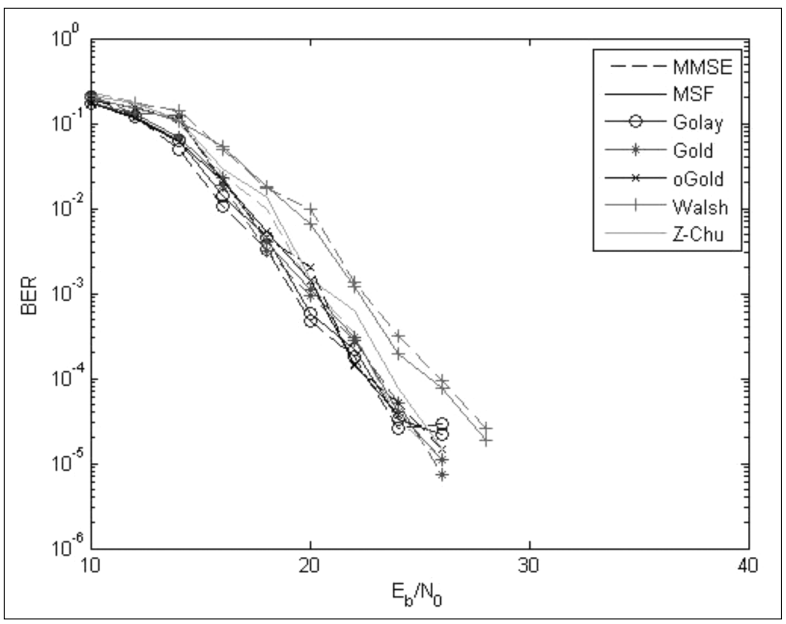


Figure 5.
Performance of the system for different spreading sequences, Rapp model of the HPA and IBO=0dB

Finally, we can describe the MC-CDMA performance for the scenario where Rapp model of HPA and $IBO = 0dB$ are set. Here, the obtained results given in *Fig. 5* show that except Walsh codes applications, the receivers and the other codes in view provide almost the same curves of *BER*. This behaviour of the Walsh code application can be explained by the same manner as in the case of the scenario analysed in the previous passage. The comparable performance of MSF-MUD and MMSE-MUD follows from the fact that the nonlinear distortion due to Rapp model of HPA at $IBO = 0dB$ is not so strong and therefore MSF-MUD cannot offer significant performance improvement over MMSE-MUD.

5. Conclusion

In this contribution, we have studied MC-CDMA transmission system properties performing through the frequency-selective fading channel for the uplink scenario. Within this study, we are focused on the analyses of the MC-CDMA performance if the different models of nonlinear HPA (Saleh and Rapp model), different spreading sequences (Walsh codes, Gold and orthogonal Gold codes, polyphase Zadoff-Chu codes and complementary Golay codes) and different multiuser receivers (MSF-MUD and MMSE-MUD) are considered.

The results of the computer simulations for a number of variations of the outlined scenario have shown that the best performance expressed by *BER vs.* E_b / N_0 is always provided if MSF-MUD in combination with Golay codes is employed. These results have confirmed fully the results obtained for the same MC-CDMA system configuration performed through AWGN channel [17]. This performance can be explained by the low level of the Golay code PAPR and by the ability

of the MSF-MUD to compensate of the nonlinear distortion due to its nonlinear piecewise structure. The performance improvement is more remarkable, if Saleh model of HPA is employed. If the Rapp model of the HPA has to be considered, the performance of the MSF-MUD and MMSE-MUD are almost the same. Following [24], this effect could be explained by the migration style of M-QAM symbols of signal constellation detected at the output of the receiver BMF.

If we sum up the results of the MC-CDMA performance analyses presented in this contribution and the outputs of [17] and [24], it can be recommended to apply MSF-MUD and Golay codes in the structure of MC-CDMA transmission systems. A possible improvement of MC-CDMA performance for small IBO values (e.g. $IBO = 0dB$) if Saleh model of HPA is used is an open question. It could be expected, that the solution of that problem could be provided by MSF-MUD exploiting a bit more complex TD, i.e. TD with more threshold level than one. The solution of the problem will be the topic of our follow-up research.

Acknowledgements

This work was supported by COST Action IC0803; "RF/Microwave Communication Subsystems for Emerging Wireless Technologies (RFCSET)" and by VEGA 1/0045/10; "Nové metódy spracovania signálov pre rekonfigurovatel'né bezdrotové senzorové siete (Advanced Signal Processing Techniques for Reconfigurable Wireless Sensor Network)".

