CTRQ 2012

The Fifth International Conference on Communication Theory, Reliability, and Quality of Service

ISBN: 978-1-61208-192-2

April 29 - May 4, 2012

Chamonix / Mont Blanc, France

CTRQ 2012 Editors

Eugen Borcoci, Politehnica University of Bucharest, Romania

Pascal Lorenz, University of Haute Alsace, France

CTRQ 2012

Foreword

The Fifth International Conference on Communication Theory, Reliability, and Quality of Service [CTRQ 2012], held between April 29th and May 4th, 2012 in Chamonix / Mont Blanc, France, continued a series of events focusing on the achievements on communication theory with respect to reliability and quality of service. The conference also brought onto the stage the most recent results in theory and practice on improving network and system reliability, as well as new mechanisms related to quality of service tuned to user profiles.

The processing and transmission speed and increasing memory capacity might be a satisfactory solution on the resources needed to deliver ubiquitous services, under guaranteed reliability and satisfying the desired quality of service. Successful deployment of communication mechanisms guarantees a decent network stability and offers a reasonable control on the quality of service expected by the end users. Recent advances on communication speed, hybrid wired/wireless, network resiliency, delay-tolerant networks and protocols, signal processing and so forth asked for revisiting some aspects of the fundamentals in communication theory. Mainly network and system reliability and quality of service are those that affect the maintenance procedures, on the one hand, and the user satisfaction on service delivery, on the other hand. Reliability assurance and guaranteed quality of services require particular mechanisms that deal with dynamics of system and network changes, as well as with changes in user profiles. The advent of content distribution, IPTV, video-on-demand and other similar services accelerate the demand for reliability and quality of service.

We take here the opportunity to warmly thank all the members of the CTRQ 2012 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to CTRQ 2012. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the CTRQ 2012 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that CTRQ 2012 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of communication theory, reliability and quality of service.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed their stay in the French Alps.

CTRQ Advisory Committee:

Eugen Borcoci, Politehnica University of Bucharest, Romania
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Pascal Lorenz, University of Haute Alsace, France
Raj Jain, Washington University in St. Louis, USA

CTRQ 2012 PROGRAM COMMITTEE

CTRQ Advisory Committee

Eugen Borcoci, Politehnica University of Bucharest, Romania
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Pascal Lorenz, University of Haute Alsace, France
Raj Jain, Washington University in St. Louis, USA

CTRQ 2012 Technical Program Committee

Gergely Acs, INRIA, Rhone-Alpes - Montbonnot, France
Himzo Bajric, Joint Stock Company BH Telecom - Sarajevo, Bosnia and Herzegovina
Jasmina Barakovic Husic, BH Telecom – Sarajevo, Bosnia and Herzegovina
Eugen Borcoci, Politehnica University of Bucharest, Romania
Bruno Checcucci, Perugia University, Italy
Karel De Vogeleer, Blekinge Institute of Technology (BTH) - Karlskrona, Sweden
Javier Del Ser Lorente, TECNALIA Research & Innovation, Spain
Michel Diaz, LAAS, France
Manfred Droste, Universität Leipzig, Germany
Ali El Masri, Troyes University of Technology -Troyes, France
Andras Farago, The University of Texas at Dallas - Richardson, USA
Julio César García Alvarez, Universidad Nacional de Colombia Sede Manizales, Colombia
Bogdan Ghita, University of Plymouth, UK
Marc Gilg, Université de Haute Alsace, France
Antti Hakkala, University of Turku, Finland
Bjarne J. Helvik, NTNU, Norway
Robert Ching-Hsien Hsu, Chung Hua University, Taiwan
Mohsen Jahanshahi, Islamic Azad University - Central Tehran Branch, Iran
Brigitte Jaumard, Concordia University, Canada
Sokratis K. Katsikas, University of Piraeus, Greece
Wojciech Kmiecik, Wroclaw University of Technology, Poland
Michal Kucharzak, Wroclaw University of Technology, Poland
Archana Kumar, Delhi Institute of Technology & Management - Haryana, India
Pascal Lorenz, University of Haute Alsace, France
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Wail Mardini, Jordan University of Science and Technology, Jordan
Rubens Matos, Federal University of Pernambuco, Brazil
Rick McGeer, HP Labs - Palo Alto, USA
Amalia N. Miliou, Aristotle University of Thessaloniki, Greece
Jean-Claude Moissinac, TELECOM ParisTech, France
Petros Nicopolitidis, Aristotle University of Thessaloniki, Greece
Shahram Nourizadeh, Domocare - AXON, France
Serban Obreja, University Politehnica of Bucharest, Romania
Jun Peng, University of Texas - Pan American - Edinburg, USA
Karim Mohammed Rezaul, Glyndwr University - Wrexham, & St. Peter's College of London, UK
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Janusz Romanik, Military Communications Institute – Warszawska, Poland

Simon Pietro Romano, University of Napoli Federico II, Italy
Giovanni Russello, CREATE-NET, Italy
Sébastien Salva, Université de Clermont Ferrand/IUT d'Aubière – Cézeaux, France
Iraj Saniee, Bell Labs, Alcatel-Lucent - Murray Hill, USA
Susana Sargento, University of Aveiro/Institute of Telecommunications, Portugal
Panagiotis Sarigiannidis, University of Western Macedonia - Kozani, Greece
Zary Segall, University of Maryland Baltimore County, USA
Dimitrios Serpanos, ISI/RC Athena & University of Patras, Greece
Adam Smutnicki, Wroclaw University of Technology, Poland
Vasco Soares, Instituto de Telecomunicações / University of Beira Interior / Polytechnic Institute of Castelo Branco, Portugal
Tae-Eung Sung, Korea Institute of Science and Technology Information, Korea
Vicraj Thomas, BBN Technologies, Inc., USA
Pierre F. Tiako, Langston University, USA
Kishor Trivedi, Duke University - Hudson Hall, USA
Elena Troubitsyna, Åbo Akademi University, Norway
Dimitrios D. Vergados, University of Piraeus, Greece
Seppo Virtanen, University of Turku, Finland
Krzysztof Walkowiak, Wroclaw University of Technology, Poland
Abdulrahman Yarali, Murray State University, USA
Nataša Živic, University of Siegen, Germany

Copyright Information

Table of Contents

QoS-based Autonomic Service Component for Service Delivery

Houda Alaoui Soulimani, Philippe Coude
Department of Architecture and R&D
Société Française de Radiotéléphonie (SFR)
Paris-France
{houda.alaouisoulimani, philippe.coude}@sfr.com

Noëmie Simoni
Department INFRES
TELECOM ParisTech - LTCI - UMR 5141 CNRS
Paris - France
simoni@telecom-paristech.fr

Abstract **- Nowadays, with the rising complexity of the service personalization in a heterogeneous and mobile context and the need to satisfy the End-to-End QoS, the service resources should be taken into account as a prominent resource as well as the network resources. Therefore, a high degree of self-sufficiency, self-management and automation is required in the service resource "service component" to enhance the service delivery. In this paper, we propose an autonomic service component "ASC" based on an integrated QoS–agent that self-controls and self-manages the service resources to dynamically adapt the service resources in response to changing situations during the user's session. To explain our proposal, we detail the mechanisms used to provision and monitor the service resource and to verify its conformity to the QoS contract established between the service providers and the customers during the exploitation. The issue of the ASC self-control and self-management is addressed according to the functional and non-functional (QoS) requirements in order to ensure the service continuity during the service delivery.**

Keywords - Autonomic Service Component; End-to-End QoS; QoS-agent; Mobility; Service Delivery.

I. INTRODUCTION

In the last years, with the fast evolution of the new generation networks and services (NGN/NGS), the user wants to access his personalized services while switching between different terminals or access networks. All these types of spatial mobility must be executed without impacting the End-to-End QoS. However, with the increasing demands on service delivery from customers, Providers are faced with the challenge of provisioning and managing their service resources in an efficient, cost-effective and flexible way. Currently, there are a set of mechanisms used to provision and monitor the network resources which caters to the specific QoS requirements of the applications at the transport network. These solutions permit to ensure and maintain the data delivery which guarantees the QoS of the transport network during the media session. These mechanisms do not take into account the QoS of the service resources knowing that there are some problems that arise during the user's session because of the services behavior in the platforms. Nowadays, QoS solutions which operate only on the network resources are no longer sufficient because the delivered QoS to a given client may be affected by many factors including the performance of the service component, the hosting platform or the underlying transport network. This is the reason why we have thought of broadening the QoS control and management in the service resource "service component" during the user's session. In fact, to guarantee the user satisfaction with regards to the service delivery in a heterogeneous and mobile context, we should provision and monitor the service resources as well as the network resources, and dynamically re-provision the service resources by benefiting from the ubiquitous services offered by different providers in the service platforms. To summarize, providing a service to the users guaranteeing the End-to-End QoS requires a horizontal QoS management at the service layer which is added to the existing QoS management at the access and transport layers.

However, the Service Oriented Architecture (SOA) [1] plays a principal role in allowing the creation of applications as a composition of independent service components offered by different providers. In addition, many available service components provide identical functionalities albeit with different quality of service capabilities. The following questions arise: How to rethink the service component to include the QoS control and management at the service level during the user's session? How a service component can enhance the service delivery? How to cover all the user's preferences and requirements (functional and non-functional QoS) and ensure a better performance in a heterogeneous and mobile context?

The purpose of this paper is to highlight the service component features that intend to decentralize and automate the QoS control and management for a flexible services composition with a better QoS performance. The benefit of this solution is to conceive a service component able to dynamically and autonomously react in real time to a change in the QoS contract during the user's session, such as availability. That's the reason why we propose in this paper an autonomic service component "ASC" based on a QoS-agent to monitor the QoS in real time during the processing. This QoS-agent triggers an event in the ASC environment in the case of a deterioration of the QoS. The receipt of an event activates the necessary mechanisms to change the current ASC used by another equivalent service component having the same functionality and QoS in order to maintain the QoS at the service level of the architecture and consequently to guarantee the service delivery.

The remainder of this paper is organized as follows: In Section II, we discuss some of the works related to the topic of this article. Section III details our proposal by explaining the service component features and mechanisms that are used to self-control and self-manage the QoS during the user's session. In Section IV, we present a scenario illustrating a utilization case of the ASC. In Section V, we evaluate the performance of the ASC through an implementation. Section VI concludes the paper.

II. RELATED WORK

In recent years, many researchers have focused their efforts on the service composition, the autonomic service component and the autonomic service architecture (ASA) especially in terms of providing the QoS to consumers in a dynamic environment.

Farha et al. [2] presented in their paper a generic Autonomic Service Architecture (ASA) to deliver applications and services over an all-IP infrastructure. Their solution is based around the concepts of SOA, virtualization and service delivery. The ASA proposes a generic framework to deal with the activation, provisioning, management and termination of network resources in an autonomic way. The ASA acts especially on the network resources to deliver a given service to customers. Cheng et al. [3] presented an approach to the autonomic service architecture (ASA) by proposing a framework for the automated management of internet services and their underlying network resources. This framework ensures the service delivery at the transport layer. The drawback of these approaches lies in the fact that they act on the network resources and do not cover the service resources which are essential to the service delivery.

Zhang et al. [4] proposed a framework to identify QoS problems in the SOA when a business process fails to deliver the quality of service. Their proposal is based on a set-covering algorithm which is used to select the locations of run-time service data collection or probes. The framework creates a dependency matrix to denote the relationships between the data recorded by probes and the service status. A diagnosis is then used to identify potential faulty services. Zhai et al. [5] presented a framework to repair failed services by replacing them with new services and ensuring that the new service process still meets the user specified QoS constraints. The drawback of these approaches lies in the fact that they present centralized solutions to identify failed services during a session. In opposition to the previous approaches, our proposition is based on a distributed self-management and self-control of the QoS to dynamically maintain the service session without impacting the end-to-end QoS.

Zambonelli et al. [6] proposed autonomic service components that are able to dynamically adapt their behavior in response to changing situations. They present mechanisms to enable the components to self-express the most suitable adaptation. The components acquire the proper degree of self-awareness to put into action the self-adaptation and the self-expression schemes. Liu and Parashar [7] presented a framework which enables the development of autonomic elements and the formulation of autonomic applications to have a dynamic composition of autonomic elements. They propose rules and mechanisms for the dynamic composition of autonomic components so that the computational behavior of the elements as well as their compositions and interactions can be managed at the runtime using dynamically injected rules. These approaches focus on the autonomic application creation paradigms and the behavior of the service component adaptation. However, these current solutions focus on self-adaptation and self-awareness of the service component and don't take into account the dynamic aspect of the session at the service level. In opposition to these approaches, we propose a distributed self-control and self-management of the QoS in each service component that is based on overcoming QoS violations without modifying any QoS parameter in the SLA. The QoS-based autonomic service component aims to enhance the service delivery in a dynamic, mobile and heterogeneous context.

III. PROPOSITION

In order to satisfy the SLA contract established between the customers and the providers, a more flexible and adapted service composition based QoS is desired to enhance the service delivery during the session mobility. Its principle task is to integrate the QoS control and management at the service level of the architecture. To do so, we should automate and distribute the QoS control and management at each service component which is involved in the delivery of a given service. That's the reason why we propose in this paper an ASC to fulfill such purpose. To explain our proposal, we firstly detail in section A the architectural and the functional aspects of the ASC. Secondly, we explain in section B our QoS model which is the basis of the service component self-management. Finally, we explain in section C the mechanism applied by the ASC during each operational step to self-manage and self-control its own resources for a dynamic reaction and adaptation during the user's session in order to maintain the service delivery with the required QoS.

A. Autonomic Service Component

Our proposal is based on two approaches: the SOA (Service Oriented Architecture (SOA) approach and the EDA (Event Driven Architecture) approach. The benefit of the SOA is the possibility to implement decentralized applications in distributed computing systems. One of the most important advantages of the SOA is its capacity to enable the rapid composition of the service components offered by various providers. The EDA complements the SOA because the service components can be activated by triggers fired on incoming events. To make the dynamic composition more effective in a heterogeneous and mobile context during the operations, we propose a novel view of the SOA (Service Oriented Architectures) based on an autonomic service component "ASC" that is generic, stateless, shareable, autonomous and self-manageable. The ASC is stateless because it performs the

 ISBN: 978-1-61208-192-2

same processing (operations) for all the requests coming from different users without storing the state or the data related to each request. This is important in the situation where it should make a dynamic replacement of a service component in the case of a deterioration of the QoS during the user's session. The ASC is shareable because it is designed for the provisioning and the processing of several users' requests according to its capacities. In order to control and manage this resources sharing, we associate a queue in the usage plan for each service component which holds all the accepted users requests. The Service component is autonomic because it is both autonomous and self-manageable. The service component is autonomous because it is functionally independent, i.e., it is self-sufficient and does not need other service components to achieve its functionality. The aim of the functional independence between the various components of service is to facilitate the change of a service component by a ubiquitous one in the case of a QoS deterioration or a malfunctioning during the processing without any impact on the global service requested by the end-user. The service component is self-manageable because it monitors its own QoS and manages its states related to the use of resources during the user's session. The service component will have at any time "t", one of the following four states: *Unavailable*, *Available*, *Activable* and *Activated*. The state *Unavailable* means that the service component is temporary or permanently inaccessible. The state *Available* means that the service component is or can be accessible. The state *Activable* means that the service component is ready to be activated, and the state *Activated* means that the resource is being used. When the service component is activated, a QoS-agent inserted in its management plan, monitors and controls the QoS contract (In contract/Out contract) and communicates the service component resources state via notification events. We mention that the QoS agent is based on two functional elements: the IQM (Internal QoS Manager) and the EQM (External QoS Manager). The IQM is in charge of the control and the management of each service component QoS. It monitors the internal resources of the service component to determine the QoS contract state (In contract / Out contract) of the ASC during the exploitation phase. The main function of the EQM is the communication and the coordination of the QoS resources between the different service components. [8]

During the user's session, a QoS agent can have one of the four following roles: Passive, Active, Interactive or Proactive. The QoS agent has a passive role when it ensures the internal processing of the QoS and does not communicate with its environment. It has an active role when it notifies the QoS resources status of the service component (In Contract/Out Contract). It has an interactive role when it interacts with other QoS-agents to negotiate the QoS parameters. The service component has a proactive role when the QoS-agent has the knowledge and the rules that enable it to make decisions on its own and send notifications to its environment. To do so, a service component instantiates a QoS model allowing a real time management of the service resources and their possible deteriorations.

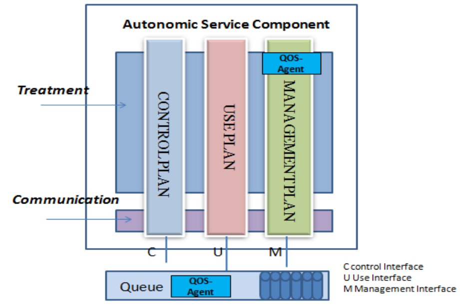


Figure 1: Structure of the Autonomic Service Component

B. QoS Model

In order to maintain the user Service Level Agreement (SLA), it is necessary to have a homogenous expression of the service component QoS to evaluate the End-to-End behavior. This is the reason why we propose a QoS model which represents the QoS in a uniform and homogeneous manner. The behavior of each component is reflected by measurable QoS parameters that can be categorized according to a vector of four criteria: *Availability, Reliability, Delay*, and *Capacity*.

Availability "$\mathcal{A}$": represents the ratio of accessibility for a service component. It indicates the number of times that the service component was accessible.

$$\mathcal{A} = 1 - \mathcal{U} / \mathcal{T}$$

$\mathcal{U}$ represents the number of times a service component has rejected a request because it was not available.

$\mathcal{T}$ represents the total number of requests sent over a period of time.

Reliability "$\mathcal{R}$": represents the ability of a service component to be executed without impacting the information. It indicates the percentage of successful invocations for a given period of measurement.

$$\mathcal{R} = 1 - \mathcal{F} / \mathcal{T}$$

$\mathcal{F}$ indicates the number of failed requests over a period of time.

Delay: represents the average time to process a request by a service component.

Capacity: represents the average number of requests processed by a service component during a unit of time.

The QoS criteria are necessary and sufficient for the self-control and self-management of the service component. These criteria are evaluated through three types of measurable values: *design values*, *current values* and *threshold values*.

Design values set the maximum processing capacity of a service component. *Current values* are used during the operation phase to monitor the behavior of the service component during the processing. *Threshold values* indicate the capacity limit that should not be exceeded by the service component in order to insure a normal treatment of the requests.

The QoS information helps to support the management, the treatment and the decision–making process of the service component. Hence, to make the right decisions during the

 ISBN: 978-1-61208-192-2

user's session at the right time and the right place, it is necessary to have an efficient representation of the real world. Therefore, we need to have a uniform information structure containing both the description of the information of the service component as well as the knowledge of the behavioral aspects of the service component (such as the QoS). The informational model contains the different profiles that are going to be solicited during the different operations phases. For the provisioning phase, we have the resource profile which contains the QoS *design values*. During the consumption phase, the "Real Time Profile" is instantiated in real time to have a dynamic management of the QoS. It contains the *QoS current values* that will be compared with the *Threshold values* to allow the control of the service component behavior during the usage.

C. *QoS Management Mechanisms*

When the user demands a global service (GS) with an SLA that can be composed as a business process invoking a variety of available ASCs, it is necessary to apply a number of mechanisms to perform the service delivery during the user's session steps. During the pre-provisioning step, each ASC proceeds by a *QoS Admission Control* "*QAC*" to verify the service capacity required to treat a new user's request. The *QAC* allows the selection of the ASCs that suit the user's QoS requirements because of the ubiquitous characteristics of the ASCs offered by different providers. During the provisioning step, each ASC reserves the necessary resources to process the user's demand. Each ASC associates a queue (Figure 2) in the usage plan to hold all the accepted users requests because of its shareable characteristic. In order to control this resource sharing, we must apply a "*QAC*" to accept a user's query in the queue. The role of the *QAC* is to determine whether a new requestor can be accepted in the ASC queue, without violating the SLAs of the already accepted requestors. Finally, the QoS

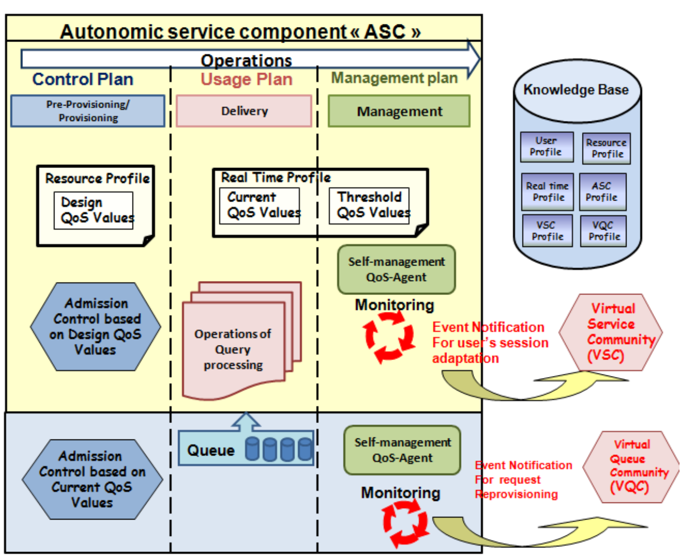


Figure 2: QoS Management Mechanisms applied by the ASC

should be dynamically and continuously managed during the consumption step. Our self-management vision is based on the QoS-agent which is integrated in the management plan of each ASC to monitor the QoS related to the processing of the query as well as the QoS related to the ASC queue. We have already proposed ubiquitous services and queues communities in order to ensure a replacement of the ASC during the service delivery when a QoS agent detects a deterioration of the QoS. In fact, each ASC belongs to a Virtual Service Community (VSC) [9], which contains a set of ASCs having the same functionality and an equivalent QoS, each ASC's queue also belongs to a Virtual Queue Community (VQC) [10], which contains ASCs queues of equivalent QoS.

The ASC is structured on three plans: the control, the usage and the management plan. The usage plan includes the main functions performed by the service component in order to process the requests. It includes all the mechanisms to process a request during the consumption phase. The control plan includes all the mechanisms used synchronously during the user's session to provision the service component resources. The management plan contains the self-management functions applied asynchronously on the data to control the service component resources. We explain in the following via an automaton the QoS management mechanisms used by the ASC in each plan during the different operational phases: the pre-provisioning, the provisioning, the delivery and the management phase.

During the pre-provisioning phase, the ASC receives in the control plan a signaling message to respond to a new user's request (Figure 3). First, it increments the "Attach" variable to "Attach+1" to keep the traceability of the users attached to a service component. Then, the QoS-agent applies a QAC mechanism to verify the statistical capabilities of its ASC. The statistical capabilities are based on the QoS *design values*. This admission control permits to determine if a component can be attached to a new user session to process its request. If the result is positive "OK", then the ASC changes its state to *Activable* in the case where it has not been previously attached to another user's session. Afterwards, it adds the user's session identifier to its profile in the knowledge base. If the result is negative "NOK", it should decrement the "Attach variable" to "Attach -1" and then the system searches in the VSC for another ASC having the same functionalities and an equivalent QoS to respond to the user's request.

During the provisioning step, the ASC receives the user's transaction in the control plan, it then increments the variable J to "J+1". Afterwards, the QoS-agent verifies the dynamic QoS based on the *QoS current values* of the queue in order to control the admission of the user's transaction in the queue. If the result is positive, it provisions the query in the queue and it changes its state from *Activable* to *Activated* if there is no query in the queue (i.e., J=1), else it decrements the variable J to"J-1".

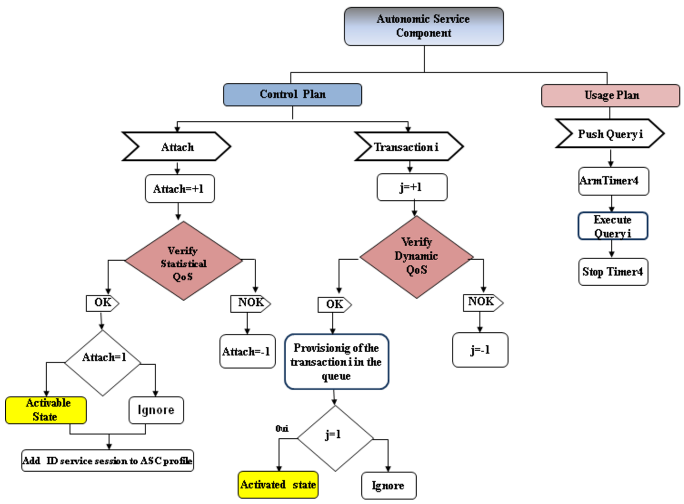


Figure 3 : Automaton of the ASC: Control and Usage Plans

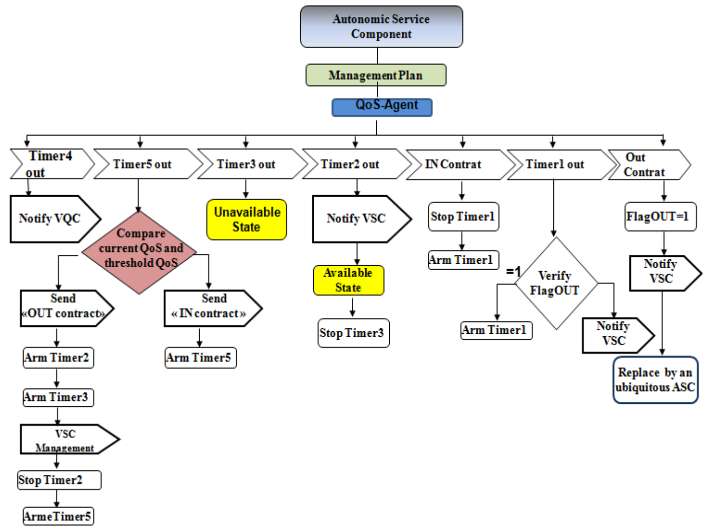


Figure 4: Automaton of the ASC: Management Plan

During the consumption step, an ASC that is in use through a user's session may not continue to function normally or to fulfill the user's QoS requirements. This is the reason why we solve such problem by proposing a number of mechanisms to monitor and manage the QoS at each ASC participating in the service delivery to the users. Our concept is based on a QoS agent integrated in the management plan of each service component. Figure 4 shows the different mechanisms used during the consumption step to maintain the service with the required QoS level. If the ASC takes a longer time to process a request (Timer 4 expired), the processing time of the requests that are provisioned in the ASC queue could be affected. The QoS-agent monitors the current QoS of the ASC queue. When it detects a QoS contract violation, it should notify the VQC to refer the request to another ubiquitous service component queue.

During the management step, The QoS–agent of each ASC composing the global service compares in real time the current QoS with the range of the QoS *threshold values*. If the result is positive, the QoS-agent sends a notification event "IN contract" to the VSC community to convey that it still respects the QoS contract. Otherwise, the QoS agent sends a notification event "Out contract" (Arm Timer 2 and Timer 3) to the VSC community in order to replace the current service component by a ubiquitous service (Stop Timer 2). At the end of Timer 3, the ASC changes its state from *available* to *unavailable* and leaves its community. Then, the ASC joins a new VSC community according to its current QoS and consequently changes its state from *unavailable* to *available*. All these operations are transparent to the end user and are done automatically by the system during the user's session.

Each ASC participates to the management of its VSC community; it shall notify regularly the other members on its QoS contract state (In Contract / Out Contract). If an ASC receives an Out contract, it sets its FlagOUT to 1 and notifies the VSC community to exclude the ASC that is out of the contract and to proceed to its replacement in order to maintain the VSC community. If the ASC receives an IN Contract, it will stop the Timer1 which has been armed in anticipation of receiving an IN Contract. In the case where Timer 1 expires (Timer 1 out) without receiving an IN Contract; the ASC should check first its FlagOut. If FlagOut=1, this means that the ASC has already received an Out Contract. Otherwise, it notifies the VSC about the defective communication of the ASC.

IV. SCENARIO

In this section, we use a case study to demonstrate the utility of the ASC during the user's session. The following scenario illustrates the case where the user wants to personalize his services to search for a property to buy in his geographical area. When the customer wants to search for a property, he selects two services from his catalog to obtain the required customized service. The first one is a global service named Search Property "SP" with an SLA. SP consists of three ASCs: FIND, LOCATION and GET. The second service is Google Maps.

The course of the scenario events is as follows: When the user is moving, he will first use "FIND" to search for properties according to his geographical position (Longitude, Latitude). The search is performed in the user's ambient zone

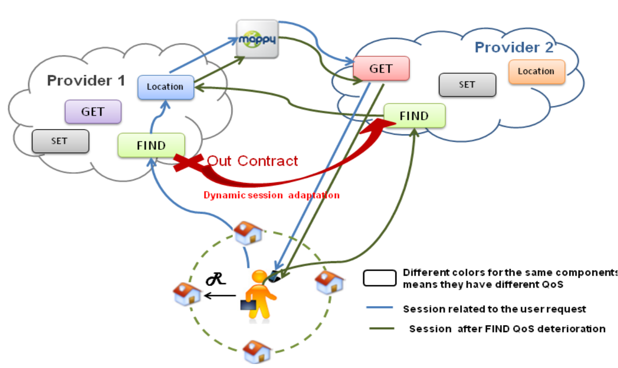


Figure 5: Scenario

 ISBN: 978-1-61208-192-2

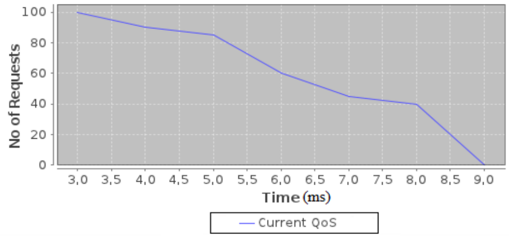


Figure 6: Basic Service component FIND

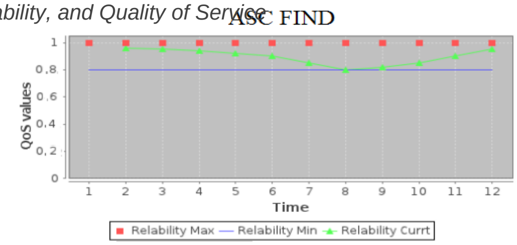


Figure 7: Autonomic Service Component FIND

which is limited by a radius R. Then, "Location" is used to obtain the GPS coordinates (Longitude, Latitude) for each property. Afterwards, he uses Google Maps to view the property on a map and finally he uses "GET" to see a detailed profile of each property (type, price and area). During the exploitation, the performance of "FIND" deteriorates and cannot provide the required QoS by the global service "Search Property". Therefore "FIND" sends an event notification to all the members of its VSC community to process its change. There are always some ubiquitous candidates to replace this component "FIND" and the selection algorithm used in the VSC community is essentially based on the user's geolocation in order to ensure the equivalent previous response time during the session. Once the VSC community finds a functional and QoS equivalent "FIND" on another platform, the connection is transferred from FIND@provider1 to FIND@provider2 (dynamic provisioning). Finally, the user's service session is maintained and the service delivery is performed. We should mention that all these operations are transparent to the user and are done dynamically and automatically by the system during the user's session.

V. IMPLEMENTATION & PERFORMANCE

In order to validate our proposal, we have used the language JAVA to develop our ASCs as independent EJBs and the JMS 1.1 queues to ensure communication between the ASCs. We have also used Oracle (V 10g) for the knowledge base and the JFreeChart API to plot the performance results.

In order to test the performance of our proposal, we have run the scenario that has been previously described with a basic service component (BSC) in one case and an ASC in the other. First case: we overload the BSC"FIND" and we notice that the data is lost because its QoS has deteriorated (Figure 6) and consequently the session is interrupted. Second case: we use ASCs with the following QoS: Find (QoS α): A_1= 0,99, R_1= 0,98, D_1=300ms, C_1=10; Location(QoS β): A_2=0,97, R_2=0,86, D_2=500ms, C_2=8; Get (QoS μ): A_4=0,98, R_4=0,90, D_4= 450ms, C_4=9. We note that if one of the four QoS criteria degrades during the user's session, it must proceed to change the service component because all the QoS criteria (A, R, D, and C) are necessary to maintain the overall QoS of the service component. During the user's session, the QoS-agent of the ASC "FIND" detects a QoS deterioration at the reliability criteria (R_1min = 0, 8 at T=8ms) (Figure 7). It invokes the VSC to change it by an equivalent ASC"FIND". We notice that the performance of the QoS is better thanks to the QoS-agent because it prohibits a complete QoS deterioration of the service component and maintains the service delivery.

VI. CONCLUSION

In order to enhance the service delivery, we have proposed in this paper an efficient solution based on an autonomic service component "ASC" which plays a key role in offering a service according to the consumer's requirements in a heterogeneous and mobile context. The ASC contains a QoS-agent to dynamically self-control and self-manage its own resources during the processing. Up to now, we have proposed a number of mechanisms to seamlessly adapt the user's session against any QoS deterioration caused during the operations. This solution is useful to resolve the problems of contemporary architectures that require more real time dynamicity, flexibility and responsiveness. Through experimentation, we have demonstrated that this proposal provides a better performance of the QoS and therefore a better service delivery to the consumers. In the future, we will try to find the best solution for the selection algorithm used in the virtual service community (VSC) and the virtual queue community (VQC).

ACKNOWLEDGMENT

The authors would like to thank all the participants in the UBIS project which is financed by the French ANR VERSO 2008.

REFERENCES

[1] "Portal for SOA and WS " www.service-architecture.com, April 2012.

[2] R. Farha, A. Leon-Garcia,"Blueprint for an Autonomic Service Architecture," Proc. IEEE. Autonomic and Autonomous Systems (ICAS 2006), Silicon Valley, CA, July. 2006, pp. 16-16.

[3] Y. Cheng, R. Farha, S.K. Myung, A. Leon-Garcia, and J.W.-K. Hong, "A Generic Architecture for Autonomic Service and Network Management," Computer Communications, November 2006, pp. 3691-3709.

[4] J. Zhang, Y. Chang, and K. Lin, "A Dependency Matrix Based Framework for QoS Diagnosis in SOA," Proc. IEEE. Service-Oriented Computing and Applications (SOCA 2009), Taipei, Taiwan , January 2009, pp. 1-8.

[5] Y. Zhai, J. Zhang, and K. Lin, "SOA Middleware Support for Service Process Reconfiguration with End-to-End QoS Constraints," Proc. IEEE. International Conference on Web Services (ICWS 2009), Los Angeles, CA, July 2009, pp. 815-822.

[6] F. Zambonelli, N. Bicocchi, G. Cabri, L. Leonardi, and M. Puviani, "On Self-adaptation, Self-expression, and Self-awareness in Autonomic Service Component Ensembles," Proc. IEEE. Self-Adaptive and Self-Organizing Systems Workshops (SASOW 2011), Michigan, Oct 2011, pp. 108 - 113.

[7] H. liu, and M. Parashar, "A component based programming framework for autonomic applications, " Systems, vol.36, May. 2006, pp. 341-352, doi:10.1109/TSMCC.2006.871577.

[8] H. Alaoui, P. Coude, and N. Simoni, "Modèle organisationnel pour le pilotage dynamique de la qualité de service de bout en bout d'une session user-centric," Gestion des réseaux et de services (Gres 2010), Montreal, Canada, Oct 2010, unpublished.

[9] H. Alaoui, N. Simoni, P. Coude, "User-centric and QoS-based service session," proc. IEEE. Asia-Pacific Services Computing Conference (APSCC 2011), Jeju, South Korea, Dec 2011, pp. 267-274.

[10] S Kessal, N Simoni, " QoS based Service Provisioning in NGN/NGS context," proc.IEEE. Network and Service Management (CNSM 2011), Paris, France, Oct 2011, pp.1-5.

Evaluation of a new Scheduling Scheme for VoIP with Mobility in 3G LTE

Richard Musabe, Hadi Larijani
Glasgow Caledonian University
Glasgow, Scotland, UK
e-mail: richard.musabe@gcu.ac.uk, h.larijani@gcu.ac.uk

***Abstract-* 3G long term evolution is an all internet protocol based network and one of its main aims is to improve mobile multimedia services. This is achieved through streamlining the system for packet services. This leads to improvements in the form of higher bit rates, lower latencies, and a variety of service offerings. However, more challenging technical difficulties may be expected to arise when voice traffic flows over a long term evolution network. There is a major change in the way voice is transmitted in long term evolution network, in that it is transmitted in packets instead of through circuits. The fact that long term evolution is designed to support mobility, it is of great importance to analyse the effect of mobility to our new scheduling scheme for voice over internet protocol in long term evolution systems. In this paper, we analyse the effect of mobility to our proposed scheduling algorithm, voice over internet protocol optimisation scheduling algorithm while taking into account the quality of service parameters of voice traffic in long term evolution network. Using long term evolution-SIM simulation software, we evaluate the performance of our proposed scheduling algorithm and compare it with other scheduling algorithms in the literature such as; exponential proportional fair and proportional fair scheduling algorithms. Simulation results showed that mobility had a significant impact on the quality of service for voice traffic on all the three scheduling algorithms. Our proposed algorithm provided the best quality of service for voice traffic compared to the other two scheduling algorithms based on the packet loss ratio, delay, and throughput metrics.**

Keywords- LTE; Mobility; Scheduling Schemes; VoIP.

I. INTRODUCTION

3G long term evolution **(**LTE) was identified by the third generation partnership project (3GPP) as the preliminary version of next generation wireless communication systems because of its high data rates [1]. This mobile cellular communications technology provides a maximum 100Mbps downlink and 50Mbps uplink when using 20 MHz bandwidth [2]. In the downlink physical layer, LTE uses Orthogonal Frequency-Division Multiple Access (OFDMA) radio technology to meet the LTE requirements for spectrum flexibility and enables cost-efficient solutions for wide carriers with high peak rates. In the uplink, LTE uses a pre-coded version of OFDMA known as Single-Carrier Frequency-Division Multiple Access (SCFDMA) in order to compensate for a drawback with normal OFDMA of a high Peak-to-Average-Power Ratio (PAPR) [3]. Wireless technology has expanded from voice only to high-speed data, multimedia applications, and wireless internet [4].

LTE requirements for high data rates are achieved by the fact that this technology is only designed for packet switched networks (PSN); hence, there is no need for the circuit switched mode. However, this design brings with it more technical challenges especially for voice services. Voice over internet protocol **(**VoIP) services are both delay and packet loss sensitive. The biggest challenge of VoIP over LTE is to deliver Quality of Service (QoS). Normally, users would expect voice with the same quality as that provided by circuit switched networks. However, traffic delivered over PSNs is subject to delay and packet loss [5]. A major issue with VoIP over LTE is that 3G LTE adopts a different method of resource transmission from other cellular systems like Code Division Multiple Access (CDMA).

3G LTE uses Physical Resource Blocks (PRB) as its transmission unit. PRBs can be defined as the basic unit with both frequency and time aspects [6]. Basically, the base station of 3G LTE, known as eNodeB has a fixed number of available PRBs according to their allocated bandwidth and it is supposed to assign PRBs repeatedly at every Transmission Time Interval (TTI) [2]. Another issue with VoIP over LTE is that LTE systems also tend to support a very high mobility of up to 350 km/h [7]. LTE aims at providing a fast and seamless handover from one cell (source cell) to another (target cell) [8]. However, the consequences of the handover procedures in LTE systems depends entirely on the type of application that is being used, for example some applications would tolerent a short interuption while others would not. In this paper, the application that is being used is VoIP, so it is crucial to evaluate the QoS of VoIP for the high mobility. Different techniques have been introduced in recent years in order to overcome the challenges of voice over LTE [2] [1] [5]. However, QoS for voice over LTE is still a big challenge taking into account the mobility features of LTE, fading channels of wireless links as well as delay and packet loss sensitive voice characteristics. Our contributions in this paper are:

- Analyse the effect of mobility to our proposed scheduling algorithm; VoIP optimisation scheduling algorithm **(**VOSA) while taking into account the QoS parameters of voice traffic in LTE
- Evaluate the performance of our proposed scheduling algorithm VOSA with mobility features and compare it with other scheduling algorithms developed in [9] such as: exponential proportional fair **(**EXP-PF) and proportional fair (PF) scheduling algorithms.

The simulation results were generated using the open source LTE system simulator called long term evolution-SIM (LTE-SIM) [9]. It models different uplink and downlink scheduling strategies in multicell/multiuser environments; taking into account user mobility, radio resource optimization, frequency reuse techniques, the adaptive modulation, and coding (AMC) module. It also includes other aspects that are relevant to the industrial and scientific communities.

The rest of the paper is organised as follows: Section II discusses the general aspects of VoIP. Section III describes different scheduling algorithms used in this paper. Section IV describes the system model, scenario setup, handovers and mobility patterns. Section IV presents the simulation results and performance evaluation. Section V reviews the main conclusions, and introduces the future work.

II. VoIP

A. *Brief Description*

VoIP is a way of transmitting voice traffic as data packets over an IP network. Voice traffic is first transformed into digital signals then it is compressed and broken into a series of packets. These series of packets will later be reassembled and decoded at the receiver. Voice digitizing and encoding can either be done before or concurrently with packetization [10]. This technology has grown rapidly due to different factors such as: low cost, the integration of voice and data traffic over the existing networking infrastructures, etc.

VoIP is transmitted over a packet-switched network rather than the circuit-switch network protocols of the PSTN [11]. Initially voice and data were transmitted using two different networks but with the introduction of VoIP technology, they can both be transmitted using the same network infrastructure. VoIP reduces costs by avoiding the use of traditional PSTN. With VoIP technology the high cost for long distance calls and international calls transported over the circuit switched network can be reduced by transporting voice calls over low cost flat pricing packet switched network [12].

The advantages of VoIP such as integrated services and flexibility has attracted more customers as well as companies from circuit switch networks to packet switch networks. The fact that VoIP uses a packet-switched network means that the QoS provided by VoIP is not as good as that of circuit-switched network. This is due to the fact that real-time traffic such as voice is affected by technical issues like end-to-end delay or latency, jitter, and packet loss, hence adversely affecting the quality of voice [13]. Unique treatment should be given to voice traffic as it is vital for it to reach the destination in the quickest time [14].

Since our main aim is to analyse the effect of mobility to the QoS of voice traffic, it is of great importance to analyse some important parameters that describe QoS of voice in LTE networks. These parameters will be investigated in the next sub section.

B. *VoIP QoS Analysis*

The adaptive multirate (AMR) voice codec is one of the most popular voice codecs used in LTE. This codec provides 32-bytes voice payload every 20 milliseconds during talk-spurt period and 7-bytes payload carries a silence descriptor (SID) frame every 160 millisecond [1]. VoIP protocol stack that utilizes the real transport protocol (RTP) is encapsulated using user datagram protocol (UDP), and in turn is carried by IP. The combination of all these protocols requires a 40 byte IPv4 header or a 60-byte IPv6 header, but the overhead brought about by these headers causes serious degrading in spectral efficiency in supporting VoIP services. To solve this problem, an efficient and robust header compression (ROHC) technique is used. This technique solves the overhead problem by minimizing the size of the IP/UDP/RTP headers as little as 2 or 4 bytes using IETF RFC 3059 [6] [15].

One of the main characteristics of voice traffic is block/packet error rate that leads to packet loss and delay [7]. According to [16], the allowed maximum mouth-to-ear delay for voice is 250ms with the assumption that the delay for the core network is approximately 100ms, while the tolerable delay for radio link control (RLC), MAC buffering, scheduling, and detection should be strictly lower than 150ms as shown in Fig.1. Hence, taking into account that both end users are LTE users, tolerable delay for buffering and scheduling must be lower than 80ms. A delay of 50ms from eNB to UE has been chosen for the 3GPP performance evaluation metric limit to better account for variability in network end-to-end delays [7]. When voice packets are transmitted over a packet switched network, packets will be dropped due to error rate and packet delay exceeding the target latency. However with the occurance of the packet loss, voice quality is not affected if the error rate is less than outage threshold [7]. This means that the Qos of VoIP in LTE is limited by an outage limit, described in TR 25.814 [6] and was later updated in R1-070674 [17]. The outage limit means that error rate of VoIP users must be kept within 2%. The overall description of the QoS for voice users in LTE can then be defined as the maximum number of VoIP users that can be supported without exceeding a given threshold. At least 95% of total VoIP users should meet the above described outage limits [1].

III. SCHEDULING ALGORITHMS

Different scheduling algorithms have been introduced in recent years in order to overcome the challenges of voice over LTE.

In [1], the authors proposed an efficient LTE scheduler to increase the capacity of VoIP in E-UTRA Uplink. The proposed scheme modified the persistent scheduling algorithm proposed in [5] such that the resources of two VoIP users can be coupled, this brought about early termination gains without the need of additional control signals. The proposed efficient scheduling method employs a resource sharing approach. It also employs the random user pairing and best user pairing method to improve the capacity of VoIP services over E-UTRA Uplink. The results showed that the employment of their proposed scheduling scheme makes a larger available capacity than that resulting from the original persistent scheduling.

In [2], a Medium Access Control (MAC) layer PRB scheduling algorithm was proposed. The key ideas of this scheme are VoIP priority mode and its adaptive duration management. The VoIP priority mode assigns PRBs first to VoIP calls and it is also able to minimize VoIP packet delay and packet loss while the adaptive duration management is able to prevent the overall system performance degradation. The proposed MAC scheduler in this paper allocates PRBs in a round robin way and the scheduling order is determined according to the following factors; the queue length and Signal to Interference Noise Ratio (SINR) of each call, so the larger the factor values are the earlier the corresponding call is scheduled. Their results show that when the VoIP priority mode is not used, the packet drop rate rises rapidly as the number of VoIP call increases. On the contrary, when using the VoIP priority mode, the drop rate remains at low level around 1 % in spite of the increase of VoIP calls. However, this proposed mode has got a possible negative effect. It might degrade the efficiency of the eNodeB resource utilization because VoIP calls are seldom able to fully utilize the allocated PRB capacity.

 ISBN: 978-1-61208-192-2

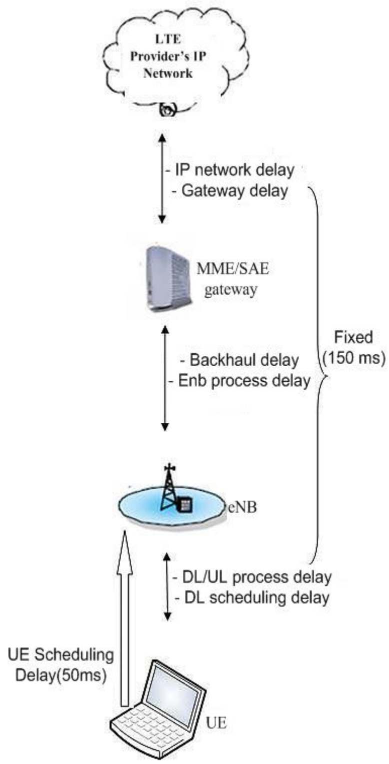


Figure 1. End To End Delay Components in E-UTRAN (LTE)

In [5], a new semi-persistent scheme of MAC scheduling, which adaptively prioritize VoIP traffic by not forcing other traffics to starve, was proposed. The proposed scheme, which combines VoIP priority mode with user coupling, allows utilizing the system capacity efficiently. The priority mode duration is adaptively controlled using the channel condition and two users are coupled to share the resources. The controlling user is also determined dynamically using minimum information as possible. The priority mode works by assigning resources to VoIP traffic in a priority basis at the same time controlling the duration of this mode dynamically according to the channel conditions in order to avoid starvation of other services in the same time. The scheme also allows user coupling where two users share the resources allocated to them by the eNodeB therefore offsetting the low resource utilization while in priority mode.