Authors

JURAJ GAZDA was born in 1984 in Kosice, Slovakia. He received his M.Sc. degree in Electronics and Telecommunications in 2007 from the Faculty of Electrical Engineering and Informatics, Technical University of Kosice and currently he is working towards his PhD at the same university. During '06-'07 he spent one semester at Delft Univ. of Technology, The Netherlands. Since 2009, he is with Research group in Electromagnetism and Communications, La Salle, Univ. Ramon Llull, Barcelona as a guest researcher working in the area of nonlinear effects in OFDM. His research interests include effects of non-linear amplification on multicarrier transmission schemes and design of advanced receivers for Beyond 3G and 4G transmission systems.

PETER DROTÁR was born in 1984 in Kosice, Slovakia. He received the M.Sc. in Electronics and Telecommunications from Technical University of Kosice in 2007 and currently he is a PhD candidate at the same university. During the years he visited and worked with Ramon Llull Univ., Barcelona and Hamburg Univ. of Technology. His research interests include power amplifier nonlinearities in MIMO–OFDM systems and receiver design.

DUSAN KOCUR was born in 1961 in Kosice, Slovakia. He received his M.Sc. (Ing.) and Ph.D. (CSc.) in Radioelectronics from the Faculty of Electrical Engineering, Technical University of Kosice, in 1985 and 1990, respectively. Nowadays, he is the full professor at the Department of Electronics and Multimedia Communications of his Alma Mater. His research interests are wireless multicarrier communication systems (OFDM, MIMO-OFDM, MC-CDMA) and UWB radar signal processing.

PAVOL GALAJDA was born in 1963 in Kosice, Slovakia. He received the Ing. (MSc.) and CSc. (PhD.) in radioelectronics from the Faculty of Electrical Engineering, Technical University (TU) of Kosice, in 1986 and 1995, respectively. At present he is an associate professor at the Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, TU of Kosice. He is a member of Czech and Slovak Radioelectronics Engineering Society. His research interest are in nonlinear systems and communications and include chaos, multicarrier communications and software defined radio.

References

[1] S. Hara, R. Prasad, "Overview of multicarrier CDMA", IEEE Communications Magazine, Vol. 35, No. 12, pp.126–133, December 1997.

[2] K. Fazel, S. Kaiser, Multicarrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX, 2nd ed., John Wiley & Sons, England, 2008.

[3] S. Nobilet, J. Helard, D. Mottier, "Spreading sequences for uplink and downlink MC-CDMA Systems: PAPR an MAI minimization" European Transactions on Telecommunications, 2002.

[4] B.-J. Choi, E.-L. Kuan, L. Hanzo, "Crest-factor study of MC-CDMA and OFDM", IEEE Vehicular Technology Conference, Vol. 1, pp.233–237, 1999.

[5] A.-M. Mourad, A. Gueguen, R. Pyndiah, "Impact of the spreading sequences on the performance of forward link MC-CDMA systems", IEEE 8th International Symposium on Spread Spectrum Techniques and Applications, Vol. 30, pp.683–687, September 2004.

[6] L. Hanzo, M. Munster, B. Choi, T. Keller, OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting, John Wiley & Sons, England, 2003.

[7] J. Gazda, P. Drotár, D. Kocur, P. Galajda, R. Blicha, "Joint Evaluation of Nonlinear Distortion Effects and Signal Metrics in OFDM based Transmission Systems," Acta Electrotechnica et Informatica, Vol. 9, No. 4, pp.55–60, December 2009.

[8] T. Jiang, Y. Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," IEEE Trans. Broadcasting, Vol. 54, No. 2, pp.267–268, June 2008.

[9] M. Deumal, A. Behravan, T. Eriksson, J.L. Pijoan, "Evaluation of performance improvement capabilities of PAPR-reduction methods", Wireless Personal Communications, Vol. 47, No. 1, pp.137–147, October 2008.

[10] A. Katz, "Linearization: Reducing Distortion in Power Amplifiers", IEEE Microwave Magazine, Vol. 2, No. 4, pp.37–49, December 2001.

[11] M. Deumal,
Multicarrier communication systems with low sensitivity to nonlinear amplification,
Ph.D. Thesis,
Universitat Ramon Llull, Barcelona, Spain, 2008.

[12] N. Hathi, I. Rodrigues, I. Darwazeh, J. O'Reilly,
"Analysis of the influence of Walsh-Hadamard code allocation strategies on the performance of the multicarrier CDMA systems in the presence of HPA non-linearities",
The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications,
Vol. 3, pp.1305–1309, 2002.

[13] J. Tellado, L.M. C Hoo, J.M. Cioffi,
"Maximum likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding",
IEEE Transactions on Communications,
Vol. 51, No. 2, pp.218–228, February 2003.