This is achieved by allowing two users who have different channel conditions to share resources. Basically, the proposed scheme consists of two parts; user coupling and link authority change. User pairing method takes care of pairing VoIP users according to the channel conditions so that the pairing results in the most efficient usage of resource. The link authority change ensures that the user in need of resource at any point of time has the authority on the link to get the fair share. The authority adaption phase is executed using the Acknowledgement / No acknowledgement (ACK/NACK) channel of the users. Each user monitors his couple's ACK/NACK channel and changes the authority status if the signal on his own channel and his pair's channel is different, i.e., 'his signal is ACK and his pair's signal is NACK or vice versa'. On the other hand, the authority remains same if the signals of both the channels are same. Their results show a great improvement over the original persistent mode.

In [9], different scheduling algorithms were developed and the main ones included; PF and EXP-PF. PF scheduling algorithm focused on maximizing the total network throughput as well as assuring fairness among flows. It allocates resources based on two factors; experienced channel quality and the past user throughput [19]. This scheduler uses the metric descrbed as the ratio between the instantaneous available data rate and the average past rate with reference to the i-th flow in the j-th flow subchannel. This can be depicted in equation 1 below.

$$W_{i,j} = \frac{r_{i,j}}{R_{i,j}} \qquad (1)$$

where W_{ij} is the scheduler metric, R_i is the estimated average data rate and r_{ij} is the instantaneous available data rate.

EXP-PF scheduling algorithm basically aimed at increasing the priority of real-time flows as opposed to non-real-time flows. In other words, the flows with head-of-line packet delay very close to the delay threshold [20]. Its metrics were calculated as follows;

$$W_{i,j} = \exp\left(\frac{\alpha_i D_{HOL,i} - X}{1+\sqrt{X}}\right)\frac{r_{i,j}}{R_{i,j}} \qquad (2)$$

and

$$X = \frac{1}{N_{r,t}}\sum_{i=1}^{N_{r,t}} \alpha_i D_{HOL,i} \qquad (3)$$

with N*r,t* being the number of active downlink real-time flow.

Considering a packet delay threshold $T_{i,}$ the probability α_i is defined as the maximum probability that the delay $D_{HOL,I}$ of the head-of-line packet delay exceeds the delay threshold. Therefore α_i is given by

$$\alpha_i = -\frac{\log \alpha_i}{T_i} \qquad (4)$$

With all these techniques in the literature, QoS for voice over LTE is still a big challenge taking into account the fading channels of wireless links as well as delay and packet loss sensitive voice characteristics. Based on the algorithm in [2], we propose a new scheduling algorithm called VOSA, with the aim of improving the performance of voice traffic over a 3G LTE network.

At the same time it reduces the negative impact that may be caused by the introduction of the new algorithm on the entire system's performance. Details of VOSA can be found in [18]. This algorithm is activated at every TTI by considering if there is a VoIP call and if the duration period of the new algorithm has not exceeded the limit. To determine the duration of our new algorithm, we use the adaptive method proposed in [2]. This method provides limits to *VOSA* by adaptively changing between a pre-specific minimum and maximum according to the ratio of dropped packets. Higher drop ratio means that there are many ongoing VoIP calls, and hence it is necessary to increase the limits to allow more consecutive TTIs to be dedicated to VoIP calls. On the other hand, low drop ratio implies that QoS of VoIP calls are satisfied at decent levels, and thus it is safe to reduce the duration of the algorithm and serve other service in the network. It should be noted that the adaptive method used here considers only dropped packets due to many ongoing VoIP calls (Call congestions). However, packet loss can also be due to different factors such as fading channels, interferences, etc and these factors were put under consideration while scheduling other kinds of traffic in our network.

Our scheduling scheme is designed by making modification to the algorithm in [2]. Basically, the *VOSA* allocates PRBs to VoIP calls based on the arrival time metric. Once the PRBs allocation is done, the scheduling order of the calls is determined by the following factors: Quality feedback (QF) and queue length (QL) of each call. The better the factor values are, the earlier the corresponding call is scheduled. In our algorithm we use the following equation:

$$D_f(i) = Q_{feedback(i)} * Q_{length(i)} \qquad (5)$$

where $Q_{feedback\ (i)}$ and $Q_{length\ (i)}$ are the quality feedback and queue length respectively. Equation (5) implies that the better the wireless link and the longer the queue length, the earlier the corresponding call is scheduled to have the PRBs. This is calculated at every TTI. The details of the proposed algorithm consist of two parts. The first part of *VOSA* describes the PRBs allocation at every TTI. The second part is the adaptive method to control the duration of the proposed algorithm, detailed in [2].

In short, scheduling starts at every TTI. The activation of the scheduling algorithm is determined by whether there exists a VoIP call or not, and whether the count of the consecutive scheduling algorithm enabled TTIs does not exceed the limit. Otherwise the normal mode is set to schedule non-real time traffic based on similar factors as in equation (5) and the count is reset. VoIP calls are assigned one PRB at a time. Calls with long queues and better wireless link are served first. It continues in the same routine as long as there are remaining PRBs and there are calls having data to send. If there are remaining PRBs after VoIP call scheduling, the remaining PRBs are allocated to other services in the same way as the normal mode. This prevents PRBs from being wasted.

VOSA scheduling algorithm is summarised in 8 different steps;

1. Indentify the traffic type whether voice or any other traffic
2. Determine the user metrics (QF,QL)
3. Find the user with the highest user metric as defined in equation 5
4. Consider the set of available resource blocks RBs N_{avail_RB}, at every start of the algorithm , $N_{avail\_RB} = \{1,2,\ldots\ldots.., N_{RB}\}$ to be allocated to number of user K
5. Assign the resource block N* to the user K* with the highest user metrics value such that $N_{RB,}K^* = N_{RB,}K^*$ U {N*}
6. Schedule the user K* first
7. Delete the user K* and resource block N* from their respective lists
8. Repeat all the steps until all users are scheduled and if more resource block exists then allocate them to other traffic types.

The fact that VOSA focuses on channel quality and queue length metrics at physical and MAC layers improves the QoS of Voice calls in LTE. Using equation (5), VOSA performs scheduling at the MAC layer and schedules Voice calls first before assigning PRBs to other traffic in the network. However, the starvation of other traffics in the network is controlled by the adaptive method used. In other words, VOSA is only deployed when there are voice calls to schedule and if it does not exceed its limits.

IV. SYSTEM MODELLING AND SCENARIO SETUP

A. PRB Characteristics

In this sub-section, we introduce the characteristics of PRBs, described as the transmission resources in LTE. LTE systems consists of both a time and a frequecy plane. The time plane is divided into 1 ms TTI which consists of two slots of 0.5 ms to form 1 ms sub frames, where each sub frame contains 7 OFDMA symbols. In each TTI, there are 14 OFDMA sysmbols, where 2 symbols out of 14 are reserved for uplink pilot transmission, while the other 12 symbols are used for data and control information transimission. TTI can be defined as the minimum allocation unit in the time domain [21].

If we consider the frequency plane, the minimum allocation unit is the PRB, where each PRB contains 12 subcarriers of 15 Khz bandwidth each. The number of OFDMA symbols in a resource block depends on a cyclic prefix being used. All these can be depicted in Fig. 2. It must be noted that VoIP packets must be transmitted per TTI and they can occupy one or more PRBs [5]. The amount of data bits that can be transmitted by one PRB depends on the link between the eNodeB and the user mobile terminal. This is due to the fact that 3G LTE uses adaptive modulation and coding (AMC), in order to change modulation and coding schemes depending on the wireless link conditions.

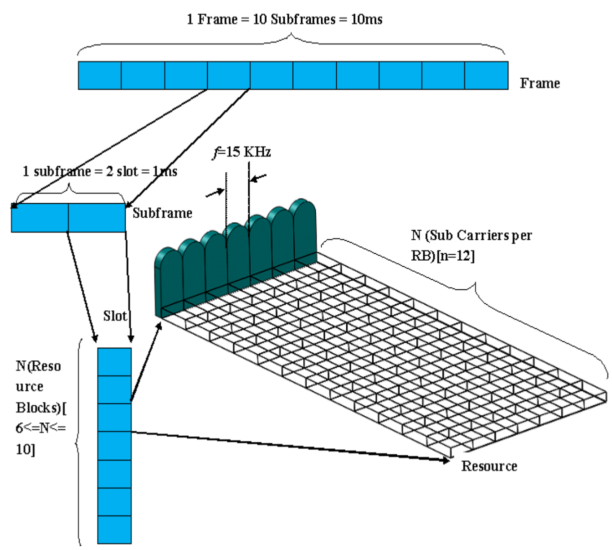


Figure 2. The structure and allocation of the eNodeB transmission resources symbols.

B. Scenario Setup

The network is made up numerous cells and different network nodes such as; the EnodeB, mobility management/gateway (MME/GW) and user equipements (UEs). All the simulations were run in a diamond-pattern scenario with 19-3-sector sites that totaled to about 57 cells. Most of the simulation parameters are presented in the table 1 below. VoIP flows are generated by the traffic generator in LTE-SIM called VoIP application. This application generates G.729 voice flows.VoIP traffic was designed with an ON/OFF Markov chain. The ON period has an exponentially distributed mean value of 3s while the OFF period has a shortened exponential probability density function with an upper limit of 6.9s as well as an average value of 3s [22]. Throughout the ON period, the source sends packets of 20bytes every 20 ms, thus implying that the source data rate is 8 kb/s. On the other hand, throughout the OFF period the rate is zero as we assume the presence of voice activity detector.

Three different scheduling algorithms were used in all simulation scenarios, these are: our proposed VOSA as well as EXP-PF and PF developed in [9]. In one simulation scenario there was no mobility features implemented so the user speed was set to zero, this is equivalent to static position. In another simulation scenario, we implemented mobility features and the user speed was set to 30 km/h which was equivalent to vehicular and the distance covered was set to 400 meters by default.

To quantitatively investigate the effect of mobility to the voice traffic including each scheduling scheme, we measured three important VoIP QoS metrics (Packet-Loss-ratio, Delay, and throughput) for all the scheduling algorithms and in both scenarios.

C. Handover and Mobility Pattern

One of the main aims of LTE systems is to provide a fast and seamless handover from one cell (source cell) to another (target cell) [8]. This is generally achieved due to the distributed nature of LTE radio access network architecture that consists of one node known as the EnodeB. According to the 3GPP release 8 specifications, handovers in LTE are hard handovers, implying that there is a minimum interruption in the services when the handover is performed.

During the handover process, VoIP users cannot be scheduled, they cannot transmit or receive any data. The only transmitted and received data is the signalling related to the handover procedure [7]. This can lead to additional delays to Packet Data Units (PDUs) hence affecting the user quality on each handover process. LTE utilises a UE assisted hard handover algorithm for mobility. The UE measures the downlink signal quality and sends the the measurement reports to the EnodeB either periodically or when an event triggers i.e., 'an interfering EnodeB becomes stronger than the current serving EnodeB'. Then the EnodeB will decide on the final handover based on the received measurement report. Normally, measurement averaging, handover margins, and timers are used to avoid excess handovers [23].

In the LTE-SIM simulator, two types of mobility models were developed, known as; random direction and random walk [24]. The user speed was selected between 0, 3, 30, 120 km/h, which were corresponding to static, pedestrian, and vehicular scenarios respectively. The user speed was mapped to a specific travel distance, i.e., 'a user travelling at 3km/h would cover 200 meters, a user travelling at 30km/h would cover 400 meters and the user travelling at 120km/h would cover 1000meters'.

If the random direction model is being used, the user chooses the speed direction at ramdom and keeps the same speed while moving towards the simulation boundary area. When the user reaches the simulation boundary area then new speed direction can be choosen. In contrary, when the random walk model is chosen, user chooses the speed direction at random but keeps moving at that speed for a specific travel distance depending on that speed. The user only changes the speed direction if the distance is convered or once the simulation boundary is reached [9].

V. SIMULATIONS AND PERFORMANCE EVALUATION

We used LTE-SIM to analyse the effect of mobility on our proposed scheduling algorithm (VOSA) and two other scheduling EXP-PF and PF. It should be noted that the details of VOSA can be found in [18]. The main contribution in this paper that was not introduced in [18] is mobility. We introduced mobility in our network and analysed it's effect on voice traffic in relation to the three scheduling algorithms in this paper.

The main reason for comparing our proposed scheduling algorithm with these other two scheduling algorithms is that, they use the same PRBs allocation as ours and similar simulation parameters except that they apply different metrics and the fairness factor. They were also used as the benchmark scheduling algorithms in the LTE-SIM simulator. This made our comparison more feasible.We measured the packet loss ratio (the rate at which VoIP packets were dropped during voice traffic transmission) while gradually increasing the number of VoIP users.

This is shown in Fig. 3. The packet drop ratio is measured and plotted on the Y axis as we increased the number of VoIP user steadily to the maximum of twenty users. As it can be seen in Fig.3, mobility had a significant impact on voice traffic in all three scheduling algorithms. There was a higher packet loss ratio in all the algorithms with mobility features compared to those without mobility features. However, there was less packet loss ratio in VOSA scheduling algorithm (with or without mobility features) compared to the other two scheduling algorithms.

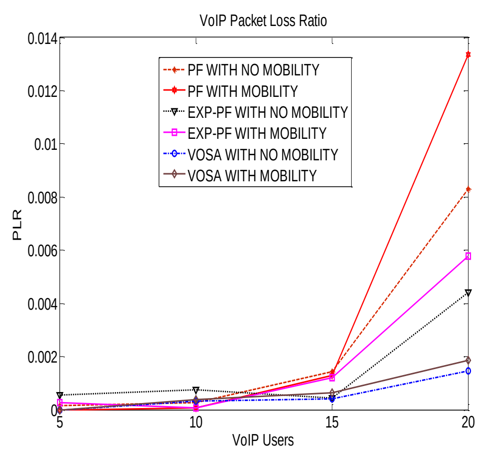


Figure 3. VoIP-Packet-Loss-Ratio Comparison

We also measured VoIP delay while gradually increasing the number of VoIP users. This is shown in Fig. 4. The VoIP delay is measured and plotted on the Y axis in seconds as we increased the number of users steadily to twenty. Similar to previous results, there was long delay in all the algorithms with mobility features and again VOSA had less delay than the other two scheduling algorithms. These two simulation results show that VOSA plays an important role in improving the QoS of Voice traffic in both scenarios.

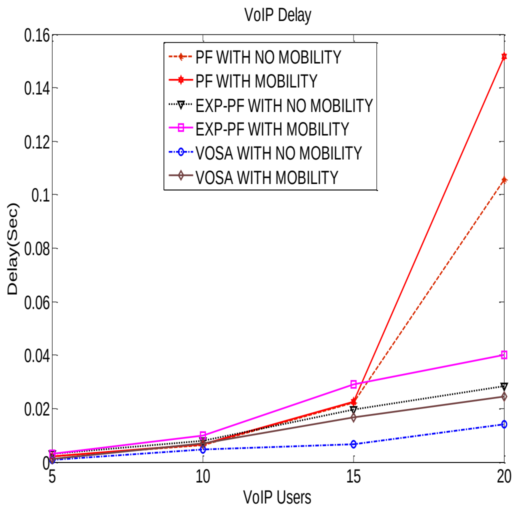


Figure 4. VoIP Delay Comparison

TABLE 1. SIMULATION PARAMETERS

Simulation Parameters	**Values**
Bandwidth	5MHZ
PRB Structure	12subcarriers,2subframes
TTI	1msec
Number of available PRBs	25
Modulations for AMC	QPSK
Number of sectors	3
Simulation time	1000 TTIs
User speeds	30km/h
Cyclic prefix	Normal
Mobility patterns	Random direction and random walk
Distance covered by user	400m
Scheduling algorithms	VOSA,EXP-PF, and PF
Cell radius	1 km

Apart from these two VoIP metrics, we also measured throughput while using all the scheduling algorithms and in both scenarios. This is shown in Fig. 5. As it can be seen, throughput decreased as the number of VoIP users increased in all algorithms and in both scenarios. This is mainly due to the fact that some VoIP packets were being dropped as the number of users were being increased, this resulted in the less utilisation of assigned PRBs.

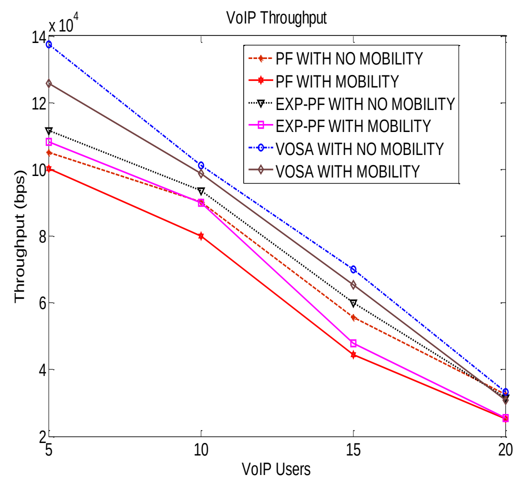


Figure 5. VoIP Throughput Comparison

VI. CONCLUSION AND FUTURE WORK

In this paper, we analysed the effect of mobility on voice traffic with three different scheduling algorithms. Using the LTE-SIM simulator, we were able to compare all the scheduling algorithms in two different scenarios. One scenario containing mobility features and another with no mobility features. Through simulations we found out that mobility had a significant impact on the QoS of voice traffic with all the three scheduling algorithms. However, our VOSA scheduling method provided better QoS of voice traffic than the other two scheduling algorithms by having short delay, less packet loss ratio, and slightly higher throughput. The main reason for the improvement in VOSA scheduling method is due to the fact that it does not cause any negative impact on the entire network's system performance. It is only deployed when there are VoIP calls to schedule and if it does not exceed its limits. These limits are determined by the adaptive method used. On top of that, the wireless link quality and the queue length factors used in VOSA played an important role in reducing the packet drop ratio and delay. In future work, we are working on determining the complexity and fairness of our proposed scheduling algorithm VOSA. Find the way of optimizing our scheduling algorithm such that we lower its complexity.

REFERENCES

[1] S. Y. Kim, "An Efficient Scheduling Scheme to Enhance the Capacity of VoIP Services in Evolved UTRA Uplink," EURASIP journal of Wireless Communications and Networking, vol 2008, Mar. 2008, pp. 1-9, doi:10.1155/2008/732418.

[2] S. Choi, K. Jun, Y. Shin, S. Kang, and B. Choi, "Mac Scheduling Scheme for VoIP Traffic Services in 3G LTE," IEEE 66th Vehicular Technology Conference, pp. 1441-1445, Oct. 2007, doi: 10.1109/VETECF.2007.307

[3] Rohde and Schwarz, "*UMTS* long term evolution (LTE) Technology introduction," A report by Rohde and Schwarz., pp. 1-30, Mar.2007.

[4] M. C. Chuah, and Q. Zhang, Introduction to Wireless Communications, US: Springer, 2006.

[5] S. Saha and R. Quazi, "Priority-Coupling-A Semi-Persistent MAC Scheduling Scheme for VoIP Traffic on 3G LTE," ConTEL, 10th International Conference on Telecommunications, pp. 325-329, Aug.2009.

[6] 3GPP, Physical layer aspects for Evolved UTRA, 3GPP Technical report 25.814, version7.1.0, pp. 1-135, Sep. 2006.

[7] T. Henttonen, K. Aschan, J. Puttonen, N. Kolehmainen, P. Kela, M. Moisio, et al, "Performance of VoIP with Mobility in UTRA Long Term Evolution," Vehicular Technology Conference, pp. 1-5, May 2008, doi: 10.1109/VETECS.2008.549.

[8] Qualcomm, LTE Mobility Enhancements, Qualcomm Incorporated, pp. 1-10, Feb.2010.

[9] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE Cellular Systems: An Open-Source Framework," IEEE Transactions on Vehicular Technology, vol.60, no.2, pp. 1-16, Feb.2011.

[10] B.Goode, "Voive over internet protocol (VoIP)," Proc.IEEE, vol.90, no.9, Dec.2002, pp. 1495- 1517, doi:10.1109/JPROC.2002. 802005.

[11] K. Gonia, Latency and QoS for Voice over IP, SANS Institute- As part of GIAC practical repository, pp. 1-21, May. 2004.

[12] G. Zhang, and M. Hillenbrand, "Implementing SIP and H.323 signalling as Webservices," 30th Euromico conference proceedings, pp. 1-7, Mar.2004.

[13] B. Khasnabish, Implementing Voice over IP. John Wiley & sons Inc, pp. 1-224, 2003.

[14] U. Varshney, A. Snow, M. McGivern, and C. Howard, "Voice over IP," ACM Journal, Jun. 2002, pp.1 –8.

[15] IETF, Attribute list extension for the service location protocol,. IETF RFC 3059, pp. 1-35, Nov.2001.

[16] ITU-T, One way transmission time. ITU-T recommendation G.114, pp. 1-20, April.2003.

[17] TSG-RAN1, LTE physical layer framework for performance verification, 3GPP TSG-RAN1 R1-070674, pp.1-30, Mar. 2007.

[18] R. Musabe, H. Larijani, B. Stewart, and T. Boutaleb, "A New Scheduling Scheme for Voice Awareness in 3G LTE," IEEE Computer Society, Sixth International Conference on Broadband and Wireless Computing, Communication and Applications, Dec.2011, pp. 1-8, doi: 10.1109/BWCCA.2011.46.

[19] G. J. Choi and S. Bahk, "Cell-throughput Analysis of the proportional fair scheduler in the single-cell environment," IEEE Transaction Vehicular Technology, vol.56, no.2, April 2007, pp. 766-778, doi: 10.1109/TVT.2006.889570.

[20] R. Basukala, H. M. Ramli, and K. Sandrasegaran, "Performance analysis of EXP/PF and M-LWDF in downlink 3GPP LTE systems," 1st AH-ICI . Kathmandu, Nepal, Nov.2009, pp.1-5, doi: 10.1109/AHICI.2009.5340336.

[21] Roke, LTE eNodeB MAC Schedular, Roke Manor Research Ltd, Siemens, pp.1-4, Jan. 2009.

[22] C. Chuah, and R. H. Katz, "Characterizing packet audio streams from internet multimedia applications," Proc.IEEE, ICC, Aug. 2002, pp. 1199-1203, doi: 10.1109/ICC.2002.997040.

[23] J. Puttonen, T. Henttonen, N. Kolehmainen, K. Aschan, M. Moisio, and P. Kela, "Voice-over-IP Performance in UTRA Long Term Evolution Downlink," Vehicular Technology Conference, pp. 2502- 2506, May 2008, doi: 10.1109/VETECS.2008.551.

[24] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," Wireless Communication and Mobile Computing (WCMC), Special issue on Mobile Ad Hoc Networking, pp. 483-502, Dec. 2002.

Quality of Service Modelling for Federated Wireless Sensor Network Testbeds Gateways

Syarifah Ezdiani Syed Nor Azlan and Adnan Al-Anbuky

Sensor Network and Smart Environment Research Centre

Auckland University of Technology (AUT), Auckland, New Zealand

{ssyednor, adnan.anbuky}@aut.ac.nz

Abstract— **The federation of WSN testbeds has brought forward the requirement for a QoS implementation, in order to ensure efficient and reliable end-to-end communication between testbeds' users and testbeds' facilities. Nowadays, the Internet has become the de facto standard for establishing remote connectivity between these testbeds. However, as the mechanism of QoS support in WSNs may be very different from that in the Internet, it is necessary to address the differences between the QoS employed in the Internet and the QoS required in WSN testbeds for the purpose of QoS provisioning. Hence, our work focuses on generating a QoS model for testbed gateways by merging the QoS perspectives from both networks. Moreover, this research will also focus on validating the modelled QoS through QoE measurements, as the network-level QoS parameters can be translated into user-level QoE perception.**

Keywords-QoS modelling; WSN testbed; federation; gateway

I. INTRODUCTION

In the recent years, there has been an increasing trend to develop wireless sensor network (WSN) testbeds which aims at providing the environment for researchers to conduct experiments [1-3]. Moreover, these testbeds which enable researchers to evaluate and validate their WSN-related work are integrated to the Internet to allow remote access to its users. Although these testbeds provide tremendous benefit to researchers, they are usually limited in size due to hardware costs. Thus, there exists a growing trend to interconnect small testbeds to provide federated testbeds with large-scale testing facilities [4]. However, while a good deal of research and development has been carried out in testbed architectural design, there is a glaring lack of studies on network performance in the environment of federating WSN testbeds.

One of the major challenges in federation of testbeds is to provide reliable and efficient connection between interconnected testbeds. Therefore, network performance indicator such as Quality of Service (QoS) must be taken into account in federating these testbeds, in order to ensure reliable and efficient real-time experimentation for testbed users.

Traditional QoS, such as employed in the Internet, mainly results from the rising popularity of end-to-end bandwidth–hungry multimedia applications, and are defined using certain parameters such as packet loss, delay, jitter and bandwidth. On the contrary, the metrics concerned such as available bandwidth and delays may not be pertinent in most WSN environment as some WSN applications could be latency-tolerant or transmitting very small packets. Moreover, in some WSNs, it is typical that the amount and quality of information is more important in WSNs [5].

The touching point between a testbed and the Internet is the gateway. Thus, it could become the main element in contributing to the performance of the federation. In order to allow for greater capabilities of fixed and simple gateways, there has been an increasing trend to turn the gateways into ones with greater capabilities. Research concerning smart gateways [6] or modular gateway may benefit the QoS significantly.

Furthermore, open federated testbeds usually allow access to multiple users to testbed resources for real-time experiment and observation. Consequently, testbeds operator's main interest would also include the way users perceived the testbeds usability, reliability, quality and time-worthiness. Hence, a further step beyond provisioning QoS is observing its users' Quality of Experience (QoE).

Therefore, in this research, we are inspired by two major goals; (1) to provide a solution for QoS mechanism to run on top of testbed gateways that satisfies stringent QoS requirements of an environment of federated testbeds, and (2) to devise a scheme for validating and verifying the network performance under the modelled QoS.

The remainder of this paper is organized as follows. Section 2 will give an overview of the research problem domains. Next, the design of our study is presented in Section 3. Finally, Section 4 concludes the paper.

II. PROBLEM STATEMENT – QOS FOR TESTBEDS FEDERATION

Due to the significant differences between WSN and the Internet, the QoS requirements generated by both networks may be very different. Indeed, existing research has concluded that the end-to-end QoS parameters employed in traditional data networks such as the Internet are not sufficient to describe the QoS in WSN [7, 8].

In this section, we distinguish the QoS requirements in WSN from the QoS requirements in the Internet, followed by an overview of integration QoS on the gateway device. This is followed with an overview of QoE for testbed users.

A. QoS Support in WSN

The two perspectives of QoS in WSNs were described in [7] to focus on the way the underlying network can provide the QoS to different application:

 ISBN: 978-1-61208-192-2

Application-specific QoS

In terms of application-specific QoS, the QoS parameters are chosen based on the way an application imposes specific requirements on sensor deployments, on the number of active sensors, or on the measurement precision of sensors. These attributes are all related to the quality of applications. The following QoS parameters may be considered to achieve the quality of applications: coverage, exposure, measurement errors, and number of active sensors.

Network QoS

From the perspective of network QoS, the QoS parameters are chosen based on how data is delivered to the sink and corresponding requirements. The main objective is to ensure that the communication network can deliver the QoS-constraint sensor data while efficiently utilizing network resources. The QoS parameters from this perspective include latency, delay, and packet loss, which are similar to traditional end-to-end Qos metrics. However, since WSNs is envisioned to be employed in diverse applications, a number of works in the literature suggested that every different application imposes different QoS requirements.

B. QoS Support in the Internet

RFC 2368 [9] definition on QoS-based routing in the Internet characterizes QoS as a set of service requirements to be met when transporting a packet stream from the source to its destination. QoS support in the Internet can generally be obtained by means of over-provisioning of resources and/or traffic engineering. While traffic bursts in the network could cause congestion, the default approach of over-provisioning which treats users at the same service class may not always provide an acceptable solution. As a QoS-enabled network should be able to handle different traffic streams in different ways, this necessitates traffic engineering approach which classifies users into classes with different priority.

IntServ model and DiffServ model are the typical QoS models employed in the Internet, which employs reservation-based and reservation-less approach, respectively. While network resources are assigned according to an application's QoS request and subject to bandwidth management policy in IntServe, QoS in DiffServe is achieved via some strategies such as admission control, traffic classes, policy managers, and queuing mechanism.

C. Integration QoS

Several studies have demonstrated network performance testing and QoS measurements of WSN-Internet integration. The performance is typically measured using several predominant QoS metrics, focussing on the QoS implementation on the gateway side of WSN integrated to the Internet.

The QoS is commonly provided by an integration controller which runs software modules and able to reconfigure the QoS parameters on the network edge router. In this typical application-level gateway approach, the performance is evaluated in terms of inter-arrival time (the time between adjacent packets), packet delay, latency or round-trip rate (RTT, the time taken by a packet to travel from the source to destination) and cumulative distribution function of the RTT.

Moreover, in a non-trivial network, a packet will be forwarded over many links via many gateways. Gateways will not commence forwarding the packet until it has been completely received. In such a network, the minimal latency is the accumulation of the minimum latency of each link, the transmission delay of each link and the forwarding latency of each gateway. In practice, this minimal latency is further augmented by processing and queuing delays. Whereas processing delay occurred while a gateway determines what to do with a newly received packet, queuing delay occurs when a gateway receives multiple packets from different sources heading towards the same destination.

D. QoE for Testbed Users

Ideally, in order to provide unifying testbed facilities for experimentations, federated testbeds allow access to multiple users over the Internet at any given time. The main objective and other major attributes of a testbed pose certain critical questions: a) how could the network ensure good users' experience? b) how does the network manage its resources to do so? c) how does the network maintain reliable connection for users' real-time experimentation? d) how does the network cope with the demand from the testbed users for a large number of nodes for their experiments? This calls for QoS mechanism that takes into consideration the QoE; tying together user perception, experience and expectation to the application combined with the network performance [10].

III. DESIGN OF THE STUDY

In this section, we discuss the tasks involved in designing the QoS model. Primarily, the QoS modelling will be carried out using a network modelling simulation tool, namely OPNET Modeler. In order to provide a means of validation, we propose that the network performance measurement under the modelled QoS shall be conducted over physical federated WSN testbeds linked to an Internet testbed.

A. QoS and QoE Requirement Analysis

The QoS and QoE mechanisms discussed in the previous section are mainly employed in separation. To the best of our knowledge, there is no precedent work that has considered merging the aforementioned QoS and QoE perspectives into one. Hence, we target to develop a solution for QoS provisioning in an environment of federated WSN testbeds by consolidating Internet QoS, WSN QoS and testbed user's QoE.

Therefore, we first conducted a QoS and QoE requirement analysis by taking into account various characteristics from the networking and users' perspectives. In this task, we map our findings on the requirement analysis from both network and users perspectives [11], in order to identify the dominant QoS and QoE metrics that are crucial for the federation.

Hence, the requirement analysis includes the following network characteristics:

1) End-to-end: In general, WSN testbeds offer resource reservation to their users. This implies that one end of the application is a testbed user, and the other end is a single node within the testbed. A QoS implementation is needed to ensure end-to-end network performance throughout an

experimentation session and any connection issues should not jeopardize the validity of the experiments in any way. In addition, the network's requirement on end-to-end performance also implies strict QoS requirements in both WSN and the Internet.

2) Criticality: Experiments on testbeds should be achieved with high information reliability. Thus, this also suggests that packet losses cannot be tolerated in any extent.
3) Bandwidth Requirements: Although bandwidth availability is not the main concern in most WSN applications, it is of great importance in the Internet. Furthermore, bandwidth may be an important factor for a group of sensors with the bursty nature of sensor traffic. Thus bandwidth plays a vital role in provisioning QoS in the testbed environment.
4) Interactivity: Testbed environments typically are based on query-driven approach, i.e., queries are made to reconfigure sensor nodes and manage resources. These commands require high reliability from the network, thus implying QoS requirements in terms of timely and reliable response from the testbed side.
5) Delay Tolerance: The network tolerance of delay may be based on the models of experiments offered by the testbed. However, in an environment of federated testbed, the end-to-end delay might be a critical factor due to the best-effort nature of the Internet.
6) Network Dimensions: When building testbeds federations for scale, the interconnection of testbeds can be of closely located testbeds, or it could expand to be of large intercontinental federations. The Internet's capability to deliver packets from one end to another plays a huge factor in the overall network performance. In addition, the total number of nodes may affect network performance as the volume of traffic is more likely to increase as the number of nodes grows.

B. Federation Network Model

Figure 1 shows the reference architecture for the federation WSN testbeds, whereby gateway devices reside in the 2nd-tier of sensor network testbeds level.

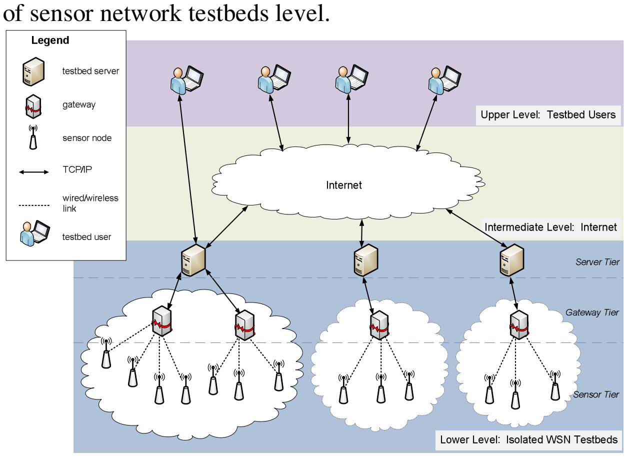


Figure 1: Reference architecture of federated WSN testbeds

Testbed server, which represents a single testbed, forwards application for users from the Internet to the gateway for node access. On the other hand, the gateway that serves as the interface between a testbed server and the sensor nodes, mainly responsible for packaging high-level user queries to network specific directives and returning filtered portions of the data to users. Therefore, being in a unique position as having the full knowledge of and control over both WSN testbed and the Internet, the gateway plays a vital role for the QoS provisioning purpose.

The main challenge concerning the interconnection between Internet and a WSN testbed is the requirements to provide access to each sensor nodes through the TCP/IP based network. Hence, we focus on distinguishing the QoS factors from various network levels, and identify those pertinent to the gateways that act as the main interface between an the Internet and a testbed. The QoS factors involves in an event of sending or receiving messages [12] over the network include the different levels as depicted in Figure 1, which may include:

1. The application on the sending/receiving sensor nodes to/from the gateway device (gateway lookup, encoding/decoding, interrupt handling, propagation)
2. Processing delay inside the gateway device (interrupt handling, propagation LAN/WLAN, packet translation)
3. Processing delay inside the testbed server (packet translation, address lookup)

C. Our Proposed Federation QoS Solution

We propose to employ a system capable of differentiating traffic classes. Different priority traffic placed in different queues could contribute greatly to the success of the QoS solution. Hence, to be able to place the traffic into a specific queue, we need to classify the traffic. Hence, classification is one of the basic functions performed on a packet when it arrives at a QoS enabled gateway, which is then followed by remarking of packet. We propose that the traffic classification is implemented on various basic testbed functionalities and experimentation tools typically offered by WSN testbeds, which may include user registration, resource/node reservation, topology selection, job scheduling/management/cancellation, files uploading, data retrieval, and nodes programming.

Classification and remarking are the basic building blocks for the QoS. Therefore, we will test the network performance by looking into test cases with combination of these two aspects, for example by implementing priority for retrieval of data and remarking for node reservation.

D. Network Modelling and Simulation on OPNET

In our preliminary testing, we have generated a network model and simulated the federation model using OPNET Modeler [13]. We have conducted simulations of testbed networks under different loads to three different scenarios:

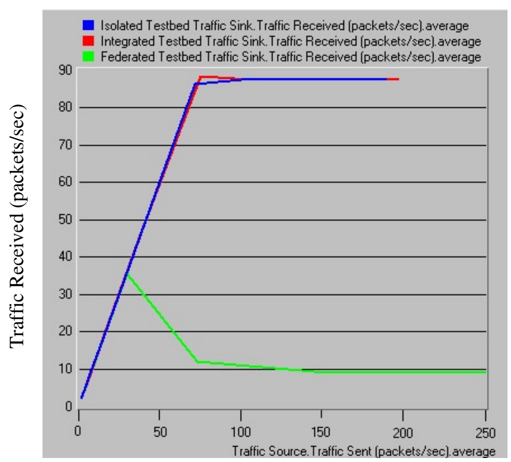
<u>Isolated Testbed</u> -This configuration represents a single testbed in isolation. Local users gain direct connection to the sensors through the gateway, which obtains access to the wireless network through an access point.

<u>Testbed Integrated to the Internet</u> - This configuration represents a single testbed connected to the Internet.

<u>Federation of Two Testbeds</u> - This configuration represents two testbeds interconnected to the Internet.

The main objective of the simulation is to compare the network's performance for isolated and integrated environment. Although this simulation did not involve any direct QoS implementation, we believe our finding from the simulation activities give an important insight pertinent to QoS provisioning, hence facilitates for further experimentation for our federation QoS modelling. We evaluated the network performance based the following QoS parameters: packet received against packet sent (reflects packet loss) and end-to-end delay.

Received Traffic: Figure 2 depicts the amount of traffic received (in packets/sec) against the amount of traffic sent (in packets/sec), for all three testbed scenarios. In both cases of an isolated testbed and a testbed integrated to the Internet, as expected, the amount of traffic received proportionally increased as more traffic is sent to the network. However, there was a point in the graph where the amount of traffic received flattened, despite having a continual increase of the traffic sent. Most importantly, at a particular point, the amount of traffic received in a WSN-Internet integrated testbed tend not to increase as rapidly as those in an isolated WSN testbed. A way to interpret this is to presume that the network has reached its full capacity in terms of link bandwidths, traffic queues or devices' processing power; therefore, it could not deliver any more traffic to its destination.

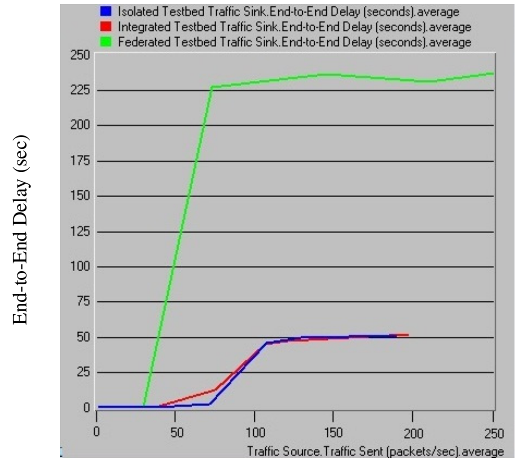


Traffic Sent (packets/sec)

Figure 2: Traffic Received as a function of Traffic Sent

On the other hand, we have observed a major difference in the resulting graph of the federated testbed scenario. The number of packet received within the federated testbed started to decrease at a certain amount of packet sent. This could be explained by the fact that single gateway or QoS-enabled gateway node may create a bottleneck over the network. Furthermore, since network capacity and number of hops are a major concern in the Internet, there is a higher probability of having packet losses as more and more packets transit on the Internet.

End-to-End Delay: Next, we run a simulation on the end-to-end delay for both scenarios, against traffic sent. Once again, the result obtained showed a significant difference between federated testbed scenario and the other two testbed topologies. It was observed that the federated testbed scenario by far demonstrate a very high end-to-end delay. In addition, as per our expectation, the testbed integrated to the Internet tend to record slightly higher delays, as compared to isolated WSN testbed. The ensuing graph is depicted in Figure 3.



Traffic Sent (packets/sec)

Figure 3: End-to-End Delay as a function of Traffic Sent

Our results have shown that due to its size, and particularly its unpredictable nature, the Internet does introduce a level of service degradation (i.e., higher packet losses and delays) when it is subjected to high traffic load. This could pose a problem for mission-critical WSN applications where there is a need for predictable performance. Similarly, this must be taken into account especially when the notion of federating WSN testbeds featuring several thousands of nodes and multiple users is of interest. Hence, the introduction of an end-to-end QoS, which ensures that packets are being delivered in a reliable and timely manner in both the WSN testbed and the Internet, could be a solution to this matter.

E. Testbed Experimentation – Linking together WISEBED and PlanetLab

One of the primary goals of this research is to assess the QoS model in a real wide area network. Hence, there is a need to analyze the performance of the QoS in a real world scenario by conducting necessary network traces and measurements. Therefore, once we have implemented the QoS on the gateway, we want to evaluate the model in a bigger, real-world environment. We need a tool to read incoming packets from the live-traffic Internet to test how well our QoS model handle users' traffic. An Internet testbed such as PlanetLab is an ideal testbed for this purpose due to its dispersion.

Therefore, in order to complete our modeling architecture, we aim to generate a federation across an open federated WSN testbed, namely WISEBED [14], and PlanetLab [15], hence providing the platform for real-time traffic simulations. WISEBED offers extension to its originally nine-federated testbeds, as WISEBED-compatible testbed can be established by researchers by employing Wisebed Runtime [11, 16].

Therefore, we propose to utilise Wisebed Runtime to be able to interconnect our network to WISEBED facilities.

As shown in Figure 4, we will devise the framework for the interconnection that will serve as the main platform to test our QoS modelling. The technical challenge is to construct the tunnelling mechanism [17] using tunnelling protocol such as Generic Routing Encapsulation (GRE) [18], to allow communication between our testbed gateway to the PlanetLab node.

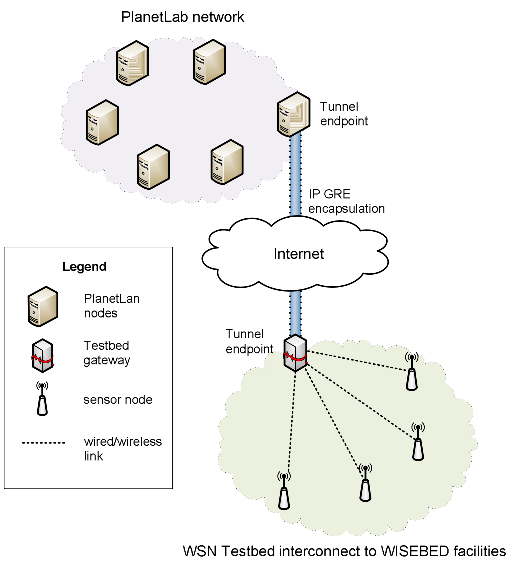


Figure 4: Network Architecture: the tunnelling communicating testbed gateway and PlanetLab node

F. *QoS Model Performance Evaluation and QoE Verification*

The QoS measurements will be based on the dominant QoS metrics which will be defined from the QoS and QoE requirement analysis. The measurement will be conducted on the tunneling platform between the PlanetLab nodes and WSN testbeds over various traffic scenarios.

Whereas in our QoS model, we propose to employ a system capable of differentiating traffic classes, we will select a specific WSN application, such as indoor/outdoor and mobile applications, for the purpose of validation and verification using objective QoE [10, 19]. In this objective QoE mechanism, we will endeavour to find a QoE solution applicable to federated testbed environment. Specifically, we expect that the investigation on the impact of QoS on QoE will present a unified formula to express dependency of federation QoS and testbed user's QoE, hence allowing for adjustment to our QoS model.

IV. CONCLUSIONS

This paper provided an overview of a work in progress of QoS modelling for federated WSN testbeds. We have presented the significant differences between WSN QoS and Internet QoS, and will conduct a thorough QoS and QoE requirement analysis to formulate a QoS model for the federation. We advocate that gateways should run QoS mechanism that merges the network-level QoS perspectives from both WSN and the Internet, as well as the user-level QoE. In order to study the proposed model performance, we utilize OPNET Modeler as the modelling and simulation software. Furthermore, to complete our modelling architecture, we plan to federate PlanetLab and the open federated WSN testbed, i.e., WISEBED, to serve as the case to our study.

REFERENCES

[1] G. Werner-Allen, P. Swieskowski, and M. Welsh, "Motelab: A Wireless Sensor Network Testbed", in the Proc. of the Fourth International Symposium on Information Processing in Sensor Networks, IPSN 2005, 25-27 April 2005, California, USA, pp. 483-488.

[2] R. N. Murty, et al., "CitySense: An Urban-Scale Wireless Sensor Network and Testbed", in 2008 IEEE Conference on Technologies for Homeland Security, May 2008, pp. 583-588, 2008.

[3] M. Gunes, B. Blywis, and J. Schiller, "A Hybrid Testbed for Long-Term Wireless Sensor Network Studies", International Workshop on Sensor Network Engineering (IWSNE '08), Santorini, Greece, 11 June 2008.

[4] I. Chatzigiannakis, S. Fischer, C. Koninis, G. Mylonas, and D. Pfisterer, "WISEBED: An Open Large-Scale Wireless Sensor Network Testbed," in SENSAPPEAL, 2009, LNICST 29, N. Komninos, Ed., ed: Institute for Computer Science, Social-Informatics and Telecommunications Engineering, 2010, pp. 68 -87.

[5] J.-F. Martinez, A.-B. Garcia, I. Corredor, L. Lopez, V. Hernandez, and A. Dasilva, "Modelling QoS for Wireless Sensor Networks," in IFIP International Federation for Information Processing. vol. 248, Wireless Sensor and Actor Networks ed Boston: Springer, 2007, pp. 143-154.

[6] D. Bimchas et al., "Middleware for Smart Gateways Connecting Sensornets to the Internet,"5th International Workshop on Middleware Tools, Services and Run-Time Support for Sensor Networks (MidSens'10), Bangalore, India, pp. 8-14, 2010.

[7] D. Chen and P. K. Varshney, "QoS Support in Wireless Sensor Networks: A Survey " in International Conference on Wireless Networks 2004 (ICWN '04), Las Vegas, Nevada, USA, 2004.

[8] B. Bhuyan, H. K. D. Sarma, N. Sarma, A. Kar, and R.Mall, "Quality of Service (QoS) Provisions in Wireless Sensor Networks and Related Challenges", Scientific Research, Computer and Communications Journals, Vol. 2 No. II, Nov. 2010, pp. 861 – 868, 2010.

[9] E. Crawley, et al., "A Framework for QoS-based Routing in the Internet," RFC 2386, http://www.ietf.org/rfc/rfc2386.txt, August 1998, <retrieved: February, 2012>.

[10] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A Generic Quantitative Relationship between Quality of Experience and Quality of Service", IEEE Network, vol. 24, no. 2, Mar./Ap. 2010, pp. 36-41

[11] H. Hellbruck, M. Pagel, A. Kroller, D. Bimchas, D. Pfisterer, and S. Fischer, "Using and Operating Wireless Sensor Network Testbeds with WISEBED," in The 10th IFIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2011), Favignana Island, Sicily, Italy, 2011, pp. 171 -178.