[14] J. Krajnak, M. Deumal, P. Pavelka, et al.
"Multi-user detection of nonlinearly distorted MC-CDMA symbols by microstatistic filtering",
Wireless Personal Communications,
Vol. 47, No. 1, pp.147–160, October 2008.

[15] P. Silva, R. Dinis,
"Turbo multiuser detection for MC-CDMA signals with strongly nonlinear transmitters",
International Symposium on Communications and Information Technologies,
pp.1171–1175, October 2007.

[16] S. Verdu,
Multi-user Detection,
Cambridge University Press, New York, 1998.

[17] P. Drotár, J. Gazda, D. Kocur, P. Galajda,
"Joint Microstatistic Multiuser Detection and Cancellation of Nonlinear Distortion Effects for the Uplink of MC-CDMA Systems Using Golay Codes,"
International Journal of Electronics, Communications and Computer Engineering ,
Vol. 1, No. 2, pp.87–93, 2009.

[18] D. Falconer, T. Kolze, Y. Leiba,
"Proposed System Imparment Models,"
IEEE 802.16.1, pp.00/15, 2000.

[19] A.M. Saleh,
"Frequency-independent and frequency dependent non-linear models of TWTA",
IEEE Transactions on Communications,
Vol. 29, pp.1715–1720, November 1981.

[20] C. Rapp,
"Effects of HPA-nonlinearity on 4-DPSK-OFDM-signal for digital sound broadcasting system",
Proc. 2nd European Conference Satellite Com.,
pp.179–184, 1991.

[21] D. Kocur, J. Krajnak, S. Marchevsky,
"Piece-Wise Linear Multi-Channel Complex Microstatistics Filters",
The 10th International Conference on Intelligent Engineering Systems 2006 (INES'2006),
London, UK, June 26-28, pp.53–56, 2006.

[22] G.R. Arce,
"Microstatistics in signal decomposition and the optimal filtering problem,"
IEEE Trans. Signal Proc.,
Vol. 40, pp.2669–2682, 1992.

[23] P. Pavelka, J. Krajcák, P. Galajda, D. Kocur,
"Efficient design procedure of Microstatistic Multi-user Detector for nonlinearly distorted MC-CDMA,"
Proc. of IEEE Int. Conf. Radioelektronika,
pp.1–6, April 2007.

[24] P. Drotár, J. Gazda, D. Kocur, P. Galajda,
"Effects of Spreading Sequences on the Performance of MC-CDMA System with Nonlinear Models of HPA",
Radioengineering,
Vol. 18, No. 1, pp.48–54, April 2009.

Call for Papers

IEEE VTS Connecting the Mobile World

VTC2011-Spring
BUDAPEST
Beyond the Generations Game

2011 IEEE 73rd Vehicular Technology Conference
15–18 May 2011 Budapest, Hungary
Submissions by 30 Sep 2010

VTC is the premier semi-annual wireless communications conference of the IEEE Vehicular Technology Society (VTS). Researchers, industry professionals and academics dedicated to innovation across the broad field of wireless systems and networks are cordially invited to contribute to the on-going scientific dialogue across this vibrant community. In 2011, the IEEE VTC Spring will be held in Budapest, Hungary, an enchanting city also often referred to as the pearl of the Danube. You are invited to submit papers in all areas of wireless communications, networks, services, and applications. Potential topics are solicited in, but are not limited to the following categories:

- **Antennas and propagation**
- **Transmission techniques**
- **MIMOs and space-time-frequency processing**
- **Cognitive radio and spectrum sensing**
- **Cooperative comms, distributed MIMOs and relaying**
- **Wireless multiple access techniques**
- **Wireless networks**
- **Ad hoc, mesh and sensor networks**
- **Mobile satellite and positioning systems**
- **Wireless applications and services**
- **Vehicular electronics and telematics**

General Chair
Lajos Hanzo
University of Southampton, UK

Technical Program Chairs
Andrea Conti *University of Ferrara, Italy*
Wei Chen *Tsinghua University, China*
Iain B. Collings *CSIRO, Australia*

Prospective authors are encouraged to submit a 5-page full paper (or a 2-page extended abstract including results) through the conference web site. Submission deadline 30 September 2010

For more information, visit www.vtc2011spring.org

Call for Papers

IEEE VTS Connecting the Mobile World

CAPS2011

Fourth Workshop on Context Awareness for Proactive Systems
15–16 May 2011, Budapest, Hungary

Following CAPS2005, 06 and 07, the fourth Workshop on Context Awareness for Proactive Systems will be held in conjunction with IEEE VTC2011-Spring in Budapest, Hungary.