[12] T. Baumgartner, et al., "Virtualising Testbeds to Support Large-Scale Reconfigurable Experiment Facilities," EWSN 2010, LNCS 5970, pp. 210-223, Springer-Verlag Berlin Heildelberg 2010.

[13] OPNET Official Website, http://www.opnet.com, <retrieved: February, 2012>.

[14] WISEBED Project Website, http://www.wisebed.eu/, <retrieved: February, 2012>.

[15] PlanetLab, An Open Platform for Developing, Deploying, and Assessing Planetary-Scale Services, http://www.planet-lab.org/, <retrieved: February, 2012>.

[16] The WISEBED Consortium, "Testbed Runtime wiki and bug tracker," 2010, https://www.itm.uni-luebeck.de/projects/testbed-runtime/, <retrieved: February, 2012>.

[17] J. Lischka and H. Karl, "RiaS - Overlay Topology Creation on a PlanetLab Infrastructure," in 2nd ACM SIGCOMM Workshop on Virtualised Infrastructure Systems and Architectures VISA 2010, New Delhi, India, 2010, pp. 9 - 15.

[18] D. Farinacci, T.Li, S. Hanks, D. Meyer, and P. Traina, "Generic Routing Encapsulation (GRE)," RFC 2784, March 2000.

[19] P. Brooks and B. Hestnes, "User Measures of Quality of Experience: Why Being Objectives and Quantitative is Important " IEEE Network, vol. 24, pp. 8 -13, March/April 2010.

System Reliability of Fault Tolerant Data Center

Montri Wiboonrat
College of Graduate Study in Management
Khon Kaen University, Bangkok Campus
Bangkok, Thailand
montwi@kku.ac.com, mwiboonrat@gmail.com

***Abstract*—A single point of failure (SPOF) in system operations is a weak point of system reliability. Mean time to failure (MTTF) of system operations is equal to the shortage component's MTTF in system. A Tier IV data center is designed to eliminate the SPOF. Data center system reliability is not only depended on the MTTF of each component in the system, but also relies on the mean time to repair (MTTR) of each component. Researcher performed Tier IV DC power distribution systems (PDS) through simulating software, BlockSim7. The research question is tried to investigate how to improve system reliability. Component's inherent characteristic (CIC) and system connectivity topology (SCT) are applied to improve the system reliability of Tier IV data center. The results demonstrated an increasing PDS reliability, site plus site, of Tier IV data center and improving survival probability of system that helps for future improvement on any critical system.**

Keywords—System Reliability; Mean Time To Failure; Mean Time To Repair; Probability density function (pdf).

I. Introduction

The redundancy represents a possible approach to enhancing system reliability. In a series-parallel design methodology, serial systems reduce reliability, while more parallel systems help increase it. The redundancy scheme helps enhance the overall system reliability. It, however, costs more. A data center consists of multiple hardware components that are bound to fail sooner or later. The Tier IV data center is a fault tolerant system that is designed to eliminate a single point of failure (SPOF) [3]. System downtime adversely affects not only recovery and lost opportunity costs, but also the company's reputation and customers' confidence. The reliability and cost trade-off becomes a controversial issue among all concerned, including top management, IT managers, and financial managers [8]. Different organizations have different levels of recovery time objective (RTO) and recovery point objective (RPO) subject to system failures and power outages to varying degrees of risks [4]. This paper employs the reliability block diagrams (RBD) with reliability information obtained from IEEE 493, or the so-called Gold Book, and component vendors' field test data. The study attempts to improve the data center system reliability by integrating 2 parallel systems of TIA 942's Tier IV data center [3]. The research question comes up with, how the reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems and which once is the most benefits to investor subject to investment, efficiency, and system reliability. Research findings suggest the system connectivity topology, i.e., parallel topology, helps increase the system reliability of DC operations.

II. Reliability Background

A. *Reliability Factors on Data Center Failures*

System failure in data centers may be caused by many sources.

1. Human error: daily operation, regular planned downtime for maintenance, and unplanned downtime [5], [9].
2. Component's inherent characteristics (CIC) and system failures are dependent on mean time to failure (MTTF), the complexity of system connectivity topology (SCT), and operational conditions. A component/system failure may propagate or activate other component/ system fault, error, or failure. This process or failure is similar to a chain-reaction that is affected from component to component, component to system, or system to system [5], [7].
3. Operational conditions are other factors that relates to system failures, e.g., humidity, temperature, altitude, and dust.
4. Natural disasters; this is beyond the control of data center to handle. A design for a parallel site needs to be considered to compensate for downtime losses [8].

B. *Reliability Determination*

All equipment reliability data is obtained from the IEEE 493 Gold Book Standard, as shown in Table I. Fig. 1 depicts a single line diagram of a representative network for the Tier IV data center. The components shown in the networks are labeled with numbers, which correspond with the reference numbers in Table I. Network reliability analysis is performed with reliability data for referenced components taken from this table.

This paper investigated the system reliability/availability of a Tier IV data center in terms of the frequency and duration of power outages. System availability depends on:

1. Reliability and maintainability of its components: including mean time between failure (MTBF) and mean time to repair (MTTR) of component's inherent characteristics (CIC) distribution, failure modes effects and criticality analysis (FMECA), and environmental effects [4], [7].

2. System design or system connectivity topology (SCT) (configuration or topology, dependency, and failure detection).
3. System operation behavior (operational characteristics, switching procedures, and maintenance services).

The following assumptions apply to the proposed Tier IV data center system networks, as shown in Fig. 1:

- Failure rates and repair times are exponentially distributed.
- Actual power cable lengths may be indicated on the drawing. The cable failure rate is thus determined per the indicated actual cable length.
- The generators are 2N redundant.
- The power grids, generators and UPSs are 2(N+1) redundant, applicable to Tier IV.
- The transformers, switchgears, automatic transfer switches (ATSs) and bus bars are redundant.
- There are two paths of power distribution systems.
- Terminations and splices, while normal for all systems, are not included on the drawing, and are not included in the calculations.
- The assumed breaker failure modes are 50% open and 50% short.

III. DATA CENTER MODEL ASSUMPTION

A data center is a complex system that consists of 16 systems. Operations on business continuity strategy mode: primary and secondary sites require 2 data centers operating at the same time to back each other up, i.e., system-of-systems (SoS) or site-plus-site [3]. To limit the scope of this investigation, the researcher focused on data center Tier IV, power distribution systems (PDS) and power distribution system-of-systems (PDSS: both primary and secondary), parallel input, from incoming utility throughout loaded consumptions, e.g. server racks, storage racks, and networking equipment racks. This research assumes the external systems e.g. utility system, communication system or integrated service provider (ISP), are described to operation 100% uptime during this research.

PDS is the most sensitive system for data center downtime. Thus, the research tested fault injection by the cutting of the main utility system as human error or unplanned downtime. UPSs will take action to recharge power back to the system immediately as long as the battery can handle the loaded points. Gen-Set will activate within 15 seconds to change power back to the UPSs to resupply on loaded points [10]. By TIA 942, Tier IV, Gen-Set availability is 96 hours for consecutive operation without interruptions [3].

On this research simulation, the research defines the fault tolerance of data center system design as reliability of 2 parallel system of Tier IV data center. The modification prototype in Fig. 2 is reproduced from original Tier IV data center from Uptime Institute, Fig.1. Each number in Fig. 2 is referred from IEEE 493 Gold Book [7]. Table I refers to characteristic of each number in Fig. 2 that identifies failure rate, MTBF, and MTTR of each number. A failure rate (λ) of Primary data center system and Secondary data center system is assumed equivalent and each system has a constant (λ).

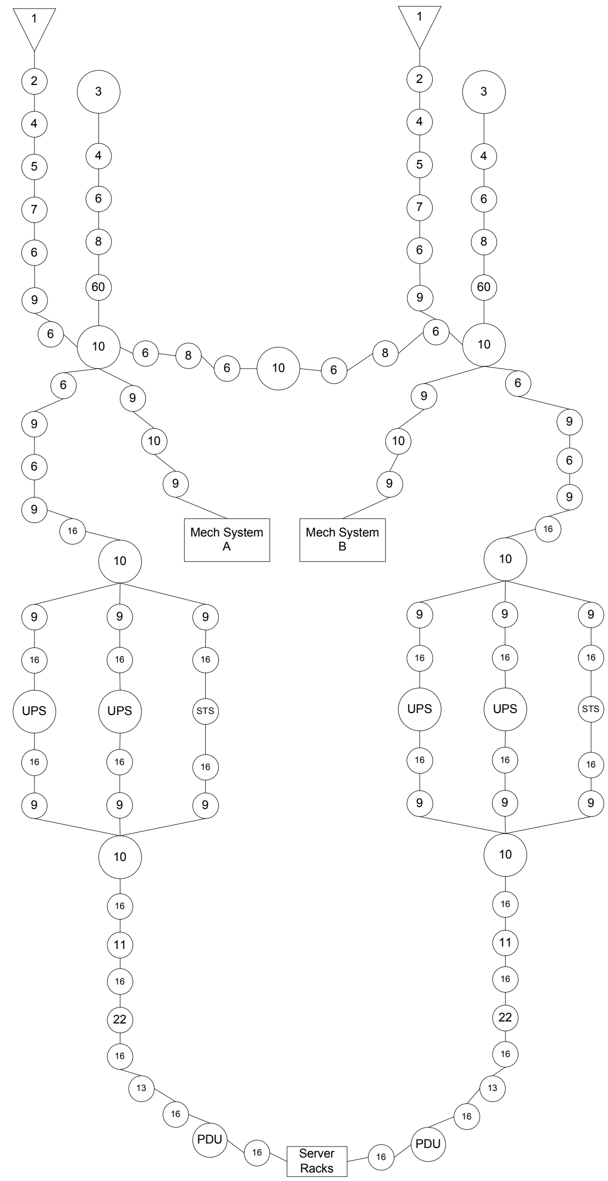


Fig. 2. PDSS of Tier IV Data Center

The reliability of data center system designs for the two parallel systems are reliable in parallel operations only when the failure of both systems results in system operation. On the other hand, for a parallel system to succeed, at least one of the two parallel systems in the whole system needs to perform successfully, or operate without failure, for the operating interval on the intended mission.

The research proposes a simulation approach applied to a reliability block diagram (RBD) by BlickSim 7. The system reliability results from Fig.1 through RBD show the MTBF of Tier IV data center is 75,434.78 hours, and the failure rate (λ) is 14.0865×10^{-6} [5]. The exponential distribution is applied in Tier IV data center reliability analysis. The *f(T)* of the exponential distribution is extremely convenient, often used to describe a steady-state hazard rate, as shown in Fig. 3 [2].

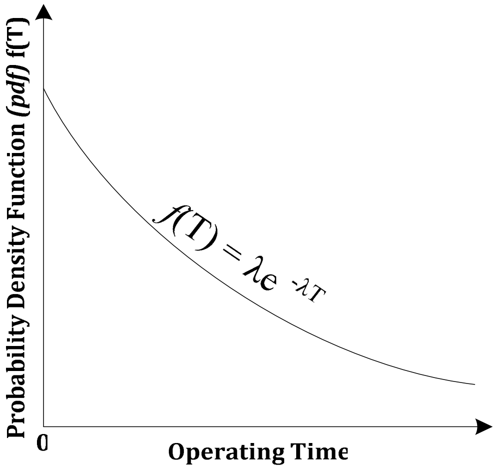


Fig. 3. Relationship of *pdf* and Operation Time [2]

The probability density function *(pdf)* of two parallel systems, that two systems are equal, when $\lambda_1 = \lambda_2 = \lambda$, is given by:

$$f_{SP}(T) = -\frac{d[R_{SP}(T)]}{dT} \tag{1}$$

The failure rate of two parallel systems is given by:

$$\lambda_{PS}(T) = \frac{2\lambda e^{-\lambda T} - 2\lambda e^{-2\lambda T}}{2e^{-\lambda T} - e^{-2\lambda T}} \tag{2}$$

The MTBF of the *PS* is given by:

$$MTBF_{PS} = \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{\lambda + \lambda}, \quad \text{or} \quad MTBF_{PS} = \frac{1.5}{\lambda}, \quad \text{or}$$

$$MTBF_{PS} = 1.5m \tag{3}$$

Where *m* is the MTBF of each DC-PDSS unit. It must be observed that for parallel systems even with data center PDSS units that have a constant failure rate (λ).

$$\lambda_{PS}(T) \neq \frac{1}{MTBF_{PS}} \tag{4}$$

The precise determination of the SoS reliability, the change of failure rate, and PDSS reliability of the surviving system need to be properly taken into account. The reliability of the SoS of two load-sharing exponential units shown in Fig. 4 is given by Kececioglu (1991) as follows:

R(t): Probability that Primary site (PS or 1) and Secondary site (SS or 2) complete their mission successful with *pdf*'s $f_1(T)$ and $f_2(T)$, respectively, or the probability that PS fails at $t_1 < t$ with *pdf* $f_1(T)$, and SS functions till t_1 with *pdf* $f_2(T)$ and the functions for the rest of the mission, on in $(t - t_1)$, with *pdf* $f_{2'}(T)$, or the probability that SS fails at $t_2 < t$ with *pdf* $f_2(T)$ and PS functions till t_2 with *pdf* $f_1(T)$ and then functions with *pdf* $f_{1'}(T)$ in $(t - t_2)$.

This scenario is presented in Table II and depicted in *pdf* form in Fig. 4. when both of PDSSs are exponential. The system model of integration of two equal system failure rates and SCT as parallel is illustrated in Fig. 4. Fig. 4 (a) and (b) shows the overall failure rate of system integration decrease under the parallel topology concept [1]. Fig. 4 (a) is illustrated chronologically and shows sequence reaction for the total system when Primary data center fails at time t_1 after starting the operation period. The indicator of changing factor is transformed by $\lambda_2 \Rightarrow \lambda_{2'}$ on Secondary data center. And, vice versa, the changing on Fig. 4 (b) is transformed by $\lambda_1 \Rightarrow \lambda_{1'}$ on Primary data center when Secondary data center failure during time t_1 of operation period, when $\lambda_{1'} = \lambda_{2'} = \lambda'$, is given.

TABLE II
MATRIX OF SOS, FUNCTION MODES, PDF'S AND TIME DOMAINS FOR ALL PDSS SUCCESS FUNCTION MODES [1]

PDSS Success function mode of DC number	SoS, function modes, *pdf*'s, and time domains	
	Primary Site (PS or 1)	Secondary Site (SS or 2)
1	G* $f_{ps}(T)$; t	G* $f_{ss}(T)$; t
2	B** $f_{ps}(T)$; $t_1 < t$	G* $f_{ss}(T)$; $t_1 < t$ $f_{ss}'(T)$; $t-t_1$
3	G* $f_{ps}(T)$; t_2 $f_{ps}'(T)$; $t-t_2$	B** $f_{ss}(T)$; $t_2 < t$

G*: SoS is good throughout the designed mission.
B**: SoS fails before the designed mission

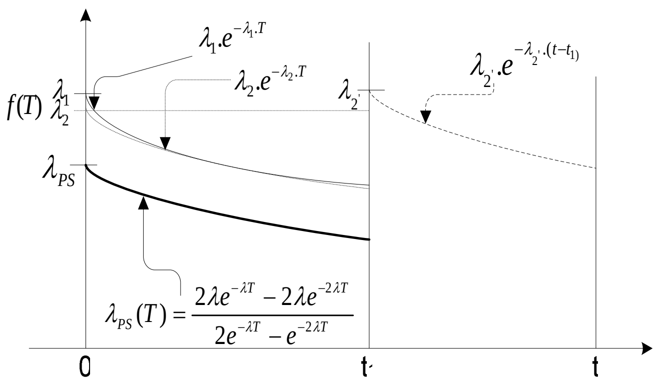
This is age and mission dependent even through the units are exponential. The first age mission $t = T$. Mathematically, the reliability of the parallel system (*PS*) is given by:

$$R_{PS}(T) = 2e^{-\lambda T} - e^{-2\lambda T} \tag{5}$$

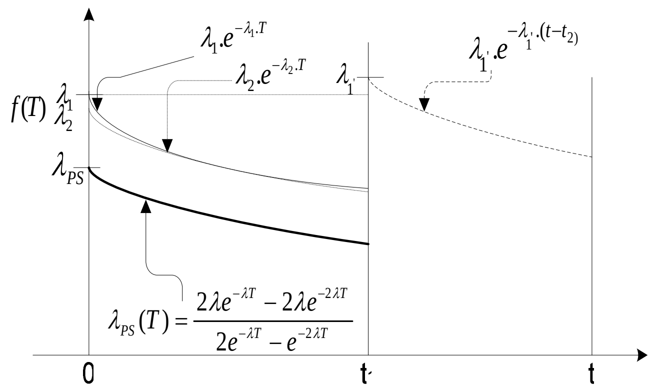
It must be observed that $\lambda_{PS}(T)$ is not constant, but a function of age, although each system of data center has a constant $\lambda_1 = \lambda_2 = \lambda$ and $\lambda_{1'} = \lambda_{2'} = \lambda'$.

System integration and operations of site plus site, normal operation of data center PDSS, of Tier IV data center is shown on Fig. 5. Utility is going to supply power throughout the PDS: Transformer, automatic transfer switch (ATS), bus-bar, UPSs, bus-bar, and loaded points. Calculation of SoS or parallel system (*PS*) of Tier IV data center is derived from Kececioglu (1991). Kececioglu defines $MTBF_{PS}$ of two systems equal to 1.5*m*, as referred to, in (3). Substitution λ = 14.0865x10^{-6} and T = 43,800 hours of Tier IV data center to (2). The result is $\lambda_{PS}(T)$ = 8.88215x10^{-6}. Substitution $\lambda_{PS}(T)$ (1), we will receive $MTBF_{PS}$= 112,585.371417 hours. The result from substitution $\lambda_{PS}(T)$ = 8.88215x10^{-6} and T = 43,800 hours to (5) shows the reliability *R(T)* of parallel system of Tier IV data center, SoS, during 5 years equal to *89.6128152%*.

 ISBN: 978-1-61208-192-2



(a). A pdf's of two parallel DC systems when primary DC fails



(b). A pdf's of two parallel DC systems when secondary DC fails

Fig. 4. Comparing between *pdf* and operation time of DC-PDSS load-sharing system [1]

The assumption on Fig. 4 shows $\lambda_{PS} < \lambda$, we need to prove $\lambda < \lambda'$ after Primary site failover to Secondary site, as shown in Fig. 5. Since in (4) derived results contrast with (1). We are given:

$MTBF_{1_t\,or\,2_t}$ is MTBF of each data center PDSS system before failure of Primary or Secondary site,

$MTBF_{1'_t\,or\,2'_t}$ is MTBF of each data center PDSS after failure of Primary or Secondary site,

λ is failure rate of each data center PDSS before failure,

λ' is failure rate of each data center PDSS after failure.

$$\lambda_{PS} < \lambda < \lambda' \text{ and } MTBF_{PS} > MTBF_{1_t\,or\,2_t} > MTBF_{1'_t\,or\,2'_t} \quad (6)$$

Now we have $MTBF_{PS}$ = 112,585.371417 hours and substitute to (3). We received m = 75,056.91 hours. When we compare with original MTBF from Tier IV data center that equal to 75,434.78 hours, that means $MTBF_{1_t\,or\,2_t} > MTBF_{1'_t\,or\,2'_t}$ and vice versa $\lambda < \lambda'$ as well for (1), as depicted in Fig. 4. Hence the assumption on (6) is correct.

$$R_{LS}(t) = e^{-2\lambda t} + \frac{2\lambda}{\lambda' - 2\lambda}(e^{-2\lambda t} - e^{-\lambda' t}) \quad (7)$$

For the two parallel load-sharing (*LS*) of data center PDSS's, if the data center PDSS's are equal, i.e., they have the same *pdf*, then $\lambda_1 = \lambda_2 = \lambda$ and $\lambda_{1'} = \lambda_{2'} = \lambda'$, we will derive (7) at t = 5 years, as illustrated in Fig. 4. The reliability result is *85.3775%*.

When one data center PDSS fails before the mission is completed in LS, and the data center PDSS is exponential. The reliability of the existing Secondary data center PDSS when Primary site fails calculates from (8), as shown in Fig. 3(a).

$$R_{2'} = (t - t_1) = \frac{R_{2'}(t)}{R_{2'}(t_1)} = \frac{e^{\lambda'_2 t}}{e^{\lambda'_2 t_1}} = e^{-\lambda'_1(t-t_1)} \quad (8)$$

Assumption, the Primary site fail at time = t_1, what is the reliability of the Secondary site during the left 3 year operation? Substitution t, t_2, and λ'; from previous calculation, and t= 43,800, t_1 = 17,520 hours to (8). We will get $R_{2'}(t) = 70.4593882\%$.

Normal operation of data center PDSS is depicted in Fig. 5 or on Primary site on left hand and Secondary site on the right hand. The utilities supply power throughout the PDSS: aerial cable, fuse, transformer, cable, automatic transfer switch (ATS), bus-duct, circuit breaker, cable, UPS, circuit breaker, cable, PDU, cable, and loaded points. After utility outage or PDSS fails on Primary or Secondary site, normal operation is degraded, when either of both sites is resume accomplishing the system will resume to normal operation, as illustrated in Fig. 6. Whenever, data center of Primary or Secondary site failure the services for external operations, end users, will not interrupt. The capacity to handle the transactions will reduce a half or increase double waiting queue or time for executing process. Absolute failure in case can be happened only when both sites are completely destroyed or malfunction at the same time [12].

IV. Discussion

A reliability of the parallel Tier IV data center system for a critical mission of T duration before the first failure or 5 years on this simulation is *89.6128152%*. This *89.6128152%* implies that the probability of survival on normal operation will still perform function continuously during mission time over 5 years or *10.3871848%* of the probability of the parallel Tier IV data center system which may fail before 43,800 hours of mission operation. The reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems; conversely the efficiency of LS system is better than the normal two parallel systems in terms of distributing transaction, service response times, and expanding capacity and capability.

The research defines the Tier IV DC parallel system failure when both Primary data center and Secondary data center fail at the same time. MTTR and maintenance systems of each data center are the keys to keeping parallel systems more reliable and available. The critical condition, which is MTTR, must be less than the maximum tolerable period of disruption (MTPD) [11]. MTTF of equipment is depended on 4 conditions; first, selected on CICs and designed on SCT of a data center system, second, operation of data center site conditions, third, related risks on daily operation of human activities, and last, procedural maintenance [5]. The most reasons of system failures come from human omission, e.g., they do not follow the manual instruction step by step, or commission, e.g., the system crashes during installing system.

 ISBN: 978-1-61208-192-2

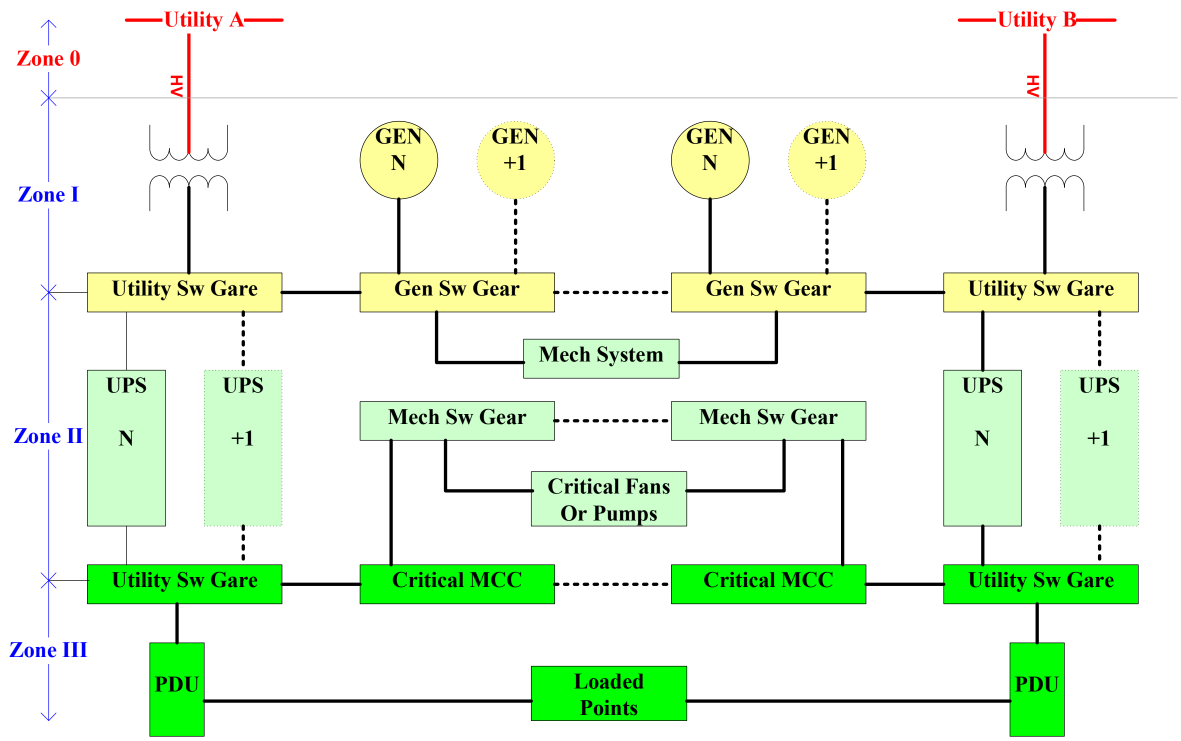


Fig. 1. Original Tier IV Data Center Diagram [3]

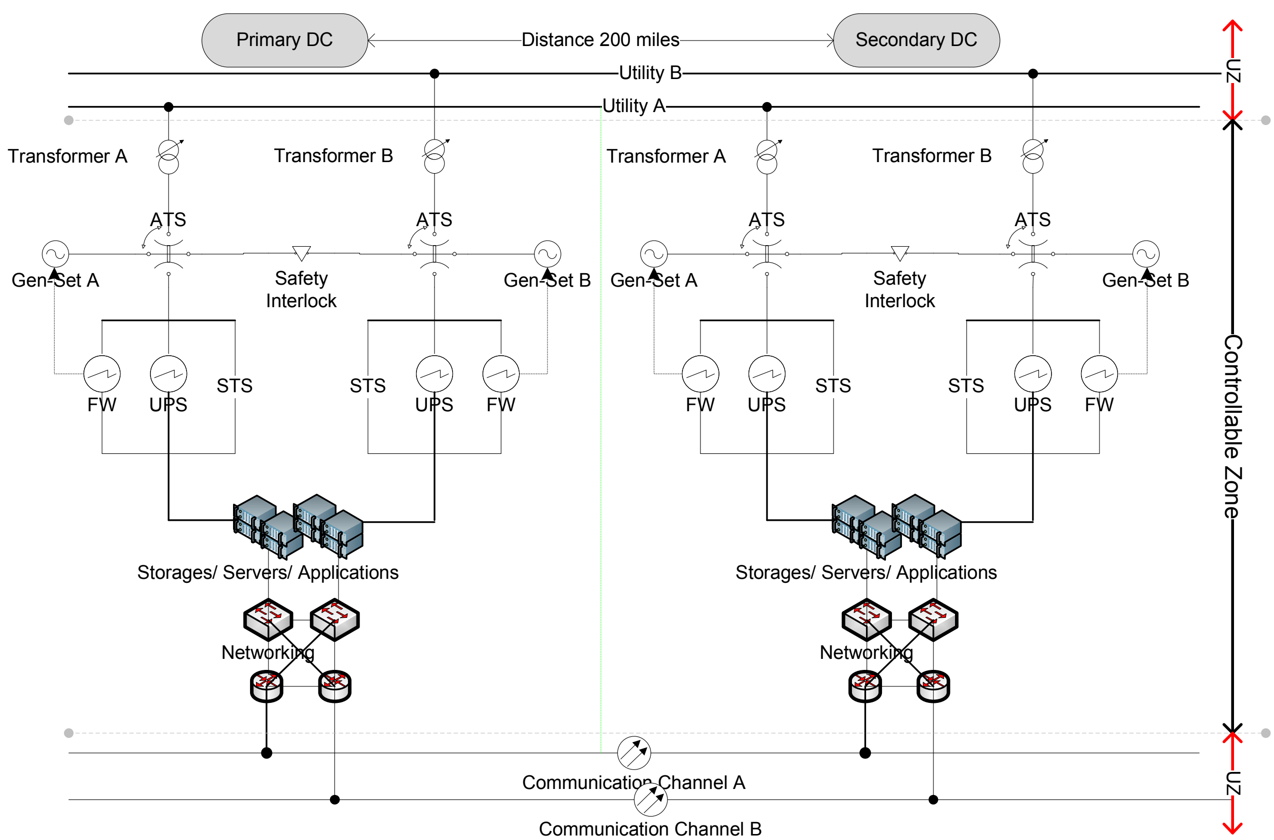


Fig. 5. Parallel System Design of Site plus Site Tier IV Data Center

 ISBN: 978-1-61208-192-2

TABLE I
EQUIPMENT RELIABILITY DATA FROM IEEE 493 GOLD BOOK [7]

Ref. #	Item Description	PREP Item #	Inherent Availability	MTTR (Hours)	Failure Rate Failure/Year	Calculated Availability	MTBF
1	Single Circuit Utility Supply, 1.78 failure/unit years, A=0.999705, Gold Book p.107	NA	0.999705	1.32	1.956		4,481.60
2	Cable Arial, ≤ 15kV, per - mile	32	0.99999022	1.82	0.04717		185,838.46
3	Diesel Engine Generator, Package, Stand-by, 1500kW	98	0.99974231	18.28	0.1235		70,979.76
4	Manual Disconnect Switch	187	0.9999998	1	0.00174		5,037,931.03
5	Fuse, 15kV	117	0.99995363	4	0.10154		86,330.51
6	Cable Below Ground in conduit, ≤ 600V - 1000 feet	47	0.99999743	11.22	0.00201		4,361,194.03
7	Transformer, Liquid, Non Forced Air, 3000kVA	208	0.99999937	5	0.00111		7,897,297.30
8	Ckt. Breaker, 600V, Drawout, Normally Open, > 600Amp	68	0.99999874	2	0.00553		1,585,171.79
9	Ckt. Breaker, 600V, Drawout, Normally Closed, > 600Amp	69	0.99999989	0.5	0.00185		4,738,378.38
10	Switchgear, BareBus, 600V	191	0.9999921	7.29	0.00949		923,709.17
11	Ckt. Breaker, 600V, Drawout, Normally Closed, < 600Amp	67	0.99999986	6	0.00021		41,742,857.14
12	Ckt. Breaker, 600V, Normally Closed, < 600 Amp, Gold Book p.40	63	0.99998948	9.6	0.0096		913,125.00
13	Ckt. Breaker, 3 Phase Fixed, Normally Closed, ≤ 600Amp	61	0.99999656	5.8	0.0052		1,685,769.23
14	Ckt. Breaker, 3 Phase Fixed, Normally Open, > 600Amp	62	0.99998532	37.5	0.00343		2,555,685.13
15	Cable, Above Ground, No Conduit, ≤ 600V, per 1000 ft.	20	0.99999997	2.5	0.00012		73,050,000.00
16	Cable, Above Ground, Trays, ≤ 600V, per 1000 ft, Gold Book p.105		0.99999831	10.5	0.00141		6,217,021.28
20	Cable Arial, ≤ 15kV, per - 300 feet	32		1.82	0.00268	0.9999994	3,270,895.52
22	Switchgear, Insulated Bus, ≤ 600V		0.99999953	2.4	0.0017	0.9999995	5,156,470.59
26	Bus Duct, Gold Book p. 206, per Circuit foot		0.99981596	12.9	0.000125	0.999816	70,080.18
60	Cable Below Ground in conduit, ≤ 600V - 300 feet			11.22	0.000603	0.9999992	14,537,313.43
80	Ckt. Breaker, 600V, Drawout, Normally Open, > 600Amp	68		2	0.002765	0.9999994	3,170,343.58
90	Ckt. Breaker, 600V, Drawout, Normally Closed, > 600Amp	69		0.5	0.000925	0.9999999	9,476,756.76
110	Ckt. Breaker, 600V, Drawout, Normally Closed, < 600Amp	67		6	0.000105	0.9999999	83,485,714.29
120	Ckt. Breaker, 600V, Normally Closed, > 600 Amp, Gold Book p.40	63		9.6	0.0048	0.9999947	1,826,250.00
130	Ckt. Breaker, 3 Phase Fixed, Nomally Closed ≤ 600 Amp, Gold Book p. 40	61		5.8	0.0026	0.9999983	3,371,538.46
140	Ckt. Breaker, 3 Phase Fixed, Nomally Open, > 600 Amp	62		37.5	0.001715	0.9999927	5,111,370.26
150	Cable, Above Ground, No Conduit, ≤ 600V, per 1000 ft.	20		2.5	0.000096	0.9999997	91,312,500.00

 ISBN: 978-1-61208-192-2

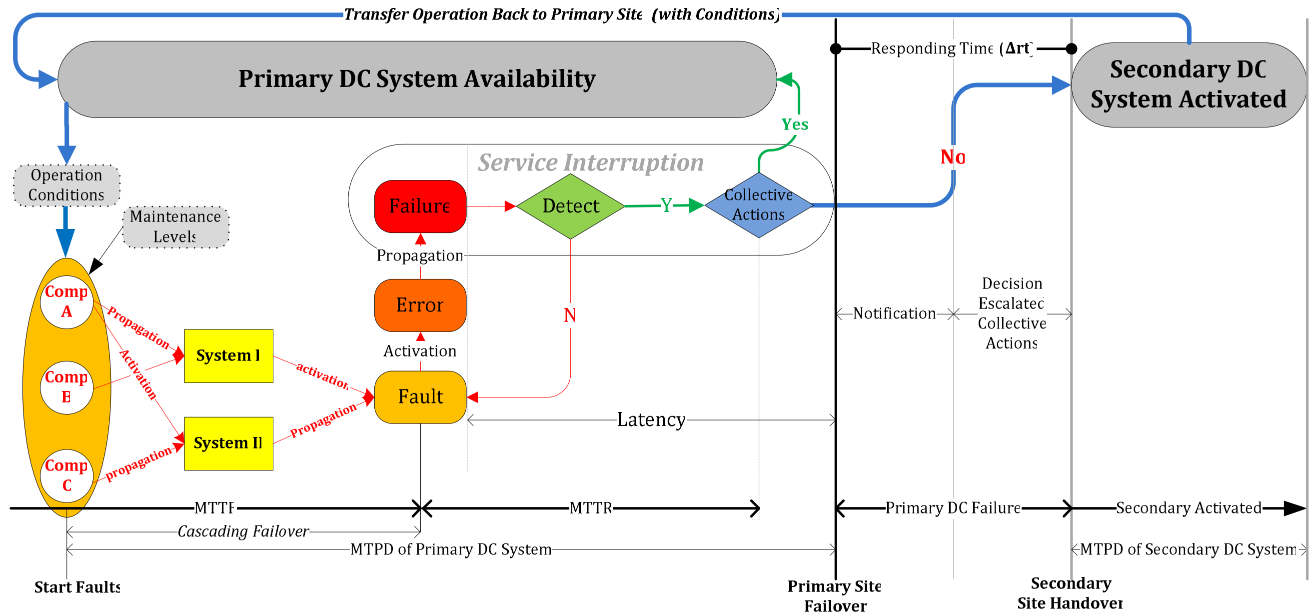


Fig. 6. System Failure Life Cycle Model.

During the design process, engineers, consultants, and designers need to understand throughout the transformation of a system failure cycle. The research results are derived from the root cause analysis of each system failure cycle to prepare the preventive actions and pre-planning for the corrective actions.

V. CONCLUSION

A simulation results from reliability block diagram (RBD) helps consultants, data center designers, IT managers, and contractors to foreseen the output of reliability system design. This helps to save time and costs from trial and error processes which in real data center operations cannot be accepted. To improve the system reliability data center designer needs to understand the MTTF, MTTR, CIC, and SCT of each type design pattern to optimize between reliability and investment. As a result from equation (6),

$$\lambda_{PS} < \lambda < \lambda' \text{ and } MTBF_{PS} > MTBF_{1_t\,or\,2_t} > MTBF_{1'_t\,or\,2'_t}$$

is shown the MTBF of each $\lambda_{PS} < \lambda < \lambda'$ presented reliability MTBF; *89.6128152%> 85.3775% > 70.4593882%* respectively, as in equation (6). This is implied that the reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems.

A regular maintenance, monitoring, and automatically control system is not only preventive and alert the system before disaster occurs but also help increase system reliability, system operations on energy conservative mode, and extending operations life-cycle of all equipments. All of these key factors are contributing to data center project operations success.

REFERENCES

[1] D. Kececioglu, "Reliability Engineering Handbook," Volume 2, Prentice Hall, Englewood, Cliffs, New Jersey, 1991.
[2] D. J. Klinger, Y. Nakada, and M. A. Menendez, "AT&T Reliability Manual," AT&T, Van Nostrand Reinhold, New York, 1st edition, February 3, 1999.
[3] Turner IV, W. P., J. H. Seader, V. Renaud, and K. G. Brill, Tier Classification Define Site Infrastructure Performance, White Paper, The Uptime Institute, Inc. 2008.
[4] M. Wiboonrat, "An Empirical IT Contingency Planning Model for Disaster Recovery Strategy Selection," IEEE International Engineering Management Conference, June, 28 to 30, 2008, Estoril, Portugal, accepted for publication.
[5] M. Wiboonrat, "An Empirical Study on Data Center System Failure Diagnosis," 3rd International Conference on Internet Monitoring and Protection, IEEE ICIMP 2008, Romania, June 29-July 5, 2008, pp. 103-108.
[6] M. Wiboonrat, "An Optimal Data Center Availability and Investment Trade-Offs," 9th International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/ Distributed Computing, IEEE SNPD 2008, Thailand, August 6-8, 2008, accepted for publication.
[7] M. Wiboonrat, "Beyond Data Center Tier IV Reliability Enhancement," Power Conversion and Power Management, Digital Power Europe, November 13-15, Munich, Germany, 2007.
[8] M. Wiboonrat, "Power Reliability and Cost Trade-Offs: A Comparative Evaluation between Tier III and Tier IV Data Centers," Power Conversion and Power Management, Digital Power Forum, 10-12 September 2007, San Francisco.
[9] M. Wiboonrat, and C. Jungthirapanich, "A Taxonomy of Causal Factors and Success Criteria on Project Management Success and Project Success," Eight Int. Conference on Opers. & Quant. Management, ICOQM 2007, October 17-20 2007.
[10] M. Wiboonrat, "Dependability Analysis of Data Center Tier III," 13th International Telecommunications Network Strategy and Planning Symposium, Networks 2008 in Budapest, Hungary, September 28 – October 2, 2008, accepted for publication.
[11] IEEE Std. 1100-1999, (Revision of IEEE Std. 1100-1992), IEEE Recommendation Practice for Powering and Grounding Electronic Equipment, 22 March 1999.
[12] M. Wiboonrat, "Transformation of System Failure Life Cycle," International Journal of Management Science and Engineering Management, Vol. 4, No. 2, 2008, pp. 143-152.

 ISBN: 978-1-61208-192-2

Content Delivery Architectures for Live Video Streaming: Hybrid CDN-P2P as the best option

Melika Meskovic
BH Telecom, Joint Stock Company
Sarajevo, Bosnia and Herzegovina
melika.meskovic@bhtelecom.ba

Himzo Bajric
BH Telecom, Joint Stock Company
Sarajevo, Bosnia and Herzegovina
himzo.bajric@bhtelecom.ba

Mladen Kos
Faculty of Electrical Engineering and Computing
Zagreb, Croatia
mladen.kos@fer.hr

Abstract: **In the last years, with the help of high-speed and broadband networking, the content delivery service has been grown up widely. There are a lot of providers for online streaming via Content Delivery Network (CDN), Peer to Peer (P2P) network or hybrid CDN-P2P system. In this paper, we analyse a hybrid solution for real-time streaming: hybrid CDN-P2P mechanism that takes the best of both CDN and P2P content delivery architectures. By adapting the best of breed of both worlds (CDN and P2P), we detail a hybrid model which, thorough quantative analysis, shows the benefits of merging the two architectures. Through simulation, we will show that a hybrid CDN-P2P approach is a much more stable platform whilst at the same time a very cost effective approach in providing a live streaming service to the masses.**

Keywords – live streaming; CDN; P2P; hybrid CDN-P2P

I. INTRODUCTION

Live video streaming has long been projected as the “killer-application” for the Internet. While this expectation has been in effect for several years now, only in recent years with the deployment of increased bandwidth at the end user’s side has this promise finally turned into reality (e.g., [14, 15, 16]).

CDN and P2P networks have been most accepted technologies in use for delivering streaming content today. In a CDN scenario, the streaming content is delivered to sibling CDN servers that are placed in various georaphical regions and used to reduce the overall load on the streaming source. When a client requests for the streaming content, the CDN server closest to that client will deliver the stream and not the CDN server acting as the main source of the stream. This architecture (CDN) is caracterized by very high bandwidth capacity and huge disk space thus making it the best option to provide the highest quality stream. However, the economical factors, such as cost of hardware, and the complexities with the scalability make this model less popular.

Multiple other solutions have been found to reduce the number of deployed servers [2][4][5], thus trying to overcome the economical factors of CDN to some extent. However, the provided service on those solutions does not match the quality of service provided by the pure CDN architecture.

On the other hand, P2P streaming network [7], [8] introduces a concept of a completley decentralized system. Once the content is received, a peer automatically becomes a source of the stream to other peers. An increase of active peers increases the quality of the stream delivery since the model is based on peer abundant bandwidth utilization. In order to get the quality of service that is provided by the CDN architecture, a P2P architecture would require a huge number of participating peers.

This brings us to an evaluation of a hybrid CDN-P2P solution, which is highly recommended in order to eliminate all the weaknesses of those two original architectures. By using this type of architecture, we can have a cost-effective streaming system. This type of delivery system (hybrid CDN-P2P architecture) benefits from the advantages of two technologies: the use of CDN servers assures the best quality of streaming service and use of the P2P network reduces the price of system thus resulting in a cost-performance content delivery service.

In a given hybrid CDN-P2P architecture, a CDN server usually acts as a component which assures the availability of requested resource (in our case a live video stream) and the speed of transmitting the requested resource.

In contrast, a peer is not only a request component but is also a support function to the CDN server in the process of content delivery to other peers. A large number of well-organized peers can reduce the server load significantly.

In this document, in Sections 2 and 3, we will cover the basic infrastructure characteristics for both CDN and P2P networks taking a look at their functionality and overall benefits as well as the shortcomings. In Section 4, we will then introduce our new proposed architecture that is made up of a hybrid solution taking both previous architectures as the basis of our new hybrid cdn-p2p system. Section 5 will cover the methodology based on which we plan to simulate the proposed architecture. This section will be followed by analysis of our results, and finally, a conclusion.

II. CONTENT DELIVERY NETWORK (CDN) ARCHITECTURE

The Content Distribution Network (CDN) is the most used technology for real-time content distribution. Grouped into sets of dedicated servers, CDN servers have a very large bandwidth and a huge capacity of storage. This enabled them to deliver data to a large amount of users. These servers are often organized at a hierarchic structure and are placed in multiple locations over multiple backbones. There are three kinds of servers in a CDN group set:

 ISBN: 978-1-61208-192-2

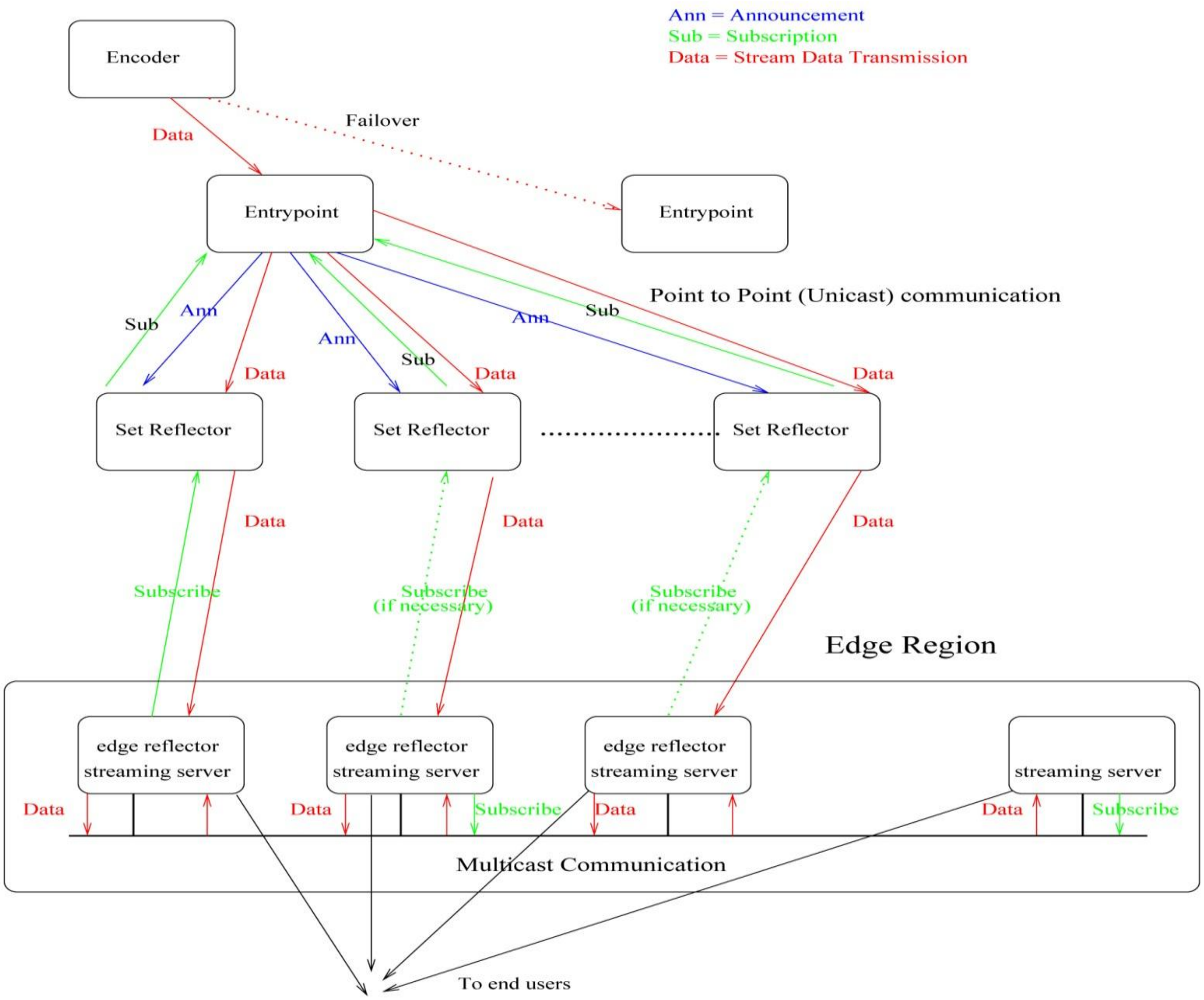


Figure 1. Akamai content delivery network [2]

- Encoder Server - gets and converts media from media source into small chunks
- Transport Server - distributes data in the network
- Edge server - transfers media to end-user

As an example, Figure 1 illustrates the architecture of a Akamai CDN.

Analyzing the architecture from the stream source and all the way to the end user, we first have an encoder server which is by default closest to the stream source (i.e. video stream or a audio stream). This positioning enables the encoder server to get the content fastest as it needs to encode it before it is sent out to the rest of the network.

The next server in the CDN architecture is the transports server which is responsible for distribution/transporting the encoded content to the edge servers. The transport server must have a very large capacity of storage because it has to store a lot of encoded data.

The rest of CDN network is made up of edge servers which are the closest to end user. In a hierarchic architecture, an edge server is a leaf which manages its end user.

For performance optimization, there also can be several tracker servers which are used to balance the server load between all the servers in the network. Sometimes, any given server in the architecture, irrelevant of its original purpose, can be used to do this task provided that it has free resources available.

The process of obtaining the stream from the end user's side starts with a tracker server detecing the edge server that needs to be used which is closest to the originating user.

Once the edge server has been identified (i.e. tracked), all the requests of media will be transferred to that edge server. Whenever an edge server cannot provide the content, it will hand-over the request to another edge server. This task of organization is done by server load balancing which uses one or more "layer 4-7 switches".

Taking Akmai as an example, a DNS forwarding mechanism is used to redirect request coming from clients in order to equilibrate server-load between CDN servers and to make the content distribution more effective.

To enhance the the quality and the reliability of the stream, a provider can also use some fault tolerant servers or backup servers [11] to assure that there is always response to any given request and that there is no sudden break of data transfer. This is one of the reasons why a CDN type network usually costs a lot. Hoever, this is also one of the factors that enable a CDN network to provide an unrivaled streaming quality.

III. PEER TO PEER (P2P) ARCHITECTURE

Whilst P2P architectures have been mainly used for file sharing in the past, today they are more and more used to delivery media content [6], [8]. In a P2P streaming network, we also have a media server which gets the stream from the source and distributes it to the network so that the content can be distributed by all the participants (referred to as audiences).

A mesh-pull P2P live streaming architecture often has three major components:

- Streaming peer node – this node includes a streaming engine and a media player
- Channel streaming server - converts the media content into smaller chunks with each chunk being composed by some piece
- The tracker server - provides the streaming channel, peer and chunk information for each peer node that joins the network.

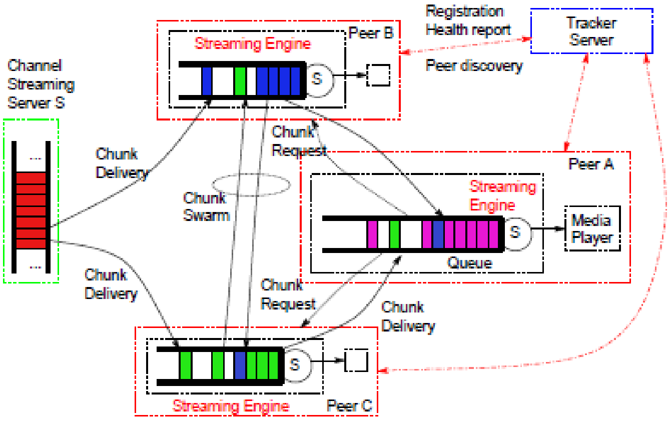


Figure 2. P2P streaming process [8]

On joining the network, a peer first downloads the list of distributed channels available on the network. After selecting a channel from the provided list, the node is registered in the tracking server. From that point onwards, like other peers that have already registered to the same channel, this node participates in the process of streaming the media to other existing nodes or new joining nodes. At the media playing stage process, the peer downloads the list of pieces available in other peers so that it will know which is the best peer that will respond the request of missing pieces in its playing buffer.

Even though P2P streaming network works in a de-centralized mode, there are also servers called „trackers" that store peers and channel information. A tracker server can be a regular computer which has a limited capacity of storage but also has a fast internet connection. That is possible since the information that needs to be stored is in simple text format which makes the requirement for hige disk space mute. In contrast, the tracker server must have a high speed internet connection so that it can send the peer and channels list as fast as possible.

PPLive is the most popular P2P streaming architecture today. It has been reported that a PPLive supported 1,480,000 nodes viewing a live media stream at the same time with 1 PC server and 10Mbps bandwidth.

IV. HYBRID CDN-P2P ARCHITECTURE

Both of the two above mentioned technologies have their advantages and disadvantages. A CDN can assure the quality of service by using distributed CDN servers with high bandwidth and large capacity of storage. But these servers often cost too much. In contrast, a P2P Live streaming system is much cheaper but the speed of media streaming depends on the number of joined peers and their availability of content resource, internet connection. PPlive can be used in cases which need to serve a huge number of audiences in the same time; however, they cannot assure the quality of service and cannot serve a special requirement of high definition content.

Therefore, hybrid CDN-P2P architecture is indispensable to have the best solution for content streaming, in particular for live streaming service.

A. System Overview

In this section, we describe the general architecture of the proposed system. As shown in Figure 1, there are three major components: (1) Management Center (MC) comprising the DNS-based Global Server Load Balance (GSLB) system, content management and configuration system, and monitoring and billing systems; (2) cache servers, referred to as Service Nodes (SN) that deliver video contents from content providers to end users; (3) end hosts which may either be legacy clients, which directly obtain the stream from the edge servers or LiveSky-enabled clients, which can additionally engage in P2P transfers.

System Management: The Management Center (MC) is responsible for efficient control and monitoring of the proposed system. The DNS-based GSLB system in the MC redirects user requests to the nearest, lightly loaded server [3].

The MC distributes configurations to the SNs using XML messages; these messages use incremental updates to reduce the communication overhead. The configuration parameters include channel information, source information, operating strategies etc.

B. CDN Overlay

The SNs are organized into several tiers, with Tier0 being closest to the content source and Tiern−1 closest to the end users as shown in Figure 3. We refer to the SNs in Tiern−1 as edge SNs since they are directly responsible for serving end users. The SNs in the remaining tiers are core SNs since their primary responsibility is to act as a distribution overlay to deliver the content to the edge SNs. This hierarchical arrangement is typical of many CDN infrastructures to effectively magnify the total system capacity, reduce the load at the content source, and also leverage the benefits of caching requested contents in higher layers.

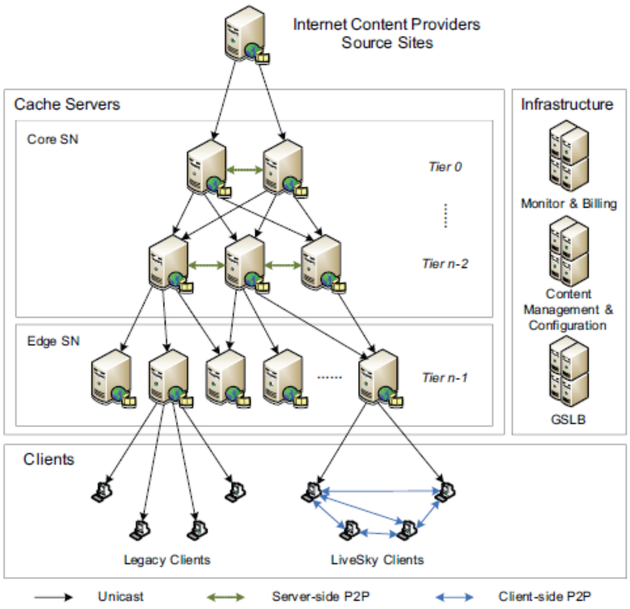


Figure 3. System architecture

Each SN is allocated a unique ID. When SNi boots up, it sends a “alive” message to the MC. The MC then broadcasts the alive message to other SNs. A different SNj can obtain the attributes of SNi (e.g., IP address and TCP port information) from the MC to establish a TCP connection with SNi if necessary.

The server-side distribution mechanism is largely tree-based. However, in order to provide greater reliability in the presence of node or network failures, we allow each SN to retrieve the content either from SNs higher up in the hierarchy (i.e., a lower numbered tier) or from peer SNs in the same tier. Since the edge SNs are responsible for serving end users, they are typically heavily loaded and we disable peering between SNs in the edge tier.

Edge SNs handle client requests and obtain the required contents from the core SNs. Requests from edge SNs are forwarded up the hierarchy until they find a node that has the desired content. To minimize the load at the content source, only Tier0 SNs retrieve content directly from it. The goal of the server-side overlay is efficient data distribution with some measures to guard against some node failures and network delays. As the CDN nodes have high availability and are stable, a tree-based overlay with additional peer edges satisfies the goals of providing reliable, yet efficient data transmission.

C. System operation

A client first obtains the URL for the live stream from the content source (e.g., http://domainname/live1). The GSLB component of the CDN takes into account the client location, the edge SN location, and the edge SN loads to find a suitable edge SN for this client. The client is then redirected to this edge SN using traditional DNS-based redirection techniques [3].

Each edge SN serves multiple roles. First, it acts as a regular server for legacy clients. Second, it serves as a tracker for the P2P operation to bootstrap new clients with candidate peers. Third, it acts as a seed for the P2P operation for the proposed system-enabled clients assigned to it. The edge SNs are pre-configured with some decision logic that decides if a new proposed system-enabled client should be served in CDN-mode or if they should be redirected to the P2P overlay. Finally, the edge SN is used for some optimizations in the P2P operation. Note that the P2P overlays are localized on a per-edge SN basis; i.e., the peers with which a LiveSky enabled node communicates in the P2P mechanism are also assigned to the same edge SN as this node. We discuss these last two roles in more detail in the next section.

D. Client distribution

Legacy Clients: As discussed earlier, there are two types of clients: legacy clients which receive contents directly from the edge SNs and LiveSky enabled clients which can either receive contents from the edge SNs or additionally use P2P mechanisms. An important distinction between the legacy and LiveSky clients is that the LiveSky clients can access a higher quality video stream whereas the legacy clients may only be able to access a lower quality stream. This incentivizes users to install the LiveSky client software and encourages widespread adoption. In our experience, we find that typically more than 50% of users have adopted LiveSky.

LiveSky’s P2P Mechanism: Recent proposals [17, 18] demonstrate that a hybrid approach combining the multi-tree [19, 16] and mesh [20, 21] schemes achieves both efficient delivery and robustness to churn. We adopt a similar scheme in the proposed system. The video stream is a single bit-rate encoding (i.e., we do not use any layered coding) and is separated into several sub-streams according to the stream frame id. For example, if the video is divided into six sub-streams, substream0 consists of frames 0, 6, 12, 18, . . . substream1 consists of frames 1, 7, 13, 19, . . . and so on. Peers are organized in a tree-based overlay on a per sub-stream basis. This ensures that all nodes contribute some upload bandwidth. Additionally, in order to be robust to network or node failures, peers also use a mesh-style pull mechanism to retrieve missing frames for continuous playback.

V. SIMULATION

In order to evaluate the effectiveness of hybrid CDN-P2P solution, we have simulated the network architecture and have given a certain test case.

The NS2 Simulator has been chosen in order to test our scenario. The NS2 is a discrete event simulator targeted at networking research. NS provides substantial support for simulation of TCP, routing, and multicast protocols over wired and wireless (local and satellite) networks. This simulator is used in the simulation of routing protocols, among others, and is heavily used in ad-hoc networking research, and supports popular network protocols, offering simulation results for wired and wireless networks alike [22].

Before we started our simlation, we had to categorize the nodes into two types. The chosen types for the simlation were

- CDN Server type - nodes that act as CDN servers
- Peer types – nodes that act as Peers

Once we have completed the categorization, we introduced a buffer to all the nodes in the network. The size of the buffer used indicates the number of slots that can store a piece of transmitted data during a playback time. Keeping that in mind, all nodes that have been categorized as CDN servers would have unlimited capacity for the buffer size since while the nodes categorizes as Peer types would have limited capacity for the buffer size assigned.

We assume that CDN server nodes are always ready to immediatley transmit, when requested by the peer nodes, any piece of stream that has been already received by the stream source without going back to the stream source.

For transmission between the CDN server node and a peer node, we will assume the same delay as between two peer nodes.

After that has been completed, each local peer's buffer is then divided into two parts as already mentioned above. The percentage of P2P part in the buffer is defined by the configuration file.

In a non-simulated environment (i.e. in an application environment), this parameter would actually be variable and able to adapt to the changing number of CDN servers and peer clients which would profit from the maximum work load of CDN servers.

Major parameters of our simulated network are:

- Network size (N): the number of nodes including CDN servers.
- Network protocol: protocol used in each node.
- Buffer length (l_{buffer}):
- Chunk size (l_{chunk})
- P2P percentage in a buffer (a)
- CDN bandwidth (b_{CDN})
- P2P bandwidth (b_{P2P})
- delay between nodes ($d_{transport}$)
- delay between CDN-source ($d_{CDN\text{-}source}$)

We have tried to simulate real-time streaming the video content of 400Kbps (or 50KB/s) which has a quality of a business video conference. In the simulation, we consider each piece of data has 5KB length (or 40Kb). Suppose that each play out is for 1 second of media. Hence, a chunk to play has a length of 400Kb. CDN servers in our simulations have an upload capacity of 10Mbps. In real world, internet connection speed of peers are usually much variable, but we suppose that each peer in our network have a connection of 512Kbps. We consider that each node can buffers up to 20s of playing time. Our network uses also an unreliable transport to make it more reality. From that, we can than adjust the rate of lost packet.

After having taken a look at some peer-to-peer applications, we found that delays between peers are from 20ms to 1500ms [13]. Therefore, we apply this interval of delay to all the transaction, not only transactions between two peers but also transactions between a peer and a CDN server. Furthermore, it takes a little delay when CDN requests content from source so we define this delay ($d_{CDN\text{-}source}$) too. Hence, each transaction j-th has a random delay $d_j = d_{CDN\text{-}source\_j} + d_{transport\_j}$.

TABLE 1 TEST PARAMETERS

Name	Value used in simulation
Video codec	400Kbps
CDN bw	10Mbps
P2P bw	512Kbps
Lost packet	Random value: 0 – 20%
Delay between CDN-P2P	Random value: 20 – 1000ms
Delay between P2P-P2P	Random value: 20 - 1000ms
Delay between CDN-source	Random value: 0 - 15ms
Buffer length	20s of playing content
Network size	Depend purpose of test
Chunk size	10 pieces

VI. RESULT

We simulated with the main parameters described in previous section and we changed the number of peers participated, rate of packet lost.

First delay is a value of time which a media player must wait for from beginning of streaming process to play out the first chunk. In the media on demand network, first delay is not very important and we can tolerate it but in the real-time streaming mode, first delay has a very important role in streaming process. Suppose that the streaming, lately does not have any delay time during the playback but the first delay is too much, so the content of media be played in the media player would be older and there is no meaning of real-time service. For example, users joined into a football match streaming session, even there are not any interruption during the session but if the first delay is long, people would see a goal lately. In our simulations, we tried to evaluate the first delay in different configurations to know if we can have a best configuration which assures the smallest first delay. As we can see in the Figure 4, the first delay is usually longer than other delays. The reason is: at the beginning of streaming, all the peers in network do not have any data in buffer. In other words, all slots in buffer map are empty. Hence, it must wait for streaming process to fill in at least first lchunk slots so that media player can play out the first chunk.

Furthermore, when a peer sends a request to other peers or to CDN server, it will receive (if possible) list of pieces responded randomly. For example, the peer named "A" need a list of piece {0, 1, 2, 3, 4, 5, 6, 10, 12, 14, 16, 18} to fill in the buffer, it sends this list to peers which have any of those piece. Because the bandwidth of a peer is limited; one peer can send only k pieces at one time then if k < size of list request, that peer would choose random k piece in his available response list to reply back to peer A. Hence, peer A will receive a chaotic response list from others peers. That is why, to fulfill all lchunk pieces of first chunk (from start point), it can take a longer time.

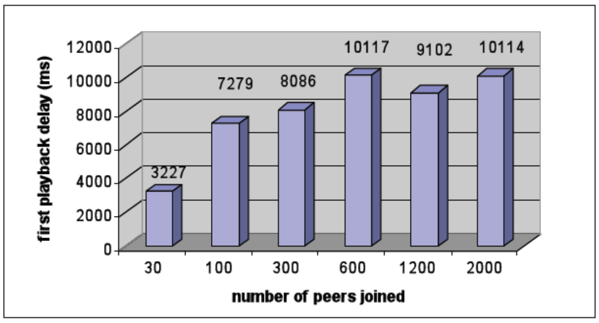


Figure 4. Playing delay of first 100 chunks

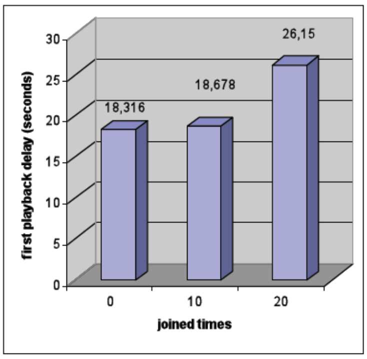


Figure 5. Playing delay in different joined times

From the second chunk, the delay time is almost reduced because after each playback, the buffer map is moved forward so there are slot in P2P part which have data will be transfer to CDN part. Therefore, the CDN part can be fulfilled even much faster. This effect also proves the important impact of P2P part: prepares and reduces waiting time of playing buffer.

However, there are also other moments that buffer must wait to play next chunk because in the simulator. The reason is the way we choose piece to send back each time is randomly like we have discussed above. There are properly pieces that could be received much longer than other. Hence, to play chunks which these pieces belong to, it takes longer time than other earlier chunks.

Joined times also has impact to the first delay of streaming process. In the figure above, we consider a peer join from beginning of stream. Now, we let audiences join at different time and see how first delay change in different nodes. We then take a test with 1000 nodes. The results show that, a peer which joins a streaming later would wait longer to play first chunk. Additionally, network size can have impact to first delay too. If the number of joined peers in the streaming increases, the first delay of a given peer who joined from the beginning of streaming could be longer. In real live streaming application, audiences usually join in at same time from the beginning of a live streaming session, for example to view a live football match. Therefore, we must reduce the first delay so that the streaming has a meaning of "real-time".

VII. CONCLUSION

By combining the best of both approaches in delivering live streaming content to end users (the CDN architecture and P2P architecture) we were able to show that a hybrid model based on the above mentioned architectures can provide a much better service than either architecture individually.

Having setup a simulation environment in NS2 simulator, we tested different cases that all had the same QoS parameters which were used to evaluate performance of the overall network. We tested the setup environment using 1000 nodes and different join times (which may not be the case for live-streaming as usually everyone would join at approximately the same time).

By concentrating on the most important factor of live-streaming service provisioning (i.e. the playout delay), we have shown that in our simulation, with the number of joining peers varying all the time and whilst the QoS factors are being fulfilled, the first playback delay's performance on a network peeks at approximately 600 peers. At this point, the QoS parameters are still met and past this point the first playout delay does not cross over 11000 ms even at a point of 2000 peers joining.

This shows us that a hybrid CDN-P2P approach is a much more stable platform whilst at the same time a very cost effective approach in providing a live streaming service to the masses.

REFERENCES

[1] Dongyan Xu, Sunil Suresh Kulkarni, Catherine Rosenberg, and Heung-Keung Chai, "A CDN-P2P hybrid architecture for cost-effective streaming media distribution," Computer Networks, vol. 44, pp. 353-382, 2004.

[2] Leonidas Kontothanassis, Ramesh Sitaramant, Joel Wein, Duke Hong, Robert Kleinberg, Brian Mancuso, David Shaw, and Daniel Stodolsky, "A transport layer for live streaming in a content delivery network," Proceedings of IEEE, vol. 92, issue 9, pp. 1408 – 1419, September 2004.
[3] ChinaCache CDN Principle, http://en.chinacache.com/index.php/products-and-services/streaming.html., December, 2011.
[4] Rajkumar Buyya, Al-Mukaddim Khan Pathan, James Broberg, and Zahir Tari, "A Case for Peering of Content Delivery Networks," IEEE Distributed System Online, vol. 7, issue 10, pp. 3 – 3, October 2006.
[5] John Dilley, Bruce Maggs, Jay Parikh, Harald Prokop, Ramesh Sitaraman, and Bill Weihl, "Globally Distributed Content Delivery," IEEE Internet Computing, vol. 6, issue 5, pp. 50-58, September 2002.
[6] Meeyoung Cha, Pablo Rodriguez, Sue Moony, and Jon Crowcroft, "On Next-Generation Telco-Managed P2P TV Architectures," Proceedings of the 7th International Conference on Peer-to-peer systems (IPTPS'08), pp. 5-5, 2008.
[7] Thomas Silverston and Olivier Fourmaux, "Measuring P2P IPTV Systems," Traffic, Retrieved from http://www.nossdav.org/2007/program.html, 2007.
[8] Xiaojun Hei, Chao Liang, Jian Liangy, Yong Liu, and Keith W. Ross, "A Measurement Study of a Large-Scale P2P IPTV System" IEEE Transactions on multimedia, vol. 9, issue 8, pp. 1672 – 1687, December 2007.
[9] Jian Ni, Danny H. K. Tsang, Ivan S. H. Yeung, and Xiaojun Hei, "Hierarchical Content Routing in Large-Scale Multimedia Content Delivery Network," IEEE International Conference on Communications (ICC '03), vol. 2, pp. 854 – 859, May 2003
[10] Afergan Michael, Leighton Thomson, and Parikh Jay, "Hybrid content delivery network and peer-to-peer network," United States Patent Application 20080155061, June 2008.
[11] Jian Ni and Danny H. K. Tsang, "Large-scale cooperative caching and application-level multicast in multimedia content delivery networks," IEEE Communications Magazine, vol. 43, issue 5, pp. 98-105, May 2005
[12] Cheng Huang, Angela Wang, Jin Li, and Keith W. Ross, "Understanding Hybrid CDN-P2P: Why Limelight Needs its Own Red Swoosh," Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV '08) Germany, pp. 75-80, May 2008
[13] http://www.bittorrent.com/, December, 2011.
[14] Eric Franchi, "Live Streaming Continues Momentum With March Madness," http://www.mediapost.com/publications/article/101750/, November, 2011.
[15] Marshall Kirkpatrick, "The Numbers Are In, Live Video Online Is Blowing Up," http://www.readwriteweb.com/archives/live_video_big.php, November, 2011.
[16] Liz Gannes, "The Obama Inauguration Live Stream Stats," http://newteevee.com/2009/01/20/the-obama-inauguration-live-stream-stats/, November, 2011.
[17] Mengjie Zhang, Yaobin Lu, Ling Zhao, and Shuiqing Yang, "A Peer-to-Peer Network for Live Media Streaming – Using a Push-Pull Approach," Proceedings of ACM International Conference on Multimedia, pp. 287-290, 2005.
[18] Susu Xie, Bo Li, Gabriel Y. Keung, and Xinyan Zhang, "Coolstreaming: Design, Theory, and Practice," IEEE Transactions on Multimedia, 9(8), pp.1661–1671, May 2007.
[19] Vidhyashankar Venkataraman, Paul Francis, and John Calandrino, "Chunkyspread: Multi-tree Unstructured Peer-to-Peer Multicast," Proceedings of IPTPS, pp. 2-11, Februar 2006.
[20] Xinyian Zhang, Jiangchuan Liu, Bo Li, and Tak-Shing, Pete Yum, "CoolStreaming / DONet: A Data-driven Overlay Network for Peer-to-Peer Live Media Streaming," Proceedings of IEEE (INFOCOM 2005), pp. 2102-2111, March 2005.
[21] PPLive. http://www.pplive.com., December, 2011.
[22] http://www.isi.edu/nsnam/ns/, December, 2011.

Enabling QoS in the Internet of Things

Marie-Aurélie Nef, Leonidas Perlepes, Sophia Karagiorgou, George I. Stamoulis
Department of Computer & Communication Engineering
University of Thessaly
Volos, Greece
e-mail: {manef, leperlep, sokaragi,georges}@inf.uth.gr

Panayotis K. Kikiras
AGT Germany
Darmstadt, Germany
e-mail: pkikiras@agtgermany.com

***Abstract*—With the emergence of the Internet of Things (IoT), it is necessary to define service models, which can categorize IoT applications and determine the Quality of Service (QoS) factors necessary to satisfy the requirements of those services. On the other hand, as Wireless Sensor Networks (WSN) constitute a main component of the IoT, they become a key factor concerning QoS provision. In this perspective, we focus our analysis on the possible WSNs integration approaches in the IoT while providing QoS and which best practices to adopt. Furthermore, regarding QoS requirements, we also define service models for the IoT and expose their feasibility through a categorization of IoT applications.**

Keywords-Internet of Things; Wireless Sensor Network; QoS; Service Models; MAC

I. INTRODUCTION

The Internet of Things (IoT) will likely be one of the most important technological breakthroughs of the years to come. IoT could be conceptually defined as a dynamic global network infrastructure with self configuring capabilities based on standard and interoperable communication protocols where physical and virtual things have identities, physical attributes, and virtual personalities, use intelligent interfaces, and are seamlessly integrated into the information network [1].

In the IoT, smart things/objects are active participants in business, information and social processes where they are enabled to interact and communicate among themselves and with the environment by exchanging data and information sensed about the environment, while reacting autonomously to the real/physical world events and influencing it by running processes that trigger actions and create services with or without direct human intervention [1]. In this perspective, it is necessary to define service models, which can categorize IoT applications and then determine which Quality of Service (QoS) factors are necessary to satisfy the requirements of those services.

Smart objects are lightweight devices with a sensor or actuator and a communication device. These devices are capable of sensing various types of incidents/parameters and communicating those with other devices. They can be battery-operated, and typically have three components: a CPU (8, 16 or 32-bit microcontroller), memory (a few tens of kilobytes) and a low-power wireless communication device (from a few kilobits/s to a few hundreds of kilobits/s). The size of these devices is very small [2][3]. These devices can work together, forming for example a wireless sensor network (WSN). As a main component of the IoT, WSNs become a key factor concerning QoS provision and therefore should be integrated in the IoT in the best possible way.

In this paper, based on the analysis of the current QoS MAC solutions in WSNs and simulation results, we focus our QoS analysis on the possible WSNs' integration approaches in the IoT. Then, regarding QoS requirements, we also define service models for the IoT and expose their feasibility through a categorization of IoT applications.

The remainder of this article is organized as follows. Section II focuses on a review of current QoS MAC solutions in WSNs. In Section III, we provide a summary of different service models and performance analysis of the IEEE 802.15.4 from [4]. The fourth Section describes a WSN integration approach in the IoT providing QoS. Then we propose best practices to adopt when using the IEEE 802.15.4 protocol for WSNs while providing the aforementioned service models. In the Section V, we extend the service models described in Section III to the IoT and we present a categorization of IoT applications according to them. Finally, conclusions are presented in Section VI.

II. MAC SOLUTIONS FOR QOS IN WSNS

Many aspects of WSNs such as routing, preservation of battery power and topology control have been studied in previous papers [5], [6], [7]. However, the area of QoS in WSNs remains largely open. The main reason is that WSNs are quite different from traditional wired and wireless networks from several points of view (e.g. energy, processing and memory constraints, heterogeneous and unevenly distributed traffic, network's dynamic changes and scalability problems).

In the following paragraphs, a summary of current QoS-aware MAC solutions for WSNs is provided. Two complete surveys of QoS-Aware MAC protocols and Real Time (RT) QoS support can be found in [8] and [9] respectively. The main characteristics of each protocol are described below:

1) Implicit prioritized access protocol (I-EDF) [10] *and dual-mode MAC protocol* [11]: they adopt a cellular backbone network and thus they are topology-dependent. They use Frequency Division Multiple Access (FDMA) and Time Division Multiple Access (TDMA) to guarantee bounded delay (HRT). Energy efficiency is not considered.

2) PEDAMACS [12]: this TDMA-based protocol that aims to achieve both energy efficiency and delay guarantee

TABLE I. A COMPARISON OF MAC PROTOCOLS

Name	MAC type	RT type	Topology dependent	Energy efficient	Scalability
I-EDF, Dual-mode MAC	FDMA-TDMA	HRT	Cell structure	N/A	Moderate
PEDAMACS	TDMA	HRT	No	High	Low
IEEE 802.15.4	Slotted CSMA/CA, GTS	Best effort/HRT	No	Moderate	Good
Saxena et al.	CSMA	Best effort /Low-Latency	No	High	Good
PQ-MAC	TDMA /CSMA	Best effort/Low-Latency	No	Moderate	Low
I-MAC	TDMA /CSMA	Best effort/Low-Latency	No	High	Moderate
Diff-MAC	CSMA	Best effort /Low-Latency	No	Moderate	Good
EQ-MAC	TDMA /CSMA	Best effort/Low-Latency	1-hop cluster based	High	Moderate
Suriyachai et al.	TDMA	HRT	Data gathering tree	Moderate	Low

(HRT). However, in order to accomplish this, it requires powerful access point (AP). This requirement has reduced its practical application and attractiveness.

3) *IEEE 802.15.4 standard* [13]: it basically uses CSMA/CA. In the beacon-enabled synchronized mode, it provides guaranteed time slots (GTS) and thus, in this case, HRT. It also provides energy saving.

4) *Saxena et al.* [14]: the autors propose a CSMA/CA protocol designed to support three types of traffic: streaming video, non-real-time and best effort. The device adjusts the duty cycle depending on the dominating traffic received in order to achieve energy saving.

5) *PQ-MAC* [15]: it uses both CSMA and TDMA. Energy saving is handled by an advanced wake up scheme, while prioritization is handled by a doubling scheme for high priority data.

6) *I-MAC* [16]: this protocol is based on Z-MAC [17] and defines three priority levels. It uses both CSMA and TDMA.

7) *Diff-MAC* [18]: it is a CSMA/CA based protocol, which provides differentiated services and hybrid prioritization very useful in multimedia applications. Its dynamic adaptation brings higher complexity.

8) *EQ-MAC* [19]: it uses both CSMA and TDMA. It achieves good energy saving and provides service differentiation but only works for bluster based single hop networks and cannot handle multi-hop transmissions.

9) *Suriyachai et al.* [20]: it is a TDMA based protocol, which can provide deterministic bounds for communication between two devices. Although, as it is based on a data gathering tree, its scalability is quite low.

Table II summarizes the QoS support and the major differences of the 10 protocols described in this section.

III. SERVICE MODELS AND PERFORMANCE EVALUATION OF IEEE 802.15.4

A complete description of service models and performance evaluation of the IEEE 802.15.4 standard is presented in [4]. In this section, we provide a summary of this analysis.

A. Service Models

The three service models are based on three factors: interactivity (yes/no), delay (Non Real-Time, Soft Real Time (SRT) and Hard Real Time (HRT)) and criticality (yes/no). The first model is the Open Services Model. It is interactive as it is based on user's queries, non-RT and non mission-critical. The second model is the Supple Services Model. This model is sometimes interactive, sometimes not, depending on the user's subscription, it is SRT and mission-critical. The third model is the Complete Services model. It is not interactive as there is a continuous flow of data, it is SRT or HRT depending on the application and is mission-critical.

B. IEEE 802.15.4 Performance Analysis

In order to provide the services described above in a WSN, we want to be able to provide services, which includes both best effort and HRT, to take into account energy saving while not being dependent of a certain topology in order to offer a solution practically applicable. From this perspective, we can conclude that the 802.15.4 standard offer the best compromise of the aforementioned characteristics and in this optic, the following performance analysis was conducted [4].

The simulations were performed with ns-2 simulator [21], using the IEEE 802.15.4 extension developed at City College of New York [22].

In our simulations we consider an 802.15.4 wireless network with one PAN coordinator and N reduced-function devices, with the nodes located close in a communication distance of 25 meters. This assumption prevents hidden terminal problem which results in data collisions. Wireless nodes are organized in star and random topologies (Fig. 1).

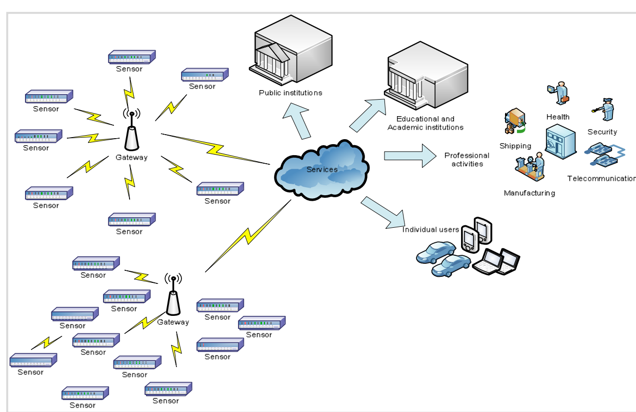


Figure 1. Simulation topologies (Adopted from [4]).

TABLE II. TRAFFIC PARAMETERS (ADOPTED FROM [4])

Parameter	Single user category		Multiple user categories	
Service Type	No Guaran-tees	Mobile Services	No Guarantees - Guarantees	Real-time - Guarantees
Traffic Type	Poisson	Constant Bit Rate (CBR)	Poisson – File Transfer Protocol (FTP)	Constant Bit Rate (CBR) – File Transfer Protocol (FTP)
Number of Nodes	25, 50, 100		25, 50, 100	
Number of Flows	1/4, 1/2, 3/4 and 4/4 number of nodes		¼ Poisson – ¾ FTP	¼ CBR – ¾ FTP
			½ Poisson – ½ FTP	½ CBR – ½ FTP
			¾ Poisson – ¼ FTP	¾ CBR – ¼ FTP
Node Position	Star / Random		Star / Random	
Traffic Direction	Node to Coordinator		Node to Coordinator	
Packet Size	40 Bytes		40 Bytes	20 Bytes – 40 Bytes

Traffic flows are generated in one-hop transmitting data directly to the coordinator, either in distributed or constant bit rate depending on the simulation model. Moreover, for each model, we distinguish two categories: no traffic differentiation (Single User) and traffic differentiation (Multiple User). A summary of traffic characteristics are presented in Table II. Additionally, multi-hop scenarios are evaluated in the random topology. Each simulation runs for 500 s and 15 times.

1) *Single User Category:* In a star topology with no guarantees, the power consumption does not exceed 0.35 and the worst case limit reaches 80% in packet delivery ratio for a loaded network. In a star topology with mobile services, better performance metrics are achieved: 2.5 times greater throughput and 5% better packet delivery ratio. In a random topology with no guarantees, we have better network performances compared to the star topology: 83% better in power consumption and doubled network lifetime. In a random topology with mobile services, the network behavior is similar compared to the star topology and better performance metrics are achieved: 17% greater throughput, 40% lower average delay and 75% lower power consumption.

2) *Service Differentiation for Multiple User Categories:* In a star topology, a grade of service differentiation can be achieved with packet delivery ratio reaching 87% and 83% for non guaranteed/guaranteed services and 83% for RT and guaranteed services respectively. There is no difference in throughput and average delay between the classes of services. In a random topology, there are no major discrepancies in network behavior. The throughput is 30% greater and average delay is 6% better compared to star topology metrics. The power consumption is reduced by about 80% compared to a large scaled star network.

3) *Priority Based Service Differentiation:* Provision of multi-level priority based services by tuning properly the size of the CW. In a star topology, priority services can be achieved with the throughput metric ranging from 4% to 20% depending on the service differentiation provided and the network load. Affected by the same factors, energy consumption ranges between 0.6% and 15%. The worst case in packet delivery ratio reaches 85% of successful delivery. In a random topology, a better delivery ratio is achieved and the gain in power consumption reaches 72%.

IV. WSNs Integration Approach in the IoT Providing QoS And Best Practices

A. *WSNs Integration Approach in the IoT Providing QoS*

In fact, one of the most important components in the IoT paradigm is WSN. The benefits of connecting both WSN and other IoT elements go beyond remote access, as heterogeneous information systems can be able to collaborate and provide common services. However, deploying WSNs configured to access the Internet raises novel challenges, which need to be tackled before taking advantage of the many benefits of such integration.

There are a lot of approaches to connect WSNs to the Internet. According to [23] and [24] the most effective, flexible and scalable approach is inspired for current WLAN structure and forms a dense 802.15.4 access point network, where multiple sensor nodes can join the Internet through the network's gateway (Fig. 2).

With gateways acting only as repeater and protocol translators, sensor nodes are also expected to contribute to QoS management by optimizing the resource utilization of all heterogeneous devices that are part of the future Internet of Things.

B. *Proposed Best Practices*

From Figure 3 [4] and the analysis presented in Section III, we can suggest some best practices about topology and traffic type, based on each application's priorities concerning quality factors.

If nor differentiation of traffic neither a strict delay bound is to be provided, the best topology is the random topology as it assures much better energy saving (83%), throughput (19%) and packet delivery ratio (6%) than the star topology. This approach mostly fits the Open Services model and in some cases the Supple Services model and Complete Services model.

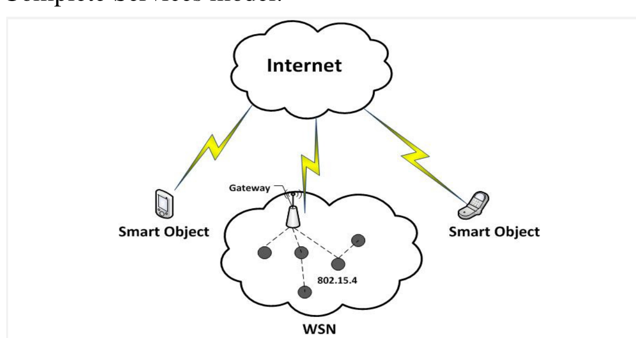
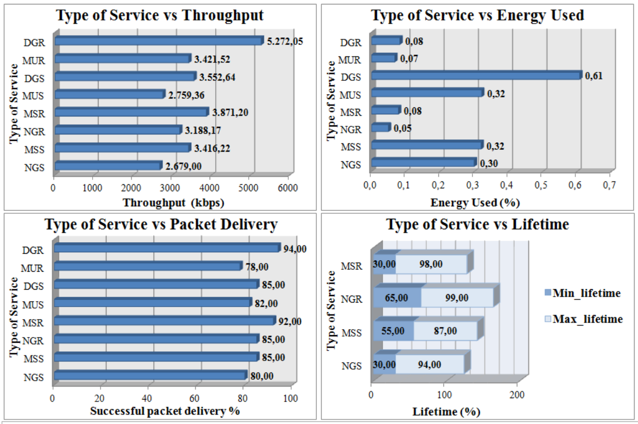


Figure 2. Integration of WSN in the IoT.



NGS:	Single User No Guarantees Star	MUS:	Multiple User No Guarantee/Guarantee Star
MSS:	Single User Mobile Services Star	DGS:	Multiple User Real-Time/Guarantee Star
NGR:	Single User No Guarantees Random	MUR:	Multiple User No Guarantee/Guarantee Random
MSR:	Single User Mobile Services Random	DGR:	Multiple User Real-Time/Guarantee Random

Figure 3. Simulation models statistics for traffic parameters described in Table II (Adopted from [4]).

On the other hand, in order to support RT applications without traffic differentiation, the star topology achieves better delay than the random topology, but as mentioned above, costs more energy (75%). This could be the case while providing Complete Services model.

When different types of traffic are required, as for the two previous cases, again the star topology consumes dramatically much more energy (Non-RT: 78% and RT: 86%) than the random topology and performs worse in terms of throughput (Non-RT: 24% and RT: 48%) and delivery ratio (RT: 10%). This scenario is more likely to fit the Supple Services Model and the Complete Services model.

We can conclude that the only reason to use the star topology is when no other topology is feasible with the specific lightweight devices or when the delay provided by the random topology is not satisfactory. Again, this could be the case while providing Complete Services model when traffic differentiation is required.

V. Applications and Service Models in the IoT

In [25], the authors present numerous applications made possible by the IoT. Each of these applications has different requirements in terms of QoS. Based on the three service models for WSNs described in Section III, we extend those services to the IoT and demonstrate their feasibility categorizing the applications described in [25] according to these services' characteristics.

In the next paragraphs, for each domain of application described in [25], that is to say Transportation & Logistics, Healthcare, Smart environment, Personal & Social and Futuristic, we categorize each specific application in one of the three services based on their characteristics.

A. Transportation & Logistic

1) Logistics: this is either interactive or non-interactive, in many cases it requires a SRT guarantee and is mission-critical, thus it belongs to the Supple Services model.

2) Assisted driving: this is obvious that this application is mission-critical and requires a continuous flow of data with HRT guarantee, for these reasons it can ben classified into the Complete Services model.

3) Mobile ticketing: if the application's purpose was only to provide information about transportation services, it would be classified into the Open Services model as it is interactive, without requirements of synchronous data. As the oportunity to buy the related tickets is provided, it belongs to the Supple Services model as there is a need of SRT guarantee and the application is now mission-critical.

4) Monitoring environmental parameters: the application is non RT or SRT, it provides periodically collected data and thus is not interactive and is mission-critical as this will influence the measures to be taken. This application can be classified in the Supple Services model.

5) Augmented maps: this application belongs to the Open Services model as it is based on interaction, doesn't require RT and isn't mission-critical.

B. Healthcare

1) Tracking: requiring a continuous flow of data, SRT or HRT guarantees and being mission-critical, this application is another example of a Complete Services model's application.

2) Identification and authentication: this application is interactive, it requires no RT or only SRT guarantees but the security provided makes it mission-critical. Therefore, the application belongs to the Supple Services model.

3) Data collection: interactive or not, it requires a SRT guarantee and is mission-critical. The application can be classified into the Supple Services model.

4) Sensing: as for tracking, it is mission-critical, requires HRT guarantees and a continuous flow of data and thus belongs to the Complete Services model.

C. Smart environments

1) Comfortable homes and offices: this application is interactive or not, it requires no RT or only SRT guarantees but as it can be used for monitoring and alarm systems, it becomes mission-critical. The best classification is thus the Supple Services model.

2) Industrial plants: the application is interactive or not, it requires SRT guarantees, it is mission critical and belongs therefore to the Supple Services model.

3) Smart museum and gym: this is another classic example of Open Services model, it is interactive, does not require RT guarantees and is not mission-critical.

D. Personal and social

1) Social networking: this application is interactive, it doesn't require RT guarantees and isn't mission-critical, therefore it can be categorized into the Open Services model.

TABLE III. APPLICATIONS IN IOT AND THEIR CORRESPONDING SERVICE MODEL

Application domain	Application	Model
Transportation & Logistic	Logistics	Supple
	Assisted driving	Complete
	Mobile ticketing	Supple
	Monitoring environmental parameters	Supple
	Augmented maps	Open
Healthcare	Tracking	Complete
	Identification & Authentication	Supple
	Data Collection	Supple
	Sensing	Complete
Smart Environments	Comfortable homes and offices	Supple
	Industrial plants	Supple
	Smart museum and gym	Open
Personal and Social	Social networking	Open
	Historical queries	Open
	Losses & Thefts	Supple
Futuristic	Robot taxi	Complete
	City information model	Supple
	Enhanced game room	Complete

2) Historical queries: as for social networking, it belongs to the Open Services model for the same reasons.

3) Losses & Thefts: this application is interactive; it requires no RT or only SRT guarantees but is mission-critical, so the best classification is the Supple Services model.

E. Futuristic

1) Robot taxi: enhanced form of assisted driving, it is obvious that this application belongs to the Complete Services model for the same reasons.

2) City information model: this set of applications is interactive or not, it is mission-critical and requires SRT guarantees, therefore it can be categorized into the Supple Services model.

3) Enhanced game room: this application belongs to the Complete Services model as it requires continuous flow of data, HRT guarantees and is mission-critical.

Table III summarizes the above analysis.

VI. CONCLUSION

As a main component of the IoT, WSNs contribute to the management of QoS by optimizing the resource utilization. In this perspective, we first presented a review of current QoS-aware MAC protocols in WSNs, and then we summarized the service models and the performance analysis of the IEEE 802.15.4 from [4]. Afterwards, we presented one of the best ways to integrate WSNs in the IoT providing QoS, using a dense IEEE 802.15.4 access point network, where multiple sensor nodes can join the Internet through the network's gateway. We proposed best practices to adopt when using this protocol in order to provide service models in WSNs. Finally, we demonstrated the feasibility of extension of those service models to the IoT and we categorized different IoT applications according to them.

REFERENCES

[1] O. Vermesan et al., "Internet of Things Strategic Research Roadmap", European Research Cluster on the Internet of Things, Cluster Strategic Research Agenda 2011.

[2] A. Dunkels Adam and J.P. Vasseur, "IP for Smart Objects", Internet Protocol for Smart Objects (IPSO) Alliance, Sepember 2008.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. 2002. Wireless sensor networks: a survey. Computer Networks, vol. 38, no. 4, March 2002, pp. 393-422, doi=10.1016/S1389-1286(01)00302-4

[4] M.-A. Nef, S. Karagiorgou, G. Stamoulis, and P. Kikiras, "Supporting Service Differentiation in Wireless Sensor Networks," Proc. 15th Panhellenic Conf. on Informatics, Sept. 2011, pp. 127-133, doi: 10.1109/PCI.2011.49.

[5] W. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks, " Proc. 5th ACM/IEEE Mobicom Conf., August 1999, pp. 174-185, doi: 10.1145/313451.313529.

[6] N. Bulusu, D. Estrin, L.Girod, and J. Heidemann, "Scalable Coordination for Wireless Sensor Networks: Self-Configuring Localization Systems, " Proc. 6th International Symposium on Communication Theory and Application (ISCTA 2001), July 2001.

[7] D. Estrin, "Comm'n Sense: Research Challenges in Embedded Networking Sensing, " Presented at the UCLA Computer Science Department Research Review, April 2001.

[8] M. A. Yigitel, O. Durmaz Incel, and C. Ersoy, "QoS-aware MAC protocols for wireless sensor networks: A survey," Comput. Netw., 2011, doi: 10.1016/j.comnet.2011.02.007.

[9] Y.J. Li, C.S. Chen, Y.-Q. Song, and Z. Wang, "Real-time QoS support in wireless sensor networks: a survey, " Proc. 7th IFAC Int Conf on Fieldbuses & Networks in Industrial & Embedded Systems (FeT'07), Nov. 2007.

[10] M. Caccamo, L.Y. Zhang, L. Sha, and G. Buttazzo, "An implicit prioritized access protocol for wireless sensor networks, " Proc. 23rd IEEE RTSS, 2002, pp. 39-48.

[11] T. Watteyne, I. Augé-Blum, and S. Ubéda, "Dual-mode real-time MAC protocol for wireless sensor networks: a validation/simulation approach, " Proc. InterSense, 2006.

[12] S. C. Ergen and P. Varaiya, "PEDAMACS: Power Efficient and Delay Aware Medium Access Protocol for Sensor Networks, " IEEE Transactions on Mobile Computing, vol.5 no.7, pp.920-930, July 2006.

[13] IEEE Std 802.15.4, "Part 15.4: Wireless medium access (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs), " IEEE-SA, 2006.

[14] N. Saxena, A. Roy, and J. Shin, "Dynamic duty cycle and adaptive contention window based QoS-MAC protocol for wireless multimedia sensor networks," Comput. Netw., 2008, doi: 10.1016/j.comnet.2008.05.009.