Proactive computing and communication systems are connected to the physical world by means of sensors and actuators which are used to both measure and manipulate the physical surroundings. The gathered environmental data serve proactive systems as stimuli to which they respond in terms of providing users with appropriate resources, information, and services.

In order to fulfil this task, proactive systems need to and benefit from taking users' contexts into account, i.e. using the gathered sensor data to infer users' state, activities, goals, and so on and to adjust their proactive behaviour accordingly. In addition, mobile and pervasive environments have turned out to be a promising application area for proactive systems. Deploying proactive systems in such rapidly changing environments enforces the need to make them context-aware.

Context awareness in proactive systems opens up a lot of novel opportunities, however, it also poses new challenges upon proactive computing technology. The major objective of the workshop is to study and explore these challenges and proposed ways of meeting them. This includes research on modelling and representing context in proactive computing systems, frameworks and architectures for context handling, sensor and actuator management, context reasoning, learning, and prediction as well as on modelling, recognising and fulfilling user demand.

Papers on following (but not limited to) are invited:

- Context information gathering and data management
- Frameworks and architectures for context-aware systems
- User demand recognition and modelling
- User demand recognition and modelling
- Context reasoning
- Sensor and actuator management
- Context modelling and representation
- Context learning and prediction techniques
- Context-based resource, information, and service provisioning
- Infrastructures for proactive systems
- Context aware applications

Submission of full papers	**15 October 2010**
Notification of acceptance	**15 January 2011**
Camera Ready Papers	**15 February 2011**

For more information, visit www.vtc2011spring.org

WMNC 2010

3rd Joint IFIP Wireless Mobile Networking Conference
13-15 October 2010 – Budapest, Hungary

The 3rd Joint IFIP Wireless Mobile Networking Conference (WMNC'2010), combining PWC (Personal Wireless Communications), MWCN (Mobile and Wireless Communication Networks Conference) and WSAN (Wireless Sensors and Actor Networks Conference) into one event, will be held between 13 and 15 October 2010 in Budapest, Hungary. The conference will be organized by Óbuda University and Scientific Association for Infocommunications Hungary, the co-organizer is Budapest University of Technology and Economics.

The WMNC 2010 papers will be electronically published by IEEE Xplore.
In addition, authors of selected papers will be invited to submit an extended version of their papers to a special issue of the Telecommunications Systems journal edited by Springer (JCR indexed).

For WMNC 2010 was approved on the last TC6 meeting (Paris, December 2009)
the following budget (TC6 will pay for it): 2 STG (Student Travel Grant), 750 Euros each
1 BPA (Best Paper Award), 500 Euros

IFIP TC6 pay directly the money to the students/authors.

13 October	14 October	15 October
08.30-09.30 Registration		
Room I (Pauli):		
09.30-10.00 Opening, Keynote presentation	09.00-10.30 **Wireless multimedia** / II	09.30-10.00 Keynote presentation
10.00-12.00 **Radio resource management** / I	*10.30 Coffee brake*	10.00-11.30 **Mobility and security management** / II
	11.00-12.30 **Wireless sensor networks** / II	11.30-12.00 Closing
12.00 Lunch	*12.30 Lunch*	*12.00 Lunch*
13.00-15.00 **Quality of Service** / I		
15.00 Coffee break		
15.30-17.00 **Wireless multimedia** / I		
17.00-18.00 Discussion		
Room II:		
13.00-14.30 **Mobility and security management** / I	09.00-10.30 **Radio resource management** / II	
14.30 Coffee break	*10.30 Coffee brake*	
15.00-16.30 **Wireless sensor networks** / I	11.00-12.30 **Quality of Service** / II	
16.30-18.00 Discussion	*12.30 Lunch*	
	Excursion to Visegrád + Gala Dinner	

More information: http://www.kvk.uni-obuda.hu/wmnc2010

REGISTRATION DEADLINE

Deadline for registration September 24, 2010, CET 24:00. After this date only on-site registration is possible.

REGISTRATION FEE (without accommodation)

Late (after July 15)

Academic/Industrie rate	*550 Euro*
IFIP/IEEE/HTE members rate	*500 Euro*
Phd rate	*450 Euro*

The registration fee includes 69.500 HUF (including VAT) catering that appears on the written invoice.

REGISTRATION SECRETARIAT

Maria TEZSLA
Scientific Association for Infocommunications
Kossuth Lajos tér 6-8, H-1055 Budapest, Hungary
Phone: +36 1 353 1027 Fax: +36 1 353 0451
e-mail: info@hte.hu
www.hte.hu