[15] H. Kim and S.-G. Min, "Priority-based QoS MAC protocol for wireless sensor networks, " Porc 2009 International Symposium on Parallel & Distributed Processing, IEEE Computer Society, 2009, pp. 1-8, doi: 10.1109/IPDPS.2009.5161184.

[16] I. Slama, B. Shrestha, B. Jouaber, and D. Zeghlache, "A hybrid MAC with prioritization for wireless sensor networks, " 33rd IEEE Conference on Local Computer Networks, 2008, pp. 274-281, doi: 10.1109/LCN.2008.4664180.

[17] I. Rhee, A. Warrier, M. Aia, and J. Min, "Z-MAC: a hybrid MAC for wireless sensor Networks, " Proc. ACM Sensys, 2005, pp. 90-101.

[18] M.A. Yigitel, O. Durmaz Incel, and C. Ersoy, "Diff-MAC: a QoS-aware MAC protocol with differentiated services and hybrid prioritization for wireless multimedia sensor

networks," Proc. 6th ACM workshop on QoS and security for wireless and mobile networks, 2010, pp. 62-69, doi: 10.1145/1868630.1868642.

[19] B. Yahya and J. Ben-Othman, "Energy efficient and QoS aware medium access control for wireless sensor networks," Concurrency and Computation: Practice and Experience, 2010, pp. 1252-1266, doi: 10.1002/cpe.1579.

[20] P. Suriyachai, U. Roedig, and A. Scott "Implementation of a MAC protocol for QoS support in wireless sensor networks, " Proc. 2009 IEEE International Conference on Pervasive Computing and Communications, 2009, pp. 1-6, doi: 10.1109/PERCOM.2009.4912839.

[21] K. Fall, K. Varadhan, The ns Manual, The VINT Project, A Collaboration between researchers at UC Berkeley, LBL, USC/ISI, and Xerox PARC, 22 May 2008.

[22] J. Zheng and M. J. Lee, "A comprehensive performance study of IEEE 802.15.4, " Sensor Network Operations, IEEE Press, Wiley Interscience, Chapter 4, 2006, pp. 218-237.

[23] R. Roman and J. Lopez, "Integrating Wireless Sensor Networks and the Internet: a Security Analysis," Internet Research: Electronic Networking Applications and Policy, vol. 19, no. 2, 2009.

[24] D. Christin, A. Reinhardt, P. S. Mogre, and R. Steinmetz, "Wireless Sensor Networks and the Internet of Things: Selected Challenges," Proc. 8th GI/ITG KuVS Fachgespräch Drahtlose Sensornetze, 2009.

[25] L. Atzori, A. Iera, G. Morabito, "The Internet of Things: A survey, " Comput. Netw., 2010, doi: 10.1016/j.comnet.2010.05.01.

Comparative Analysis of Scheduling Algorithms for UMTS Traffic in Case of DiffServ Network Congestion

Sladjana Zoric
Faculty of Electrical Engineering
and Computing, Zagreb, Croatia
e-mail: sladana.zoric@gmail.com

Mladen Kos
Faculty of Electrical Engineering
and Computing, Zagreb, Croatia
e-mail: mladen.kos@fer.hr

Himzo Bajric
BH Telecom, Joint Stock Company
Sarajevo, Bosnia and Herzegovina
e-mail: hbajric@bhtelecom.ba

Abstract — **Having in mind picture of future all-IP network, as well as fixed-mobile convergence, interworking between Universal Mobile Telecommunication System (UMTS) and external networks is crucial for providing end-to-end Quality of Service (QoS). For the purpose of achieving QoS in all-IP mobile networks, particularly for broadband multimedia services, Differentiated Services (DiffServ) mechanism should be applied to UMTS technology. This paper proposes using of Low Latency Queuing (LLQ) scheduler, with the main idea of mapping voice and video telephony in two different QoS classes and virtual queues, but at the same time using Priority Queuing (PQ) within LLQ for both voice and video telephony over all other QoS classes. To proof the concept, a simulation study was performed using Network Simulator version 2 (ns-2). Evaluation results of simulation study are presented for UMTS traffic that passes through UMTS core network and overloaded external IP backbone network. Performances of LLQ scheduler are compared with other most widely used scheduling algorithms such as Weighted Fair Queuing (WFQ), Weighted Round Robin (WRR) and Priority Queuing (PQ). The main objective of simulation was to provide QoS parameters such as IP Transfer Delay (IPTD), IP Delay Variation (IPDV) and IP Loss Rate (IPLR) for conversational and streaming traffic classes below standard defined values but not to completely exhaust bandwidth for interactive and background traffic classes. The obtained results were statistically processed using Statistical Package for the Social Sciences (SPSS) version 17.0.**

Keywords - QoS; DiffServ; Scheduling; LLQ.

I. INTRODUCTION

Within core network of Universal Mobile Telecommunication System (UMTS), IP Multimedia Subsystem (IMS) [1] presents foundation of network which is completely based on Internet Protocol (IP) and provides support for multimedia services like Voice over IP (VoIP) and Video Streaming. As new multimedia services require higher restrictions in network parameters and have different requirements, the support for Quality of Service (QoS) is necessary. Prominent advantage of UMTS is its ability to provide diverse services with QoS guarantees. This paper will focus on the analysis of ensuring QoS for UMTS real-time traffic (Conversational and Streaming traffic class) in a mixed network environment, composed of the UMTS core network and IP external domain. For providing QoS in IP networks, IETF has developed different QoS mechanisms, like Differentiated Services (DiffServ) [2] and Integrated Services (IntServ) [3]. IntServ have scalability and complexity problems, while DiffServ can be implemented in UMTS network with little or no management complexity. DiffServ architecture is based on a simple model where traffic entering a network is classified and conditioned at the boundaries of the network according to the Differentatied Services Code Point (DSCP) field in IP header, and assigned to different behavior aggregates. Within the core of the network, packets are forwarded according to the Per-Hop Behaviours (PHB) associated with the DSCP. PHB definitions do not specify any particular implementation mechanism and therefore the problem of PHB implementation has recently gained significant attention. According to 3GPP specifications, mapping of UMTS traffic classes into PHB can be done in gateway GPRS Support Node (GGSN), in order to get efficient PHB configurations [4]. Standard QoS mapping authorizes both voice and video telephony to be mapped to the same QoS class. Video traffic has larger packet sizes than voice traffic and can cause significant delay of voice packets when aggregating both together to the same QoS class. Authors in paper [5] have analyzed refined mapping between voice and video telephony but do not take other UMTS traffic classes into account. On the other side, in paper [6] are discussed QoS aspects both for real-time and non real-time traffic in UMTS simulation environment, but only in case of Priority Queuing (PQ) and Weighted Round Robin (WRR) schedulers. Dekeris, et al., [7] combine WFQ and LLQ, but the main drawback of this idea is the property that delay of high priority class (Video conferencing) could be reduced, but at the same time Voice traffic got the highest delay time. In our paper is presented concept of mapping voice and video telephony to different QoS classes and idea of implementing Low Latency Queuing (LLQ) traffic scheduler on network elements. Within LLQ, Priority Queuing (PQ) is used for scheduling of both voice and video telephony with respect to other traffic classes. Performances of LLQ scheduler are compared with other most available traffic scheduling algorithms such as Weighted Fair Queuing (WFQ), Weighted Round Robin (WRR) and Priority Queuing (PQ). In this paper, network congestion effect on QoS parameters for real-time traffic is investigated, that is conversational and streaming traffic classes. The aim of our proposed model was to provide IP Transfer Delay (IPTD), IP Delay Variation

(IPDV) and IP Loss Rate (IPLR) to be below standard defined values, even in case of high network overload.

This paper is organized as follows. Section II makes a brief presentation of UMTS QoS model, the most commonly used PHB and our suggestion of mapping between UMTS and DiffServ domain. Section III provides overview of scheduling algorithms which will be used in simulation study. Section IV presents ns2 simulation model and simulation results, together with their discussion and analysis to show the conclusions that are warranted. Section V concludes this paper and describes direction for the future work.

II. UMTS TO DIFFSERV QOS MAPPING

3GPP standard proposes a layered architecture for the support of end-to-end QoS. To realize a certain network QoS, a Bearer Service (BS) with clearly defined functionalities has to be set up from the source to the destination of a service and includes all aspects to enable the provision of a contracted QoS.

UMTS BS attributes form a QoS profile and define the grade of service provided by the UMTS network to the user of the UMTS bearer service. UMTS specification [8] defines four traffic classes and they are: conversational, streaming, interactive and background. The main difference between these classes is how delay sensitive the traffic is. Applications of conversational and streaming classes are the most delay sensitive and intended for real-time traffic, while applications of interactive and background classes require higher reliability. Examples of applications are voice and video telephony for Conversational class and Video Streaming for the Streaming class. Interactive class is used by interactive applications like interactive web browsing, while Background class can be used for background download of e-mails.

Since the UMTS packet switched core network is based on an IP, DiffServ can be used for QoS provisioning. Figure 1 shows the example of how end-to-end QoS may be accomplished for a significant number of scenarios. In this paper, first scenario from 3GPP specification has been chosen, where the GGSN supports DiffServ Edge function and the IP network is DiffServ enabled. The application layer identifies QoS requirements, which are mapped into Packet Data Protocol (PDP) context parameters in UE. Local mechanism in the UE uses the PDP context for QoS over the UMTS access network, and the IP backbone network uses DiffServ to provide QoS guarantees. According to [5] IP BS manager is located in GGSN and uses standard IP mechanisms to manage IP bearer service. Provision of IP BS manager is optional in User Equipment (UE) and mandatory in the GGSN. Translation/Mapping function in GGSN provides interworking between the mechanisms and parameters used within the UMTS bearer service and those used within the IP bearer service. It is operator's choice to define the mechanisms for the provisioning of resources among the different DiffServ PHB classes, as well as the mapping from the UMTS QoS classes, to the DSCP. The DiffServ working group of IETF has defined different PHB groups for different applications.

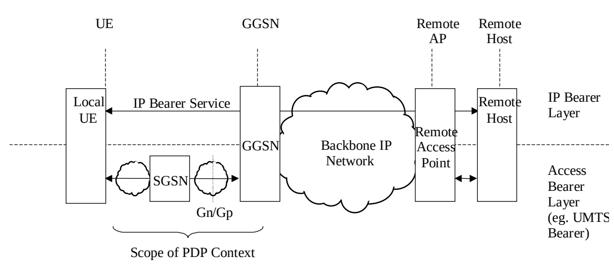


Figure 1. Network architecture for QoS conceptual model [5]

The EF-PHB [9] is intended to support low-loss, low-delay and low-jitter services. The EF guarantees that traffic is serviced at a rate that is at least equal to a configurable minimum service rate (regardless of the offered load and non-EF traffic) at both long and short intervals. IETF defines AF-PHB group in [10]. AF allows the operator to provide assurance of delivery as long as the traffic does not exceed some subscribed rate. Traffic that exceeds the subscription rate faces a higher probability of being dropped if congestion occurs. AF PHB defines four independent PHB classes, each with three dropping precedence level. Each corresponding PHB is known as AF_{ij}, where i represents AF class, while j is the drop precedence. Within an AF class, packets of drop precedence p experience a level of loss lower (or equal to) than the level of loss experienced by packets of drop precedence q if $p<q$. Each AF class is configured with separate buffer and bandwidth. Default PHB has best effort forwarding characteristics.

Service provider should consider all the UMTS QoS classes that are defined in network, aggregate these classes into a manageable set of new groups, based on their QoS requirements. Correspondingly, a set of available PHB that have similar characteristics should be chosen, and a one-to-one mapping assigned. Traditionally, all traffic in Conversational class, i.e. voice and video telephony, should be mapped to the same EF class which is intended for critical voice traffic. In fact, voice packets have short and constant packet size while video packets have large and variable packet size. When injecting voice telephony traffic together with bursty video telephony traffic, video traffic can cause degradation as well as delay of voice service. Therefore we suggest the mapping of voice and video telephony to two different DiffServ virtual queues and DiffServ classes. Voice telephony is mapped to EF class and video telephony to AF11 class. Aggregate of Streaming class is mapped into AF21, as it requires low variation of delay and has higher delay constraint than the Interactive class but less constraint than the Conversational class Aggregate of Interactive class is mapped to AF31, and do not have special requirements, except reliability. Aggregate of Background class is mapped to default PHB.

TABLE I. QOS MAPPING TABLE

TRAFFIC CLASS	PHB	DSCP VALUE
Conversational voice	EF	46
Conversational video	AF11	10
Streaming video	AF21	18
Interactive	AF31	26
Background	BE	0

III. TRAFFIC SCHEDULING ALGORITHMS

PHB simply characterizes the externally observable forwarding behavior of a DiffServ router to the corresponding traffic stream. PHB definitions do not specify any particular implementation mechanism. To instantiate a particular PHB, network administrator activates and tunes an appropriate combination of specific packet-scheduling algorithms and Active Queue Management (AQM) mechanisms supported by the DiffServ router. The choice of a traffic scheduling algorithm is important for the implementation of behavior aggregates in a DiffServ network. When multiple queues are sharing common transmission media, there must be a scheduler to decide how to pick up packets from each queue to send out and is responsible for enforcing resource allocation to individual flows. If there is no congestion on the interface, packets are transmitted as they arrive. If the interface is experiencing congestion, scheduling algorithms are engaged. Scheduler performances have the highest impact on the level of service a packet receives [11]. The most popular and available scheduling algorithms in IP routers, and used in our simulation are: WFQ, WFF, PQ and LLQ.

A. *Priority queuing (PQ)*

In classic PQ, packets are first classified by the system and then placed into different priority queues. Packets are scheduled from the head of the given queue only if all queues of higher priority are empty. Within each of the priority queues, packets are scheduled in FIFO order. Benefit of PQ is relatively low computational load on the system. The biggest problem of using PQ is if the volume of higher-priority traffic becomes excessive, lower priority traffic can be dropped as the buffer space allocated to low-priority queues starts to overflow.

B. *Weighted Round Robin (WRR)*

In WRR, packets are first classified into various service classes and then assigned to a queue that is specifically dedicated to that service class. Each of the queues is then serviced in a round robin (RR) order. The weight indicates how many packets have to be sent in each cycle from each queue. The WRR scheduler doesn't take the size of the transmitted packets into account. As a result, it is difficult to predict the actual bandwidth that each queue obtains, but it ensures that all service classes have access to at least some configured amount of network bandwidth.

C. *Weighted Fair Queuing (WFQ)*

WFQ supports flows with different bandwidth requirements by giving each queue a weight that assigns it a different percentage of output port bandwidth. WFQ supports the fair distribution of bandwidth for variable-length packets by approximating a generalized processor sharing (GPS) system. GPS [12] assumes that the input traffic is infinitely divisible and that all sessions can be served at the same time. GPS is a theoretical model and in reality it cannot be implemented. WFQ schedules packets according to their arrival time, size, and the associated weight. Upon the arrival of a new packet, a "virtual finish time" is calculated which represents time at which the same packet would finish to be served in the GPS system. WFQ outputs packets in the ascending order of the virtual finish time.

D. *Low Latency Queuing (LLQ)*

LLQ is a combination of PQ and Class-Based Weighted-Fair Queuing (CBWFQ). CBWFQ extends the standard WFQ functionality to provide support for user-defined traffic classes. The LLQ, like PQ checks the low-latency queue first and takes a packet from that queue. If there are no packets in the low-latency queue, the normal scheduler logic applies to the other non-low-latency queues, giving them their guaranteed bandwidth. LLQ allows delay-sensitive applications such as voice to be given preferential treatment over all other traffic classes [13].

IV. SIMULATION MODEL

Simulation is performed using network simulator ns2, which is an event-driven simulator targeted at networking research [14] and independent developed module for scheduling algorithms used in this paper [15]. Default implementation of LLQ within ns2 simulator, which supports scheduling of only one queue with PQ is changed in order to support scheduling of two queues with PQ (for voice and video telephony). The aim of simulation is to evaluate performances of our proposed idea in terms of QoS parameters and perform comparison with other schedulers such as WFQ, WRR and PQ where voice and video telephony are mapped to the same virtual queue and traffic class. According to ITU-T Recommendation Y.1541 [16], QoS parameters for conversational traffic class should be:

- IP Transfer Delay, IPTD $\leq$ 100 ms,
- IP Delay Variation, IPDV $\leq$ 50 ms,
- IP Loss Rate, IPLR $\leq 10^{-3}$,

while for streaming traffic class these parameters should be:

- IP Transfer Delay, IPTD $\leq$ 400 ms,
- IP Delay Variation, IPDV $\leq$ 50 ms,
- IP Loss Rate, IPLR $\leq 10^{-3}$.

Simulation model is presented in Figure 2. UMTS infrastructure is not fully simulated (radio interface), only the core network between SGSN and GGSN which is not congested. This is not inconsistent with the concept of UMTS architecture which specifies that access and core networks are independent [17].

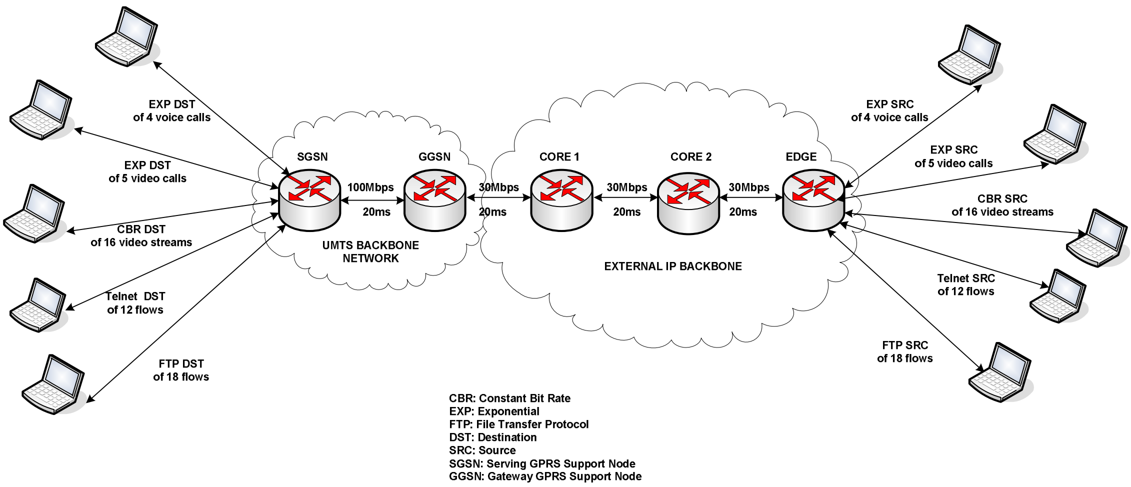


Figure 2. Simulation model

The capacities of links in external IP backbone are dimensioned in a way to implement network configuration whose load equals 20%. Starting from this configuration, we decrease the capacity of appropriate links gradually to 200% according to the amount of traffic passing through external IP backbone, which is 6Mb/s and is constant during all simulation. Voice telephony traffic sources are represented with Exponential (EXP) traffic generator, which generates 4 flows with packet size of 80 bytes and 100 kb/s rate, while video telephony sources generate 5 flows with 1000 bytes and 100kb/s rate. Constant Bit Rate (CBR) generator is used in order to generate 16 video traffic flows with packet size of 1000 bytes and 300 kb/s rate. Interactive traffic sources are simulated with Telnet application, and generate 12 flows with 500 bytes packets size with 100 kb/s rate. Background data sources are configured with File Transfer Protocol (FTP) traffic generator with 1000 bytes packet length and 10 kb/s rate from 18 flows. CBR and EXP traffic generators are attached to UDP agents, while FTP and Telnet traffic generators are attached to TCP agents. Conversational traffic (voice and video telephony together) produces 15%, streaming traffic produces 80%, while interactive and background traffic produce 2% and 3% of overall generated traffic. Time Sliding Window 2 Color Marker (TSW2CM) is used as a policer to determine how to mark and prioritize the packet according to user requirements. Weighted Random Early Detection (WRED) is used as Active Queue Management mechanism for Streaming and Interactive traffic, while for Conversational and Background traffic is used Drop Tail. In case that PQ scheduler is used, Conversational traffic has the highest priority, while Background traffic has the lowest priority. Weights of other schedulers are configured in such a way, that weight represents percentage of output port bandwidth: 15 for Conversational, 80 for Streaming, 2 for Interactive and 3 for Background traffic class. The queue lengths are constant and are defined with 30 packets for Conversational class and 50 packets for all other traffic classes. Simulation results in this paper are depicted only for second flow generated from all traffic sources, which is chosen randomly, but could be for any of the generated flows.

As we can see from Figures 3a and 3b, average end-to-end delay for Conversational traffic class (for both voice and video telephony) stays within 100 ms only when PQ and LLQ are implemented on network nodes. Considering the effect of different schedulers on average end-to-end delay for Streaming traffic, which is depicted in Figure 3c, we can notice that all schedulers have almost the same performances and provide satisfactory level of QoS according to reference [16]. Results from Figures 3d and 3e show that jitter stays below 30 ms for voice telephony and below 35 ms for video telephony in all experiments. The same behavior is also observed in Figure 3f, for Streaming traffic, where jitter is lower than 12ms for all schedulers. Results for jitter show non-monotonic behavior: increasing with the network load, reaching some maximum and the decreasing. More network latency is necessary in order to deliver a stream due to network congestion. In Figures 3g and 3h packet loss rate for voice and video telephony is presented. Network congestion has the greatest influence on WRR scheduler, which does not perform satisfactory when network is overloaded more than 100%. On the other side, there is no packet loss for PQ and LLQ schedulers. Similarly, in Figure 3i, packet loss rate for Streaming traffic is presented; we see that it remains below 0.2% for all schedulers except WRR, even under heavy network congestion of 200%.

Table II depicts the average values of the link throughput between nodes CORE2 and EDGE for Interactive and Background traffic classes. As expected, when PQ scheduling algorithm is used, lower priority classes are starving and throughput is equal zero for both Interactive and Background traffic classes when network congestion in higher than 120%. LLQ on the other side provides fair level of bandwidth for lower priority classes, and at the same time fulfills QoS requirements for Conversational and Streaming traffic classes. From all these experiments, it can be concluded that our approach of using LLQ scheduler and mapping of voice and video into two different classes

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-192-2

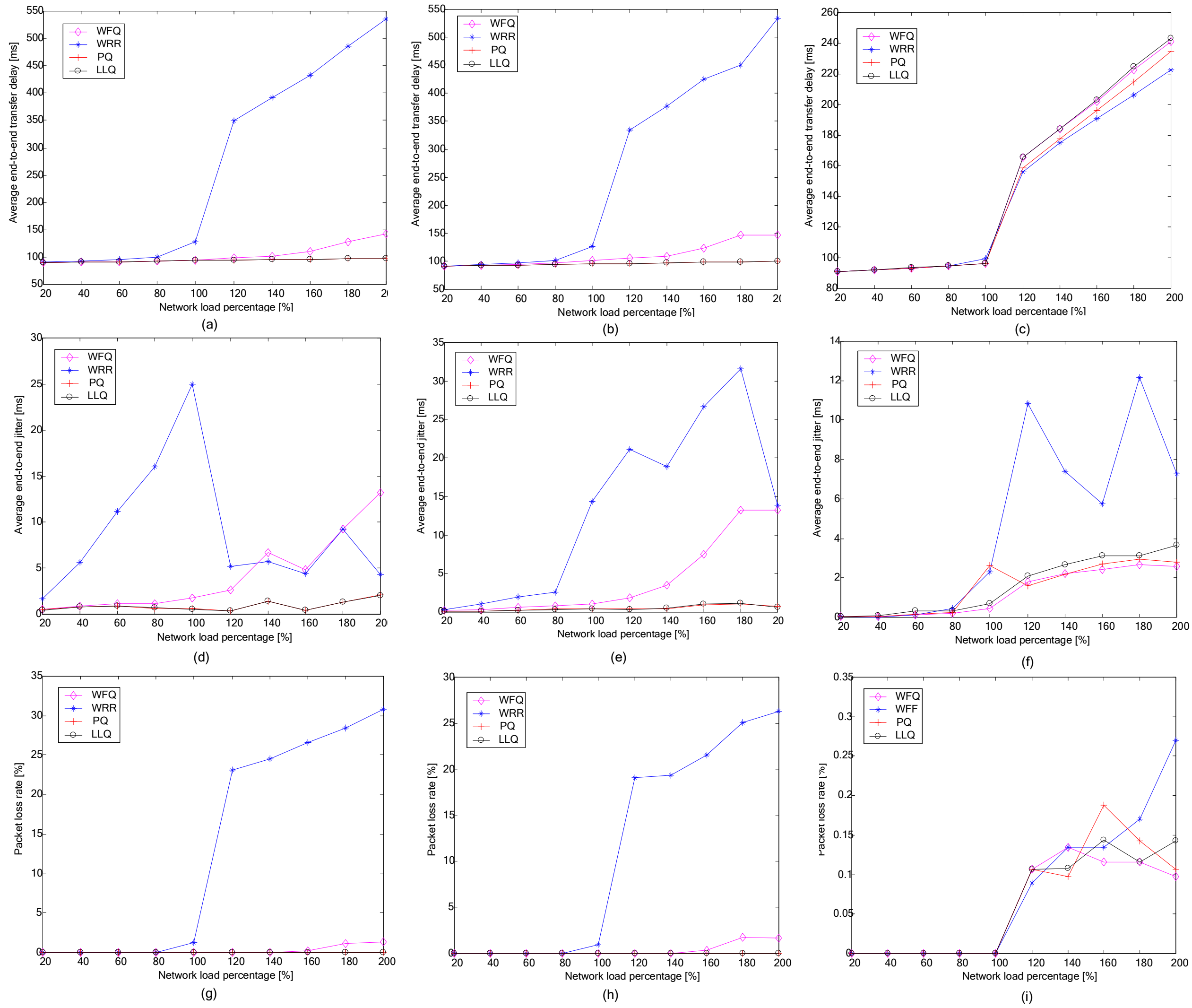


Figure 3. Simulation results: (a) End-to-end delay of voice telephony; (b) End-to-end delay of video telephony; (c) End-to-end delay of video streaming; (d) Jitter of voice telephony; (e) Jitter of video telephony; (f) Jitter of video streaming; (g) Packet loss rate of voice telephony; (h) Packet loss rate of video telephony; (i) Packet loss rate of video streaming

provides critical performance parameters for voice and video telephony below standard defined values. These results are consequence of handling Conversational traffic with strict priority over all other traffic classes.

The obtained results are statistically processed using Statistical Package for the Social Sciences (SPSS) version 17.0. The null hypothesis states that QoS parameters can be guaranteed when using LLQ. Against the null hypothesis is setup the alternative hypothesis. The 95% confidence interval is chosen, which relates to level of statistical significance of $p<0.05$. The regression analysis is conducted to find the relationship that explains how the variation in IPTD/IPDV/IPLR values for Conversational and Streaming traffic classes, depends on the variation in network overload. Coefficient of correlation (r) is measured to give the true direction the correlation, while the coefficient of determination (r^2) is measured to give the strength of correlation. Using comparative analysis of different regression models in SPSS, we have decided to use cubic polynomial regression model since it has the highest value of determination coefficient r^2. The results of regression analysis have been summarized in Table III only for our approach of using LLQ scheduler as it provides the best performances among all other simulated schedulers.

TABLE II. AVERAGE THROUGHPUT FOR INTERACTIVE AND BACKGOUND TRAFFIC CLASSES

TRAFFIC CLASS	WFQ	WRR	PQ	LLQ
Interactive	8.93	7.79	4.62	8.71
Background	37.20	33.07	27.40	35.41

TABLE III. STATISTICAL ANALYSIS OF QOS PARAMETERS FOR CONVERSATIONAL AND STREAMING TRAFFIC CLASSES – LLQ SCHEDULER

	Regression model	b_0	b_1	b_2	b_3	r^2	r	p
Voice telephony IPTD	$y=b_0+b_1t+b_2t^2+b_3t^3$	-325.36	0	-0.120	0.002	0.997	0.997	>0.05
Video telephony IPTD	$y=b_0+b_1t+b_2t^2+b_3t^3$	-934.68	0	0.138	0	0.999	0.999	>0.05
Video streaming IPTD	$y=b_0+b_1t+b_2t^2+b_3t^3$	-22.713	0.867	0	$7.36*10^{-7}$	0.891	0.944	>0.05
Voice telephony IPDV	$y=b_0+b_1t+b_2t^2+b_3t^3$	367.128	-977.614	945.438	-248.369	0.781	0.609	>0.05
Video telephony IPDV	$y=b_0+b_1t+b_2t^2+b_3t^3$	-9.78	229.924	148.92	-200.7	0.941	0.885	>0.05
Video streaming IPDV	$y=b_0+b_1t+b_2t^2+b_3t^3$	26.883	141.885	-69.228	12.068	0.987	0.974	>0.05
Voice telephony IPLR	-	-	-	-	-	-	-	-
Video telephony IPLR	-	-	-	-	-	-	-	-
Video streaming IPLR	$y=b_0+b_1t+b_2t^2+b_3t^3$	59.96	484.788	2610.032	0	0.902	0.814	>0.05

From Table III, we can see that regression dependency between packet loss and link load percentage for conversational traffic is not possible to determine. Value of packet loss rate is constant and is always zero for conversational traffic class no matter how high is traffic overload in external IP backbone network. For all other results depicted in Table III, p value (significance) is greater than 0.05, which means that there is almost no statistical relationship between QoS parameters and link load percentage when LLQ scheduler is used.

V. CONCLUSION AND FUTURE WORK

QoS as an end-to-end concept has to be satisfied through the interworking of all the entities the UMTS traffic is passing through. In order to achieve desired end-to-end performances, it is crucial to define efficient QoS mapping scheme between UMTS services and IP QoS classes in case of DiffServ based network. This paper presented one example of mapping which was implemented on GGSN and EDGE nodes, as they perform DiffServ edge function in our simulation model. That approach suggested mapping of voice and video telephony into two different QoS classes and virtual queues. Other important problem which was pointed out in our work concerns the implementation of PHB and the choice of traffic scheduling algorithm. We proposed the idea of using LLQ scheduler, with PQ scheduling for both voice and video telephony over all other traffic classes. Default implementation of LLQ scheduler in ns2 has been changed in order to support scheduling of two virtual queues with Priority Queuing. The results from our simulation study indicate that using LLQ provides better performances than using WFQ, PQ and WRR schedulers in terms of QoS parameters such as IPTD, IPDV and IPLR for real-time UMTS traffic. Results obtained from statistical analysis indicate that there is almost no statistical relationship between the performance metrics of Real-time services and the network load when novel approach is used. Future work will focus on hierarchical traffic scheduling in order to perform refined scheduling between voice and video telephony.

REFERENCES

[1] S. Zorić, J. Baraković, and H. Hodžić, „QoS Architecture in IP Multimedia Subsystem of UMTS“, 50th International Symposium ELMAR-2008, Zadar, Croatia, September 2008.

[2] S. Blake, et al., „An architecture for Differentiated Services", Technical Report RFC 2475, IETF, December 1998.

[3] R. Braden, D.Clark and S.Shenker, „Integrated Services in the Internet Architecture: an Overview“, Technical Report RFC 1633, IETF, June 1994.

[4] TS 23.207 Release 7, „End-to-end Quality of Service, concept and architecture", 3GPP, June 2007.

[5] R.B. Ali, S. Pierre, and Y. Lemieux, „UMTS-to-IP QoS mapping for voice and video telepfony services“, IEEE Network, vol. 19, no. 2, pp. 26-32, 2005.

[6] N.V.Lopes, M.J. Nicolau, and A. Santos, „Efficiency of PRI and WRR DiffServ Scheduling Mechanisms for Real-Time Services on UMTS Environment", NTMS, November 2008.

[7] B. Dekeris, T. Adomkus, and A. Budnikas, „Analysis of QoS assurance using weighted fair queuing (WFQ) scheduling discipline with low latency queue (LLQ)", 28th Int. Cof. Information Technology Interfaces ITI, Cavtat, Croatia, June, 2006.

[8] TS 23.107 Release 7, „Quality of Service, concept and architecture", 3GPP, September 2007.

[9] V. Jacobson, et al., „An Expedited Forwarding PHB“, Technical Report RFC 2598, IETF, June 1999.

[10] J. Heinanen, et al., „Assured Forwarding PHB Group“, Technical Report RFC 2597, IETF, June 1999.

[11] S. Zoric and M. Bolic, „Fairness of Scheduling Algorithms for Real-Time Traffic in DiffServ Based Networks“, 15th IEEE Mediterranean Electrotechnical Conference MELECON, Valletta, Malta, April 2010.

[12] A. Parekh and R. Gallager, „A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case“, ACM/IEEE Transactions on Networking, pp. 344-357, June 1993.

[13] S. Zoric, M. Kos, and H. Bajric, „Fairness of scheduling algorithms for Real-time UMTS traffic in case of IP Link Congestion“, ICUMT 2011, Budapest, Hungary, October 2011.

[14] Network Simulator, http://www.isi.edu/nsnam/ns, April 2012

[15] http://sergioandreozzi.com/research/network/diffserv4ns/, April 2012

[16] ITU-T Recommendation Y.1541, „Network performance objectives for IP-based services“, February 2006.

[17] S.I. Maniatis, E.G. Nikolouzou, and I.S. Venieris, „QoS issues in the converged 3G wireless and wired networks“, IEEE Communications Magazine, Volume 40, Issue 8, 2002

A Reputation System Model Based on Indication to Combat Pollution in P2P Networks

Fabíola Bento Soares
College of Computing
Federal University of Uberlandia - UFU
Uberlândia/MG, Brazil
e-mail: fabiolasb@gmail.com

Luís Fernando Faina
College of Computing
Federal University of Uberlandia - UFU
Uberlândia/MG, Brazil
e-mail: faina@facom.ufu.br

Newarney Torrezão da Costa
College of Computing
Federal University of Uberlandia - UFU
Uberlândia/MG, Brazil
e-mail: newarney@gmail.com

Jamil Salem Barbar
College of Computing
Federal University of Uberlandia - UFU
Uberlândia/MG, Brazil
e-mail: jamil@facom.ufu.br

***Abstract*— P2P Networks are compound by nodes, servers and suppliers of services or resources. This kind of system allows us as part of it to supply and ask for resources in easy manner as well as for fake or corrupted content resources. Some mechanisms based on resources or based on nodes, named reputation systems, are developed to decreasing the pollution in P2P Network. This article presents a reputation system model based on nodes and also means to defeat the pollution issue by expelling malicious nodes, underprivileging selfish nodes and improving the honest ones. Simulations prove the effectiveness of the model in question.**

Keywords-reputation; poisoning; pollution; peer-to-peer networks

I. INTRODUCTION

A network architecture Peer-to-Peer (P2P) is made of elements that may perform as clients and suppliers of resources [1]. Those networks are spread systems, which have interconnected nodes with self-organization capability, aimed at sharing diverse resources such as, music, video, document, among others. Another skill to be stand out is their capability to adapt at the same time they keep the connectivity with acceptable performance, without mediators or support of a central control office [2].

They also make feasible a great inlet and outlet of members with most diverse intentions, therefore, there are no means to avoid poisoning, that is, the availability of resources with corrupted or useless content by P2P network. Due to this poisoning occurs the pollution, which configures the network breakdown. Such mainly happens for the action of bad-intentioned users that poison a specific resource. This poisoning of resources may take place in several manners such as adding it a invalid content, supplying it with false information able to corrupt files, also by changing frequencies of music files, linking Trojans, among other techniques.

Resources, whether corrupted or not, have the same set of information, thus generating an issue, for it becomes hard to find those not poisoned [3]. Such issue may be detracted by a system of reputation that help the requestor nodes to acquire resources with valid content, besides avoiding dishonest nodes in the network [4]. In this context, there are nodes or selfish users which never share resources or narrow its permission to access, along with generous users that always release valid resources and cooperate for the right maintenance and good working of P2P network by supplying reliable information.

In this perspective, great issues in reputation systems are: discriminate valid resources from invalid ones in distributed and decentralized systems in reason of the dynamic behavior it presents, stand out selfish users to repress them for they don't add resources to the network. This work proposes to minimize these problems by reducing the pollution in P2P networks by means of a methodology which benefits generous users to the detriment of bad-behaved users, whether malicious or selfish.

Therefore, instead of consulting directly the node reputation which hosts the recourse, a consulting to partner nodes is made. These nodes are addressed as supernodes and have already used the recourse requested, or they know the node reputation that hosts this recourse. Thus, the model privileges the good reputation nodes with the information shared with partner nodes. Therefore, the quality of the content offered by network evolves according to the distributed and regular policy. Accordingly, with gradual use of the net, generous users broaden their part in the system due to its quality of resources, giving priority to the proposition of groupware model. Despite the existence of selfish and bad-intentioned users, their performance in the net shall be reduced as partner nodes narrow their reputation.

In order to develop the proposition, this article is structured as follows: the Introduction previously presented; Section 2 approaches the works related; Section 3 describes the model in question; Section 4 presents the simulation to

 ISBN: 978-1-61208-192-2

validate the hypothesis, and eventually, Section 5 brings final considerations.

II. Related Works

The performance of systems of reputation is based on two main concepts, which are, confidence and reputation. Confidence is about leaving the analysis whether something is a fact or not by delivering this study to the source where this information came from, and simply take this into consideration. For an individual to be considered trustful, it is necessary that it has positive, honest and cooperative attitudes in relation to the entities dependent on it. On the other hand, reliability is one's capacity to be trustful, which means that confidence is a consequence of reliability [5].

Reputation, by one's turn, is what one knows about the character or position of an individual before the judgment of a community. Therefore, reputation reflects the community's vision over an individual, while trustfulness is about a subjective opinion.

The systems of reputation represent a important alternative to help users themselves settling confident relationships through Internet, allowing them to make personal evaluations over individuals performance and identify the reputations estimated before the opinion of a community. Thus, those systems present mechanisms to stand out and manage reliability relations among users [6].

In this work, two systems of reputation were approached that guided the new model proposed: the Credence System [7] and the System Based upon Resources [8]. Both were proposed for P2P environment purely decentralized, in which information of nodes reputation is spread over the net.

A. Credence system of reputation

This system allows users to classify the resources obtained concerning its authenticity whether polluted or not. It works based on a protocol of research by voting, used to disseminate the rank of those resources by the system and a correlation scheme of votes that gives more weight to the ones came from pairs that tend to have the same opinion [7].

Before the acquisition of a resource, one computes a correlation among the nodes of the net, that is, applicant and supplier nodes. It is natural that bad-intentioned users lie about their reputation to poison the net. This correlation presents two strategies to protect the statistics of confidence, which is locally stored. In one of them only locally computed correlations are changed, that is, the client may apply for the auditing of the correlation choosing one of the nodes involved on it, keeping its integrity.

In the second strategy, in practice, the local confidence statistics has significant amount of redundant information, densely connected, forming cycles and raising maliciously the reputation before its mates. The auditing might identify such behavior and somehow punish the responsible ones.

In this aspect, the relationship between two nodes is expressed by the correlation of their vote record, checking if the nodes tend to vote in similar manner, which we call positive correlation, in different manner, which is named negative correlation, or if their records of votes are not correlated.

B. System of reputation based on resource

In this model, before making the choice of a resource for download, the requestor applies to other nodes the score of the candidate resource, and weighs these resources according to the reliability and information received from partner nodes. When a node requires a resource it consults all the nodes in the net. Once it has the response, the applicant node may choose among the replies and get what it wants. After nodes interaction, the applicant ratify if what was required is in fact what it wanted [8].

This result is not always optimistic. As a common user, it cannot distinguish between the authentic and polluted resources unless it's possible to verify the content after getting the resource, or make a remote evaluation, which is practically impossible. In this system of reputation, each node validates the authenticity of the resource it gets and records its result. Thus, the requestor either receives a authentic resource or a polluted one.

Such mechanism aims at restraining the nodes from diffusing polluted resources in the following situations: a) damage by some kinds of bad behavior, as the sharing of invalid resources; b) false information given to other nodes in the net, thus sending misleading content over the resources; c) collusion of nodes that give right opinions about some resources by pretending to be an honest node in order to gain confidence of another ones.

A distinction is made in relation to the resource and to the nodes concerning the reputation and reliability, in which the reputation of a resource from a node point of view, is used to evaluate the expectancy of a node in relation to this resource. The reliability of a node, from another node point of view, is a subjective expectation which believes that the evaluation of the resource is true.

Each participant node keeps the record of identification of a resource in a set classified as *RS*, and the node identification is classified in a set named *NS*, which is compound by nodes that has been publicly evaluating at least one resource in *RS* set. Each node has a local storage defined by *L* to be saved in a matrix, and the reliability value of other nodes is saved in another matrix *R*.

Each node *P* has a set of resources identifiers RS_P, and a set of node identifiers NS_P. RS_P is compound of resources locally authenticated by *P*. NS_P is compound by nodes which has publicly evaluated at least one of its resources in RS_P, and $RS_P[0]$ is fitted to be *P*.

In addition, each node stores locally the information in which *P* takes part in a matrix L_P. Other reliability values are stored in a vector R_P.

 ISBN: 978-1-61208-192-2

C. Analysis of the reputation systems P2P networks

In both systems analyzed, the polling for a resource brings overhead in P2P net. This occurs because every node which possesses the requested information shall be able to respond to the consult. Also, there isn't a hierarchy model that distributes the nodes according to its reputation. In these systems, it is not considered the validity of the repository. This implies in risks to the security about the broadcast of reputation data in the net. In these models, once the user is considered to be trustful, it shall have the permission to spread the reputation obtained among the network's nodes.

However, for a network with thousands of transactions, in many cases, the access of a same file shall be done by many users at the same time. In case there is a group of malicious users that propagate a information of positive reputation to a set of files with doubtful integrity, this act may compromise the quality of network resources. Therefore, it is interesting that, even with the spread of reputation information, only the average shall be considered. Thus, if 10% of bad-intentioned users pollute the net and forge the reputation of its files, one expects that the evolution itself of the net with interactions among well-behaved, shall be able to judge this poisoning attempt.

The sequence of events to find a resource in both systems is the same. The node requestor, *Node_Requestor*, does the query using keywords, the message *Query(keyword)* is sent to the whole subset of nodes in P2P net, according to the employed protocol. The nodes which don't possess the recourse and are willing to share it, reply with the message *QueryHit(nodeID, resourceID,* see Figure 1.

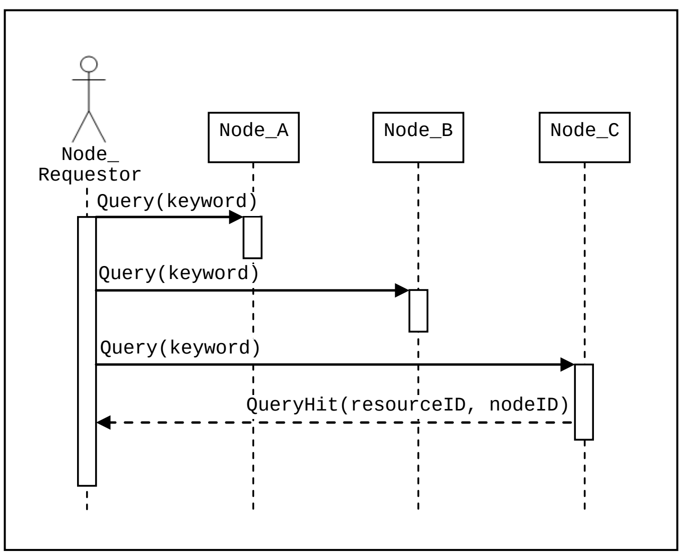


Figure 1. Sequence of events to consulting resource.

After obtaining the responses, proceeds the evaluation of these. The applicant node, repeatedly chooses a recourse by sending a message to request the confidence punctuation for the nodes different from those that have the recourse with the message *RatingQuery(resourceID)*. The expert node

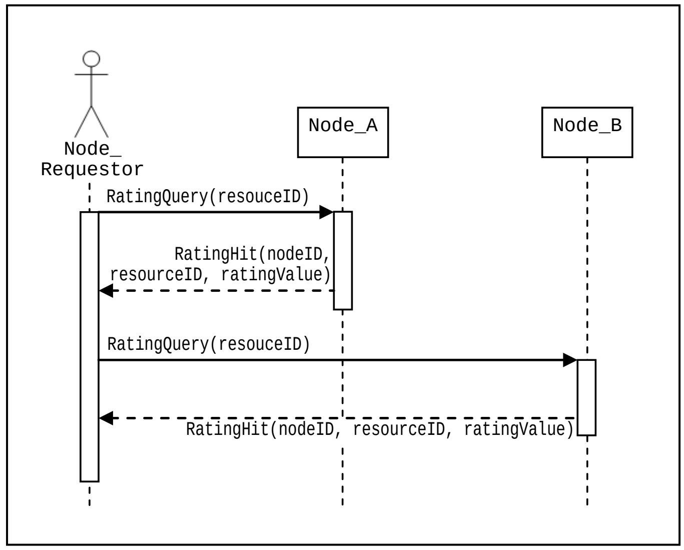


Figure 2. Sequence of events to validate the trust.

responds with a message *RatingHit(nodeID, resourceID, ratingValue)*, see Figure 2.

At each of the systems, a validation of confidence is done in distinct manner, this correlation is presented next. After the acquisition of the recourse, the nodes which supplied information about the recourses are notified about the achievement of the resource, and the reputation of the recourse origin is updated for higher, when its valid, or otherwise lower.

III. MODEL BASED ON INDICATION

Society represents a model of reputation and confidence evolved, considering that, during the interactions, the behavior of elements belonging to this system is now considered and at each need for interaction, one makes a analysis by which one infers if it is trustful or not to interact with the node. When one needs a specific recourse, the requestor, in its seek of reference data to infer if a behavior is acceptable or not, it consults other elements that may have already acquire such recourse and have good reference of concluded negotiations. This approach differs from current ones, for it proposes to receive a indication of a recourse from a reliable source. Still there is a model of economy that allows the net to stretch, enabling new members to acquire a good reputation in the course of their interactions.

From this information, one applies one's policy of analysis and infers in which node it is possible to acquire such recourse with success. The analogy is established when someone who needs a certain service, a *baby-sitter* for instance, consults somebody who knows a qualified professional that was previously hired and with who one had no issues. This model tends to classify a individual according to its behavior converging to the "true" of the society and not to a pre-established threshold of the system.

The model proposed acts in the same manner of the society, searching recourses by indication of nodes to others they had already interacted. With this information, it's possible to rank the nodes which has similar recourses, and

 ISBN: 978-1-61208-192-2

the requestor policy is uncharged to infer whether one should or should not acquire such recourse, and in addition, classify the transaction and generate the record of interactions, ranking each time more precisely the nodes of the society. The evaluations are made as for the supplier node as to the recourse stored on it, and both has a reputation measured by the system.

Another strategy of the proposed model is to encourage the sharing of recourses. In case it doesn't happen, the malicious of selfish nodes are purged and not allowed to take part in P2P network.

A. Hierachy structure

The nodes of the net are classified in two levels, which are nodes and supernodes. Nodes are the elements that may share and request recourses, and any entrant node in the P2P net fits initially in this category. The supernode, besides being a node, is also able to make indications about the reputation of nodes and recourses, and make public the information of the data repository of its subordinates with the rest of the network. The supernode is a trustful element of the net. Compulsorily, every node is related to a supernode, and this association occurs in the occasion of the inlet of the node in the P2P net. The supernode, chosen to support the inlet node, is the one that possess the least amount of subordinate nodes. A node is associated solely to a supernode [9].

The node is promoted to a supernode when it reaches a pre-established threshold of points. This score is initially pre-defined, but alters dynamically with the evolutions of supernodes, that is, the least score to become a supernode is the average of the current punctuation of a group of supernodes. Besides the scores, one must consider several basic characteristics to promote a node to a supernode, for instance, storage capability, band width, or still, the period of participation in P2P net. After the promotion, it is possible to be lowered to a node, according to its behavior. At each affiliation of a new node to a supernode, this receives a punctuation for publishing the information in P2P network. Every supernode has another contingency supernodes, in case some unavailability, the node may affiliate itself to the contingency supernode [10].

B. Repositories

Locally, each node stores a repository containing information of its shared recourses. In each shared recourse is given a punctuation to the supplier node. Also, it is given a initial punctuation to the recourse. Similarly to the node punctuation, each recourse takes a punctuation at each transaction, which may increase or decrease it. Each local information of each node is replicated in the supernode, and the supernode, with a certain frequency, replicates to the contingency supernodes.

C. Seek for recourses

The seek process for recourses is unique for avoiding the consult to every node of the net, but only to the supernodes, thus reducing a lot the amount of requests sent. When a requestor element, *Node_Requestor*, wishes s certain recourse, it sends to every supernode, including its superior, a requisition *Query(keyWord)*, and every supernode that know the holder of this recourse respond with the identification of it. The supernode, according to the previously mentioned, is a element of reliability that knows either the one who possess a certain recourse or its respective reputation, node and recourse. Thus, the supernode sends such reputations to the requestor node in the message *QueryResponse(resourceID, nodeID, resourceReputation, nodeReputation)*, see Figure 3.

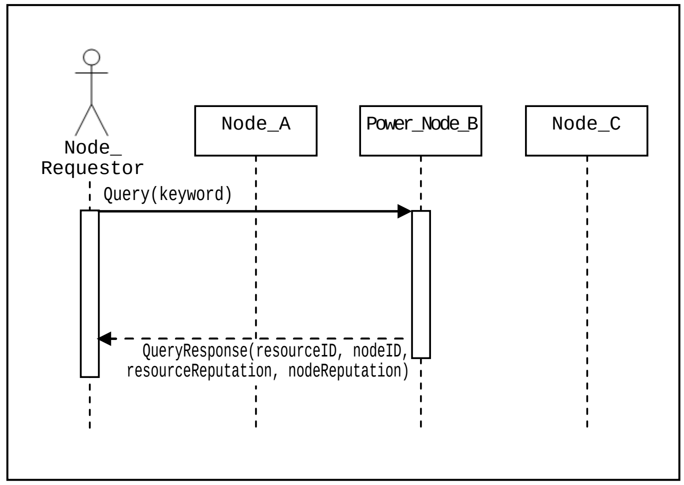


Figure 3. Demand for resources.

Either the score of recourses or the nodes score has a ranking character, since it has positive values, otherwise, it becomes eliminatory. In the course of interactions, in case the node or file zero out its scores, they shall be no longer indicated by the supernode. The nodes which have a score lower or equal to zero may not request for resources as well. This last rule restricts selfish users from staying in the network.

D. Acquisition

With the answers from supernodes, the requestor node chooses the right recourse ordered by the reputation and picks it. The recourses with higher reputation along with the nodes of higher reputation shall probably succeed in the acquisition of the recourse. Each recourse acquired generates a cost in scores for the requestor, independently whether the file is valid or not.

E. Qualification

After the acquisition of the recourse, the node requestor qualifies the transaction, scoring the same way either the supplier node or the recourse acquired. Such interaction is locally stored and sent to the supplier node. In case of conflict of score information, the one that prevails is always the lower, this inhibits bad-intentioned users that lie about its reputation to gain reliability.

 ISBN: 978-1-61208-192-2

IV. Simulation

The model proposed assumes that the system may the parameterized in order to attend to the dynamism of the behavior of the P2P networks. To simulate this, values were defined to the score of behavior actions, such as sharing a recourse, consume recourse, among others, according to Table I.

TABLE I. VALUES FOR BEHAVIORAL ACTIONS

Behavioral action	Value
Score to node entrant	100 points for the node.
Scoring for the shared resource	50 points for the resource.
Share resource	1 point for sharing the node.
consuming Resource	-3 Point for the consumer.
Being well qualified to provide resource	10 points for the resource 10 points for the provider node.
Be badly qualified to provide resource	-20 Points for the resource -20 Points for the provider node.
Super-node join node	100 points for super-node.
Score for promotion to the super-node	1000 points

Due to the rule of each inlet node to compulsorily associate to a supernode, it was created a supernode element to start the activities of the P2P networks.

One of the propositions of the model is to permit that new honest nodes get into the net and reach good reputation. After the creation of the supernode to simulate this scenario, one executed 4 steps applied to the values of Table II at each of them.

TABLE II. DISTRIBUTION OF NODES AND RESOURCES APPLIED

Element	Quantity
Nodes generous	35%
Malicious nodes	35%
Nodes Selfish	30%
Total of nodes	2000
Valid files	50%
Invalid files	50%
Total resources shared	56,000
Total interactions made	270,000

The execution considered the selection of requestor nodes and recourses supplied at chance. One of the restrictions imposed was that a node may not take a recourse of itself, for instance, the node itself does not consume its own recourse nor scores its own recourse. The average of reputation of honest nodes raises mainly by the sharing of valid files. On the other hand, the average of reputation of malicious nodes tends to zero, for sharing invalid files. Selfish nodes follow the same trend of the malicious ones, but, with least score, for they only consume recourses, according to Figure 4.

Next, in Figure 5, it is possible to see the evolution of the amount of excluded nodes due to bad behavior and the amount of nodes promoted due to sharing of valid files during interactions. Frequently, nodes with good behavior are promoted once they are never despised, tending to promote every node with this behavior. Malicious or selfish nodes are despised with least proportion to the promoted nodes, still permitting to be promoted in case they become honest nodes.

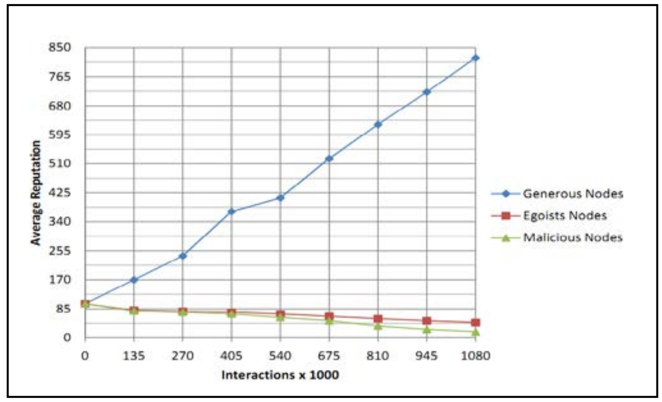


Figure 4. Reputation average x Interactions.

It is important to stand out that threshold of promotion of supernodes are dynamic, and still the scores given to the behavior actions are parameterized, but can be fitted. A node which is promoted to a supernode may be lowered to a node once more in case of a bad behavior, ensuring that the supernodes are trustful elements in the net.

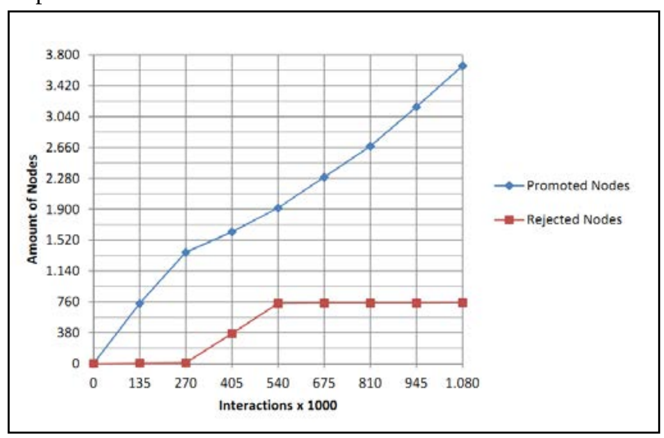


Figure 5. Comparative Interactions x Nodes.

By its turn, evaluations in simulated environment show that the model combats pollution, for attacks mainly the malicious nodes that share invalid files, promoting generous nodes and underprivileging selfish nodes. The benefits that the system promotes were validated.

V. Conclusion and Future Works

We presented a model of reputation able to combat the poisoning of recourses and, consequently, the pollution of the P2P network. Still it was showed that, from some interactions with the net, malicious nodes are despised and can no longer share not even consume recourses of the net. Selfish users are also underprivileged and follow the same tendency but with a less steep curve, according to which was proposed in the simulation section.

The hierarchy characteristic applied to the distributed and decentralized net, allows a uniform growth, at the same time that the distribution weighs the supernodes that have less recourses associated. The overhead is drastically mitigated in such a way that, in the worse of cases, it would

 ISBN: 978-1-61208-192-2

reach to be equal to correlated works. The capability to parameterize the system helps finding the best threshold of promotion of supernodes and the purge of malicious and selfish nodes.

REFERENCES

[1] S. Marti, and H. Garcia-Molina, "Taxonomy of trust: Categorizing P2P reputation systems", Computer Networks: The International Journal of Computer and Telecommunications Networking - Management in peer-to-peer systems, vol. 50, Elsevier, August 2005, pp. 472–484, doi:10.1016/j.comnet.2005.07.011.

[2] M. P. Barcellos, and L. P. Gaspary, "Segurança em Redes P2P: Princípios, Tecnologias e Desafios", Proceedings of short courses of the 24th Brazilian Symposium on Computer Networks, pp. 1-50, Brazil, Mai 2006.

[3] C. Costa, "Disseminação de Conteúdo Poluído em redes P2P", Proceedings of the 24th Brazilian Symposium on Computer Networks, pp. 1-13, Brazil, Mai 2006.

[4] J. S. P. Guedes, "Framework para Cálculo de Reputações de Agentes de Software Baseado em Testemunhos", Pontíficia Universidade Católica do Rio de Janeiro, pp. 45-82, Brazil, March 2007.

[5] A. Josang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision", Decision Support Systems, vol. 43, Elsevier, The Netherlands, July 2007, pp. 618-644, doi: 10.1016/j.dss.2005.05.019.

[6] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara, "Reputation Systems: Facilitating Trust in Internet Interactions" Communications of the ACM, Vol. 43, USA, December 2000, pp. 45-48, doi:10.1109/IWQoS.2010.5542754.

[7] K. Walsh, and E. G. Sirer, "Experience with an Object Reputation System for Peer-to-Peer Filesharing", In Proceedings of the 3rd conference on Networked Systems Design & Implementation, Vol. 3, pp. 1-14, USENIX Association Berkeley, USA, 2006.

[8] H. Chen, and G. Chen, "A Resource-based Reputation Rating Mechanism for Peer-to-Peer Networks", Proceedings of the Sixth International Converence on Grid and Cooperative Computing, IEEE Computer Society, USA, August 2007, pp. 535-541, doi:10.1109/GCC.2007.23.

[9] N. Christin, A. S. Weigend, and J. Chuang, "Content Availability, Pollution and Poisoning in File Sharing Peer-to-Peer Networks", Proceedings of the 6th ACM conference on Electronic commerce, ACM, Canada, June 2005, pp. 68-77, doi:10.1145/1064009.1064017.

[10] F. Benevenuto, C. Costa, M. Vasconcelos, V. Almeida, J. Almeida, and M. Mowbray, "Impact of Peer Incentives on the Dissemination of Polluted Content", Proceedings of the ACM Symposium on Applied Computing, France, April 2006, pp. 1875-1879, doi:10.1145/1141277.1141720.

[11] J. Liang, N. Naoumov, and K. W. Ross, "The Index Poisoning Attack in P2P file-sharing Systems", Proceedings of the 25th IEEE International Conference on Computer Communications, Spain, April 2007, pp. 1-12, doi:10.1109/INFOCOM.2006.232.

[12] R. Gupta, and A. K. Somani, "Reputation Management Framework and its use as Currency in Large-scale Peer-toPeer Networks", Proceedings of the Fourth International Conference on Peer-to-Peer Computing, USA, 2004, pp. 124-132, doi:10.1109/P2P.2004.44.

[13] M. Feldman, K. Lai, I. Stoica, and J. Chuang, "Robust incentive Techniques for Peer-to-Peer Networks", Proceedings of the 5th ACM Conference on Electronic Commerce, USA, May 2004, pp. 102-111, doi:10.1145/988772.988788.

[14] S. D. Kanvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust Algorithm for Reputation Management in P2P Networks", Proceedings of the 12th international conference on World Wide Web, USA, 2003, pp. 640-651, doi:10.1145/775152.775242.

[15] A. C. F. Lopes, "Um método para geração de estimativas de reputação mais precisas perante a oscilação de comportamento das entidades avaliadas". Universidade Federal Fluminense, Instituto de Computação, Brazil, July 2006.

[16] C. Costa, "Combatendo a Disseminação de Conteúdo Poluído em redes P2P", Porceedings of the 25th Brazilian Symposium on Computer Networks and Distributed Systems, pp. 1-13, Brazil, Mai 2007

Inter-domain Peering in Content-Aware Networks for Multimedia Applications

Eugen Borcoci
Telecommunication Dept.
University POLITEHNICA of Bucharest
Bucharest, Romania
eugen.borcoci@elcom.pub.ro

Radu Dinel Miruta
Telecommunication Dept.
University POLITEHNICA of Bucharest
Bucharest, Romania
radu.miruta@elcom.pub.ro

Serban Georgica Obreja
Telecommunication Dept.
University POLITEHNICA of Bucharest
Bucharest, Romania
serban.obreja@elcom.pub.ro

***Abstract* — The Content Aware Networking is an emerging architectural solution, responding to the significant increase in Internet content orientation. This paper is a continuation of a previous work and refines a management framework, for inter-domain peering in overlay Virtual Content Aware Networks (VCAN), QoS enabled, built over multi-domain, multi-provider IP networks. An overlay inter-domain topology service and negotiation protocols are defined in this paper, based on cooperation of the CAN Managers belonging to network domains. The scalability and efficiency is preliminary analyzed. The work is part of the research effort, inside a FP7 European Information and Communication Technologies (ICT) research project, ALICANTE, oriented to multimedia distribution based on CAN approach.**

Keywords — Content-Aware Networking, Network Aware Applications, Multi-domain, Inter-domain peering, Management, Multimedia distribution, Future Internet.

I. Introduction

The current Internet limitations are recognized, related to the needs of today world and the global spread of this technology. High research efforts are spent to find enhanced architectural solutions or "clean slate" ones, to solve the limitations thus leading to the Future Internet (FI) architectures. Sample of works are presented in [1] - [8]. The work [1] emphasizes the strong orientation of the FI towards content and services and shows the importance of management. Network virtualization is seen as an important "tool" to overcome the ossification of the current Internet [2] - [5]. The overview paper [5], identifies the inefficiency of the current Internet for time-sensitive multimedia content delivery and analyses new solutions based on *Content Oriented Networking* (CON) with decoupling of contents from hosts at networking level. A major trend is the shift from the traditional TCP/IP stack concepts with agnostic network layer to more intelligency in the network layer. New network nodes process the data, based on *content type* recognition or, even more, treating the data objects based on their *name* and not based on *location address*, [6][7]. Inline with this, a new concept is the Content-Awareness at Network layer (CAN) and Network-Awareness at Applications layers (NAA). These new solutions are hopefully able to better support the development of the networked media systems and also the market orientation towards content. The approach is claimed by many studies to bring new benefits for both, Service and Application Layer and Network layer, thus creating a powerful *cross-layer optimization loop* between the transport and applications and services.

The European FP7 ICT running research project, "Media Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments", ALICANTE, [9][10][11], adopted the NAA/CAN approach. It targets to define an architecture, and then to fully specify, design and implements a Media Ecosystem, on top of multi-domain IP networks, to offer a large variety of services for different business actors playing roles of consumers and/or providers.

Architecturally, ALICANTE is a "middle-way" solution: it adopted content-type recognition at network level and light virtualization (separation in the Data Plane of the virtual networks but a single management and control plane). This solution is believed to offer seamless deployment perspectives and tries to avoid the scalability problems (still open research issues) of the full CON approaches.

Several cooperating environments are defined, including several business entities/actors: *User Environment (UE),* containing the End-Users; *Service Environment (SE),* containing High Level Service Providers (SP) and Content Providers (CP); *Network Environment (NE),* where a new CAN Provider exists (CANP - managing and offering Virtual Content Aware Networks- VCANs); traditional Network Providers (NP/ISP) - managing the network elements at IP level. By "environment", it is understood a generic grouping of functions working for a common goal and which possibly vertically span one or more several architectural (sub-) layers.

A VCAN can span several network domains, where each one is managed independently (a realistic business constraint), while offering different levels of QoS guarantees for media flows needs. Therefore an architectural decision have to be taken, on how to manage the peering in the Data Plane (including inter-domain routing) and in the Management and Control Plane (M&C signaling) in order that they cooperate to the realization of a shared VCAN. This is the subject of this paper. Several solutions are analyzed for M&C (cascade, hub, mixed) and finally the so called " hub model" has been selected. A M&C negotiation protocol is proposed to run between domain managers. The scalability aspects are preliminary discussed.

The CANP offers to the upper layers enhanced VCAN-based connectivity services, unicast and multicast (QoS enabled) over multi-domain, multi-provider IP networks. The VCAN resources are managed quasi-statically by provisioning and also dynamically by using adaptation procedures for media flows. The management is based on vertical and horizontal Service Level Agreements (SLAs) negotiated and concluded between providers (e.g SP-CANP). In the Data Plane, content/service description information (metadata) can also be inserted in the media flow packets by the Content Servers and treated appropriately by the intelligent routers of the VCAN.

The paper continues the starting work on VCAN presented in [12] [13]. It is organized as follows. Section II presents samples of related work. Section III summarizes the overall ALICANTE architecture. Section IV shortly presents the content awareness features of the system and QoS assurance solutions. Section V is the main one, dedicated to the peering solution selected and associated negotiations aiming to extend a VCAN over several domains. Section VI contains some conclusions and future work outline.

II. RELATED WORK

The paper objective is to develop management solutions to govern the construction of VCANs, QoS capable over several independent network domains which should be peered and assure guaranteed QoS enabled transport of real-time and media traffic.

For inter-domain QoS enabled domain peering, there exist basically two kinds of approaches. The first one [14] [15], proposes QoS enhancements for the Border Gateway Protocol (BGP). The BGP advertises QoS related information between network domains – seen at limit as autonomous systems (ASes), and then a QoS aware routing table is built. However, the notion of content awareness at domains level is absent there.

Other solutions for inter-domain QoS peering and routing are based on the overlay network idea [16] [17] [18]. An overlay network is defined, which first, abstracts each domain with a node, represented by the domain resource manager, or more detailed with several nodes represented by the egress routers from that domain. There exist protocols to transport QoS and other information between nodes and, based on this information, QoS routing algorithms are used to choose the QoS capable path. In [16] a Virtual Topology (VT) is defined by a set of virtual links that map the current link state of the domain without showing internal details of the physical network topology. Then *Push* and *Pull* models for building the VT at each node are considered and analyzed. In the *Push* model each AS advertises its VT to their neighbor ASes. This model is suited for small topologies. In the Pull model the VT is requested when needed, and only from the ASes situated along the path between given source and destinations; the path itself is determined using BGP.

After routes are found, a negotiation protocol should be run [12]-[15] [18], to establish inter-domains Service Level Specification (SLS) agreements (SLS is the SLA technical part) containing clauses for QoS guarantees.

Related to management of inter-domain peering, several solutions are examined and compared (cascade, hub, mixed-mode) [16][17][18]. However, neither solution considers the content awareness capabilities of the multiple domain infrastructure, nor the virtualization aspects. This paper takes these into account. Also, ALICANTE architecture realizes parallel Internet planes as in [19], but mapped onto VCANs, and additionally achieves cooperation between the network layer and applications and services layers, thus realizing a traffic optimization loop (OL), similar to [20].

III. ALICANTE SYSTEM ARCHITECTURE AND VCAN MANAGEMENT

A. General Architecture

The general ALICANTE architecture is already defined in [9][10][11]. A set of business actors is defined, composed of traditional SP, CP, NP - Providers and End-Users (EU). New business actors are introduced: CAN Provider (CANP) offering virtual layer connectivity services and the Home-Box (HB)- partially managed by the SP, the NP, and the end-user, located at end-user's premises and gathering content/context-aware and network-aware information. The HB can also act as a CP/SP for other HBs, on behalf of the EUs. Correspondingly, two novel virtual layers exist: the CAN layer and the HB layer. The novel CAN routers are called *Media-Aware Network Elements (MANE)* to emphasize their additional capabilities: content and context - awareness, controlled QoS/QoE, security and monitoring features, etc.

The CAN layer M&C is partially distributed; it supports CAN customization to respond to the SE needs, including 1:1, 1:n, and n:m communications and also allow efficient network resource exploitation. The interface between CAN and the upper layer supports *cross-layer optimizations* interactions, e.g., offering network distance information to HBs to help collaboration in P2P style, [20]. A hierarchical monitoring subsystem supervises several points of the service distribution chain and feeds the adaptation subsystems with appropriate information, at the HB and CAN Layers. Figure 1 presents a partial view on the ALICANTE architecture, with emphasis on the CAN layer and management interaction. The network contains several Network Domains (ND), belonging to NPs (they can be also seen as Autonomous Systems - AS) and access networks (AN). The ANs are out of scope of VCANs. One *CAN Manager* (CANMgr) exists for each IP domain to assure the consistency of VCAN planning, provisioning, advertisement, offering, negotiation installation and exploitation. Each domain has an *Intra-domain Network Resource Manager* (IntraNRM), as the ultimate authority configuring the network nodes. The CAN layer cooperates with HB and SE by offering them CAN services.

B. *VCAN Management*

The VCAN Management framework has been already defined in [12]. Here only a short summary is recalled for sake of clarity. At the Service Manager SM@SP, the CAN Network Resources Manager (CAN_RMgr) performs all actions needed for VCAN support on behalf of SP. It performs, at SP level, *VCAN* planning, provisioning (negotiation with CANP on behalf of the SP) and then VCAN operation supervision. The *CANMgr@CANP* performs, at the CAN layer, VCAN provisioning and operation. The two entities interact based on the SLA/SLS contract initiated by the SP. The interface implementation for management is based on Simple Object Access Protocol (SOAP)/Web Services. The contracts/interactions of SLA/SLS types performed in the M&C Plane are shown in Fig. 1:

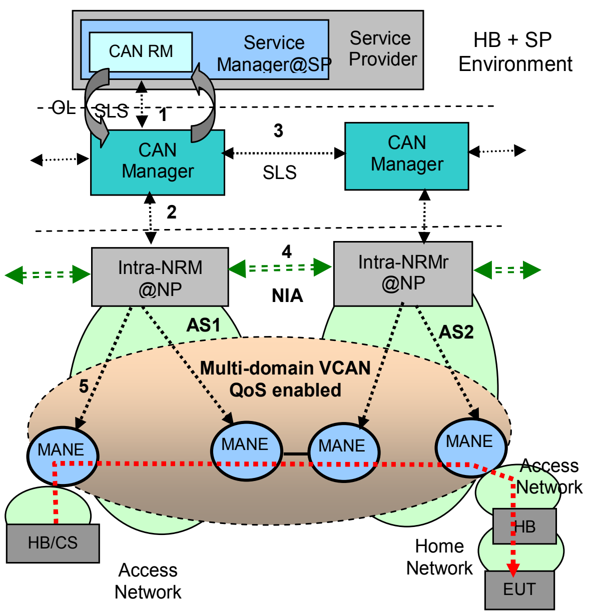


Figure 1. High level ALICANTE architecture: multi-domain VCANs and main management and control interactions

Notations: RM – Resource Management; HB-Home Box; CS- Content Server; EUT- End User Terminal; OL - Optimization Loop; NIA-Network Interconnection Agreements; SP, NP – Service, Network Providers

Interactions in the Fig. 1 are described as follows:*SP-CANP(1)*: the SP requests to CANP to provision/ modify/ terminate VCANs while CANP says yes/no; also CANPmight advertise existent VCANs to SP; *CANP-NP(2)* : CANP negotiates resources with NP;; *CANP-CANP(3)* – negotiations are needed to extend a VCAN upon several NP domains; *Network Interconnection Agreements (NIA) (4)* between the NPs or between NPs and ANPs; these are not new ALICANTE functionalities but are necessary for NP cooperation.

After the SP negotiates a desired VCAN with CANP, it will issue the installation commands to CANP, which in turn configures, via Intra-NRM (action 5), the MANE functional blocks (input and output).

IV. CONTENT AWARENESS AND QOS AT CAN LAYER

The content awareness (CA) is realized in three ways:

(i) by concluding a SP - CANP SLA concerning different VCAN construction. The content servers are instructed by the SP to insert some special *Content Aware Transport Information (CATI)* in the data packets. This simplifies the media flow classification and treatment by the MANE; (ii) SLA is concluded, but no CATI is inserted in the data packets (legacy CSs). The MANE applies packet inspection for data flow classification and assignment to VCANs. The flows treatment is still based on VCANs characteristics defined in the SLA; (iii) no SP–CANP SLA exists and no CATI. The flows treatment can still be CA, but conforming to the local policy at CANP and IntraNRM.

The DiffServ and/or MPLS technologies support splitting the sets of flows in QoS classes (QC), with a mapping between the VCANs and the QCs. Several levels of QoS granularity can be established when defining VCANs. The QoS behavior of each VCAN (seen as one of the parallel Internet planes) is established by the SP-CANP.

Generally a 1-to-1 mapping between a VCAN and a network plane will exist. Customization of VCANs is possible in terms of QoS level of guarantees (weak or strong), QoS granularity, content adaptation procedures, degree of security, etc. A given VCAN can be realized by the CANP, by combining several processes, while being possible to choose different solutions concerning routing and forwarding, packet processing, and resource management.

The definitions of local QoS classes (QC) and extended QCs and meta-QoS classes were adopted in ALICANTE, [14][15][18][19] to allow capturing the notion of QoS capabilities across several domains. Each domain may have its local QoS classes and several local QCs can be combined to form an extended QC..The types of VCANs defined for different QoS granularities based on QCs are described in [12]: VCANs based on meta-QCs, [14], VCANs based on local QC composition and hierarchical VCANs based on local QC composition. The last case is the most efficient but also the most complex. Inside each VCAN, several QCs are defined corresponding to platinum, gold, silver, etc. In such a case, the mapping between service flows at SP level and CANs can be done per type of the service: VoD, VoIP, Video-conference, etc.

V. CAN MULTI-DOMAIN PEERING

A. *Horizontal M&C VCAN Negotiation*

A given VCAN may span one or several IP domains. In a multi-domain context, one should distinguish between two topologies (in terms of how the domains are linked with each others): *Data plane topology* and *M&C topology*. The first can be of any kind (depending on SP needs and including the domains spanned by a given VCAN). In a general case, one may have a mesh/graph of domains. The *M&C topology* defines how the CAN Managers associated to different domains inter-communicate for multi-domain VCANs construction. The VCAN initiating CANMgr has to negotiate with other CAN Managers. There exist two main models to organise this communication at management level: *hub* model and *cascade* model [14][15][18][19].

The *hub model* was selected; it has the advantage that initiating CANMgr can know, each VCAN component (network) and its status. A drawback is that each CANMgr should know the inter-domain topology (complete graph) of network domains. They could be of lower tier grade or be Autonomous Systems (AS), involved in a VCAN. Given the tiered hirarchy of the Internet, the number of Network Domains (ND) involved in an E2E chain is not too high (actually is lower than 10, [8]), scalability problem is not so stringent. Two functional components are needed: (1) inter-domain topology discovery protocol; (2) overlay negotiation protocol for SLA/SLS negotiations between CAN Managers.

The *cascade* model, [15], [18] is more advantageous for initiating CAN Manager if a chain of domains is to form the VCAN. However, for an arbitrary mesh topology of the NDs composing the VCAN, and for multicast enabled VCAN, this model offers less efficient management capabilities.

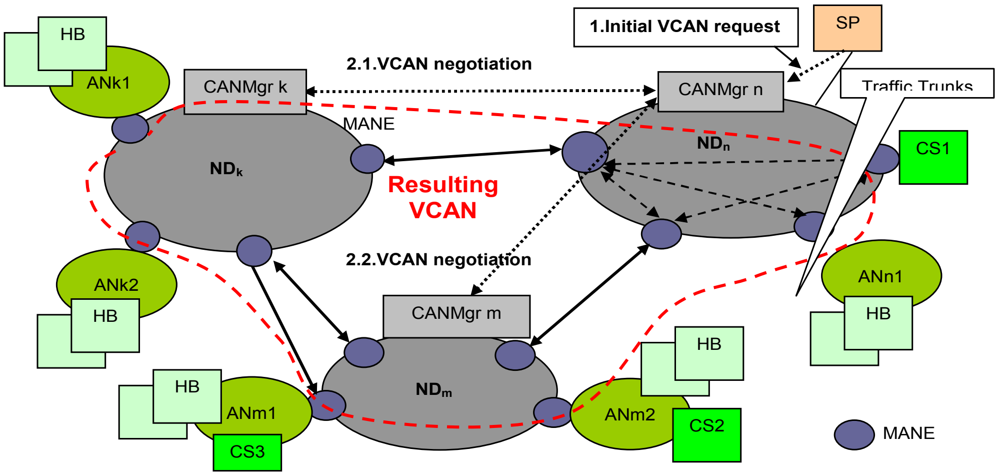


Figure 2 Example of a multi-domain VCAN (hub model for management plane)

Fig. 2 shows an example of a multi-domain VCAN. It is supposed that the inter-domain discover protocol has already produced its results, so each CANMgr knows about the inter-domain graph and have inter-domain routing information, including link capacities and QoS related capabilities. The SP asks for a VCAN to a CANMgr (Initiator) – see action 1. It was supposed that the SP knew the edge points of this VCAN, i.e. the MANEs IDs where different sets of HB currently are, or they will be connected. The initiator *CANMgr_n* determines all network domains (ND) involved (from the SP information and its inter-domain knowledge) and then negotiate in parallel with all other CAN Managers (actions 2.1, 2.2) to establish the VCAN = {VCANn **U** VCANm **U** VCANk}. The split of the SLS parameters (if it is the case) should be done at the initiator (e.g. for delay). In a successful scenario, the multi-domain VCAN is agreed and then it is later instantiated in the network.

B. Overlay Virtual Topology

Constructing VCAN over one or multiple domains is a main target of the CAN Manager. Each ND has complete autonomy w.r.t its network resources including network dimensioning, off-line traffic engineering (TE), and also dynamic routing. The CANMgr cooperating with Intra-NRM is supposed to know about its network resources.

Given that in ALICANTE each ND has associated the IntraNRM and CANMgr, one could abstract the both under the name of NDMgr. This entity should have an abstract view of its network domain and output links towards neighbors in a form of a set of virtual pipes (called Traffic Trunks). A set of such pipes can belong to a given QoS class. As already stated, a multiple domain VCANs should also belong to some QoS class and therefore inter-domain QoS aware routing information is necessary in order to increase the chances of successful SLS establishment, when negotiating the multi-domain VCAN. The multi-domain VCANs deployment needs knowledge on a virtual multi-domain topology.

Each ND can assure QoS enabled paths towards some destination network prefixes while implementing its own network technology: DiffServ, MPLS, etc. Also, each ND can be seen in an abstract way as an *Overlay Network Topology (ONT)* expressed in terms of *TTs (traffic trunks)* characterized by of bandwidth, latency, jitter, etc. One TT is belonging to a given QoS class QCi.

We define an *Overlay Network Service (ONS)* responsible for getting the ONTs related to NDs belonging to

 ISBN: 978-1-61208-192-2

a multi-domain VCAN. The CANMgrs will then inter-negotiate the SLS contracts in order to reserve VCAN resources and finally ask installation of them. The overlay topology can be hierarchised on several levels.

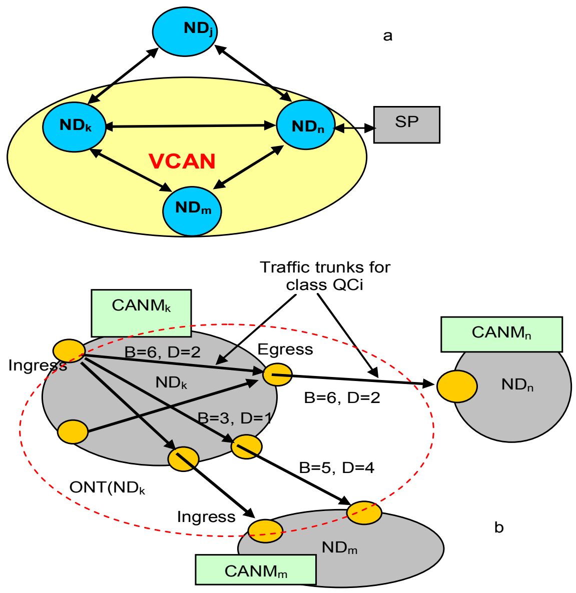


Figure 3. a. Inter-domain level overlay network topology (ONT); b. ONT of the domain NDk (B= bandwidth, D= max delay- generic figures)

Fig. 3.a presents a first level (inter-domain) ONT, in which, each domain ND is seen as a node. The overlay graph of the NDs belonging to the VCAN is composed of the nodes NDk, NDn, NDm. Then, a second order ONT can be defined for one ND. Figure 3.b shows the ONT for the domain NDk, composed of TTs, each one characterized by a bandwidth and a delay. If the initiating CAN Manager knows the ONT graph of the NDs involved, then provisioning of QoS enabled VCANs can be done.

The ONS can be act in two ways: a *proactive* (*push*) mode and a *reactive* (also called *pull* or *on demand*) mode in order to obtain the overlay (virtual) topologies of other NDs.

In the *proactive case*, every ND advertises its ONT to other NDs without being requested for. The advantage is the same as in IP proactive routing protocols: the ONTs of other NDs are already available at a given ND because they are periodically or event-triggered advertised among ND managers. The the advertisement can executed at an initiative of each ND manager, so this model allows promotion of some routes to other domains. This can be subject of policies. The dynamicity is high (event driven advertisements), but the complexity is also high. Scalability problems exist, because of high control traffic volume and also flooding the neighbour NDs with (maybe) not needed information.

In the *reactive* (*on-demand*) mode the ONTs are obtained on demand by an ND interested to reach a given destination prefix. The ND will query each domain of a given path to get the ONTs. No advertising mechanism is necessary. The scalability is higher because only the ONTs of the chosen routes will be obtained. Studies [8] show that the mean End to End (E2E) communication in the Internet usually involves few domains (less than 8). Therefore, the number of domains to be queried to obtain the ONTs is small. The pull model latency is higher (need time for queries and calculations). The updates of ONT knowledge is not event driven w.r.t other NDs, because lack of advertisements. For ALICANTE we have chosen the reactive model.

In ALICANTE case if a CANMgr wants to build an ONT it will query its directly linked (at data plane level) neighbour domains (i.e the corresponding CAN Managers). It is supposed that it has the knowledge of such neighbours. There two possibilities of a querry:

a. non-selective querry/demand- the asking CANMgr wants to know all neighbourhood of the asked neighbours

b. selective demand- the asking CANMgr wants to know answers only from those AS neighbours which have paths to a given set of destinations.

In case a. each queried CANMgr can return – in a first most simple approach only its list of neighbours. At receipt of such information, the interrogating CANMgr updates its topology data base. Then it queries the new nodes learned and so on. The process continues until the interrogating node CANMgr learns the whole graph of "international" topology.The extension of such a zone can be determined by local policies. Because the topology structure changes events are not very frequent (weeks, months), the topology construction process could be run at large time intervals (once a day, for example). Consequently the amount of messages used to build the ONT will not overload the significantly the network. In the case b. the query process is similar, but the answers will be selective, i.e., filtered conforming the required set of destinations.

The area of knowledge desired by a given CANMgr can be determined by policies and can be enlarged if needed. The Fig. 4 shows such an area for two levels of extension.

We summarize the ALICANTE design decision: *on-demand model based on overlay network topology service; based on non-selective and selective queries and simple answers* (i.e., no internal ONT information of a domain is made public). Advantages are: less complexity, higher speed, preserving intra-domain information. Drawbacks are: higher probability of failures of QoS enabled paths at first attempt.

VI. Conclusions and Future Work

The paper proposed a management solution for inter-domain peering, in Content Aware Networks for a multi-domain and multi-provider environment. The management is based on horizontal SLAs negotiated and concluded between CAN providers (represented by CAN Managers) the result being a set of parallel VCANs offering different classes of services to multimedia flows, based on CAN/NAA concepts. The inter-domain approach is to develop an overlay topology service to support VCAN construction, thus obtaining several parallel QoS planes. A CAN Manager is initiating the multi-domain VCAN realization by using the overlay topology service. The system is currently under complete design and implementation in the framework of the FP7 research project ALICANTE. Validation and performance evaluation results will be shown in a future work.

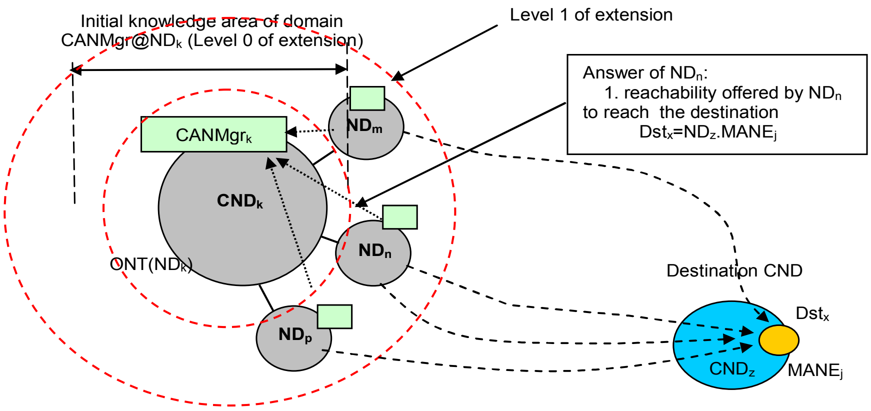


Figure 4 Different areas of ONT knowledge for NDk (Level 0, Level 1, …) in selective –query mode

Acknowledgments

This work was supported partially by the EC in the context of the ALICANTE project (FP7-ICT-248652) and partially by the project POSDRU/89/1.5/S/62557.

References

[1] Schönwälder, J., Fouquet, M., Dreo Rodosek, G., and Hochstatter, I.C., "Future Internet = Content + Services + Management", IEEE Communications Magazine, vol. 47, no. 7, Jul. 2009, pp. 27-33.

[2] Baladrón, C., "User-Centric Future Internet and Telecommunication Services", in: G. Tselentis, et. al. (eds.), Towards the Future Internet, IOS Press, 2009, pp. 217-226.

[3] Turner, J. and Taylor, D., "Diversifying the Internet," Proc. GLOBECOM '05, vol. 2, St. Louis, USA, Nov./Dec. 2005, pp. 760-765.

[4] Anderson, T., Peterson, L., Shenker, S., and Turner, J., "Overcoming the Internet Impasse through Virtualization", Computer, vol. 38, no. 4, Apr. 2005, pp. 34–41.

[5] 4WARD, "A clean-slate approach for Future Internet", http://www.4ward-project.eu/.

[6] J.Choi, J.Han, E.Cho, T. Kwon, and Y.Choi, A Survey on Content-Oriented Networking for Efficient Content Delivery, IEEE Communications Magazine, March 2011.

[7] V. Jacobson et al., "Networking Named Content," CoNEXT '09, New York, NY, 2009, pp. 1–12.

[8] W.K. Chai, et.al., CURLING: Content-Ubiquitous Resolution and Delivery Infrastructure for Next-Generation Services , IEEE Communications Magazine, March 2011.

[9] FP7 ICT project, "MediA Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments", ALICANTE, No248652, http://www.ict-alicante.eu/ (last accessed: Dec. 2010).

[10] Borcoci, E., Negru, D. and Timmerer, C., "A Novel Architecture for Multimedia Distribution based on Content-Aware Networking" Proc. of. CTRQ 2010, Athens, June 2010, pp. 162-168.

[11] ALICANTE, Deliverable D2.1, ALICANTE Overall System and Components Definition and Specifications, http://www.ict-alicante.eu, Sept. 2011.

[12] Eugen Borcoci, Mihai Stanciu, Dragoş Niculescu, Daniel Negru, George Xilouris, Connectivity Services Management in Multi-domain Content-Aware Networks for Multimedia Applications, " Proc. of. INTERNET 2011, Luxembourg, June 2011.

[13] Borcoci, E. and Iorga, R., "A Management Architecture for a Multi-domain Content-Aware Network" TEMU 2010, July 2010, Crete.

[14] Levis, P., Boucadair, M., Morrand, P., and Trimitzios, P., "The Meta-QoS-Class Concept: a Step Towards Global QoS Interdomain Services", Proc. of IEEE SoftCOM, Oct. 2004.

[15] Howarth, M.P. et al., "Provisioning for Interdomain Quality of Service: the MESCAL Approach", IEEE Communications Magazine, June 2005, pp. 129-137.

[16] Fabio L. Verdi, Maurıcio F. Magalhaes "Using Virtualization to Provide Interdomain QoS-enabled Routing", Journal of Networks, April 2007, pp. 23-32.

[17] 14] Z. Li, P. Mohapatra, and C. Chuah, Virtual Multi-Homing: On the Feasibility of Combining Overlay Routing with BGP Routing, University of California at DavisTechnical Report: CSE-2005-2, 2005.

[18] ENTHRONE, End-to-End QoS through Integrated Management of Content, Networks and Terminals, FP6 project, URL: http://www.ist-enthrone.org/.

[19] Boucadair, M. et al., "A Framework for End-to-End Service Differentiation: Network Planes and Parallel Internets", IEEE Communications Magazine, Sept. 2007, pp. 134-143.

[20] Aggarwal, V., Feldmann, A., "Can ISPs and P2P Users Cooperate for Improved Performance?", ACM SIGCOMM Computer Communication Review, vol. 37, no. 3, Jul. 2007, pp. 29-40.

Software Digital-Down-Converter Design and Optimization for DVB-T Systems

Shu-Ming Tseng
Graduate Institute of Computer and Communication Engineering
National Taipei University of Technology
Taipei, Taiwan
shuming@ntut.edu.tw

Jian-Cheng Yu
Graduate Institute of Computer and Communication Engineering
National Taipei University of Technology
Taipei, Taiwan
g931303@hotmail.com

Zheng-Hau Lin
Graduate Institute of Computer and Communication Engineering
National Taipei University of Technology
Taipei, Taiwan
steve76813@hotmail.com

***Abstract* –In this paper, a novel Digital-Down-Converter (DDC) architecture for PC-based software Digital Video Broadcasting-Terrestrial (DVB-T) receiver is proposed. The sampling rate of A/D is 192/7 MHz and the order of the lowpass FIR filter is seven (8 coefficients) in DDC. Furthermore, the Combining of the Mixer and Filter (CMF) is also proposed, including the pre-processing of mixer and filter coefficients and storing the results in a look-up table. If the input data of length is N, the proposed CMF scheme has 2N multiplications, while the previous architecture, which does not pre-process the mixer and filter coefficient together in advance, has 3N multiplications. Finally, the algorithms also are optimized in assembly code to satisfy DVB-T real-time reception requirement.**

Keywords- DVB-T; DDC; Decimate; Real-time.

I. INTRODUCTION

The structure of the receiver end of Digital Video Broadcasting –Terrestrial (DVB-T) system [1] is shown in Figure 1. The Digital Down Converter (DDC) is an important part of DVB-T system; it converts the Intermediate Frequency (IF) signal into baseband, reduces the signal sampling rate and then makes it easy for the later real time demodulation. The traditional DDC in DVB-T system was made by hardware circuits [2] [3]. However, it is not easy to be integrated and modified. The most important benefit of Software Radio (SR) research is that people can modify and change the signal processing procedure, the algorithm, and the result can be easily tested. Hence, there are some researches implemented DDC by software [4] [5].

In Figure 1, the IF generated by the tuner is 32/7MHz, defined by the DVB-T standard [1]. Besides, the sampling rate of DDC output in DVB-T standard must be 48/7MHz [1]. The previous DDC structure is shown in Figure 2, it composes of mixer、lowpass Finite Impulse Response (FIR) filter and down-sampling.

A Combining the Mixer and Filter (*CMF)* method is proposed. In other words, the mixer coefficients and filter coefficients are multiplied and the results are stored in a look-up table, as shown in (5). Furthermore, The *CMF* algorithms also are optimized in assembly code. The proposed *CMF* scheme combines the mixer and filter operations. Hence, it can reduce the computational complexity by one third. Assume the input data of length N, the previous scheme has 3*N* multiplications, but the proposed *CMF* only has 2*N*.

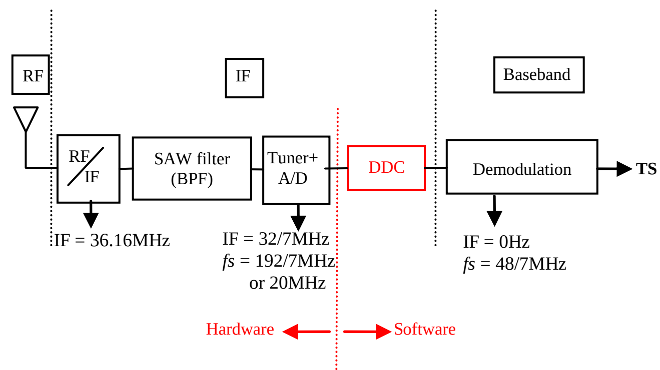


Figure 1. The structure of DVB-T receiver

The rest of this paper is organized as follows: In Section II, the system model of DDC is presented. The decision of A/D sampling rate is described in Section III. The proposed CMF scheme for DDC is described in details in Section IV. The CMF optimization is presented in Section V. Performance discussion is described in Section VI. Finally, we conclude in Section VII.

II. SYSTEM MODEL

The previous of DDC is introduced in Figure 2. According to [6], received signal s *(n)* is transformed to baseband by mixer first:

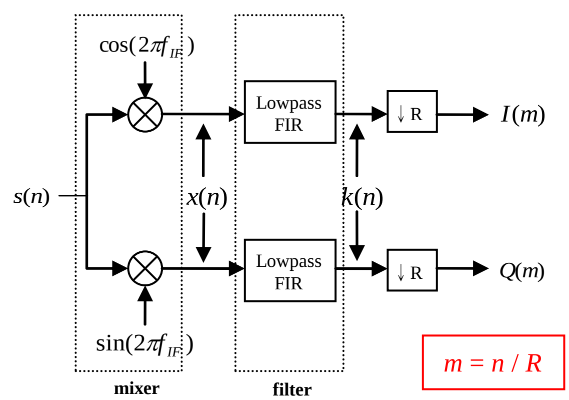


Figure 2. The DDC of previous structure

$$x(n) = s(n) * \exp(-j2\pi * f_{IF} * (n / f_{AD})), \quad 1 \le n \le N \qquad (1)$$

where N is the length of DDC input data, f_{IF} and f_{AD} is IF and sampling rate of A/D, respectively. The IF power level must be in the sampling rate range. Once it is lower than the A/D

 ISBN: 978-1-61208-192-2

sampling rate range, the resolution performance will be poor. On the contrary, if the IF power level is higher than the A/D sampling rate range, then it will produce distortion in the system. And then employ a low-pass FIR to avoid aliasing effect after down-sampling:

$$k(n) = h(0)x(n) + h(1)x(n-1) + \cdots + h(M)x(n-M)$$
$$= \sum_{m=0}^{M} h(m)x(n-m) = h(n) \otimes x(n) \qquad (2)$$

where M is the order of FIR filter, $h(n)$ is the filter coefficients, and "$\otimes$" is linear convolution. Final step is down-sampling; R means the down-sampling rate. After down-sampling, the data stream will satisfy the required sampling rate of DVB-T standard.

III. CHOICE OF A/D SAMPLING RATE

In Figure 3, most of f_{AD} is following the commercial specification: 20MHz [5] [7] to record data. According to DVB-T standard [1], the IF of tuner is 32/7MHz. Besides, the sampling rate of DDC output in DVB-T standard must be 48/7MHz [1].

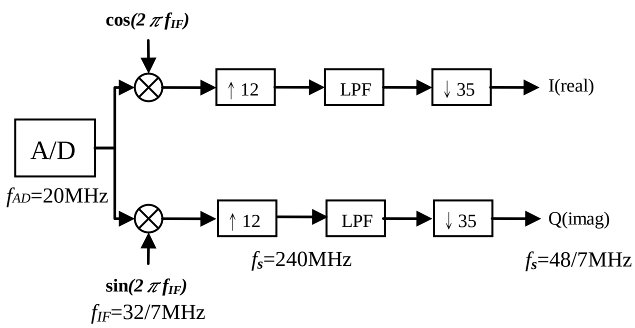


Figure 3. Digital-Down-Converter in [5] [10]

Because these three (20, 32/7, 48/7) are not in a multiple relationship, it will increase computational complexity.

In [8], f_{AD} is 4 times as much as IF; it could simplify the calculation of mixer. Moreover, integer decimation is proposed in [9]. According to [8] [9], changing the f_{AD} to be multiple of IF or the sampling rate of DDC output would simplify the DDC computation. Hence, the 192/7MHz of f_{AD} is chosen. It will match the multiple relation in [8] [9] simultaneously. For architecture of f_{AD} is 192/7MHz, as shown in Figure 4. This architecture avoids up-sampling calculation compared with Figure 3. The structure of f_{AD} = 192/7MHz is simpler than f_{AD} = 20MHz's. Thus, the 192/7MHz is chosen as f_{AD} in our structure.

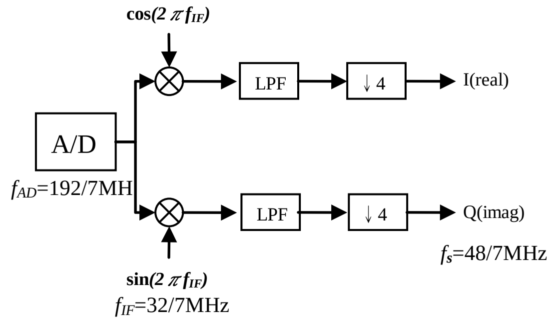


Figure 4. Digital-Down-Converter using f_s=192/7MHz

The environment of the hardware whose specification is listed in Table I. As we know, the case A has lots of additional up-sampling computations. From Table II, the case B can save much more time than case A.

TABLE I. HARDWARE LIST

Item	Model
CPU	Intel® Core™ i7-2600K (3.40GHz)
Memory	DDR2 800 2GB x 2
Main board	ASUS P8H67-M PRO Rev 1.xx
Graphic card	n.a

TABLE II. THE ELAPSED TIMES OF DIFFERENT F_{AD} (IN MATLAB)

	Elapsed time(s)
A. f_{AD} = 20MHz	441.28
B. f_{AD} = 192/7MHz	175.55

IV. PROPOSED CMF SCHEME

In this section, the new *CMF* scheme is proposed to simplify the DDC computation. The architecture without proposed CMF scheme is shown in Figure 4. Because f_{IF} / f_{AD} = 32/7 ÷ 192/7 = 1/6, we have:

$$x(n) = s(n) * \exp(-j2\pi * (n * 1/6)), \quad 1 \le n \le N \qquad (3)$$

From (3), the $\exp(-j2\pi*(n*1/6))$ only have possible 6 values. Furthermore, the filter only has M+1 coefficient. Substitute (3) into (2) and we get:

$$k(n) = \sum_{m=0}^{M} \{[s(n-m) * w((n-m) \bmod 6)] * h(m)\}, \quad 0 \le n \le N-1 \qquad (4)$$

where $w(n)$ = $\exp(-j2\pi*(n*1/6))$ is defined. In order to save the elapsed time from mixer calculation, the formula (4) will be modified as below:

$$
\begin{aligned}
k(n) &= \sum_{m=0}^{M} \{ s(n-m) * [w((n-m) \bmod 6) * h(m)] \}, \quad 0 \le n \le N-1 \\
&= \sum_{m=0}^{M} \{ s(n-m) * c(m) \}, \quad 0 \le n \le N-1 \qquad (5)
\end{aligned}
$$

The major contribution of this paper is combining the $w(n)$ and $h(m)$ in (5) is a look-up table in advance, then $s(n)$ calculates the linear convolution with $c(m)$ will achieve the mixer and filter calculations. Hence, it can save the time for computation. Because the number of mixer coefficients is not equal to filter's, we choose the least common multiple of these two numbers: 24. Figure 5 shows the combination of look-up table between mixer and filter. Here we assume $M = 7$. We also set $M = 7$ in our architecture, the detail reason will be described in next section.

Observing Figure 5, The 4 groups of mixer coefficients and 3 groups of filter coefficients are used to form the look-up table. The look-up table will be divided into three parts: ①、②and ③, so the received signal $s(n)$ operate (5) with ①、② and ③ circularly. The next step is down-sampling 4 times. This part could be combined with (5). For each 4 data input in DDC, it will generate 1 data. Thus, It could avoid 3/4 calculations form FIR.

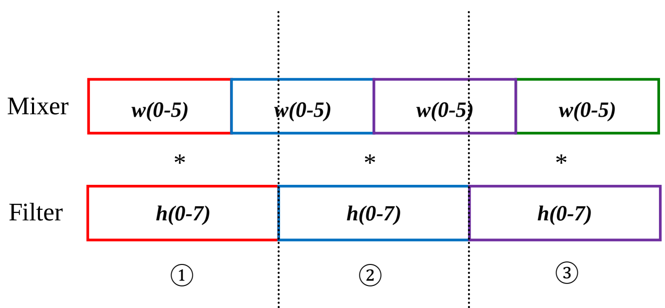


Figure 5. Look-up table of mixer and filter combination

According to (5), the block diagram of proposed algorithm: *CMF* is shown in Figure 6.

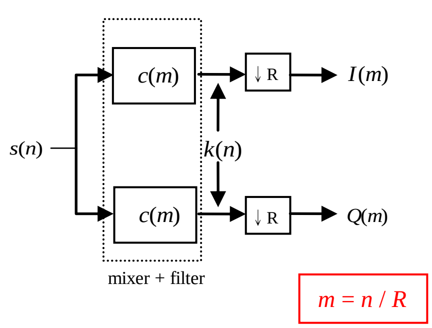


Figure 6. The block diagram of proposed *CMF*

V. *CMF* OPTIMIZATION AND IMPLEMENT IN ASSEMBLY

A. Choice of filter order

The architecture is implemented on a personal computer by software; the most important goal is fast enough to process the data. As we know, the XMM register is formed 128 bits. Hence, it could contain 16 signed byte data. In order to utilize the XMM registers efficiently, 7 orders FIR is chosen. In other words, there are 8 filter coefficients. The data type of designed filter coefficients is float. The look-up table coefficients to be integer which data type is byte data. The two groups of look-up table coefficients could be built in a XMM register. Thus, we achieve two linear convolutions in a XMM register. For example, a filter instance design for DDC in [10] is 9 orders (10 coefficients); a XMM register can only hold one group of filter coefficients. As a result, it can reduce the computational complexity.

B. Assembly implement

The parallel processing instructions is used in assembly and XMM registers. The assembly computation is mainly between the XMM registers. Besides, the new DDC calculation is based on the linear convolution in (5). Hence, the two Supplemental Streaming SIMD Extensions 3 (SSSE3) instructions are chosen to achieve PMADDUBSW and PHADDSW [11] which described as below.

PMADDUBSW: Multiply and Add Packed Signed and Unsigned Bytes.

PHADDSW: Packed Horizontal Add and Saturate Words.

These two instructions are shown in Figure 7 and Figure 8:

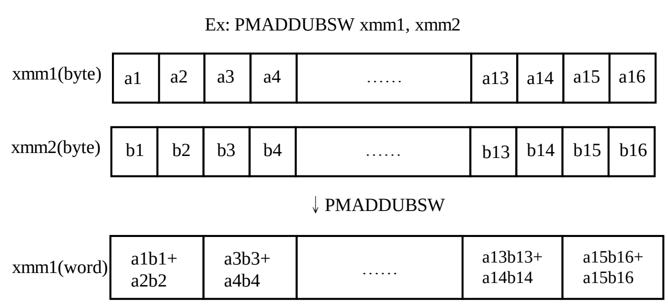


Figure 7. SIMD instruction: PMADDUBSW

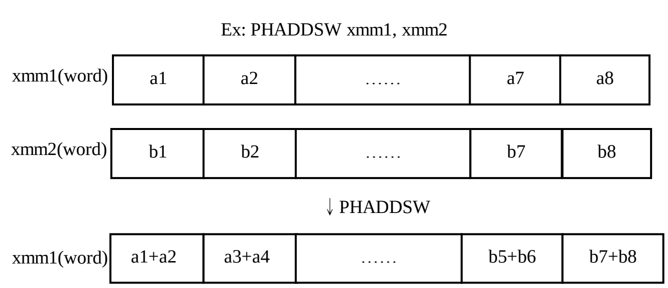


Figure 8. SIMD instruction: PHADDSW

The "PMADDUBSW" is used to multiply s (n) and c (m) in (5) together. The "PHADDSW" is used twice to sum the results in all. The data type of s (n) and c (m) in (5) is unsigned byte and signed byte, respectively. Because of the data type, only "PMADDUBSW" of SSSE3 instructions could achieve the multiplication between s (n) and c (m). The data type of "PMADDUBSW" outputs is signed word. Furthermore, the "PHADDSW" of SSSE3 instructions are used to add the outputs horizontally. The detail linear convolution example achieved by these two instructions is shown in Figure 9.

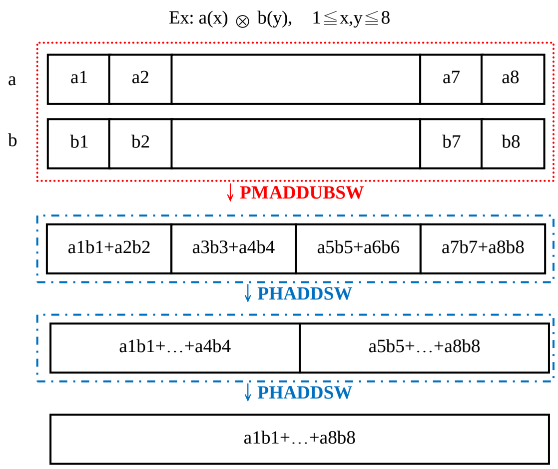


Figure 9. Linear convolution by "PMADDUBSW","PHADDSW"

VI. PERFORMANCE OF OUR DDC

The new DDC algorithm will be compared with the previous architecture. First, the DDC calculation is presented by Matlab. Then, we could know the improvement of *CMF* method. The quantity of input data is 84Mbytes; Table I is our hardware simulation environment.

Table III presents the elapsed time about f_{AD} = 192/7MHz use *CMF* or not. The mixer and filter coefficient is multiplied in advance; DDC algorithm could be simpler in (5). In our case, if the number of DDC input data is N, there will be N multiplications in (3). Moreover, our FIR filter is 7 orders, so there has 8N multiplications generated by linear convolution of (5). In fact, the down-sampling part usually combines with the FIR. The 8N multiplications will reduce to 2N because of down-sampling: 4 times. The number of multiplications in the previous DDC is 3N (=N+2N). The CMF combines mixer and filter, so the number of multiplications could be reduced to 2N additionally. Thus, the CMF can save much more elapsed time than the previous DDC algorithm. As a result, the new DDC algorithm can save about 40% elapsed time.

TABLE III. THE ELAPSED TIMES OF F_{AD} =192/7MHZ USE *CMF* OR NOT (IN MATLAB)

	Elapsed time(s)	Multiplication
f_{AD} =192/7MHz (No)	175.55	3N
f_{AD} =192/7MHz (Yes)	104.89	2N

The optimization result of proposed DDC algorithm is shown in Table IV. In order to use all 16 XMM registers, the Windows 7 (64 bits) is chosen as system OS. Besides, the Microsoft Visual Studio 2010(Team Suite edition) is used as development tool. It has the Performance Explorer to analyze the elapsed time of proposed DDC in assembly code and C code.

TABLE IV. THE ELAPSED TIMES OF F_{AD} =192/7MHZ (IN C CODE AND ASSEMBLY CODE)

Function name	Elapsed time(ms)
ddc (C)	88.07
ddc_asm (Assembly)	20.81

The DDC takes 20.81ms to decode 84Mbytes data. 84Mbytes/20.81ms = 4.04Gbytes/sec. According to f_{AD} = 192/7MHz, the real time DVB-T signal in Taiwan has 27Mbytes/sec. Hence, the proposed DDC only takes 0.67% CPU loading.

For the DVB-T software radio implement, the elapsed time of demodulation part is shown in Table V. The length of decoding data is 3 sec. However, the total elapsed time of demodulation part is 1541.89ms. Thus, the DVB-T real-time computation can be implemented (1.562sec < 3sec).

TABLE V. THE ELAPSED TIME OF DEMODULATION PART

Block	Elapsed Time (ms)
Time & Frequency Synchronize	63.66
Remove CP & FFT	67.82
Channel estimation	145.87
Deinner & Depuncher	54.49
Deoutter interleave	17.93
Demodulator	8.93
Viterbi decoder	1096.79
RS Decoder	30.67
Descrambler	0.92
Frame Synchronize	8.58
Program Initialization	27.52
Phase Compensation	8.44
C++ standard library	10.27
Total elapsed time	1541.89

VII. CONCLUSION

In this paper, the new DDC architecture is designed for f_{AD} = 192/7MHz and the algorithms are optimized by multiplying the

 ISBN: 978-1-61208-192-2

mixer and filter recorded in a look-up table in advance. The proposed *CMF* can save the additional *N* multiplications. Moreover, it is also optimized by assembly code. As a result, the decoding rate of the proposed system is greater than the required bit rate of the real-time DVB-T. The new system is fast enough to decode the DVB-T signal in real-time.

Finally, the new method can save more 40% time than the previous architecture. It only takes 20.81ms to decode 84Mbytes data. To sum up, the new DDC coding rate is 4.04Gbytes/sec.

ACKNOWLEDGMENT

The authors would like to thank all colleagues and students who contributed to this study.

REFERENCES

[1] Digital Video Broadcasting (DVB); Frame structure, channel coding and modulation for digital terrestrial television, European Standard (EN) 300 744 V1.5.1, European Telecommunications Standards Institute (ETSI), Nov. 2004.

[2] Mitchell, J. and Sadot, P., “Development of a digital terrestrial front end,” International Broadcasting Convention 12-16 Sept. 1997, IEE Conference Publication No. 447, pp. 519-524, ISSN 0537-9989.

[3] Makowitz, R. Anikhindi, S. Gledhill, J. Mayr, M. and Drozd, M., “A Single-chip Front End for DVB-T Receivers,” in Proc Consumer Electronics Con, Jun. 1998, pp. 390-391.

[4] F. Harris and R. W. Lowdermilk, “Software defined radio: Part 22 in a series of tutorials on instrumentation and measurement,” IEEE Instrum. Meas. Mag., vol. 13, no. 1, pp. 23–32, Feb. 2010

[5] Fang-Hsu Lu and You, S.D., “Channel Estimation for DVB-T Receiver with Asynchronous Sampling Clock,” IEEE 13th International Symposium on Consumer Electronics, May. 2009, pp.513-517.

[6] Yih-Min Chen, “On the Design of Farrow Interpolator for OFDM Receivers with Asynchronous IF Sampling,” Fourth International Conference on Communications and Networking in China, Aug. 2009, pp. 1-5.

[7] Karim Medhat Nasr, John P. Cosmas, Maurice Bard, and Jeff Gledhill, “Performance of an Echo Canceller and Channel Estimator for On-Channel Repeater in DVB-T/H Networks,” IEEE Trans. on Broadcasting, vol. 53, issue:3, pp. 609-618, Sep. 2007.

[8] Ji-yang Yu, and Yang Li, “An Efficient Digital Down Converter Architecture for Wide Band Radar Receiver,” in Proc 2009 IET International Radar Conf., Apr. 2009, pp. 1-4.

[9] M. J. Zhao, P. L. Qiu, and J. H. Tang, “Sampling rate conversion and symbol timing for OFDM software receiver,” IEEE 2002 International conference on Communications, Circuits and Systems and West Sino Expositions, vol. 1, pp. 114-118, May. 2002.

[10] Yih-Min Chen, and I-Yuan Kuo, “Design of Lowpass Filter for Digital Down Converter in OFDM Receivers,” International Conference on Wireless Networks, Comm. and Mobile Computing, vol. 2, Jun. 2005, pp. 1094-1099.

[11] “Intel® Advanced Vector Extensions Programming Reference,” http://software.intel.com/file/19151

[12] “Intel® 64 and IA-32 Architectures Software Developer's Manual Volume 1,” http://download.intel.com/design/processor/manuals/253665.pdf

[13] “Intel® 64 and IA-32 Architectures Software Developer's Manual Volume 2A,” http://download.intel.com/design/processor/manuals/253666.pdf

[14] “Intel® 64 and IA-32 Architectures Software Developer's Manual Volume 2B,” http://download.intel.com/design/processor/manuals/253667.pdf

 ISBN: 978-1-61208-192-2

An Unequal Error Protection Scheme for JPEG Image Transmission Using Enhanced Duo-Binary Turbo Codes

Tulsi Pawan Fowdur
Dept of Electrical and Electronic Engineering
University of Mauritius
Reduit, Mauritius
e-mail: p.fowdur@uom.ac.mu

Yashvin Beni
Dept of Electrical and Electronic Engineering
University of Mauritius
Reduit, Mauritius
e-mail: yashvinbeni@gmail.com

***Abstract*— JPEG is a widely deployed image compression standard used in several applications. However, JPEG image transmission is challenging and sophisticated strategies are required for reliable transmission. This paper investigates the performance of JPEG image transmission using duo-binary Turbo codes with Unequal Error Protection (UEP). UEP is achieved by applying a lower code-rate to protect the DC-layer of the image more efficiently and a higher code-rate for protecting the AC-layer. Additionally, the duo-binary Turbo code is enhanced by scaling its extrinsic information to improve performance and by using a stopping criterion to limit the number of iterations required for decoding. The proposed UEP scheme provides a gain of at least 10 dB in Peak Signal to Noise Ratio (PSNR) over an Equal Error Protection (EEP) scheme for a range of Eb/No values. Moreover, the gain in PSNR increases as the couple length of the duo-binary code is increased.**

Keywords- JPEG; UEP; Duo-Binary Turbo Codes.

I. INTRODUCTION

JPEG is a Discrete Cosine Transform (DCT) based image compression algorithm, which employs Huffman coding to generate a compressed bit-stream [1]. It is a widely adopted standard and forms an integral part of several applications such as web browsing and telemedicine [2]. However, the use of Huffman coding renders the JPEG coded bit-stream very sensitive to error propagation because a single bit in error can cause a complete loss of synchronisation. As such, sophisticated coding solutions are required to ensure reliable transmission. One solution is to use powerful error-correcting codes such as Turbo codes, which are well suited to protect image data as recently demonstrated in [3]. Error resilient and concealment techniques also provide a significant improvement in transmission fidelity [4,5]. Moreover, a highly efficient strategy for achieving robust JPEG image transmission is UEP. UEP consists of exploiting the fact that the DCT operation in JPEG, segments the image into layers of unequal importance. Hence, by allocating different levels of protection to these layers, a significant gain in the overall quality of the received image can be obtained.

Several efficient UEP schemes have been developed for JPEG image transmission using Turbo codes. For example, in [6], UEP and joint source channel decoding with a-priori statistics were combined and applied to JPEG image transmission. Both Turbo codes and Turbo and Turbo Trellis Coded Modulation were used and major gains in PSNR were obtained over conventional JPEG image transmission schemes. An error resilient wireless JPEG image transmission scheme, which employed product Turbo or Reed Solomon codes alongside an optimal UEP algorithm was proposed in [7]. Moreover, in [8], an UEP scheme, which employs s-random odd-even interleaving with odd-even puncturing, as well as a new UEP scheme for the soft output Viterbi algorithm, were proposed. Improved BER and PSNR performances in JPEG image transmission were obtained with these UEP schemes [8]. Finally, in [9], the performance of three UEP schemes for progressive JPEG image transmission using delay-constrained hybrid ARQ, with iterative bit and symbol combining was proposed. Gains of over 9 dB in PSNR were obtained with the UEP schemes as compared to their corresponding Equal Error Protection (EEP) schemes.

In contrast with previous works, which considered binary Turbo codes, this paper proposes an UEP scheme based on duo-binary Turbo codes. These codes provide better convergence of iterative decoding, have reduced latency, lower sensitivity to puncturing, larger minimum distance and lower memory requirement [10]. Also, the duo-binary code is modified with a scale factor [11,12] and stopping criterion [13] to further improve the performance of the UEP scheme. The proposed UEP scheme allocates more protection to the DC layer, which contains the most significant part the image after the DCT operation, and less protection to the AC layer. This is achieved by using the puncturing matrices specified for the duo-binary Turbo code of the DVB-RCS standard [14]. The UEP scheme outperforms the EEP scheme by at least 10 dB in PSNR over a range of Eb/No values. Moreover, the gain in PSNR increases as the couple length of the duo-binary code is increased.

The organization of this paper is as follows. Section II describes the complete system model. Section III presents the simulation results and analysis. Section IV concludes the paper.

II. SYSTEM MODEL

The complete encoding process is shown in Figure 1. The input image is fed to the JPEG encoder, which operates on blocks of 8x8 pixels and performs DCT, quantization and zig-zag ordering [1]. The AC and DC coefficients are then separated into the AC and DC layers. The DC layer regroups the first coefficient from all 8x8 blocks obtained after zig-zag ordering and the AC-layer is the concatenation of the 63 coefficients from all 8x8 blocks. For example with a 256x256 image, there are 1024 blocks of size 8x8 and each block has one DC coefficient and 63 AC coefficients. The DC layer hence contains 1024 coefficients and the AC layer contains 1024x63 coefficients. To prevent error propagation, the AC and DC layers are divided into blocks of 63 and 64 coefficients respectively. The blocks of the DC layer undergo Differential Pulse Code Modulation (DPCM) and DC-Huffman coding. Each block is encoded separately and after Huffman coding a header is inserted to indicate the size in bits of the resulting DC-packet. The blocks of the AC-layer undergo Run-Length Encoding (RLE) followed by AC-Huffman coding and a header is appended to indicate the size of each AC-packet. Each DC and AC packet can be decoded independently and errors within a packet do not propagate throughout the DC or AC layer.

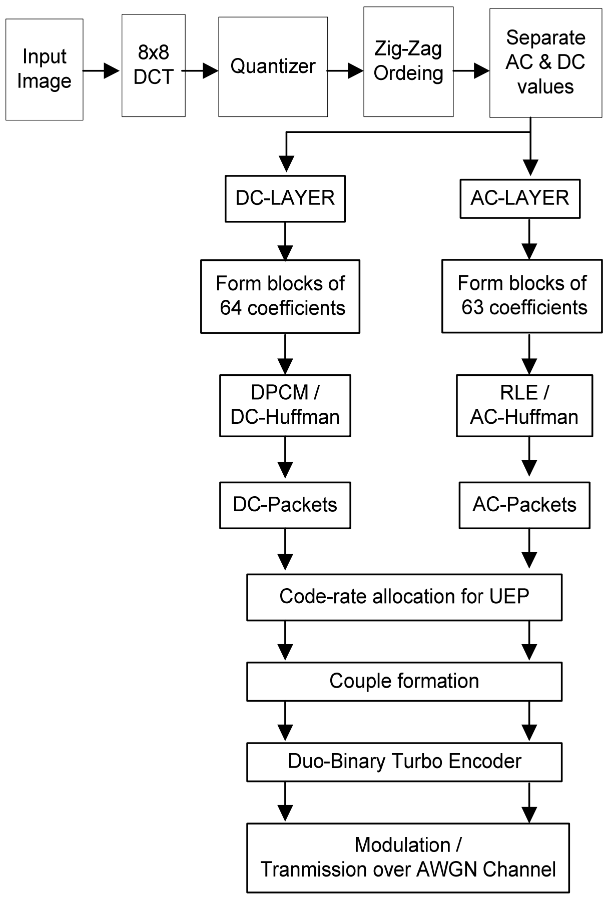


Figure 1. Complete encoding system with UEP.

A code-rate allocation is performed to provide UEP to the DC and AC packets. The DC packets are given the lowest code-rate while the AC-packets are allocated a higher code-rate. The packets are then converted into couples of length N before being sent to the duo-binary Turbo encoder. The headers are transmitted error-free through a side-channel. Finally, the encoded couples are QPSK modulated and sent over an Additive White Gaussian Noise (AWGN) channel.

The decoding system is shown in Figure 2. The received noisy array, R_t is de-punctured as appropriate and sent to the first duo-binary decoder, DEC1. DEC2 receives the interleaved counterpart of R_t. The decoders employ the Max-Log-MAP algorithm to compute the following parameters:

(a) $\overline{\gamma_t^{q(i)}}(l',l)$: The branch transition probability from state *l'* to *l* of symbol *i* at time instant *t* where *i* ϵ(0,1,2,3) for decoder q, where q = 1 or 2. It is computed as follows for the first decoder.

$$\overline{\gamma_t^{1(i)}}(l',l) = \log\left[p(u_t^2 = i).\exp\left(-\frac{[R_t - x_t^i(l)]^2}{2\sigma^2}\right)\right] = \log[p(u_t^2 = i)] - \left(\frac{[R_t - x_t^i(l)]^2}{2\sigma^2}\right) \tag{1}$$

where,

$p(u_t^2 = i)$ is the a-priori probability of symbol i obtained from the second decoder,

$x_t^i(l)$ is the modulated complex symbol at time *t*, associated with the transition from state $S_{t-1} = l'$ to $S_t = l$ and input symbol *i*,

$R_t(l)$ is the received systematic and parity complex symbols at time *t.*

For the second decoder, the computation of $\overline{\gamma_t^{2(i)}}(l',l)$ is similar to equation (1) except that it uses the interleaved version of the systematic symbols of R_t, a different set of complex parity symbols and the a-priori probability of symbol *i* obtained from the first decoder i.e., $p(u_t^1 = i)$.

(b) $\overline{\alpha_t^q}(l)$: The forward recursive variable at time *t* and state *l.* It is computed according to the following equation for a decoder with *Ms* states:

$$\overline{\alpha_t^q}(l) = \log\sum_{l'=0}^{Ms-1} e^{\overline{\alpha_{t-1}^q}(l') + \overline{\gamma_t^{q(i)}}(l',l)} \tag{2}$$

(c) $\overline{\beta_t^q}(l)$,which is the backward recursive computed at time t as follows:

$$\overline{\beta_t^q}(l) = \log \sum_{l'=0}^{Ms-1} e^{\overline{\beta_{t+1}^q}(l') + \overline{\gamma_{t+1}^{q(i)}}(l,l')} \qquad (3)$$

(d) $\Lambda_{q(i)}(t)$,which is the Log Likelihood Ratio (LLR) of symbol i where i ϵ(1,2,3), and the LLRs are normalized to the symbol '0'. This parameter is computed as follows:

$$\Lambda_{q(i)}(t) = \log \left[\frac{\sum_{l'=0}^{Ms-1} e^{\overline{\alpha_{t-1}^q}(l') + \overline{\gamma_t^{q(i)}}(l',l) + \overline{\beta_t^q}(l)}}{\sum_{l'=0}^{Ms-1} e^{\overline{\alpha_{t-1}^q}(l') + \overline{\gamma_t^{q(0)}}(l',l) + \overline{\beta_t^q}(l)}} \right] \qquad (4)$$

(e) $\Lambda_{1es(i)}(t)$ and $\Lambda_{2es(i)}(t)$: The extrinsic information of symbol i where i ϵ (1,2,3), are generated by DEC1 and DEC2 respectively. They are computed as follows:

$$\Lambda_{1es(i)}(t) = \Lambda_{1(i)}(t) - \overline{\Lambda}_{2es(i)}(t) - \Lambda_{1in(i)}(t) \qquad (5)$$

$$\Lambda_{2es(i)}(t) = \Lambda_{2(i)}(t) - \overline{\Lambda}_{1es(i)}(t) - \Lambda_{2in(i)}(t) \qquad (6)$$

$\Lambda_{1in(i)}(t)$ and $\Lambda_{2in(i)}(t)$ are the intrinsic information of symbol i where i ϵ(1,2,3), are generated by DEC1 and DEC2 respectively.

Further details on the computation of these parameters are given in [14,15,16]. In the enhanced Duo-binary decoder, the extrinsic information produced by both decoders are multiplied by a scale factor S as shown in Figure 2. The application of the scale factor improves performance because the extrinsic information value output by the Turbo decoder is most of the time too optimistic, hence by scaling it, better performance is achieved [11-12]. The controller unit accepts the extrinsic information from both decoders and uses a stopping criterion [13] to stop the iterative decoding process. At the start of the iterative decoding process, switches S1 and S2 are ON and when a given condition is met, the controller unit turns OFF both switches to stop the iterative decoding process. In this way, the decoder avoids the use of extra iterations and reduces the decoding complexity. This technique also reduces the power consumption of the decoder.

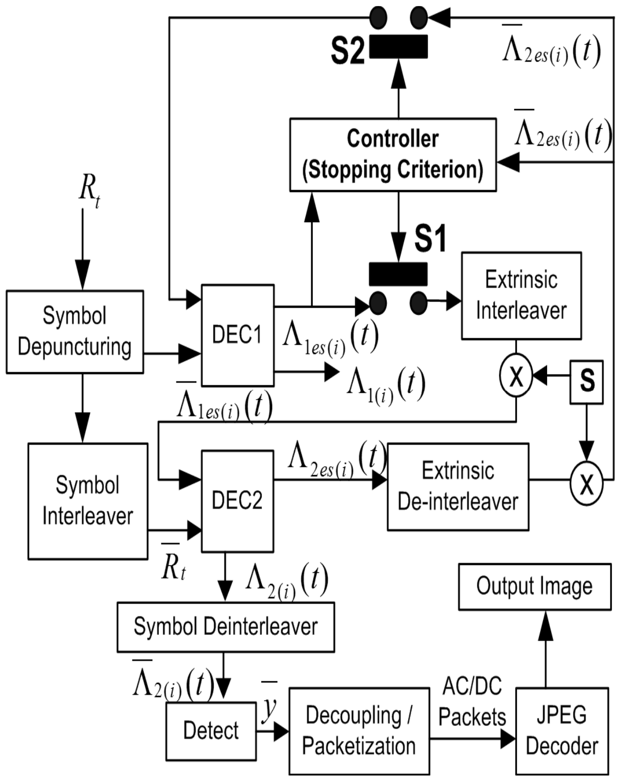


Figure 2. Decoding system with enhanced duo-binary Decoder.

A detailed algorithm for the decoding process is now presented. In this algorithm, steps 4-12 correspond to the operations of DEC1 and steps 13-21 of DEC2. The parameters $M_{11}^r(t), M_{12}^r(t), M_{21}^r(t), M_{22}^r(t)$ are used in the stopping criterion and the function ***f( )*** counts the number of sign changes between the two arguments that are passed to it. The function ***detect( )*** determines the maximum of the LLR values and outputs either symbol 0,1,2 or 3. The variable j increases by 1 because the decoder processes one couple at a time up to a maximum of N_c, which is the total number of couples in the image. The variable and r also increases by 1 up to a maximum limit of r_{max}. However, the variable num_iterations, which is used to count the number of iterations consumed by the decoder, is incremented by 0.5. This is because the stopping criterion can stop the decoding process after either DEC1 or DEC2 whereby each decoder consumes 0.5 iterations. For example, if for a given couple, the decoding process completes 2 full iterations and then at the third iteration i.e., r=3, after passing through DEC1, the stopping criterion is satisfied, then only 0.5 additional iteration is consumed and hence num_iterations will be 2.5 and not 3. The complete decoding algorithm is as follows:

 ISBN: 978-1-61208-192-2

1. num_iterations = 0

2. for j = 1:N_c

3. for r = 1:r_{max}

4. Compute: $\overline{\gamma_t^{1(i)}}(l',l)$, $\overline{\alpha_t^1}(l)$, $\overline{\beta_t^1}(l)$, $\Lambda_{1es(i)}(t)$, $\Lambda_{1(i)}(t)$
5. num_iterations = num_iterations + 0.5.

6. $M_{11}^r(t) = \max(\Lambda_{1es(2)}(t), \Lambda_{1es(3)}(t)) - \max(\Lambda_{1es(0)}(t), \Lambda_{1es(1)}(t))$
7. $M_{12}^r(t) = \max(\Lambda_{1es(1)}(t), \Lambda_{1es(3)}(t)) - \max(\Lambda_{1es(0)}(t), \Lambda_{1es(2)}(t))$

8. if (r>1)
9. if ($f(M_{11}^r(t), M_{11}^{r-1}(t)) \leq \frac{1}{N}$ or $f(M_{12}^r(t), M_{12}^{r-1}(t)) \leq \frac{1}{N}$)
10. Goto: line 23
11. endif

12. endif

13. Compute: $\overline{\gamma_t^{2(i)}}(l',l)$, $\overline{\alpha_t^2}(l)$, $\overline{\beta_t^2}(l)$, $\Lambda_{2es(i)}(t)$, $\Lambda_{2(i)}(t)$
14. num_iterations = num_iterations + 0.5.

15. $M_{21}^r(t) = \max(\Lambda_{2es(2)}(t), \Lambda_{2es(3)}(t)) - \max(\Lambda_{2es(0)}(t), \Lambda_{2es(1)}(t))$
16. $M_{22}^r(t) = \max(\Lambda_{2es(1)}(t), \Lambda_{2es(3)}(t)) - \max(\Lambda_{2es(0)}(t), \Lambda_{2es(2)}(t))$

17. if (r>1)

18. if ($f(M_{21}^r(t), M_{21}^{r-1}(t)) \leq \frac{1}{N}$ or $f(M_{22}^r(t), M_{22}^{r-1}(t)) \leq \frac{1}{N}$)
19. Goto: line 23
20. endif

21. endif

22. Endfor

23. Decoded couple, $\overline{y} = \text{detect}(\overline{\Lambda}_{2(i)}(t)))$

24. Endfor

25. Convert the received couples into AC and DC packets.
26. Perform JPEG decoding on the received packets.

III. SIMULATION RESULTS AND ANALYSIS

The performances of the following four schemes for JPEG image transmission are compared:

Scheme 1- UEP with scale factor: This scheme employs UEP to provide different levels of protection to the AC and DC packets of the image. It also uses a scale factor, S, to enhance the performance of the duo-binary Turbo code by scaling the extrinsic information, as depicted in Fig.2. The value of S has been set to 0.75 in this simulation.

Scheme 2 - UEP without scale factor: This scheme is similarly to Scheme 1 but the extrinsic information is not scaled and the value of S is set to 1.0 in Fig.2.

Scheme 3 - EEP with scale factor: It is similar to Scheme 1 but equal protection is given to the AC and DC packets.

Scheme 4 - EEP without scale factor: This scheme is similar to Scheme 3 but the scale factor, S, is set to 1.0.

In all simulations, the DVB-RCS standard duo-binary Turbo code [14] has been used with a stopping criterion. The generator polynomials in octal are g1 = 15 for the feedback branch, g2 = 13 and g3 = 11 for the parity branches. Couple lengths of N=64 and N=212 are used and the maximum number of iterations, r_{max} = 12. Puncturing matrices are chosen as per the DVB-RCS standard and the 256x256 Lena image is used as input. Moreover, it is assumed that the headers are transmitted error free over a strongly protected side channel.

The overall coding rate, O_c, was limited to $O_c < 0.97$ bits/pixel and to ensure a fair comparison the overall coding rate for UEP was kept below that of EEP. However, with UEP the DC couples are more strongly protected with a code-rate of 1/3 while the AC couples are allocated a code-rate of 4/5. On the other hand the EEP schemes allocates a fixed code-rate of 2/3 to both DC and AC couples. The overall coding rate, O_c, is computed as follows:

$$O_c = \frac{1}{T}\left(\frac{T_{DC}}{R_{DC}} + \frac{T_{AC}}{R_{AC}}\right) \qquad (7)$$

where,

T is total number of pixels in the image,
T_{DC} is the total number of bits in the DC couples,
R_{DC} is the code-rate allocated to the DC couples,
T_{AC} is the total number of bits in the DC couples,
R_{AC} is the code-rate allocated to the DC couples.

The source coding rate, S_c, and O_c, vary with the couple length because different numbers of padding bits are required to convert the bit stream from the JPEG encoder into couples of length N.

Table I gives the values of O_c and S_c for different couple lengths, N.

TABLE I
CODING RATES FOR DIFFERENT VALUES OF N

N	T_{DC}	T_{AC}	S_C	O_C	
				UEP	EEP
64	5760	36096	0.639	0.952	0.958
212	5936	36040	0.641	0.959	0.961

Figure 3 shows the graph of PSNR versus Eb/No for the four schemes with N = 64. The UEP scheme with scale factor provides a gain of 7dB in PSNR over the EEP schemes at Eb/No = 1.5dB and a major gain of 12dB in PSNR in the range 2dB ≤ Eb/No ≤ 3dB. It also outperforms the UEP scheme without scale factor by 1dB in PSNR in the range 2dB ≤ Eb/No ≤ 3dB. This gain is achieved because with the UEP the DC layer is recovered with fewer errors than the AC-layer and hence, the image can be reconstructed with much less distortions. However, it is observed that at high Eb/No values, the gain obtained with UEP over EEP decreases because the overall errors introduced in the image is considerably less.

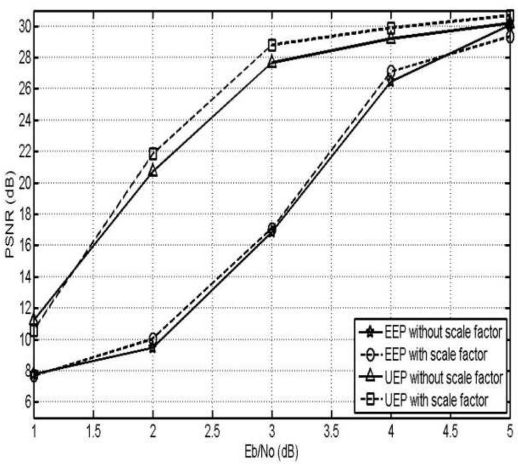


Figure 3. Graph of PSNR against Eb/No for N = 64.

The graph of number of iterations versus Eb/No for N = 64 is shown in Figure 4. The stopping criterion allows the number of iterations and hence the decoding complexity to decrease progressively as the Eb/No is increased. Interestingly, the UEP scheme with scale factor requires less iterations than the EEP schemes in the range 1dB ≤ Eb/No ≤ 3dB and provides an impressive reduction of 5.5 iterations over the EEP scheme without scale factor at Eb/No = 1dB.

Figure 5 shows the graph of PSNR versus Eb/No for the four schemes with N = 212. The UEP scheme with scale factor provides a gain of 10dB in PSNR over the EEP schemes at Eb/No = 1dB and a major gain of 15 dB in PSNR in the range 1.5 dB ≤ Eb/No ≤ 2dB. It also outperforms the UEP scheme without scale factor by about 1dB in PSNR in the range 1dB ≤ Eb/No ≤ 2dB. Moreover, with N = 212 the UEP scheme with scale factor outperforms the UEP scheme with scale factor for N = 64, by an average of 5 dB in PSNR. The gain is greater with a couple length of N=212 because the performance of the duo-binary Turbo code improves with increase in couple length.

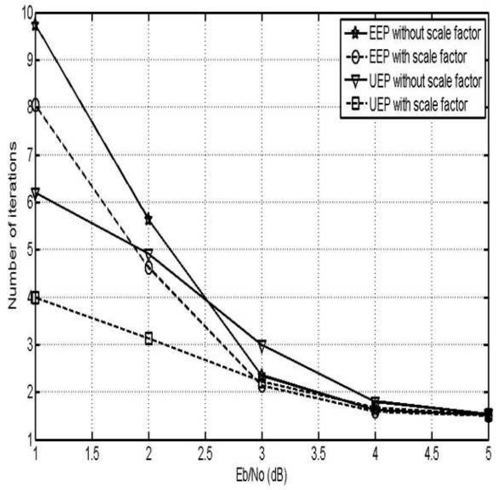


Figure 4. Number of iterations against Eb/No for N = 64.

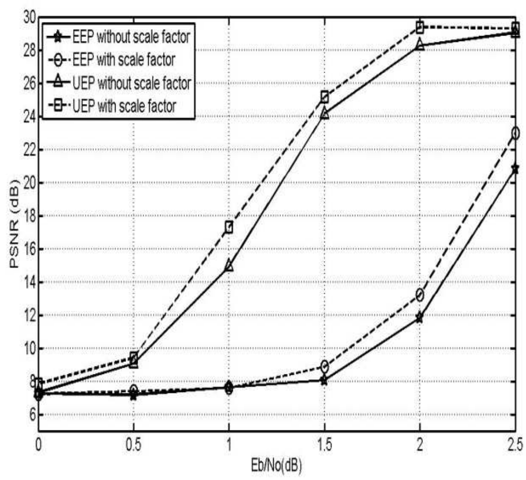


Figure 5. Graph of PSNR against Eb/No for N = 212.

The graph of number of iterations versus Eb/No for N = 212 is shown in Figure 6. It is observed that when N=212, the UEP scheme with scale factor takes less iterations than the EEP schemes only in the range 0dB ≤ Eb/No ≤ 1dB. For Eb/No > 1.5 dB the EEP scheme with scale factor requires significantly less iterations than the UEP schemes, for example, at Eb/No = 2.5dB it requires 5 iterations less than the UEP scheme without scale factor. A possible explanation for that is that the threshold used in the stopping criterion was not optimized for N = 212 and was maintained at 1/N.

There are two ways in which the UEP scheme can lead to an increase in complexity with respect to the EEP scheme. Firstly, over a certain Eb/No range, as observed in Figure 5, the UEP scheme requires more iterations than the EEP scheme. Secondly, with the UEP scheme, the duo-binary

Turbo encoder must treat the AC and DC packets separately and use different code-rates, hence different puncturing patterns are required for each of them. The same applies for the duo-binary Turbo decoder, whereby a different de-puncturing process must be used for the AC and DC packets.

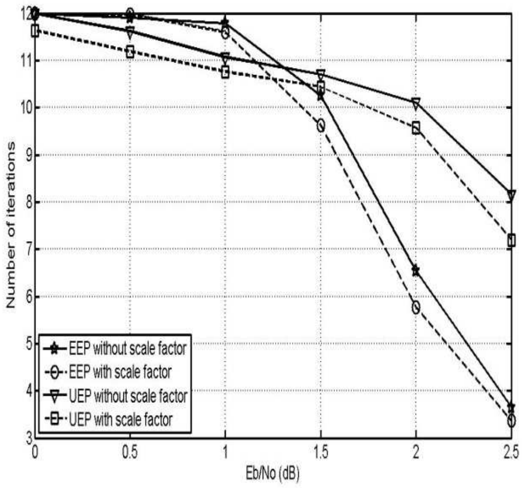


Figure 6. Number of iterations against Eb/No for N = 212.

IV. CONCLUSION

This paper proposed an efficient UEP scheme for JPEG image transmission with enhanced duo-binary Turbo codes whereby an extrinsic scale factor and stopping criterion were incorporated. The performances of four schemes were compared with couple lengths of 64 and 212. The results showed that major gains of the order of 10 dB in PSNR are obtained with the UEP scheme over conventional EEP schemes and this gain increases when the couple length is increased from 64 to 212. Furthermore, the use of the scale factor improved the PSNR performance and reduced the number of iterations required, hence the complexity. Interestingly, with a couple length of 64, the UEP scheme required fewer iterations than the EEP schemes. However, with a couple length of 212, at higher Eb/No values, the EEP schemes required less iterations. An interesting future work would be to optimize the threshold used in the stopping criterion for couple lengths greater than 64, so as to reduce the number of iterations required by the UEP schemes.

ACKNOWLEDGMENT

The authors would like to thank the University of Mauritius for providing the necessary facilities for conducting this research and the reviewers for their valuable comments.

REFERENCES

[1] R.C. Gonzalez, R.E Woods, and S.L Eddins, Digital Image Processing in Matlab, 2nd ed., Gatesmark Publishing, (2009).

[2] L.G. Yamamoto, “Using JPEG image transmission to facilitate telemedicine,” The American Journal of Medicine, vol. 13, pp. 55-57, 2009.

[3] Ersin GOSE, “Adaptive Wiener-turbo systems with JPEG & bit plane compressions in image transmission,” Turk J Elec Eng & Comp Sci, vol. 19, no.1, 2011. (doi:10.3906/elk-1003-461).

[4] L.W. Kang and J.J. Leou, “An error resilient coding scheme for JPEG image transmission based on data embedding and side-match vector quantization,” J. Vis. Commun. Image, R.17, pp. 876-891, 2006.

[5] K.S. Kim, H.Y. Lee, and H.K. Lee, “Supplementary loss concealment technique for image transmission through data hiding,” Multimed Tools Appl, vol. 44, pp. 1–16, 2009. (DOI 10.1007/s11042-009-0265-0).

[6] T.P. Fowdur and K.M.S. Soyjaudah, “Hybrid Joint Source Channel Decoding for Progressive JPEG Image Transmission,” Circuits Systems and Signal Processing, vol. 29, pp. 857-879, 2010. (DOI 10.1007/s00034-010-9188-2).

[7] M. Padmaja, M. Kondaiah and K.S. RamaKrishna, “Error Resilient Image Transmission over Wireless Fading Channels,” International Journal of Engineering Science and Technology,vol. 2(5), pp. 1242-1249, 2010.

[8] A.M. Lakhdar, R.Méliani, and M. Kandouci, “Research on Unequal Error Protection with Punctured Turbo Codes in JPEG Image Transmission System,” SJEE,vol.4, no.1, pp. 95-108, June 2007.

[9] T.P Fowdur and K.M.S Soyjaudah, “Highly Scalable JPEG Image Transmission with Unequal Error Protection and Optimal Feedback,” Telecommunication Systems, vol.49, Issue 4, pp. 355-377, 2012. (DOI: 10.1007/s11235-010-9379-y). (Published online: 25 June 2010)

[10] C.Douillard and C.Berrou,"Turbo codes with Rate-m/(m+1) Constituent Convolutional Codes," in IEEE Trans. Commun, pp. 1630-1638, October 2005.

[11] J.Vogt and A.Finger, “Improving the MAX-Log-MAP Turbo decoder,” Electronics letters, vol. 36, no. 23, November 2000.

[12] T.Gnanasekaran and K. Duraiswamy, “Performance of Unequal Error Protection Using MAP Algorithm and Modified MAP in AWGN and Fading Channel,” Journal of Computer Science, vol. 4 (7), pp. 585-590,2008.

[13] B.S. Shim, D.H. Jeong, and S.J. Lim, "A new stopping criterion for turbo codes," in Adv Communication Technology, Ihksan, pp. 1111-1116, 2006.

[14] M.R. Soleymani, Y.Gao and U.Vilaipornsawai,Turbo coding for satellite and Wireless Communication, Kluwer Academic Publishers, 2002.

[15] Y. Ould-Cheikh-Mouhamedou, S. Crozier and P. Kabal, “Distance measurement method for double binary turbo codes and a new interleaver design for DVB-RCS,” Global Telecommunications Conference 2004, (GLOBECOM '04), vol. 1, pp. 172 – 178,2004 .

[16] Youssouf Ould Cheikh Mouhamedou, On Distance Measurement Methods for Turbo Codes, McGill University, Montreal, PhD Thesis 2005. Available Online: http://www-mmsp.ece.mcgill.ca/MMSP/Theses/2005/Ould-Cheikh-MouhamedouT2005.pdf [Accessed: August 2011]

A Hierarchical Routing Algorithm for Small World Wireless Networks

Juhani Latvakoski
VTT Technical Research Centre of Finland
Oulu, Finland
Juhani.Latvakoski@vtt.fi

***Abstract*—Embedded devices are tomorrow working in dynamic wireless networks, which requires novel solutions for routing because of heterogeneity, long communication paths, long delays and weak performance. As a contribution of this research, a hierarchical routing algorithm called as, *Hi-Search*, is provided. The algorithm relies on the wireless short-cut based solution for small world wireless networks and hierarchical neighbor discovery. The provided hierarchical search algorithm is based on graph theoretical system model and network search tree analysis both on overlay and physical levels. The efficiency of the *Hi-Search* algorithm is analytically evaluated in terms of search path depths, number of control messages, and delay of the search, which are compared against the flat physical routing approach. The evaluation indicates that the search path depths for the *Hi-Search* algorithm are lower than the search path depths for the end to end physical routes. In addition, the number of control message send actions and search delays are lower compared with physical routing.**

Keywords-dynamic wireless networks; small world; routing.

I. INTRODUCTION

The number of wirelessly communicating embedded devices has continuously been increasing in recent years. It is expected that the networks consisting of such devices will be highly dynamic wireless networks. Therefore, it is here assumed that dynamic self-configurable (ad hoc) routing solutions are required. The ad hoc networking protocols, such as, e.g., Topology Dissemination Based on Reverse-Path Forwarding (TBRPF) [1], Ad hoc On-Demand Distance Vector (AODV) [2], and Dynamic MANET On-demand (DYMO) [3] are not optimal for specific operating environments due to differences in delay requirements, reaction times for route changes, the power capabilities of the routing devices, and the limitations of the bandwidth usage, quality of service level and security. A possible solution approach for these challenges is overlay networking [4, 5]. Delay-Tolerant Networks (DTN) and opportunistic networking [6, 7, 8] solutions enable communication also when the source and destination nodes not necessarily reachable at the time of communication need. However, the *heterogeneity* of nodes, radio links, and dynamic topologies still triggers challenges for them.

The selected approach in this research for solving the heterogeneity problem is application small world paradigm for wireless networks [9, 10, and 11]. By adding a few short-cut links, path length of wireless networks can be reduced drastically. Our previously published simulations have indicated that increasing the number of referred wireless short-cuts lowers the end to end delays, and makes the physical routes shorter, and also improves throughput [12]. In addition, the key enabling mechanisms for neighbor discovery are provided in [13] to solve the elementary challenges related to multilevel routing and creation of wireless short-cuts. Relying on these solutions, we have provided here, a hierarchical search algorithm, called as *Hi-Search*. The provided hierarchical search algorithm is based on graph theoretical system model and network search tree analysis both on overlay and physical level. The efficiency of the *Hi-Search* algorithm is evaluated in terms of search path depths, number of control messages, and delay of the search, which are compared against the flat physical routing approach. In physical routing, the neighboring nodes have always only direct physical communication links between each other, and in overlay routing there can also be logical communication links between logically neighboring nodes, which are not necessarily physically adjacent.

Related research is shortly analyzed in the following. The hierarchical routing schemes with distributed hash tables (DHT) are discussed in [14]. The essential difference compared with our solution is that DHT based hierarchical routing schemes do not take into consideration of physical route discovery at all. Small world based routing, called as SWER, dedicated to supporting sink mobility and small transfers have been provided in [15]. The hierarchy is based on clustering and cluster heads, and short cuts are applied for long range links between clusters. The cluster head selects a sensor node to act as agent node to form the short-cut. The difference compared with our solution, is that the weak sensor nodes and radio links are still applied in realizing the short-cut. Strategies for adding long-ranged links to centrally placed gateway node in wireless mesh networks are provided in [16]. Based on the strategies, 43% of reduction in average path length was reported. Hierarchical routing based on clustering using adaptive routing using clusters (ARC) protocol is provided in [17]. They represent a new algorithm for cluster leader revocation to eliminate the ripple effect caused by leadership changes. Other examples of related routing solutions are a contact-based architecture for resource discovery [18], Mobile router nodes used as data mules [19], and P2P-based SWOP protocol [20]. There are also quite a many solutions for neighbor discovery such as [21, 22]. However, route discovery is usually executed in flat manner, e.g., [23]. The problem in such a search is that the search queries are forwarded also into the deep leafs of the

 ISBN: 978-1-61208-192-2

search trees. Our approach is different in the sense that only the nearest logical overlay nodes are searched in the physical route level, and the network can be optimized by removing non-optimal radio links and physical routers from the path acting as a wireless short-cut.

The system mode, reasoning and the hierarchical search algorithm is described in section II. The efficiency of the hierarchical search is evaluated in section III. Finally, in section IV, conclusions are provided.

II. Hierarchical Search Algorithm

A. *Hierarchical System Model*

The network consists of different types of nodes such as *U* node, which is a user interface (UI) node, which is able to host the network, set of services and it may visualize the services for a user. *S* node is service node, which may provide set of services, act as super peer (cluster head) for services and overlay router. *R* node is a physical router node, which can route data traffic between different interfaces of the node. T node can be a sensor (*Ts*), actuator (*Ta*) or camera (*Tc*). Each of the referred nodes may not be always on and they may be mobile and can apply whatever wireless/wired access means for communication with the neighbor nodes. An example system is shown in the form of graph in Figure 1. In the example physical network graph (G_{PN}), vertice ($V_{PN}=0$) i.e. node (0) represent the User node. Each vertise has certain characteristics such as location (*L*), overlay routing capabilities (*OR*), physical routing capabilities (*PR*), radio capabilities (*R*), power capabilities (*P*) and computing power (*Cp*), $V_{PN}.\{L, OR, PR, R, P, Cp\}$. The edges ($E_{PN}$) represent the possible physical communication links between two or more nodes. Each edge has certain characteristics such as, e.g., distance (*D*) and delay (Δt), $E_{PN}.\{D, \Delta t\}$. In the example the overlay network graph (G_{ON}) are established by the U, and S vertices (V_{ON}). The dotted red lines represent the edges of the overlay network (E_{ON}). The overlay network graph is here said to be *virtual* graph of the physical network graph (G_{ON} c G_{PN}). Respectively, we can define radio network graph (G_{RN}) which shows the radio network below the physical network (G_{ON} c G_{PN} c G_{RN}). Therefore, the system model is here said to be hierarchical.

The G_{PN} can be represented in the form of a (search) tree (T_{PN}) shown in Figure 2. Such tree does not have cycles and the source of the search is represented as the root of the tree (T_{PN} $(V_{PN}=0)$). A search path is a route from the root of the tree to the leaf of the tree representing the destination of the search.

Respectively, G_{ON} can be represented in the form of a tree (T_{ON}) shown in Figure 3. It is easy to see that the height of the overlay network tree is smaller than the height of the physical network tree. It means that the overlay network path from source to the destination usually contains smaller amount of hops.

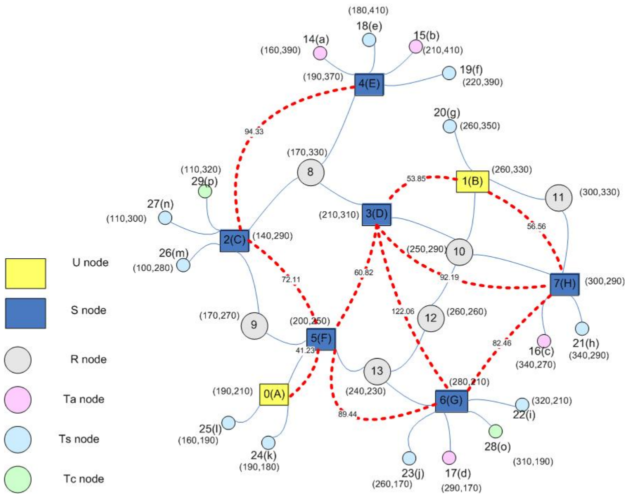


Figure 1. Example System Graph (G_{PN} and G_{ON}).

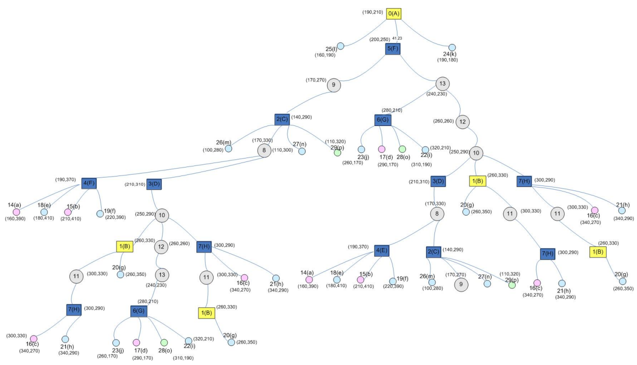


Figure 2. Example System physical network Tree (T_{PN}).

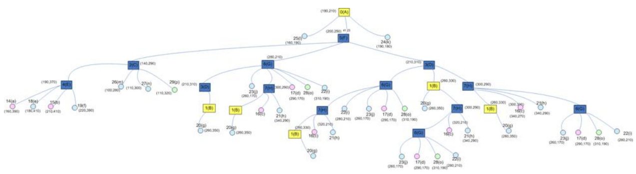


Figure 3. Example System Overlay Network Tree (T_{ON}).

B. *Reasoning*

The reasoning of the hierarchical search algorithm is represented in the following:

- Each edge in the search path means additional communication delay for the search. Therefore, the amount of levels in the search tree needs to be minimized. For example, if the search proceed into deep sub trees, which does not have the destination, the search unnecessarily disturbs the vertices and consumes the radio bandwidth in the area of the leaf sub tree.
- Each vertice in the search tree processes the search, and it adds processing delay (Δt_p) for the search. Therefore,

the amount of vertices in the search path needs to be minimized. It can be claimed that the search unnecessarily disturb all vertices in the search path, if the node is not the destination. Unnecessary disturbing any node should be minimized.

- Let us call the minimization of the search tree levels, minimization of the amount of vertices in the search path and minimization of vertice disturbing as *search tree minimization.*
- The number of levels in the search tree is lower for the T_{ON} than for T_{PN}. Therefore, it is assumed that the search tree can be minimized by relying on hierarchical search, in which the search is executed in the overlay level (T_{ON}) and the physical level search is limited to the discovery of the physical paths between each pair of neighboring S nodes (T_{PN} is split into sub trees). This means that the hierarchical search is executed in T_{ON} (Figure 3.) and in the split sub trees of T_{PN} visualized in Figure 4. In this way, the physical level search results a local optimum physical path between neighboring S/U nodes, and the overlay level search results optimum path between source and destination (S/U or T_* nodes).
- Some of the vertices are more powerful than the others, e.g., they can have good power sources and good computing platform while the others may be battery operated. It is clear that powerful vertices are better nodes for routing. Therefore, they are preferable nodes in the search path, and the usage of limited capability nodes (bottlenecks) shall be minimized.
- When looking at different search paths in G_{PN} , T_{PN} it is assumed that removing the bottleneck nodes from the search path lowers the total communication delay (Δt_c) of the search. Let's call here the removal process as *network optimization*.
- The network optimization process is focused into the split sub trees of T_{PN}, see Figure 4. Because, the R nodes are assumed to be the bottleneck nodes, the S/U nodes actively tries to remove them from the local physical communication paths, and create a shortcut between the neighboring S/U nodes. As a result from the successful network optimization, the search tree is like T_{ON} visualized in Figure 3.
- Summarizing, the hierarchical search with search tree minimization and network optimization processes results a situation, where the search path consist only powerful and well-connected S/U nodes and not any bottleneck nodes

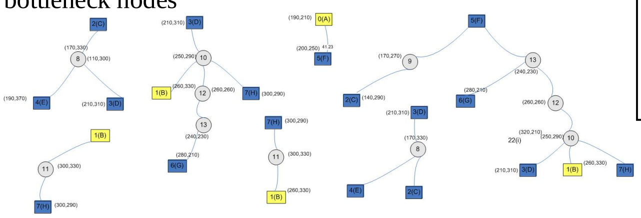


Figure 4. Split sub trees of T_{PN}.

C. Hierarchical Search Algorithm

The hierarchical search algorithm is represented in the Figure 5. When the power of a U/S node is switched on, the device will broadcast *DiscoverReq* to all of its neighbors with the information of the node itself. Each overlay node receiving the *DiscoverReq*, stores the key contents of it and replies with *DiscoverRep*, which is sent in unicast manner to the source of the DiscoverReq. The *DiscoverRep* contains overlay level routing and service information, which will be delivered to the original source of the *DiscoverReq*. Sending of the *DiscoveRep* triggers searching for the physical route between the neighboring overlay nodes, e.g., using AODV RREQ/RREP procedure. After the original source receives *DiscoverRep* via the discovered physical route, the system is ready to provide messaging services for applications.

When an *ApplicationMsg* is received from the upper layer and the overlay route is known, it is forwarded towards the intended destination. Otherwise, an overlay route is searched first, and then the message is forwarded towards it.

```
Algorithm HI-Search                         /* Hierarchical Search */

1. WHEN n(OFF) →n(ON) THEN
2.          send (DiscoveryReq, Bcast)
3. WAIT until receive (Msg)
4.          SWITCH Msg
4.                    CASE DiscoverRep (ucast)
5.                              store (DiscoverRep)
6.                              start (timer, Net-Opt)
7.                    CASE DiscoverReq (Bcast)
8.                              IF n == ON THEN
9.                                store (DiscoverReq)
10.                               send (DiscoverRep, Ucast)
11.                             ELSE
12.                               update (DiscoverReq)
13.                               forward (DiscoverReq, Bcast)
14.                   CASE applicationMsg
15.                             IF no route THEN
                                  send (ON-RouteReq)
16.                             IF route THEN send
applicationMsg
17.                   CASE ON-RouteReq
18.                             IF n == destination THEN
19.                               send ON-RouteRep
20.                             ELSE forward ON-RouteReq
21.                   CASE Timeout (Net-Opt)
22.                             optimize (network)
23.                             start (timer, Net-Opt)
24.         ENDSWITCH
25. ENDWAIT
```

Figure 5. Hierarchical Search Algorithm.

In this manner, the ApplicationMsg has been delivered to the destination using hierarchical search with minimized search tree. At any time after system is ready, the network optimization procedure can be initiated. In the network optimization, direct wireless communication links for the neighboring overlay nodes are created as communication short-cuts in the cases where it is physically possible with the available radio access technologies of the overlay nodeswhich mean lower search paths.

III. EFFICIENCY OF THE HIERARCHICAL SEARCH

The depths of the search paths, for the example graph shown in Figure 1. are visualized in Figure 6. There are 37 possible search paths for both physical and overlay networks, see Figure 2. and Figure 3. respectively. Each search path is shown in the x-axis, and the depth of the search path is shown in the y-axis in the Figure 6. For example, for the search path number 11, the depth of physical search path is 10 and the depth of the overlay search path is 5. In general, the search path depths for the overlay routes are lower than the search path depths for the end to end physical routes. The provided *Hi-Search* algorithm applies overlay route search, which mean lower search paths.

The physical search path depths of overlay hops are shown in Figure 7. (see also Figure 4.). The y-axis shows the physical search path depths, and x-axis shows the number of their required searches in Figure 1. in physical routing situation. For example, the physical search path 5-9-2, which depth is 2, happens 17 times in physical routing situation. The referred physical search paths seem generally happen multiple times in the example network in physical routing situation. This is not very efficient, and therefore the proposal is that the network optimization creates shortcuts between the neighboring overlay vertices. Then there is need to execute referred physical search paths only once for the network, and optimization can be based on it. The referred action is initiated in the row 22 of the *Hi-Search* algorithm.

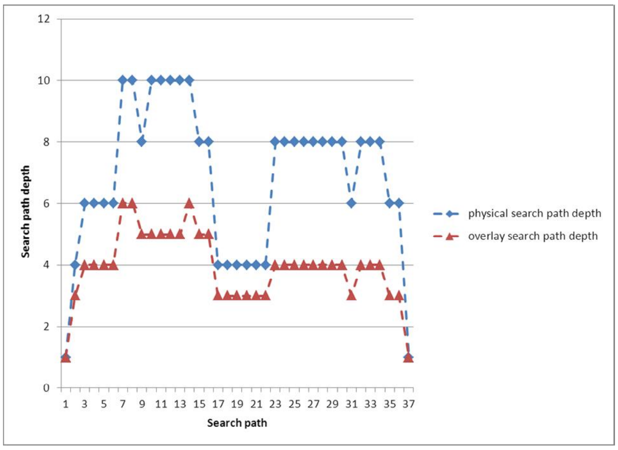


Figure 6. Search path depths.

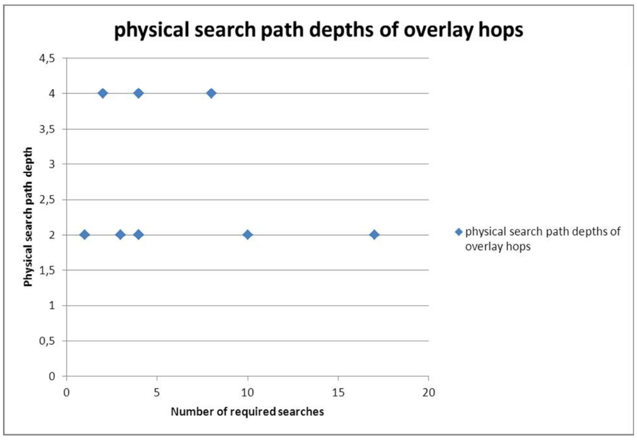


Figure 7. Physical search path depths of overlay hops.

The number of control message sending actions is shown in y-axis of Figure 8. When the physical route between neighbouring overlay nodes are searched initially and optimised, the number of control message send actions are about the same level as in physical routing (x=7). However, after the optimisation has been executed, then the number of control message send actions lowers significantly, because there isn't need to repeat optimisation. It can be seen that the number of control message send actions is lower when applying *Hi-Search* algorithm compared with physical routing.

The total delay of the search is shown in y-axis of Figure 9. It is here assumed that the delay in each physical hop, i.e. radio link, is 10ms, the optimization happen in parallel manner and the processing delay in each node is zero. The peaks of the delay for the *Hi-Search* algorithm are related to optimization of the network. After the optimization, the delays are in lower level. As a result, it is seen that the *Hi-Search* algorithm is better because it has lower search delays than physical routing.

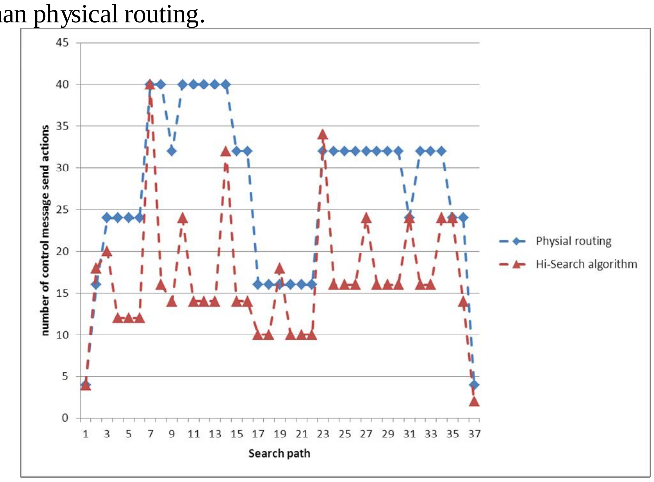


Figure 8. The number of control message sending.

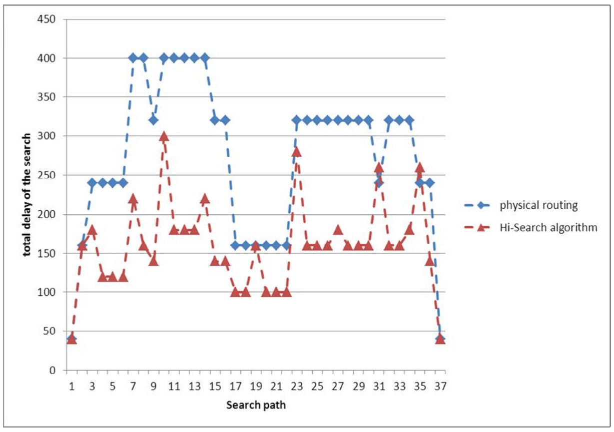


Figure 9. Delay of the search.

In practical situations, the physical characteristics including the delay in each edge vary according to applied radio access technology. The network optimization removes weak and high delay edges from the path, which may make the delay difference between physical search and *Hi-Search* algorithm even bigger than what is shown in Figure 9. In addition, the processing delay of each vertise is usually bigger than zero. When applying *Hi-Search* algorithm, the number of intermediate hops in the path is minimized in such a manner that weak vertices are removed from the path. Therefore, in practical situation the delay difference between physical search and *Hi-Search* algorithm is even bigger than what is shown in Figure 9.

IV. CONCLUSIONS

As a contribution of this research, the hierarchical routing algorithm called as, *Hi-Search*, is provided. The algorithm relies on the wireless short-cut based solution for small world wireless networks and for the key enablers for hierarchical neighbor and route discovery, network optimization and service discovery. The provided hierarchical search algorithm is based on graph theoretical system model and network search tree analysis both on overlay and physical level. The efficiency of the *Hi-Search* algorithm is evaluated in terms of search path depths, number of control messages, and delay of the search, which are compared against the flat physical routing approach.

The evaluation indicates that the search path depths for the *Hi-Search* algorithm are lower than the search path depths for the end to end physical routes. The physical routes between logically neighbouring vertices are searched only once, and then the network optimization creates shortcuts between them. It can be seen that the number of control message send actions is lower when applying *Hi-Search* algorithm compared with physical routing. In addition, it has lower search delays than physical routing. When the physical characteristics including the radio specific delays of each edge, and processing delay of each vertice in the path is taken into concern, then the delay difference between *Hi-Search* and physical routing is even bigger. This is because; network optimization removes high delay edges, and weak vertices from the path.

ACKNOWLEDGMENT

Acknowledgements are offered for Tekes and VTT for funding this work.

REFERENCES

[1] Orier R., Templin F., and Lewis M. 2004. Topology Dissemination based on reverse-Path Forwarding (TBRF). IETF RFC 3684. Feb 2004.

[2] Perkins C., Royer-Belding E., and Das S. RFC 3561. Ad hoc On-Demand Distance Vector Routing. IETF. Jul 2003.

[3] Ian D. Chakeres and Charles E. Perkins Dynamic MANET On-Demand Routing Protocol. IETF Internet Draft, draft-ietf-manet-dymo-12.txt, February 2008 (Work in Progress). Retrieved 2, 2012.

[4] Latvakoski, J. and Aapaoja, T. Towards a Routing Overlay for a Mobile Ad hoc Network. 6p. First International Workshop on Convergence of Heterogeneous Wireless Networks (ConWiN) 10th Jul 2005 Budapest, Hungary. 8p.

[5] Latvakoski J., Aapaoja T., and Kärnä J. Evaluation of routing overlay solution for a Hybrid Mobile ad hoc networks. 12p. ERCIM Emobilty workshop. 28-30 May 2008 Tampere, Finland. 12p.

[6] Luciana Pelusi, Andrea Passarella, and Marco Conti. (2006). "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad hoc Networks." IEEE Communications Magazine Nov 2006: Pp. 134-141.

[7] V.Cerf et al. (2007) Delay-Tolerant Networking Architecture. IETF RFC 4838. Apr 2007.

[8] Latvakoski, J., and Hautakoski, T. Situated Message Delivery for Opportunistic Networks. 9p. ICT Mobile and Wireless Communications Summit 2008. 10-12 Jun 2008. Stockholm/Sweden. 8p.

[9] Milgram S. The small world problem. Psychol. Today 2, pp 60-67. 1967

[10] Helmy, A. Small Worlds in Wireless Networks. IEEE Communications Letters, Vol 7. No 10. October 2003. Pp 490-492.

[11] Watts D., and Strogatz S. Collective Dynamics of small world networks. Nature Vol 393. Pp 440-442. 1998.

[12] Latvakoski, J. Towards hierarchical routing in small world wireless networks. The Fifth International Conference on Wireless and Mobile Communications ICWMC 2009. August 23-29, 2009 - Cannes/La Bocca, French Riviera, France. 6p.

[13] Latvakoski, J. Hierarchical Routing Concept for Small World Wireless Networks. 6 p. ICWMC 2010: The Sixth International Conference on Wireless and Mobile Communications September 20-25, 2010 - Valencia, Spain. 6p.

[14] Korzun, D. and Gurtov, A. 2011, "Survey on hierarchical routing schemes in "flat" distributed hash tables", Peer-to-Peer Networking and Applications, vol. 4, no. 4, pp. 346-375

[15] Liu, X., Guan, J., Bai, G. & Lu and H. 2009, "SWER: small world-based efficient routing for wireless sensor networks with mobile sinks", FRONTIERS OF COMPUTER SCIENCE IN CHINA, vol. 3, no. 3, pp. 427-434.

[16] Verma, C.K., Tamma, B.R., Manoj, B.S. and Rao, R. 2011, "A Realistic Small-World Model for Wireless Mesh Networks", *IEEE Communications Letters*, vol. 15, no. 4, pp. 455-457

[17] Belding-Royer, E.M. 2002, "Hierarchical routing in ad hoc mobile networks", Wireless Communications and Mobile Computing, vol. 2, no. 5, pp. 515-532.

[18] Helmy A., Garg S., and Nahata N. CARD: A contact-based Architecture for Resource Discovery in Wireless Ad hoc Networks.

Mobile networks and applications 10, pp. 99-113. 2005. Springler-Verlag.

[19] Jiang C-J, Chen C., Chang J-W., Jan R-H., and Chiang T. C.. Construct Small Worlds in Wireless Networks using Data Mules. Pp 28-35. Proceedings of IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing. Pp. 28-35.

[20] Ken Y.K. Hui, John C.S. Lui, and David K.Y. Yau. Small world overlay P2P Networks. Quality of Service, 2004. IWQOS 2004. Twelfth IEEE International Workshop on 7-9 June 2004. Pp. 201 – 210

[21] Narten T., Nordmark E., and Simpson W. Neighbor Discovery for IP version 6. IETF RFC 2461. http://www.ietf.org/rfc/rfc2461.txt. Available 10th Jun 2010. Retrieved 2, 2012

[22] Clausen T., Dearlove C., and Dean J. MANET Neighborhood Discovery Protocol (NHDP), http://tools.ietf.org/html/draft-ietf-manet-nhdp-04, Expires December 31, 2007. Retrieved 2,2012.

[23] Perkins C., Royer-Belding E., and Das S. RFC 3561. Ad hoc On-Demand Distance Vector Routing. IETF. Jul 2003. http://tools.ietf.org/html/rfc3561, Retrieved 2, 2012.

Problem of IMS modeling – Solving Approaches

Mesud Hadžialić, Mirko Škrbić, Nerma Šečić, Mirza Varatanović, Elvedina Zulić, Nedim Bijedić
University of Sarajevo
e-mail: mhadzialic@etf.unsa.ba, mskrbic@etf.unsa.ba, nsecic@etf.unsa.ba, mvaratanovic@etf.unsa.ba,
elvedina.zulic@bstelecom.ba, nedim.bijedic@bstelecom.ba

Abstract **– IMS (IP Multimedia Subsystem) network has high demands from perspective of multimedia, flexible and interactive communications. Fulfillment of those demands, with appropriate levels of quality, is not a simple task. The primary objective of this paper is to set emphasis on the proper dimensioning of IMS network, and the need to find a methodology applicable in planning of IMS networks, that will be able to provide quantitative results on basis of the initial request. In addition, the goal is to provide insight to the basic problems of IMS network, and ways of solving them. This primarily refers to SIP (Session Initiation Protocol) servers, which are the key part of the IMS structure, which have to deal with overload due to improper dimensioning of the network. This paper will attempt to present systematically the existing solutions and to provide guidance for the future resolution of this problem.**

Keywords-IMS; modeling; overload; SIP.

I. INTRODUCTION

With the basic idea of integrating different networks into one multifunctional IP - based network, IMS (IP Multimedia Subsystem) aims to create a unified communication environment for fixed and mobile users, by offering enriched and integrated services that were fragmented before. IMS needs to offer a high level of interaction for users - enriched calling and enhanced messaging - sharing videos, images, and other multimedia during a voice call (during one communication session), with high level of service personalization. However, this level of communication flexibility significantly complicates the management, and among other things directly influences the increase of the number of signaling messages that must be exchanged and processed.

SIP (Session Initiation Protocol) represents the core of IMS architecture. SIP servers are the main components responsible for processing and routing all signaling messages. The problem that imposes itself is the problem of dimensioning IMS network, i.e., dimensioning IMS components, which should be able to process large amounts of messages, in order to avoid server overload and therefore degradation of QoS (Quality of Service) and QoE (Quality of Experience).

SIP provides limited built-in overload control mechanism - the 503 Service Unavailable response. However, as the price of rejecting *SIP* sessions typically cannot be ignored, this mechanism cannot prevent server's congestions. When a SIP server rejects a large amount of incoming sessions its performance degrades and, additionally, the impact of overload increases through the network - this is a key observation that distinguishes SIP server overload problem from other overload problems [8]. In order to eliminate or at least partially reduce the problems of SIP server overload, various approaches have been proposed in accordance with the opinion of what is the dominant problem that leads to uncontrolled overload.

On the other hand, if the network is not initially properly dimensioned, the application of congestion control will not be enough. In short time the nodes will fall into a state of congestion, which will lead to performance degradation of the entire network, and at some point, entire network will come to an outage and collapse. Thus, the network dimensioning must be understood as the primary problem, and a congestion control as an additional factor that can improve efficiency of well dimensioned network. Network modeling or a search for a loyal representation which reflects the behavior of IMS nodes and provides required output value in dependency on the input parameters, is a key of good methodologies for planning and dimensioning IMS network.

This paper systematically presents key problems of IMS and provides a review of previous results that deal with overload problem; also provides a review of current achievements in the field of modeling and behavioral analysis of IMS network in order to optimize the same. During the process of result analysis some disadvantages were observed, and every author provided a unique guideline for solving this complex problem. The second section provides an overview of papers dealing with overload problem, while third section gives an overview of previous work in field of modeling, behavioral analysis of the IMS network in order to optimize the same. In last section, we provide guidelines for solving this complex problem.

II. IDENTIFIED CONTRIBUTIONS AND EXPERIENCES IN SOLVING THE PROBLEM OF SIP OVERLOAD

Numerous papers deal with overload problems on SIP servers. Some of the papers include a detailed analysis of possible overload causes, while others contain

 ISBN: 978-1-61208-192-2

suggestions and explanations of the different mechanisms and algorithms, which should have the ability to manage and control overload.

Guided by representative papers for the stated problem, authors observed and exposed few potential aspects of SIP overload classification.

- One of the observed aspects involves creation of new protocol or changes on existing protocol. According to existing achievements this aspect could be the most complete solution but IMS, SIP servers and SIP protocol are so widespread in commercial use that there is no sense to try to change main standards in that area.
  For example, Whitehead [9] 2005 described the framework independent of the protocol, GOCAP, but his mapping in SIP has not yet been defined. Even if this framework becomes mapped; questionable is if it will be accepted by manufacturers of equipment.
- Another approach implies the use of new network elements or applications which would predict congestions or overloads. This means additional investments in HW (Hardware) and SW (Software); it demands additional time, resources and efforts on existing applications to send some performance indicators.
  Luca Monacelli [11] describes the overload problems and offers stabilization system, STBZ (STaBiliZer), which protects all network elements of IMS. STBZ is a software application which collect measurements from network, processes it by appropriate stabilization policies and controls traffic shapers in order to avoid congestion on network elements
- One of most commonly used approaches is the creation of new algorithms on existing network elements.
  This approach directly influences source code of network elements, but upgrades and patches are standardized processes and are something that network operators often do which makes this solution acceptable.
  C. Shen and H. Schulzrinne [10] are among the first to deal with overload of TCP - based SIP server. They suggest new mechanisms for SIP overload control - which relies on the existing TCP flow control and congestion control. The algorithm consists of three components: Connection Split, Buffer Minimization and Smart Forwarding. Due to the specific nature of SIP protocol (session based) there is a need for separation of INVITE messages processing - requests which start a session, and other non-INVITE messages, in order to prevent opening of new sessions that could lead to overload and that, on the other hand, will preserve the existing sessions. This part of the algorithm is called Connection Split, which allows that INVITE and non-INVITE messages are treated differently. Smart Forwarding is enforced on an INVITE connection. When an INVITE message arrives, decision about the forwarding of INVITE request is made in relation to the current state of buffer. This way a session that should not be established can be canceled as soon as possible. The paper further examines the impact of buffer size (buffer at the transmitter and the receiver side and applications buffer at the receiver side) to bandwidth and processing time. It was concluded that the best results can be achieved through minimization of the buffer size on the receiving side, which is the third part of the algorithm. This algorithm shows very good results for classical SIP scenarios in core networks, where small number of transmitting servers simultaneously creates overload on the receiving server. But in the case of edge networks where the overload is prevalent, the described algorithm does not provide satisfactory results.
- Whatever solution is used to prevent overload it will not provide satisfactory results if network is not optimized and well dimensioned.
  There is no one disadvantage of this approach and this network optimization step must be applied on any professional network (of any kind).

We concluded that solving SIP overload problem requires combined use of exposed aspects. We propose combination of new algorithm on existing network elements and optimization and well dimensioning of network.

Guidelines stated by C. Shen and H. Schulzrinne [10] will provide a start point in further discussion of SIP overload problems.

III. IDENTIFIED CONTRIBUTIONS AND EXPERIENCES IN SOLVING THE PROBLEMS OF IMS NETWORK OPTIMIZATION

As stated earlier, this chapter provides a review of the previous works in the field of modeling and behavioral analysis of IMS network, in order to provide the guideline for optimization of the same.

The problem identified in IMS networks is the existence of bottleneck nodes at different network layers of IMS architecture. The authors used a variety of scientific methods to investigate this issue of which the largest contributions were provided by analysis and modeling methods.

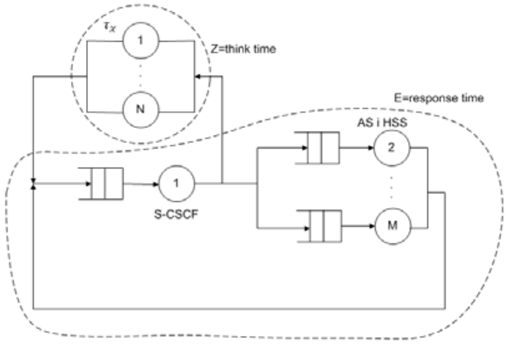


Figure 1. Model of central server for IMS [2]

Mkwawa and Kouvatsos [2], using the modeling method, have proven that exactly S-CSCF (Serving Call Session Control Function) is the throat of the IMS network that processes great number of messages.

They have compared the analytical and experimental results of IMS application modeling, user registration and the establishment of multimedia sessions, and have showed that the registration of users and establishment of multimedia session corresponds to a well-know central server model QNM (Queuing Network Model) Fig. 1.

Using Buzen's algorithm [13] and Little's theorem [12], service rates for S-CSCF, AS (Autonomous System) and HSS (Home Subscriber Server) were calculated, as well as transition probabilities, server utilization and throughput of the model.

Simulation results for server utilization and throughput for the same service speed and probability transitions were proven to coincide with the analytical model Fig. 2. [2] represents one of the ways for modeling the IMS system. This representation only indicates the network issues, but does not provide any kind of solution.

Because of the bottleneck problems additional load on S-CSCF module is not recommended, so overload problems on service layer and dynamic interaction of services are solved using Service Brokers, or SCIM (Service Capability Interaction Management) [4] module Fig. 3.

This module is not an integrated part of S-CSCF, instead it is being realized as an application server. Organization in this way allows better application server utilization in order to shift overload boundaries on the application layer. This opens new issues in the service interaction management field.

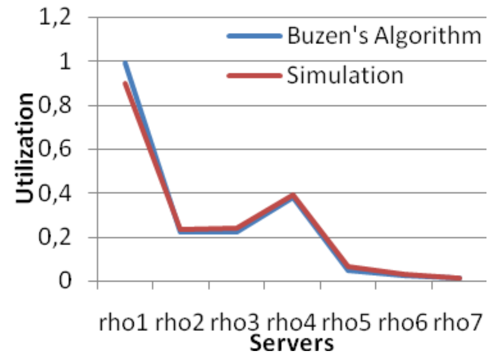


Figure 2. Comparative study of simulations and analytical results of the central server model [2]

The best overview of the results achieved by using modeling techniques are presented in [6] [7]. Both papers indicate the effects of adding servers to the networks assuming M/M/r model with Erlang C formula. It is shown in Fig. 4 and Fig. 5 that, by adding servers to the network, total waiting time in system is reduced, which has direct impact on system overload.

Two approaches were used for the simulation of M/M/r model: hyper-threading and physical server adding. It is shown that prioritizing calls affect waiting time. Fig. 6 and Fig. 7 show that with the increase of the prioritizing calls from 25 % to 50%, waiting time for all calls is increased; so, it is necessary to define a threshold value of the amount of priority calls, which will truly give better QoS and increase the performance.

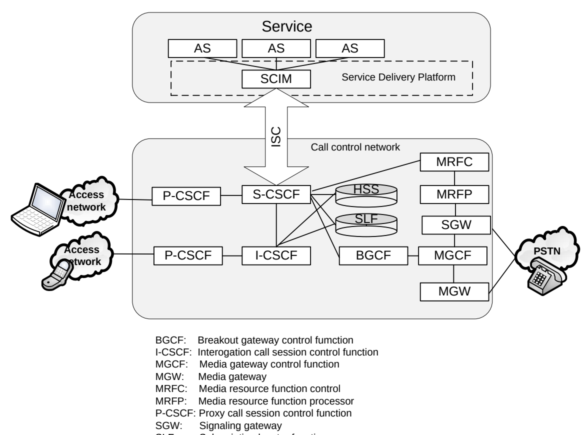


Figure 3. IMS architecture with SCIM [5]

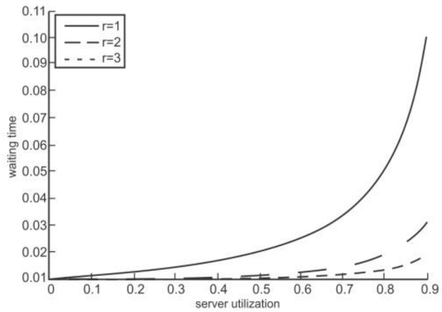


Figure 4. Mean waiting time in multi-server M/M/r system (μ=100) [6]

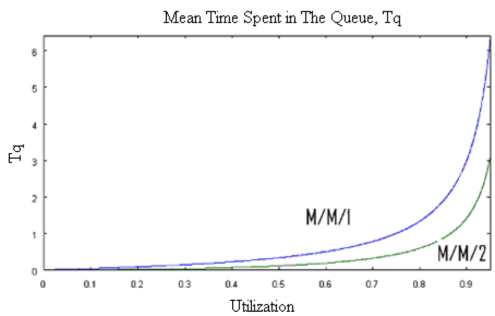


Figure 5. Mean waiting time in queue (Tq) [7]

Analysis was made with several assumptions; CSCF (Call Session Control Function) servers are replaced with one that has unlimited queue, while SIP requests and requested serving time have exponential distributions which considerably simplifies the process computation.

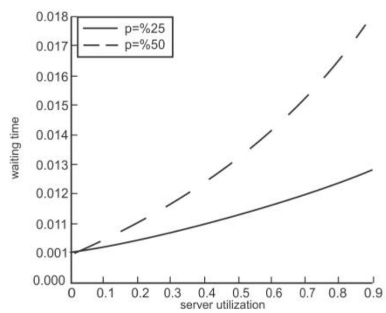


Figure 6. Comparison of prioritizing calls waiting time for two different amount of prioritizing packets [6]

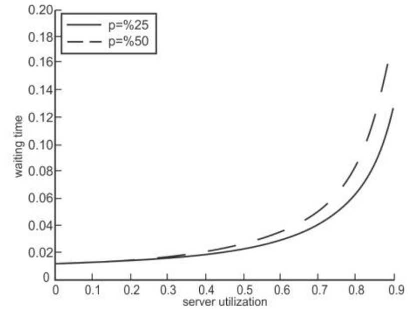


Figure 7. Comparison of non-prioritizing calls waiting time for two different amount of prioritizing packets [6]

IV. PROPOSAL FOR SOLVING OVERLOAD ISSUES AND IMS NETWORK OPTIMIZATION

Analyzing the papers in the domain of SIP servers overload and IMS network optimization can be concluded that there is no unified position on what mechanism successfully prevents overload or which traffic model gives the accurate picture of the actual IMS architecture.

Analyses that have been done so far include certain assumptions, whereby the obtained results do not provide the complete picture of the observed problems.

The complete resolution of defined problems is much more complex and must be considered in environment that is adequate for real systems. This environment assumes the use of multiple SIP servers, limited waiting lines and service times which usually do not have exponential character.

Our proposal for solving the problem of overload and optimization of IMS network is as follows:

- The prioritization of NON-INVITE over INVITE messages on S-CSCF node. This would prevent the accumulation of new sessions, which leads to efficient management and cleaning of active sessions.
- Message prioritizing implies separation of the incoming flow into two queues on S-CSCF. Service time distribution of processing an INVITE message will no longer have exponential character.
- It is necessary to experimentally determine distribution of service time of an INVITE message in case mentioned above.
- By the determination of service time distribution, preconditions are made for calculation of analytical dependencies between parameters which directly affect overload: queue length for INVITE messages and service time.
- Following step is the analytical and experimental analysis with systems that have more physically separated S-CSCF servers with previously mentioned methods applied. This analysis will provide results which should show dependency between system load and number of servers.

- Additional optimization requires application layer modeling which implies classification of applications, e.g., by real-time, non-real-time characteristics. For each group of applications, traffic management model should be defined with theoretically and experimentally determination of service requests distributions, in order to get minimal response time for every group of services. Request priority and the corresponding service prices will have important role in traffic management model. One of possible approaches is the use of multiple conditional optimal paths.

V. CONCLUSION

In order to ensure that IMS concept has the opportunity to be used in real conditions and to provide enriched services with the promising QoS, it is necessary to continue research and to make new progress towards solving these complex issues. The need to work on modeling is still persistent, and it is needed to pursue the faithful “behavior“ mapping of the IMS architecture into a model which will be able to provide analytical dependence of the output parameters with the input ones.
Only well-modeled systems can provide valid results on which can be based further mechanisms of load control, and then the steps that will contribute to the optimization of the overall architecture.

The paper gives an overview of all current achievements and provides a guideline for future work. All of the proposed methods should be the subject of the future research in order to solve the defined problem.

REFERENCES

[1] 3GPP TS 23.517, “IP Multimedia Subsystem; Functional architecture“, 2008., retrieved: December, 2011.

[2] I. M. Mkwawa and D.D. Kouvatsos, “Performance Modelling and Evaluation of IP Multimedia Subsystems“, HET-NETs08, pp. 67-79, February 2008.

[3] 3GPP TS 23.002, “Network Architecture”, 2009., retrieved: December, 2011.

[4] Nicholas S. Huslak and A.C. McQuaide Jr., “Service Brokering: Opportunities and Challenges“, AT&T Knowledge Ventures, 2007., retrieved: December, 2011.

[5] Kenichi Sakura, Soichiro Tange and Hisayuki Sekine, “Service Delivery Platform Implmenting IP Multimedia Subsystem“, FUJITSU Sci. Tech, Vol. 45, No. 4, pp. 409-414, October 2009.

[6] A.M. Amooee and A. Falahati, “Overcoming Overload in IMS by Employment of Multiserver Nodes and Priority Queues“, 2009 International Conference on Signal Processing Systems, pp. 348-352, May 2009.

[7] Mlindi Mashologu, “Performance Optimization of IP Multimedia Subsystem“, Dissertation.com, 2010., retrieved: December, 2011.

[8] Rosenberg, J., “Requirements for Management of Overload in the Session Initiation Protocol“, IETF (SIPPING, Internet Draft), Decembar 2008., retrieved: December, 2011.

[9] M. Whitehead, “GOCAP - one standardized overload control for next generation networks”, BT Technology Journal, Vol. 23, Issue 1, pp. 147-153, January 2005.

[10] C. Shen and H. Schulzrinne, “On TCP - based SIP server overload control“, IPTComm '10 Principles, Systems and Applications of IP Telecommunications, pp. 71-83, August 2010.

[11] Luca Monacelli, “Including Overload Control in Existing IMS Compilant Networks by Using Traffic Shapers“, Mobimedia '09 Proceedings of the 5th International ICST Mobile Multimedia Communications Conference, Article No. 56, September 2009.

[12] John D.C. Little, “A Proof of the Queuing Formula: L=λW", Operations Research, Vol. 9, No. 3, pp. 383-387, May-June 1961.

[13] Jeffrey P. Buzen, “Computational algorithms for closed queueing networks with exponential servers”, Communications of the ACM, Vol. 16, No. 9, pp. 527-531, September 1973.

 ISBN: 978-1-61208-192-2